Aula número 2 (08/03)

Definição. *Um subconjunto $U \subset \mathbb{R}^n$ é chamado* aberto *se para todo $x \in U$ existe $r > 0$ tal que $\mathrm{B}(x;r) \subset U$.*

`Notação`: para $x \in \mathbb{R}^n$ e $r > 0$ denotamos por $\mathrm{B}(x;r)$ a *bola aberta* de centro x e raio r definida por:

$$\mathrm{B}(x;r) = \big\{ y \in \mathbb{R}^n : \|x - y\| < r \big\}$$

e por $\mathrm{B}[x;r]$ a *bola fechada* de centro x e raio r definida por:

$$\mathrm{B}[x;r] = \big\{ y \in \mathbb{R}^n : \|x - y\| \leq r \big\}.$$

Por $\| \cdot \|$ denotamos a *norma Euclideana*: $\|x\| = \Big(\sum_{i=1}^n x_i^2 \Big)^{\frac{1}{2}}$.

Definição. *Seja $U \subset \mathbb{R}^m$ um aberto e considere uma função $f : U \to \mathbb{R}^n$. Dizemos que f é* diferenciável *num ponto $x_0 \in U$ se existe uma transformação linear $T : \mathbb{R}^m \to \mathbb{R}^n$ de modo que a aplicação r definida por:*

$$f(x_0 + h) = f(x_0) + T(h) + r(h),$$

satisfaz a condição $\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$.

De modo abreviado: f é diferenciável em $x_0 \in U$ se existe $T : \mathbb{R}^m \to \mathbb{R}^n$ linear tal que:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - T(h)}{\|h\|} = 0.$$

(1) **O caso $\mathbf{n = m = 1}$.**

Recordar que as transformações lineares $T : \mathbb{R} \to \mathbb{R}$ são todas da forma $T(h) = Lh$, onde L é um número real.

Teorema. *Uma função $f : U \subset \mathbb{R} \to \mathbb{R}$ é diferenciável num ponto $x_0 \in U$ se e somente se f é derivável no ponto x_0 (no sentido do Cálculo I), i.e., se existe o limite:*

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Demonstração. Observe que f é diferenciável no ponto x_0 se e somente se existe $L \in \mathbb{R}$ tal que:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - Lh}{h} \, \frac{h}{|h|} = 0;$$

notando que a quantidade $\frac{h}{|h|} = \pm 1$ é limitada, vemos que a condição acima equivale a:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0) - Lh}{h} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} - L = 0.$$

Isso completa a demonstração. ■
`Observação`: note que a transformação linear T que aparece na definição de função diferenciável é $T(h) = f'(x_0)h$ nesse caso.
`Exemplo`: Considerando a função $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$, e o ponto $x_0 = 4$ então escrevemos:

$$f(x_0 + h) = (4 + h)^3 = 64 + 48h + 12h^2 + h^3 = f(x_0) + T(h) + r(h),$$

onde escolhemos $T(h) = 48h$ (que é linear!) e daí $r(h) = 12h^2 + h^3$. Observe que de fato $\lim_{h\to 0} \frac{r(h)}{|h|} = \lim_{h\to 0} \frac{r(h)}{h} = 0$.

(2) **Relacionando diferenciação com derivadas direcionais e derivadas parciais**.

Definição. *Dada uma função $f : U \to \mathbb{R}^n$ definida num aberto $U \subset \mathbb{R}^m$, um ponto $x_0 \in U$ e um vetor $v \in \mathbb{R}^m$ então a* derivada direcional de f no ponto x_0 e na direção v *é definida por (se existir):*

$$\frac{\partial f}{\partial v}(x_0) = \lim_{t\to 0} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

Se e_i denota o i-ésimo vetor da base canônica *de $\mathbb{R}^m$ então a derivada direcional $\frac{\partial f}{\partial e_i}(x_0)$ é denotada também por $\frac{\partial f}{\partial x_i}(x_0)$ e é chamada a* i-ésima derivada parcial *de f no ponto x_0.*

Suponha que $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável no ponto $x_0 \in U$, de modo que existe $T : \mathbb{R}^m \to \mathbb{R}^n$ linear tal que:

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0) - T(h)}{\|h\|} = 0.$$

Fixe $v \in \mathbb{R}^m$ não nulo; vamos fazer a mudança de variável $h = tv$ no limite acima, onde $t \to 0$:

$$\lim_{t\to 0} \frac{f(x_0 + tv) - f(x_0) - T(tv)}{\|tv\|} = 0.$$

Obtemos então:

$$\lim_{t\to 0} \frac{1}{\|v\|} \left[\frac{f(x_0 + tv) - f(x_0) - T(tv)}{t}\right] \frac{t}{|t|} = 0;$$

pela linearidade de T, $T(tv) = tT(v)$ e portanto:

$$\lim_{t\to 0} \frac{1}{\|v\|} \left[\frac{f(x_0 + tv) - f(x_0)}{t} - T(v)\right] \frac{t}{|t|} = 0.$$

Como $\frac{1}{\|v\|}$ é constante e $\frac{t}{|t|} = \pm 1$ é limitado, a igualdade acima é equivalente a:

$$\lim_{t\to 0} \frac{f(x_0 + tv) - f(x_0)}{t} - T(v) = 0,$$

o que mostra que a derivada direcional $\frac{\partial f}{\partial v}(x_0)$ existe e é igual a $T(v)$.

Teorema. *Se $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável num ponto $x_0 \in U$ então f admite derivadas direcionais no ponto x_0 em qualquer direção e vale a fórmula:*

$$\frac{\partial f}{\partial v}(x_0) = T(v),$$

para todo $v \in \mathbb{R}^m$, onde $T : \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear como a que aparece na definição de função diferenciável.

Demonstração. O caso $v \neq 0$ foi provado acima e o caso $v = 0$ é trivial. ■

Corolário. *A transformação linear T que aparece na definição de função diferenciável é* única.

Demonstração. Se T e T' são duas transformações lineares que "servem" para a definição de função diferenciável então $T(v) = \frac{\partial f}{\partial v}(x_0) = T'(v)$ para todo $v \in \mathbb{R}^m$, pelo Teorema acima. ■

Definição. *A transformação linear T que aparece na definição de função diferenciável (que é única, pelo Corolário acima) é denotada por* $\mathrm{d}f(x_0)$ *e é chamada* a diferencial da função f no ponto x_0.

Quando f é diferenciável no ponto x_0 podemos escrever agora:

$$f(x_0 + h) = f(x_0) + \mathrm{d}f(x_0) \cdot h + r(h),$$

com $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$.

`Observação`: provamos a seguinte **fórmula importante** para uma função diferenciável $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ num ponto $x_0 \in U$:

$$\mathrm{d}f(x_0) \cdot v = \frac{\partial f}{\partial v}(x_0), \quad \text{para todo } v \in \mathbb{R}^m.$$

(3) **Matriz Jacobiana**.

Recorde da álgebra linear que se $T : V \to V'$ é uma aplicação linear, com V, V' espaços vetoriais, $\mathfrak{B} = (b_1, \ldots, b_m)$ uma base de V e $\mathfrak{B}' = (b'_1, \ldots, b'_n)$ uma base de V' então a *matriz que representa T nas bases $\mathfrak{B}$ e $\mathfrak{B}'$*, denotada por $[T]_{\mathfrak{B}\mathfrak{B}'}$, é a matriz $n \times m$ cuja entrada na linha i, coluna j, é a i-ésima coordenada do vetor $T(b_j)$ na base $\mathfrak{B}'$ (veja exercícios).

Definição. *Se $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ é uma função diferenciável num ponto $x_0 \in U$ então* a matriz Jacobiana *de f no ponto x_0, denotada por $\mathrm{J}f(x_0)$, é a matriz que representa a aplicação linear $\mathrm{d}f(x_0) : \mathbb{R}^m \to \mathbb{R}^n$ nas bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$.*

Para calcular as entradas de $\mathrm{J}f(x_0)$ fazemos o seguinte:

$$\mathrm{J}f(x_0)_{ij} = \big[\mathrm{d}f(x_0) \cdot e_j\big]^i = \left[\frac{\partial f}{\partial e_j}(x_0)\right]^i = \left[\frac{\partial f}{\partial x_j}(x_0)\right]^i = \frac{\partial f^i}{\partial x_j}(x_0),$$
$$i = 1, \ldots, n, \ j = 1, \ldots, m;$$

para a última passagem acima veja os exercícios.

Provamos o seguinte:

Teorema. *Se $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ é uma função diferenciável num ponto $x_0 \in U$ então a matriz Jacobiana de f no ponto x_0 é dada por:*

$$\mathrm{J}f(x_0) = \begin{pmatrix} \frac{\partial f^1}{\partial x_1}(x_0) & \cdots & \frac{\partial f^1}{\partial x^m}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f^n}{\partial x_1}(x_0) & \cdots & \frac{\partial f^n}{\partial x^m}(x_0) \end{pmatrix}.$$

`Exemplo`: Se $f : U \subset \mathbb{R}^2 \to \mathbb{R}^2$ é dada por

$$f(x,y) = \left(\frac{1}{xy}, \cos x\right), \quad U = \big\{(x,y) \in \mathbb{R}^2 : xy \neq 0\big\} \ \ (U \text{ é aberto!})$$

então a diferencial de f num ponto $(x_0, y_0) \in U$ é dada por (use a matriz Jacobiana!):

$$\mathrm{d}f(x_0,y_0) \cdot (h,k) = \left(-\frac{h}{x_0^2 y_0} - \frac{k}{x_0 y_0^2}, -h \operatorname{sen} x_0\right),$$

para $(h,k) \in \mathbb{R}^2$. Para ser rigoroso, *ainda não temos teoria suficiente para ver rapidamente que f é diferenciável*! Admitindo esse fato por enquanto, obtemos a fórmula acima para $\mathrm{d}f(x_0, y_0)$.

`Observação`: Quando $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável em todos os pontos de U então dizemos que f é *diferenciável em U*. Para cada $x_0 \in U$, $\mathrm{d}f(x_0)$ é uma aplicação linear de $\mathbb{R}^m$ em $\mathbb{R}^n$; que tipo de objeto seria $\mathrm{d}f$? Considere o espaço vetorial:

$$\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n) = \text{aplicações lineares de } \mathbb{R}^m \text{ em } \mathbb{R}^n;$$

recorde da álgebra linear que $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ pode ser identificado naturalmente com o espaço vetorial de todas as matrizes reais $n \times m$ (que tem dimensão nm). A *diferencial* de f nos dá então uma aplicação:

$$\mathrm{d}f : U \subset \mathbb{R}^m \longrightarrow \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n),$$

que leva cada $x_0 \in U$ em $\mathrm{d}f(x_0) \in \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$. No caso $m = n = 1$ então $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ pode ser identificado com $\mathbb{R}$ (matrizes 1×1!) e daí $\mathrm{d}f$ vira uma aplicação de $U \subset \mathbb{R}$ em $\mathbb{R}$; essa aplicação é a clássica derivada f' de f do Cálculo I!

(4) **O caso n=1: gradientes**.

Esse assunto será retomado depois com mais detalhes.

Se $f : U \subset \mathbb{R}^m \to \mathbb{R}$ é uma função *a valores reais* diferenciável num ponto $x_0 \in U$ então a diferencial de f em x_0 define um elemento do dual de $\mathbb{R}^m$, i.e., $\mathrm{d}f(x_0) \in \mathbb{R}^{m*}$ (recorde que o *dual* de um espaço vetorial real V é o espaço vetorial $\mathrm{Lin}(V, \mathbb{R})$).

O produto interno canônico de $\mathbb{R}^m$ induz um isomorfismo de $\mathbb{R}^m$ sobre seu dual $\mathbb{R}^{m*}$ dado por $v \mapsto \langle v, \cdot \rangle$. Esse isomorfismo leva o vetor $v = (v_1, \ldots, v_m)$ de $\mathbb{R}^m$ sobre o elemento de $I\!R^{m*}$ que é representado pela matriz $(v_1 \ \cdots \ v_m)$ (é uma matriz $1 \times m$!). Já vimos que a matriz que representa $\mathrm{d}f(x_0)$ (ou seja, $\mathrm{J}f(x_0)$) é dada por:

$$\mathrm{J}f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0) \ \cdots \ \frac{\partial f}{\partial x_m}(x_0)\right).$$

O vetor de $I\!R^m$ correspondente a $\mathrm{d}f(x_0) \in I\!R^{m*}$ é portanto o vetor

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_m}(x_0)\right)$$

conhecido como o *gradiente* de f no ponto x_0. Obtemos então a fórmula:

$$\frac{\partial f}{\partial v}(x_0) = \mathrm{d}f(x_0) \cdot v = \langle \nabla f(x_0), v \rangle,$$

para todo $v \in I\!R^m$.

(5) **Funções diferenciáveis são contínuas**.

Vamos usar aqui que transformações lineares $T : I\!R^m \to I\!R^n$ são contínuas; isso segue do fato que as mesmas podem ser escritas em termos de somas e produtos (isso será rediscutido depois).

Teorema. *Se $f : U \subset I\!R^m \to I\!R^n$ é diferenciável no ponto $x_0 \in U$ então f é contínua no ponto $x_0 \in U$.*

Demonstração. Escrevemos:

$$f(x_0 + h) = f(x_0) + \mathrm{d}f(x_0) \cdot h + r(h),$$

com $\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$; *a fortiori* $\lim_{h \to 0} r(h) = \lim_{h \to 0} \frac{r(h)}{\|h\|} \|h\| = 0$. Fazendo $h \to 0$ na fórmula acima obtemos, usando a continuidade da transformação linear $\mathrm{d}f(x_0)$:

$$\lim_{h \to 0} f(x_0 + h) = f(x_0) + \mathrm{d}f(x_0) \cdot 0 = f(x_0).$$

Isso completa a demonstração. ■

`Observação`: A existência das derivadas parciais $\frac{\partial f}{\partial x_i}(x_0)$, $i = 1, \ldots, m$, não garante que f é diferenciável em x_0. Nem a existência de todas as derivadas direcionais de f em x_0, junto com a continuidade de f em x_0 (e até mesmo junto com a linearidade de $v \mapsto \frac{\partial f}{\partial v}(x_0)$) implicam a diferenciabilidade de f em x_0. Para contra-exemplos estranhos procurem o Curso de Análise Vol. II (capítulo 3) ou o livro de Cálculo II do Guidorizzi. Não daremos importância a tais exemplos.

`Observação`: Toda teoria desenvolvida até agora pode ser feita trocando $I\!R^m$ e $I\!R^n$ por espaços vetoriais reais de dimensão finita arbitrários! Para isso precisamos apenas fazer alguns esclarecimentos sobre normas em espaços vetoriais quaisquer (fica para depois). Observamos também que a noção de derivada parcial não faz sentido quando trocamos $I\!R^m$ por um espaço vetorial qualquer, a não ser que uma base seja fixada em tal espaço; similarmente, para dar sentido à matriz Jacobiana, precisamos fazer escolhas de bases nos espaços vetoriais em questão.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Álgebra Linear`.

1. Seja $T : V \to V'$ uma transformação linear, onde V e V' são espaços vetoriais de dimensão finita. Sejam $\mathfrak{B}$ e $\mathfrak{B}'$ bases de V e V' respectivamente. Denote por $[T]_{\mathfrak{B}\mathfrak{B}'}$ a matriz que representa T com respeito às bases $\mathfrak{B}$ e $\mathfrak{B}'$; por $[x]_{\mathfrak{B}}$ denotamos o vetor coluna que contém as coordenadas com respeito à base $\mathfrak{B}$ de um vetor $x \in V$.
   (a) Mostre que
   $$[T]_{\mathfrak{B}\mathfrak{B}'}[x]_{\mathfrak{B}} = \big[T(x)\big]_{\mathfrak{B}'}.$$
   (b) Se $S : V' \to V''$ é uma outra transformação linear, com V'' um espaço vetorial de dimensão finita e $\mathfrak{B}''$ uma base de V'', mostre que
   $$[S]_{\mathfrak{B}'\mathfrak{B}''}[T]_{\mathfrak{B}\mathfrak{B}'} = [S \circ T]_{\mathfrak{B}\mathfrak{B}''}.$$

`Diferenciação`.

2. Seja $f : U \to \mathbb{R}^n$ uma função com $U \subset \mathbb{R}^m$ aberto. Escreva $f = (f^1, \ldots, f^n)$ com cada $f^i : U \to \mathbb{R}$. Para $x_0 \in U$, mostre que:
   (a) f é diferenciável em x_0 se e somente se cada f^i é diferenciável em x_0, sendo nesse caso $\big[\mathrm{d}f(x_0)\big]^i = \mathrm{d}f^i(x_0)$, $i = 1, \ldots, n$ ("a i-ésima coordenada da diferencial é a diferencial da i-ésima coordenada").
   (b) para $v \in \mathbb{R}^m$, mostre que $\frac{\partial f}{\partial v}(x_0)$ existe se e somente se $\frac{\partial f^i}{\partial v}(x_0)$ existe para cada $i = 1, \ldots, n$; nesse caso, mostre que $\big[\frac{\partial f}{\partial v}(x_0)\big]^i = \frac{\partial f^i}{\partial v}(x_0)$, $i = 1, \ldots, n$ (" a i-ésima coordenada de uma derivada direcional é a derivada direcional da i-ésima coordenada").

Aula número 3 (13/03)

(1) **O caso m = 1: curvas em $\mathbb{R}^n$; relacionando vetores tangentes e diferenciais**.

Quando $m = n = 1$, vimos que a noção de diferenciabilidade coincide com a noção clássica de derivada do Cálculo I: mais precisamente, $f : U \subset \mathbb{R} \to \mathbb{R}$ é diferenciável em $x_0 \in U$ se e somente se existe a derivada $f'(x_0) \in \mathbb{R}$ (do Cálculo I) e nesse caso a diferencial de f no ponto x_0 é dada por $\mathrm{d}f(x_0) \cdot h = f'(x_0)h$. Vamos agora ver o que acontece quando apenas $m = 1$; estamos falando então de curvas em $\mathbb{R}^n$. Como mostraremos a seguir, o caso $m = 1$ com n qualquer é essencialmente idêntico ao caso $m = n = 1$ discutido na aula anterior (na verdade, só separamos o caso $m = n = 1$ por motivos didáticos).

Definição. *Dada uma aplicação $f : U \subset \mathbb{R} \to \mathbb{R}^n$ então o* vetor tangente *a f no ponto $x_0 \in U$ é definido pelo limite (se existir):*

$$f'(x_0) = \lim_{t\to 0} \frac{f(x_0 + t) - f(x_0)}{t} \in \mathbb{R}^n.$$

Observe que $f'(x_0)$ é a mesma coisa que a derivada direcional $\frac{\partial f}{\partial v}(x_0)$ com $v = 1$ (parece estranho? números reais também são vetores!).

Recorde que uma transformação linear $T : \mathbb{R} \to \mathbb{R}^n$ é sempre da forma $T(h) = zh$, onde $z \in \mathbb{R}^n$ é um vetor (veja os exercícios).

Teorema. *Uma aplicação $f : U \subset \mathbb{R} \to \mathbb{R}^n$ é diferenciável num ponto $x_0 \in U$ se e somente se o vetor tangente $f'(x_0)$ existe; as duas seguintes fórmulas relacionam a diferencial $\mathrm{d}f(x_0) : \mathbb{R} \to \mathbb{R}^n$ com o vetor tangente $f'(x_0) \in \mathbb{R}^n$:*

$$f'(x_0) = \mathrm{d}f(x_0) \cdot 1, \qquad \mathrm{d}f(x_0) \cdot h = f'(x_0)h, \quad h \in \mathbb{R}.$$

Demonstração. É igual a que fizemos quando $n = m = 1$. A função f é diferenciável no ponto $x_0 \in U$ se e somente se existe $z \in \mathbb{R}^n$ tal que:

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0) - zh}{|h|} = 0;$$

como já fizemos várias vezes agora, trocamos $\frac{1}{|h|}$ por $\frac{1}{h}\frac{h}{|h|}$ e observamos que $\frac{h}{|h|} = \pm 1$ é limitado. Daí, f é diferenciável em x_0 se e somente se:

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0) - zh}{h} = \lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h} - z = 0.$$

Isso completa a demonstração. ■

`Observação`: Quando $m = 1$, flexibilizamos um pouco a hipótese que U seja um aberto; admitimos também que $U \subset \mathbb{R}$ seja um intervalo não necessariamente aberto. Para quem se sentir incomodado com isso, observe que todas as demonstrações que fizermos funcionam quando $m = 1$ e U é um intervalo qualquer. Também no caso $U \subset \mathbb{R}^m$ (m arbitrário) é possível flexibilizar a hipótese que U seja aberto na maioria dos teoremas, mas é *bem* mais difícil descrever quais são os subconjuntos $U \subset \mathbb{R}^m$ "admissíveis" para o desenvolvimento da teoria e por isso preferimos trabalhar só com abertos quando $m \neq 1$ (voltaremos um pouco a esse assunto quando discutirmos variedades com bordo).

`Observação`: Quando $m = 1$ é mais comum pensar nas coisas da seguinte maneira: em vez de escrever $f : U \subset \mathbb{R} \to \mathbb{R}^n$ escrevemos $\gamma : I \to \mathbb{R}^n$, onde $I \subset \mathbb{R}$ é um intervalo (tipicamente, $I = [a, b]$). Dizemos que γ é uma *curva* em $\mathbb{R}^n$; os elementos de I são tipicamente denotados por letras como t e s e são chamados de *instantes* (em vez de pontos). É comum também dizer que t é o *parâmetro* da curva γ; dizemos aí que $\gamma'(t)$ é o *vetor tangente a γ no instante t*.

(2) **Dois exemplos fáceis**.

`Exemplo`. Se $f : \mathbb{R}^m \to \mathbb{R}^n$ é uma função constante então f é diferenciável e sua diferencial é zero em qualquer ponto.

`Exemplo`. Se $f : \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear então f é diferenciável e $\mathrm{d}f(x_0) = f$ para todo $x_0 \in \mathbb{R}^m$.

(afinal de contas, a melhor aproximação linear de uma transformação linear é ela mesma!).

Em ambos os exemplos acima, quando escrevemos a definição de diferenciabilidade vemos que o resto $r(h)$ é zero, de modo que $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$ é trivial.

(3) **Soma e produto de aplicações diferenciáveis**.

Teorema. *Se $f, g : U \subset \mathbb{R}^m \to \mathbb{R}^n$ são ambas diferenciáveis em $x_0 \in U$ então $f + g$ é diferenciável em x_0 e $\mathrm{d}(f + g)(x_0) = \mathrm{d}f(x_0) + \mathrm{d}g(x_0)$.*

Intuitivamente, se $\mathrm{d}f(x_0)$ é uma boa aproximação linear para f perto de x_0 e $\mathrm{d}g(x_0)$ é uma boa aproximação linear para g perto de x_0 então $\mathrm{d}f(x_0) + \mathrm{d}g(x_0)$ é uma boa aproximação linear para $f + g$ perto de x_0. A prova segue basicamente esse padrão:

Demonstração. Escrevemos:

$$f(x_0 + h) = f(x_0) + \mathrm{d}f(x_0) \cdot h + r_1(h),$$
$$g(x_0 + h) = g(x_0) + \mathrm{d}g(x_0) \cdot h + r_2(h),$$

com $\lim_{h\to 0} \frac{r_1(h)}{\|h\|} = \lim_{h\to 0} \frac{r_2(h)}{\|h\|} = 0$. Somando as duas igualdades obtemos:

$$(f + g)(x_0 + h) = (f + g)(x_0) + \big[\mathrm{d}f(x_0) + \mathrm{d}g(x_0)\big] \cdot h + r(h),$$

com $r = r_1 + r_2$. Obviamente, $\lim_{h\to 0} \frac{r(h)}{\|h\|} = \lim_{h\to 0} \frac{r_1(h)+r_2(h)}{\|h\|} = 0$ e $\mathrm{d}f(x_0) + \mathrm{d}g(x_0)$ é uma transformação linear. ■

`Observação`: O seguinte truque (parece bobagem, mas não é) facilita às vezes o trabalho com a definição de diferenciabilidade: usamos

$$f(x_0 + h) = f(x_0) + \mathrm{d}f(x_0) \cdot h + \rho(h)\|h\|,$$

com $\lim_{h\to 0} \rho(h) = 0$, em vez da formulação dada inicialmente com o resto r.

Como é de se esperar, a diferenciação de um produto dá um pouco mais de trabalho. Nesse caso, só faz sentido considerar aplicações a valores reais.

Teorema. *Se $f, g : U \subset I\!R^m \to I\!R$ são diferenciáveis num ponto $x_0 \in U$ então o produto fg é diferenciável em x_0 e vale a fórmula:*

$$\mathrm{d}(fg)(x_0) \cdot h = \big[\mathrm{d}f(x_0) \cdot h\big] g(x_0) + f(x_0)\big[\mathrm{d}g(x_0) \cdot h\big],$$

para todo $h \in I\!R^m$.

Demonstração. O lado direito da fórmula acima define de fato uma transformação linear em h, o que já é um bom começo. Escrevemos:

$$\begin{aligned} f(x_0 + h) &= f(x_0) + \mathrm{d}f(x_0) \cdot h + \rho_1(h)\|h\|, \\ g(x_0 + h) &= g(x_0) + \mathrm{d}g(x_0) \cdot h + \rho_2(h)\|h\|, \end{aligned}$$

com $\lim_{h \to 0} \rho_1(h) = \lim_{h \to 0} \rho_2(h) = 0$. Multiplicando as duas igualdades acima obtemos:

$$(fg)(x_0 + h) = (fg)(x_0) + \big[\mathrm{d}f(x_0) \cdot h\big] g(x_0) + f(x_0)\big[\mathrm{d}g(x_0) \cdot h\big] + \rho(h)\|h\|,$$

onde ρ é dado por:

$$\begin{aligned} \rho(h) = {} & \frac{\big[\mathrm{d}f(x_0) \cdot h\big]\big[\mathrm{d}g(x_0) \cdot h\big]}{\|h\|} + g(x_0)\rho_1(h) + \big[\mathrm{d}g(x_0) \cdot h\big]\rho_1(h) \\ & + f(x_0)\rho_2(h) + \big[\mathrm{d}f(x_0) \cdot h\big]\rho_2(h) + \rho_1(h)\rho_2(h)\|h\|. \end{aligned}$$

Devemos mostrar que $\lim_{h \to 0} \rho(h) = 0$. Para isso, observe que os seis termos aparecendo na expressão para ρ são o produto de alguma coisa que tende a zero por alguma coisa limitada (quando h está próximo de zero) — alguns termos têm até mais de um fator que vai para zero. Só o primeiro termo é um pouco mais complicado: sua análise depende da observação que $\mathrm{d}f(x_0) \cdot \big(\frac{h}{\|h\|}\big)$ é limitado; isso segue do fato que a aplicação contínua $\mathrm{d}f(x_0) : I\!R^m \to I\!R$ é limitada na esfera unitária $\big\{v \in I\!R^m : \|v\| = 1\big\}$ de $I\!R^m$, já que a mesma é compacta (voltaremos a esse assunto na revisão de topologia). ■

(4) **Mais álgebra linear: aplicações bilineares e multi-lineares**.

Definição. *Sejam V_1, V_2, W espaços vetoriais. Uma aplicação $B : V_1 \times V_2 \to W$ é chamada* bilinear *quando B for linear em cada variável, i.e., para todos $v_1 \in V_1$, $v_2 \in V_2$ as aplicações $B(v_1, \cdot) : V_2 \to W$ e $B(\cdot, v_2) : V_1 \to W$ são lineares. De maneira mais explícita, B é bilinear quando valem as seguintes condições:*

(i) $B(v_1 + v_1', v_2) = B(v_1, v_2) + B(v_1', v_2)$;
(ii) $B(cv_1, v_2) = c\,B(v_1, v_2)$;
(iii) $B(v_1, v_2 + v_2') = B(v_1, v_2) + B(v_1, v_2')$;
(iv) $B(v_1, cv_2) = c\,B(v_1, v_2)$;
para todos $v_1, v_1' \in V_1$, $v_2, v_2' \in V_2$ e $c \in I\!R$. Quando $W = I\!R$ dizemos também que B é uma forma bilinear.

`Observação`: Acustumem-se com a notação $\cdot$ porque ela é muito prática. Por exemplo, na definição acima, tínhamos uma aplicação B de duas variáveis; fixando uma das variáveis, sobra uma aplicação de uma única variável. Essa aplicação obtida de B fixando uma variável, é denotada colocando um $\cdot$ na variável que ficou livre. Assim, por exemplo, se $B : V_1 \times V_2 \to W$ e $v_1 \in V_1$ então $B(v_1, \cdot)$ denota a aplicação de V_2 em W que leva v_2 em $B(v_1, v_2)$.

`Exemplo`: A *multiplicação de números reais* é uma aplicação (e na verdade uma forma) bilinear, i.e., a aplicação:

$$m : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$$

dada por $m(x, y) = xy$ é bilinear.

`Exemplo`: O *produto interno canônico* de $\mathbb{R}^n$ definido por $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ é uma aplicação bilinear $\langle\cdot,\cdot\rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ (e também uma forma bilinear).

`Exemplo`: O *produto vetorial* em $\mathbb{R}^3$, i.e., a aplicação $\mathbb{R}^3 \times \mathbb{R}^3 \ni (x, y) \mapsto x \times y \in \mathbb{R}^3$ é bilinear.

Em vista dos exemplos acima, observe que uma aplicação bilinear $(x, y) \mapsto B(x, y)$ pode ser visualizada como uma "operação binária que satisfaz a propriedade distributiva" (na verdade, as propriedades (i) e (iii) são a propriedade distributiva, (ii) e (iv) dizem um pouco a mais).

`Exemplo`: Se $\mathrm{M}_{n\times m}(\mathbb{R})$ denota o espaço vetorial de todas as matrizes reais $n \times m$ então a *multiplicação de matrizes*:

$$\mathrm{M}_{n\times m}(\mathbb{R}) \times \mathrm{M}_{m\times p}(\mathbb{R}) \ni (A, B) \longmapsto AB \in \mathrm{M}_{n\times p}(\mathbb{R})$$

é uma aplicação bilinear.

`Exemplo`: Se V, V' e V'' são espaços vetoriais então a aplicação de *composição de transformações lineares*:

$$\mathrm{Lin}(V', V'') \times \mathrm{Lin}(V, V') \ni (T, S) \longmapsto T \circ S \in \mathrm{Lin}(V, V'')$$

é bilinear. Se $V = \mathbb{R}^p$, $V' = \mathbb{R}^m$, $V'' = \mathbb{R}^n$ então este exemplo é essencialmente o mesmo que o anterior, identificando aplicações lineares e matrizes através das bases canônicas.

`Não-Exemplo`: A aplicação *soma de vetores* $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ definida por:

$$s(x, y) = x + y, \quad x, y \in \mathbb{R}^n,$$

não é bilinear. Na verdade, s é *linear* quando identificamos $\mathbb{R}^n \times \mathbb{R}^n$ com $\mathbb{R}^{2n}$!

É bastante natural generalizar o conceito de bilinearidade para aplicações de mais de duas variáveis: obtemos então as noções de trilinearidade, quadrilinearidade, ..., ou em geral n-linearidade: todas essas noções são denominadas por *multi-linearidade*, quando não se quer especificar o número de variáveis na sentença em questão.

Definição. *Sejam V_1, V_2, ..., V_n, W espaços vetoriais e seja $B : V_1 \times V_2 \times \cdots \times V_n \to W$ uma aplicação. Dizemos que B é* multi-linear *(ou n-*linear*) quando B for linear em cada variável; mais explicitamente, B é multi-linear quando dado $i = 1, \ldots, n$ e dados vetores $v_j \in V_j$, $j = 1, \ldots, i-1, i+1, \ldots, n$ então a aplicação*

$$B(v_1, \ldots, v_{i-1}, \cdot, v_{i+1}, \ldots, v_n) : V_i \longrightarrow W$$

é linear. Quando $W = \mathbb{R}$ dizemos também que B é uma forma multi-linear.

`Exemplo`: A aplicação:

$$\det : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ fatores}} \longrightarrow \mathbb{R}$$

que associa a cada n-upla $(v_1, \ldots, v_n) \in (\mathbb{R}^n)^n$ o determinante da matriz $n \times n$ que possui os vetores v_i em suas colunas (ou linhas, tanto faz aqui) é multi-linear; det é portanto uma forma n-linear em $\mathbb{R}^n$.

Recorde que uma aplicação linear $T : V \to W$ fica univocamente determinada quando especificamos seu valor sobre os vetores de uma base de V: de maneira mais explícita, se $\mathfrak{B} = (b_1, \ldots, b_m)$ é uma base de V então dados vetores $w_1, \ldots, w_m \in W$, existe uma única aplicação linear $T : V \to W$ tal que $T(b_i) = w_i$, $i = 1, \ldots, m$. Como se calcula T num vetor arbitrário $v \in V$? Escreva $v = \sum_{i=1}^m \alpha_i b_i$ e daí:

$$T(v) = \sum_{i=1}^m \alpha_i T(b_i) = \sum_{i=1}^m \alpha_i w_i;$$

quem trabalhou com os exercícios de álgebra linear da aula anterior (ou quem já tinha uma formação básica de álgebra linear) não vai ter problemas com as afirmações acima!

Quando escolhemos também uma base $\mathfrak{B}'$ em W então os vetores $T(b_i) \in W$ podem ser descritos através de suas coordenadas na base $\mathfrak{B}'$ e daí obtemos o fato básico que a aplicação linear T pode ser descrita completamente através da matriz $[T]_{\mathfrak{B}\mathfrak{B}'}$.

Quando trabalhamos com aplicações bilineares e multi-lineares em geral, a situação não é muito diferente (a diferença é que a notação fica mais feia e começam a aparecer um monte de índices). Veja os exercícios!

`Observação`: Aplicações multi-lineares não serão muito importantes agora. Voltaremos a esse assunto quando estudarmos *formas diferenciais* na parte final do curso.

(5) **Diferenciando aplicações bilineares e multi-lineares**.

Vamos assumir agora que aplicações multi-lineares $B : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_r} \to \mathbb{R}^n$ são contínuas (veja os exercícios).

Teorema. *Seja $B : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ uma aplicação bilinear. Então B é diferenciável e sua diferencial é dada por:*

$$\mathrm{d}B(x, y) \cdot (h, k) = B(h, y) + B(x, k),$$

para todos $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ e $(h, k) \in \mathbb{R}^m \times \mathbb{R}^n$.

Demonstração. Observe primeiro que a expressão acima define uma aplicação linear com respeito a (h, k). Pela bilinearidade de B temos:

$$B(x + h, y + k) = B(x, y) + B(h, y) + B(x, k) + r(h, k),$$

onde $r(h, k) = B(h, k)$. Devemos mostrar que:

$$\lim_{(h,k)\to 0} \frac{B(h, k)}{\|(h, k)\|} = 0,$$

onde $\|(h,k)\| = \sqrt{\|h\|^2 + \|k\|^2}$. Temos:

$$\frac{B(h,k)}{\|(h,k)\|} = B\left(\frac{h}{\|h\|}, \frac{k}{\|k\|}\right) \cdot \frac{\|h\|}{\|(h,k)\|} \cdot \|k\|;$$

o terceiro fator à direita da igualdade acima tende a zero e o segundo é limitado (menor ou igual a 1). O primeiro também é limitado, pois B é contínua e o conjunto

$$\big\{(v,w) \in I\!R^m \times I\!R^n : \|v\| = \|w\| = 1\big\}$$

é compacto (retomaremos esse assunto na revisão de topologia).

Segue então que $\lim_{(h,k)\to 0} \frac{B(h,k)}{\|(h,k)\|} = \lim_{(h,k)\to 0} \frac{r(h,k)}{\|(h,k)\|} = 0$. ■

`Observação`: Com um pouco de paciência é possível demonstrar (seguindo as idéias da demonstração anterior) que toda aplicação multi-linear $B : I\!R^{m_1} \times \cdots \times I\!R^{m_r} \to I\!R^n$ é diferenciável e que sua diferencial é dada por:

$$\mathrm{d}B(x_1, \ldots, x_r) \cdot (h_1, \ldots, h_r) = \sum_{i=1}^{r} B(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_r),$$

para todos $x_i \in I\!R^{m_i}$, $h_i \in I\!R^{m_i}$, $i = 1, \ldots, r$.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Álgebra Linear`.

1. Seja V um espaço vetorial real. Mostre que todas as transformações lineares
$$T : I\!R \to V$$
são da forma $T(h) = zh$ para algum vetor $z \in V$ (*dica*: $T(h) = T(h \cdot 1)$). Conclua que $\mathrm{Lin}(I\!R, V)$ é isomorfo a V; mais precisamente, mostre que a aplicação:
$$\mathrm{Lin}(I\!R, V) \ni T \longmapsto T(1) \in V$$
é um isomorfismo.

2. Sejam V, V', V'' espaços vetoriais e sejam $\mathfrak{B} = (b_1, \ldots, b_n)$, $\mathfrak{B}' = (b'_1, \ldots, b'_m)$ e $\mathfrak{B}'' = (b''_1, \ldots, b''_p)$ bases de V, V' e V'' respectivamente. Mostre que:
   (a) Se $B : V \times V' \to V''$ é uma aplicação bilinear então para todos $v = \sum_{i=1}^n \alpha_i b_i \in V$, $v' = \sum_{j=1}^m \alpha'_j b'_j \in V'$ vale a identidade:
   $$B(v, v') = \sum_{i=1}^n \sum_{j=1}^m \alpha_i \alpha'_j B(b_i, b'_j).$$
   (b) Mostre que uma aplicação bilinear $V \times V' \to V''$ fica univocamente determinada por seus valores em vetores de uma base de V e de uma base de V', i.e., mostre que dados vetores $v''_{ij} \in V''$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ então existe uma única aplicação bilinear $B : V \times V' \to V''$ tal que $B(b_i, b'_j) = v''_{ij}$ para todos $i = 1, \ldots, n$ e $j = 1, \ldots, m$.
   (c) Se $V = I\!R^n$, $V' = I\!R^m$ e $V'' = I\!R^p$, conclua do item (a) que toda aplicação bilinear $B : V \times V' \to V''$ é contínua (*dica*: somas e produtos de aplicações contínuas são contínuas).
   (d) Dada uma aplicação bilinear $B : V \times V' \to V''$, denote por B^k_{ij} a k-ésima coordenada na base $\mathfrak{B}''$ do vetor $B(b_i, b'_j) \in V''$. Obtemos então uma "matriz tri-dimensional" $(B^k_{ij})_{n \times m \times p}$. Conclua do item (b) que existe uma correspondência biunívoca entre aplicações bilineares $B : V \times V' \to V''$ e matrizes tri-dimensionais $n \times m \times p$ de números reais. Dizemos que a matriz tri-dimensional $(B^k_{ij})_{n \times m \times p}$ *representa* a aplicação bilinear B com respeito às bases $\mathfrak{B}$, $\mathfrak{B}'$ e $\mathfrak{B}''$.

`Observação`: Quando $V'' = I\!R$, i.e., no caso de formas bilineares $B : V \times V' \to I\!R$ então é possível omitir o índice k nas matrizes tri-dimensionais do item (d), de modo que formas bilineares $B : V \times V' \to I\!R$ são representadas por matrizes comuns $(B_{ij})_{n \times m}$.

   (e) Generalize (se você tiver muita paciência com notação chata, somatórias e índices) os itens anteriores para o caso de aplicações multi-lineares.

`Diferenciação`.

3. Uma *transformação afim* $A : I\!R^m \to I\!R^n$ é uma aplicação da forma $A(x) = T(x) + v$, com $T : I\!R^m \to I\!R^n$ uma aplicação linear e $v \in I\!R^n$ um vetor fixado. Mostre que $\mathrm{d}A(x) = T$ para todo $x \in I\!R^m$. Conclua que se A é uma *translação* (i.e., se $m = n$ e T é a identidade) então $\mathrm{d}A(x) = \mathrm{Id}$, para todo $x \in I\!R^m$.

Aula número 4 (15/03)

Enunciamos agora a **regra da cadeia**:

Teorema. *Sejam* $f : U \subset I\!R^m \to I\!R^n$ *e* $g : V \subset I\!R^n \to I\!R^p$ *aplicações com* $U \subset I\!R^m$*,* $V \subset I\!R^n$ *abertos e* $f(U) \subset V$*. Se* f *é diferenciável num ponto* $x_0 \in U$ *e* g *é diferenciável no ponto* $f(x_0) \in V$ *então a composta* $g \circ f : U \to I\!R^p$ *é diferenciável no ponto* $x_0 \in U$ *e sua diferencial é dada por:*

$$\mathrm{d}(g \circ f)(x_0) = \mathrm{d}g\big(f(x_0)\big) \circ \mathrm{d}f(x_0).$$

Numa frase: "a composta de aplicações diferenciáveis é diferenciável e a diferencial da composta é a composta das diferenciais".

A regra da cadeia pode ser pensada de modo intuitivo da seguinte forma: se $\mathrm{d}f(x_0)$ é uma boa aproximação linear para f perto de x_0 e $\mathrm{d}g\big(f(x_0)\big)$ é uma boa aproximação linear para g perto de $f(x_0)$ então a composta $\mathrm{d}g\big(f(x_0)\big) \circ \mathrm{d}f(x_0)$ é uma boa aproximação linear para a composta $g \circ f$ perto de x_0.

A demonstração da regra da cadeia, embora seja simples, será deixada para depois da revisão de topologia. O objetivo dessa aula é compreender vários aspectos do enunciado da regra da cadeia.

`Observação`: Nesse ponto já temos várias ferramentas para mostrar que uma aplicação é diferenciável. A saber:

(i) aplicações constantes, lineares e multi-lineares são diferenciáveis;
(ii) a soma e o produto de aplicações diferenciáveis é diferenciável (no caso do produto, só podemos considerar aplicações a valores reais);
(iii) quando o domínio é unidimensional, diferenciabilidade equivale à derivabilidade do Cálculo I; em particular, aplicações familiares como seno, cosseno, exponencial, etc, etc, etc, são diferenciáveis;
(iv) uma aplicação a valores em $I\!R^n$ é diferenciável se e somente se cada uma de suas n coordenadas é diferenciável;
(v) a composta de aplicações diferenciáveis é diferenciável.

Obtemos em particular que aplicações definidas por fórmulas são diferenciáveis (com exceção das que envolvem raízes de expressões que podem se anular, já que $\sqrt{\cdot}$ não é derivável na origem).

(1) **Interpretando a regra da cadeia em termos de matrizes Jacobianas: a regra da cadeia clássica**.

Como vimos nos exercícios da aula número 2, a matriz que representa a composta de duas transformações lineares é obtida fazendo o produto das matrizes que representam cada transformação linear envolvida. Em termos de matrizes, a regra da cadeia nos diz então que:

$$\mathrm{J}(g \circ f)(x_0) = \mathrm{J}g\big(f(x_0)\big)\mathrm{J}f(x_0);$$

expandindo o produto de matrizes obtemos:

$$\frac{\partial (g \circ f)^i}{\partial x_j}(x_0) = \sum_{k=1}^{n} \frac{\partial g^i}{\partial x_k}\big(f(x_0)\big)\frac{\partial f^k}{\partial x_j}(x_0).$$

Alternativamente, aplicando os dois lados da igualdade $\mathrm{d}(g\circ f)(x_0)=\mathrm{d}g\big(f(x_0)\big)\circ\mathrm{d}f(x_0)$ no j-ésimo vetor e_j da base canônica de $I\!R^m$ obtemos:

$$\begin{aligned}\frac{\partial(g\circ f)}{\partial x_j}(x_0)&=\mathrm{d}g\big(f(x_0)\big)\cdot\frac{\partial f}{\partial x_j}(x_0)=\mathrm{d}g\big(f(x_0)\big)\cdot\left(\sum_{k=1}^n\frac{\partial f^k}{\partial x_j}(x_0)\,e_k\right)\\&=\sum_{k=1}^n\Big[\mathrm{d}g\big(f(x_0)\big)\cdot e_k\Big]\frac{\partial f^k}{\partial x_j}(x_0)=\sum_{k=1}^n\frac{\partial g}{\partial x_k}\big(f(x_0)\big)\frac{\partial f^k}{\partial x_j}(x_0),\end{aligned}$$

ou seja:

$$\frac{\partial(g\circ f)}{\partial x_j}(x_0)=\sum_{k=1}^n\frac{\partial g}{\partial x_k}\big(f(x_0)\big)\frac{\partial f^k}{\partial x_j}(x_0).$$

É muito comum denotar pontos de "ambientes diferentes" por letras diferentes; por exemplo, os pontos do domínio de f podem ser denotados com x e os pontos do domínio de g com y (não por motivos matemáticos, apenas porque muitas vezes facilita a leitura). Daí, intuitivamente pensa-se que f é "função de x" e g é "função de y" (embora isso não queira dizer nada, formalmente) e reescreve-se a regra da cadeia sob a forma:

$$\frac{\partial(g\circ f)}{\partial x_j}(x_0)=\sum_{k=1}^n\frac{\partial g}{\partial y_k}\big(f(x_0)\big)\frac{\partial f^k}{\partial x_j}(x_0).$$

Muitas vezes também os matemáticos resolvem aderir àquela notação (às vezes, confusa, para dizer a verdade) que é encontrada principalmente em textos de física, onde não se fala explicitamente em funções e composição de funções, mas apenas em "variáveis dependendo umas das outras" e as composições são todas subentendidas (essa notação também tem seus méritos). Pensando então em "variáveis" $x\in U\subset I\!R^m$, $y\in V\subset I\!R^n$ e $z\in I\!R^p$, com y dependendo de x (i.e., temos a função f), z dependendo de y (i.e., temos a função g) e daí z também depende de x (i.e., subentende-se a composição $g\circ f$) então a regra da cadeia fica:

$$\frac{\partial z}{\partial x_j}=\sum_{k=1}^n\frac{\partial z}{\partial y_k}\,\frac{\partial y_k}{\partial x_j},$$

onde também os pontos onde as derivadas são calculadas são omitidos (e devem ser "adivinhados" pelo contexto).

(2) **A visão funtorial da regra da cadeia**.

Algumas vezes em matemática precisamos descrever igualdades entre composições de certas funções que seriam difíceis de serem escritas e de serem visualizadas usando a "notação bola" padrão $(g\circ f)$. Em vez disso, utiliza-se um recurso gráfico mais eficiente, conhecido como *diagrama comutativo*. Abaixo temos um exemplo:

$$\begin{tikzcd}
\mathbb{R}^m \arrow[r, "f"] \arrow[dd, "g"'] \arrow[rrd, bend left, "\rho"] & \mathbb{R}^n \arrow[dr, "h"] & \\
 & & \mathbb{R}^k \\
\mathbb{R}^p \arrow[r, "\mathcal{F}"'] \arrow[uur, "\sigma"] & \mathbb{R}^m \arrow[ur, "\alpha"'] &
\end{tikzcd}$$

O diagrama acima codifica as igualdades $\sigma \circ g = f$, $\alpha \circ \mathcal{F} = h \circ \sigma$ e $\rho = h \circ f$. A vantagem da formulação da regra da cadeia na forma "a diferencial da composta é a composta das diferenciais" é que podemos diferenciar todas as flechas que aparecem num diagrama comutativo obtendo um novo diagrama comutativo. No nosso exemplo:

$$\begin{tikzcd}
\mathbb{R}^m \arrow[r, "\mathrm{d}f"] \arrow[dd, "\mathrm{d}g"'] \arrow[rrd, bend left, "\mathrm{d}\rho"] & \mathbb{R}^n \arrow[dr, "\mathrm{d}h"] & \\
 & & \mathbb{R}^k \\
\mathbb{R}^p \arrow[r, "\mathrm{d}\mathcal{F}"'] \arrow[uur, "\mathrm{d}\sigma"] & \mathbb{R}^m \arrow[ur, "\mathrm{d}\alpha"'] &
\end{tikzcd}$$

no diagrama acima omitimos os pontos onde as diferenciais são calculadas para simplificar a exposição, mas não é difícil adivinhar quais são tais pontos: por exemplo, se f é diferenciada no ponto $x_0 \in \mathbb{R}^m$ então h é diferenciada no ponto $f(x_0)$ (complete o resto!).

`Observação`: O nome "funtorial" usado nessa seção vem da *teoria das categorias*: um *funtor* é essencialmente uma regra que pode ser aplicada em um diagrama comutativo produzindo um novo diagrama comutativo (mas esse assunto não tem nada a ver com o nosso curso).

(3) **Compondo funções e curvas**.

Seja $\gamma : I \subset \mathbb{R} \to \mathbb{R}^m$ uma curva (onde $I \subset \mathbb{R}$ é um intervalo) e $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$ contendo a imagem de γ. Suponha que $t \in I$ é tal que γ é diferenciável em t e f é diferenciável em $\gamma(t)$ (isso acontece, por exemplo, se γ e f forem diferenciáveis em todo seu domínio). A aplicação composta $f \circ \gamma : I \subset \mathbb{R} \to \mathbb{R}^n$ é uma curva em $\mathbb{R}^n$ que é diferenciável em t pela regra da cadeia; a diferencial de $f \circ \gamma$ no instante t é dada por:

$$\mathrm{d}(f \circ \gamma)(t) = \mathrm{d}f\big(\gamma(t)\big) \circ \mathrm{d}\gamma(t).$$

Recorde que o vetor tangente a uma curva é obtido aplicando sua diferencial ao vetor $1 \in \mathbb{R}$ (vide início da aula número 3); aplicando os dois lados da igualdade acima a $1 \in \mathbb{R}$ obtemos:

$$(f \circ \gamma)'(t) = \mathrm{d}f\big(\gamma(t)\big) \cdot \gamma'(t).$$

`Exemplo`: Se $T : \mathbb{R}^m \to \mathbb{R}^n$ é uma transformação linear (por exemplo, imagine o caso que $m = n = 3$ e T é uma rotação) e se $\gamma : I \subset \mathbb{R} \to \mathbb{R}^m$ é uma curva diferenciável no instante t então:

$$(T \circ \gamma)'(t) = T\big(\gamma'(t)\big).$$

(quando T é uma rotação a igualdade acima exprime o fato que "o vetor tangente de uma rotação de γ é a rotação do vetor tangente de γ"). Se $T : \mathbb{R}^n \to \mathbb{R}^n$ é uma *translação*, i.e., se $T(x) = x + v$ para algum $v \in \mathbb{R}^n$ então $\mathrm{d}T(x) = \mathrm{Id}$ para todo $x \in \mathbb{R}^n$ e:

$$(T \circ \gamma)'(t) = \gamma'(t),$$

ou seja, transladando uma curva não alteramos seu vetor tangente (veja os exercícios da aula número 3).

`Exemplo`: Seja $B : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ uma aplicação bilinear e sejam $\gamma : I \subset \mathbb{R} \to \mathbb{R}^m$, $\mu : I \subset \mathbb{R} \to \mathbb{R}^n$ duas curvas diferenciáveis num instante $t \in I$. Temos então que a curva $(\gamma, \mu) : I \to \mathbb{R}^m \times \mathbb{R}^n$ é diferenciável em t e $(\gamma, \mu)'(t) = \big(\gamma'(t), \mu'(t)\big)$ (veja os exercícios da aula número 2). Segue da regra da cadeia que a curva $B \circ (\gamma, \mu) : I \to \mathbb{R}^p$ é diferenciável em t e que seu vetor tangente no instante t é dado por:

$$\frac{\mathrm{d}}{\mathrm{d}t} B\big(\gamma(t), \mu(t)\big) = \mathrm{d}B\big(\gamma(t), \mu(t)\big) \cdot \big(\gamma'(t), \mu'(t)\big);$$

usando a fórmula deduzida na aula número 3 para a diferencial de aplicações bilineares obtemos:

$$\frac{\mathrm{d}}{\mathrm{d}t} B\big(\gamma(t), \mu(t)\big) = B\big(\gamma'(t), \mu(t)\big) + B\big(\gamma(t), \mu'(t)\big).$$

Interprete a fórmula acima quando substituímos B por cada uma das aplicações bilineares discutidas nos exemplos da aula número 3 (veja também os exercícios).

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Diferenciação`.

(1.) Rededuza a fórmula para a diferenciação de um produto de aplicações:

$$f, g : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$$

usando a regra da cadeia como no último exemplo da aula. Generalize o resultado para obter a diferencial do produto interno $U \ni x \mapsto \langle f(x), g(x) \rangle \in \mathbb{R}$ de aplicações $f, g : U \subset \mathbb{R}^m \to \mathbb{R}^n$.

Complemento à Aula número 4 (15/03)

Faremos um exemplo adicional de aplicação da regra da cadeia. Considere o espaço vetorial real $\mathrm{M}_{n\times n}(I\!R)$ de todas as matrizes reais $n\times n$ (que é isomorfo a $I\!R^{n^2}$). O conjunto

$$\mathrm{GL}(n, I\!R) = \big\{A \in \mathrm{M}_{n\times n}(I\!R) : A \text{ é inversível}\big\}$$

formado por todas as matrizes inversíveis é aberto em $\mathrm{M}_{n\times n}(I\!R)$; isso segue do fato que $A \in \mathrm{GL}(n, I\!R) \Leftrightarrow \det(A) \neq 0$ e do fato que a função determinante det é contínua (esse argumento será revisto na revisão de topologia).

`Observação`: A notação $\mathrm{GL}(n, I\!R)$ vem do fato que o conjunto das matrizes reais inversíveis $n \times n$ é conhecido como o *grupo linear geral de* $I\!R^n$.

A *função inversão* $\mathcal{I} : \mathrm{GL}(n, I\!R) \to \mathrm{M}_{n\times n}(I\!R)$ definida por $\mathcal{I}(A) = A^{-1}$ é diferenciável. De fato, cada entrada de A^{-1} é uma função racional das entradas de A de modo que cada coordenada de $\mathcal{I}$ é um quociente de uma soma de produtos das funções coordenadas de $\mathrm{M}_{n\times n}(I\!R)$. O nosso objetivo é calcular a diferencial de $\mathcal{I}$.

Seja $\mathcal{C} : \mathrm{M}_{n\times n}(I\!R) \times \mathrm{M}_{n\times n}(I\!R) \to \mathrm{M}_{n\times n}(I\!R)$ a *função multiplicação de matrizes* definida por $\mathcal{C}(A, B) = AB$. Sabemos que $\mathcal{C}$ é bilinear e portanto diferenciável. Temos:

$$\mathcal{C}\big(A, \mathcal{I}(A)\big) = \mathrm{I},$$

para toda matriz $A \in \mathrm{GL}(n, I\!R)$, onde I denota a matriz identidade $n \times n$. Denote por $i : \mathrm{GL}(n, I\!R) \to \mathrm{M}_{n\times n}(I\!R)$ a *função inclusão* (i.e., $i(A) = A$) e por

$$(i, \mathcal{I}) : \mathrm{GL}(n, I\!R) \longrightarrow \mathrm{M}_{n\times n}(I\!R) \times \mathrm{M}_{n\times n}(I\!R)$$

a função $A \mapsto (A, A^{-1})$. Obviamente i é diferenciável e $\mathrm{d}i(A)$ é a identidade para todo $A \in \mathrm{GL}(n, I\!R)$ (já que i é a restrição da identidade de $\mathrm{M}_{n\times n}(I\!R)$ a um aberto – veja os exercícios). A igualdade $\mathcal{C}\big(A, \mathcal{I}(A)\big) = \mathrm{I}$ nos diz que a função composta $\mathcal{C} \circ (i, \mathcal{I})$ é constante (e igual a I), donde:

$$\mathrm{d}\big(\mathcal{C} \circ (i, \mathcal{I})\big)(A) = 0, \quad A \in \mathrm{GL}(n, I\!R).$$

Usando a regra da cadeia obtemos para todos $A \in \mathrm{GL}(n, I\!R)$, $H \in \mathrm{M}_{n\times n}(I\!R)$:

$$\mathrm{d}\mathcal{C}(A, A^{-1}) \cdot \big(H, \mathrm{d}\mathcal{I}(A) \cdot H\big) = 0;$$

como $\mathcal{C}$ é bilinear obtemos:

$$\mathcal{C}\big(A, \mathrm{d}\mathcal{I}(A) \cdot H\big) + \mathcal{C}(H, A^{-1}) = A\,\mathrm{d}\mathcal{I}(A) \cdot H + HA^{-1} = 0$$

e portanto:

$$\mathrm{d}\mathcal{I}(A) \cdot H = -A^{-1}HA^{-1}.$$

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Diferenciação`.

1. Mostre que diferenciabilidade é uma *propriedade local*, ou seja, dados $U, V \subset \mathbb{R}^m$ abertos e $x \in \mathbb{R}^m$ com $x \in V \subset U$ então uma função $f : U \to \mathbb{R}^n$ é diferenciável em x se e somente se sua restrição $f|_V$ é diferenciável em x; nesse caso as diferenciais são iguais, i.e., $\mathrm{d}(f|_V)(x) = \mathrm{d}f(x)$.
2. Usando a fórmula deduzida para a diferencial da função $\mathcal{I}$, mostre que se $t \mapsto A(t)$ é uma curva em $\mathrm{M}_{n\times n}(\mathbb{R})$ diferenciável em $t = t_0$ e tal que $A(t)$ é inversível para todo t então a curva $t \mapsto A(t)^{-1}$ é diferenciável em $t = t_0$ e seu vetor tangente é dado por:

$$\frac{\mathrm{d}}{\mathrm{d}t}A(t)^{-1}\bigg|_{t=t_0} = -A(t_0)^{-1}A'(t_0)A(t_0)^{-1}.$$

Aula número 5 (20/03)

Definição. *Seja M um conjunto. Uma* métrica *em M é uma função $d : M \times M \to \mathbb{R}$ satisfazendo as seguintes propriedades:*

(EM1) $d(x, y) \geq 0$ para todos $x, y \in M$ e $d(x, y) = 0$ se e somente se $x = y$;

(EM2) $d(x, y) = d(y, x)$ para todos $x, y \in M$ (simetria)*;*

(EM3) $d(x, z) \leq d(x, y) + d(y, z)$ para todos $x, y, z \in M$ (desigualdade triangular)*.*

Um espaço métrico *é um par (M, d) onde M é um conjunto e d é uma métrica em M.*

`Exemplo`. A aplicação $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ dada por $d(x, y) = \|x - y\| = \big(\sum_{i=1}^n (x_i - y_i)^2\big)^{\frac{1}{2}}$ é uma métrica em $\mathbb{R}^n$. As propriedades (EM1) e (EM2) são triviais; quanto à desigualdade triangular (embora provavelmente vocês todos já conheçam) será demonstrada mais adiante quando discutirmos produtos internos e normas. A métrica d é conhecida como a *métrica Euclideana*.

`Exemplo`. Se (M, d) é um espaço métrico e $S \subset M$ é um subconjunto de M então $d|_{S\times S}$ é uma métrica em S (verifique!) e portanto $\big(S, d|_{S\times S}\big)$ é um espaço métrico; dizemos que $\big(S, d|_{S\times S}\big)$ é um *subespaço métrico* de (M, d) ou que $d|_{S\times S}$ é a métrica em S *induzida* por d. Observe em particular que se S é um subconjunto de $\mathbb{R}^n$ e d é a métrica Euclideana então $\big(S, d|_{S\times S}\big)$ é um espaço métrico.

Os dois exemplos acima são essencialmente os únicos que interessam nesse curso; uma pequena exceção: às vezes precisaremos também considerar o $\mathbb{R}^n$ (e, mais geralmente, espaços vetoriais reais de dimensão finita) munidos de métricas induzidas por normas diferentes da norma Euclideana (isso será discutido mais adiante). Veremos, no entanto, que todas as normas num espaço vetorial real de dimensão finita são *equivalentes* num sentido que será esclarecido depois, de modo que basicamente os dois exemplos acima são de fato os que interessam.

O espaço $\mathbb{R}^n$ é o produto de n cópias de $\mathbb{R}$. Mais geralmente, existe uma maneira mais ou menos canônica de tornar o produto de espaços métricos um espaço métrico. Decidimos introduzir essa construção mais geral abaixo:

Definição. *Sejam (M_i, d_i), $i = 1, \ldots, n$, espaços métricos. Definimos no produto cartesiano $M = M_1 \times \cdots \times M_n$ uma métrica d fazendo:*

$$d\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = \left(\sum_{i=1}^n d_i(x_i, y_i)^2\right)^{\frac{1}{2}},$$

para todos $x_i, y_i \in M_i$, $i = 1, \ldots, n$. A métrica d é chamada a métrica produto *em M.*

As propriedades (EM1) e (EM2) para d são imediatas. A desigualdade triangular pode ser mostrada facilmente usando a desigualdade triangular para a métrica Euclideana de $\mathbb{R}^n$ (veja Exercício 18). É possível também definir métricas computacionalmente mais simples no produto $M = \prod_{i=1}^n M_i$; tais métricas são *equivalentes* num sentido que será esclarecido depois. No momento, usaremos a métrica produto introduzida acima, de modo que a métrica Euclideana de $\mathbb{R}^n$ é obtida fazendo o produto de n cópias de $\mathbb{R}$.

Mesmo não sendo importante para esse curso, vamos mostrar dois exemplos de espaços métricos essencialmente diferentes de $\mathbb{R}^n$, só para cultura geral.

`Exemplo`. Se M é um conjunto qualquer então a aplicação $d : M \times M \to \mathbb{R}$ definida por $d(x, y) = 0$ para $x = y$ e $d(x, y) = 1$ para $x \neq y$ é uma métrica em M (verifique!). A métrica d é conhecida como *a métrica zero-um*.

`Exemplo`. Seja X um conjunto e seja $M = \mathcal{B}(X, \mathbb{R})$ o conjunto de todas as funções limitadas $f : X \to \mathbb{R}$ (uma função é *limitada* se existe $k \in \mathbb{R}$ com $|f(x)| \leq k$ para todo $x \in X$). Definimos uma métrica d em M fazendo:

$$d(f, g) = \sup_{x \in X} \big| f(x) - g(x) \big|,$$

para todas $f, g \in M$. A verificação das propriedades (EM1)–(EM3) não é difícil (para quem sabe trabalhar com sup).

(1) **Terminologia Básica da Teoria dos Espaços Métricos**.

Seja (M, d) um espaço métrico (às vezes a gente cansa de dizer (M, d) e diz só M, quando d está subentendido; mas cuidado, num mesmo conjunto M podemos definir muitas métricas). Vamos listar várias definições.

Definição. *Dado $x \in M$ e $r > 0$ então a* bola aberta *de centro x e raio r é definida por:*

$$\mathrm{B}(x; r) = \big\{ y \in M : d(x, y) < r \big\}$$

e a bola fechada *de centro x e raio r é definida por:*

$$\mathrm{B}[x; r] = \big\{ y \in M : d(x, y) \leq r \big\}.$$

Definição. *Sejam $A \subset M$ um subconjunto e $p \in M$ um ponto. Dizemos que:*

(i) p é um ponto interior *de A se existe $r > 0$ com $\mathrm{B}(p; r) \subset A$;*

(ii) p é um ponto exterior *de A se existe $r > 0$ com $\mathrm{B}(p; r) \subset A^{\mathrm{c}} = M \setminus A$, i.e., se existe $r > 0$ com $\mathrm{B}(p; r) \cap A = \emptyset$;*

(iii) p é um ponto de fronteira *de A se p não é um ponto interior nem um ponto exterior de A, i.e., se para todo $r > 0$ temos $\mathrm{B}(p; r) \cap A \neq \emptyset$ e $\mathrm{B}(p; r) \cap A^{\mathrm{c}} \neq \emptyset$;*

(iv) p é um ponto de aderência *de A se para todo $r > 0$ temos $\mathrm{B}(p; r) \cap A \neq \emptyset$;*

(v) p é um ponto de acumulação *de A se para todo $r > 0$ a bola $\mathrm{B}(p; r)$ contém pontos de A diferentes de p (p pode ou não pertencer a A).*

Definição. *Um subconjunto $A \subset M$ é dito:*

(i) aberto, *se todos os pontos de A são interiores;*

(ii) fechado, *se todo ponto de aderência de A está em A;*

(iii) denso, *se todo ponto de M é um ponto de aderência de A;*

(iv) limitado, *se A está contido em alguma bola de M.*

Definição. *Associados a um subconjunto $A \subset M$ estão os seguintes subconjuntos de M:*

(i) *o* interior *de A, denotado* $\text{int}(A) \subset M$*, é o conjunto dos pontos interiores de A;*

(ii) *o* fecho *de A, denotado* $\overline{A} \subset M$*, é o conjunto dos pontos de aderência de A;*

(iii) *o* exterior *de A, denotado* $\text{ext}(A) \subset M$*, é o conjunto dos pontos exteriores de A;*

(iv) *a* fronteira *de A, denotada $\partial A \subset M$, é o conjunto dos pontos de fronteira de A;*

(v) *o* conjunto derivado *de A, denotado A', é o conjunto dos pontos de acumulação de A.*

É difícil se acostumar com a terminologia acima de uma hora para outra. Para se familiarizar com isso tudo, o jeito é fazer algumas figuras com $M = \mathbb{R}^n$ $(n = 1, 2, 3)$ e fazer montes de exercícios para entender os inter-relacionamentos que existem entre todos esses termos. Várias sugestões legais (e fáceis) de exercícios estão listadas no final desta aula (vide Exercícios 1–13).

Teorema. *Se $x \in M$ é um ponto de acumulação de $A \subset M$ então para todo $r > 0$ a bola aberta* $\text{B}(x; r)$ *contém infinitos pontos de A.*

Demonstração. Suponha por absurdo que $\text{B}(x; r) \cap A$ é finito; daí existe $s > 0$ com $s < r$ e $s < d(x, y)$ para todo $y \in \text{B}(x; r) \cap A$, $y \neq x$. Segue que a interseção $\text{B}(x; s) \cap A$ não contém pontos diferentes de x, uma contradição. ■

A seguinte terminologia costuma ser útil também:

Definição. *Uma* vizinhança *de um ponto $x \in M$ é um subconjunto $V \subset M$ que contém um aberto que contém x; equivalentemente, $V \subset M$ é uma vizinhança de x se x é um ponto interior de V.*

`Observação`: Uma vizinhança de um ponto x *pode não ser aberta*, embora alguns (poucos) autores definam vizinhança de x como sendo um aberto que contém x (eu detesto isso!); no nosso curso *vizinhanças não precisam ser abertas* e quando eu quiser me referir a uma vizinhança que também é aberta vou dizer explicitamente *vizinhança aberta*.

Definição. *Seja (M, d) um espaço métrico. A coleção τ de todos os subconjuntos abertos de M é chamada a* topologia *de M.*

Definição. *Seja X um conjunto. Uma* topologia *no conjunto X é uma coleção τ de subconjuntos de X tal que as seguintes propriedades valem:*

(TP1) $\emptyset \in \tau$ *e* $X \in \tau$;

(TP2) *se $(U_i)_{i \in I}$ é uma família de elementos de τ então a união $\bigcup_{i \in I} U_i$ está em τ;*

(TP3) *se $U, V \in \tau$ então $U \cap V \in \tau$.*

Um espaço topológico *é um par (X, τ) onde X é um conjunto e τ é uma topologia em X. Os elementos de τ são chamados os* abertos *do espaço topológico (X, τ).*

A topologia associada a um espaço métrico de fato satisfaz as propriedades (TP1), (TP2) e (TP3) (veja o Exercício 15). Não temos interesse em estudar a teoria geral de espaços topológicos (na verdade, mesmo com respeito a espaços métricos só precisaremos de subespaços de $\mathbb{R}^n$ na prática); mas é bom ter em mente a noção geral de espaço topológico de qualquer forma.

Um conceito definido para espaços métricos (M, d) é dito um *conceito topológico* quando o mesmo pode ser redefinido usando apenas a topologia de (M, d). Por exemplo, "conjunto aberto", "conjunto fechado", "conjunto denso" são noções topológicas enquanto que "conjunto limitado" não é (isso será demonstrado posteriormente). Quando o enunciado de um teorema envolve apenas noções topológicas então esse enunciado faz sentido no contexto geral de espaços topológicos; isso *não* significa que tal teorema continue verdadeiro para espaços topológicos quaisquer (a maioria requer alguma hipótese adicional). Esse assunto não será importante no nosso curso.

(2) **Seqüências**.

No que segue (M, d) denota um espaço métrico.

Definição. *Seja $(x_n)_{n\geq 1}$ uma seqüência em M. Dizemos que $(x_n)_{n\geq 1}$* converge *para um ponto $x \in M$ se dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $d(x_n, x) < \varepsilon$ para todo $n \geq n_0$; escrevemos $\lim_{n\to\infty} x_n = x$ ou, mais resumidamente, $x_n \to x$.*

Nos exercícios pedimos para vocês mostrarem que o limite de uma seqüência é único (se existir).

Definição. *Dizemos que $(x_n)_{n\geq 1}$ é uma* seqüência de Cauchy *em M se dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $d(x_n, x_m) < \varepsilon$ para todos $n, m \geq n_0$.*

Nos exercícios pedimos para vocês mostrarem que toda seqüência convergente é de Cauchy.

`Terminologia`. Dizemos que alguma propriedade é satisfeita para todo número natural *n suficientemente grande* se existe $n_0 \in \mathbb{N}$ tal que a propriedade é satisfeita para todo $n \geq n_0$ (i.e., se a propriedade deixa de ser satisfeita no máximo para um número finito de n's). Dizemos que a propriedade é satisfeita para todo número natural *n arbitrariamente grande* se dado $n_0 \in \mathbb{N}$ então existe $n \geq n_0$ que satisfaz a propriedade (i.e., se existem infinitos n's para os quais a propriedade é satisfeita).

Na terminologia acima, uma seqüência $(x_n)_{n\geq 1}$ converge para x se e somente se dado $\varepsilon > 0$ então x_n fica na bola $\mathrm{B}(x; \varepsilon)$ para todo n suficientemente grande.

Definição. *Um espaço métrico M é dito* completo *quando toda seqüência de Cauchy em M é convergente.*

Teorema. *Sejam (M_i, d_i), $i = 1, \ldots, n$, espaços métricos e considere $M = \prod_{i=1}^n M_i$ munido da métrica produto d. Seja $(x_k)_{k\geq 1}$ uma seqüência em M e escreva $x_k = (x_k^1, \ldots, x_k^n)$ para todo $k \geq 1$. Então:*

(a) *$x_k \to x \in M$ se e somente se $x_k^i \to x^i$ em M_i para todo $i = 1, \ldots, n$ ("uma seqüência converge se e somente se converge coordenada por coordenada");*

(b) *$(x_k)_{k\geq 1}$ é de Cauchy em M se e somente se $(x_k^i)_{k\geq 1}$ é de Cauchy em M_i para todo $i = 1, \ldots, n$ ("uma seqüência é de Cauchy se e somente se for de Cauchy coordenada por coordenada").*

Demonstração. Segue facilmente observando que, para todo $\varepsilon > 0$ e todos $x, y \in M$:

$$\left(\sum_{i=1}^n d_i(x_i, y_i)^2\right)^{\frac{1}{2}} = d(x, y) < \varepsilon \Longrightarrow d_i(x_i, y_i) < \varepsilon, \quad i = 1, \ldots, n,$$

$$d_i(x_i,y_i) < \frac{\varepsilon}{\sqrt{n}},\ i=1,\ldots,n \Longrightarrow d(x,y) = \left(\sum_{i=1}^{n} d_i(x_i,y_i)^2\right)^{\frac{1}{2}} < \varepsilon. \ \blacksquare$$

Corolário. *O produto de espaços métricos completos é completo.* $\blacksquare$

Usando o fato que $I\!R$ é completo (vide qualquer livro de Análise Real) obtemos também:

Corolário. *$I\!R^n$ é completo.*

O fato que $I\!R^n$ é completo seguirá também quando mostrarmos mais adiante que todo limitado e fechado em $I\!R^n$ é compacto.

Os seguintes teoremas relacionam a topologia de M com a noção de limite de seqüência.

Teorema. *Sejam $A \subset M$ e $x \in A$. O ponto x está no interior de A se e somente se dada uma seqüência $(x_n)_{n\geq 1}$ convergente para x então $x_n \in A$ para todo n suficientemente grande. Em particular, A é aberto se e somente se para toda seqüência $(x_n)_{n\geq 1}$ que converge para um ponto de A temos $x_n \in A$ para todo n suficientemente grande.*

Demonstração. Se $x \in \mathrm{int}(A)$ então existe $\varepsilon > 0$ com $\mathrm{B}(x;\varepsilon) \subset A$. Como $x_n \to x$ então $x_n \in \mathrm{B}(x;\varepsilon)$ e portanto $x_n \in A$ para todo n suficientemente grande. Reciprocamente, suponha que toda seqüência que converge para x fica em A para n suficientemente grande. Se x não fosse um ponto interior de A então (negue a definição!) para todo $n \geq 1$, não pode ser $\mathrm{B}\big(x;\frac{1}{n}\big) \subset A$ e portanto existe $x_n \in \mathrm{B}\big(x;\frac{1}{n}\big)$ com $x_n \notin A$. Daí $x_n \to x$ mas $x_n \notin A$ para *todo* $n \geq 1$, uma contradição. $\blacksquare$

Teorema. *Sejam $A \subset M$ e $x \in M$. Então x é um ponto de aderência de A se e somente se x é o limite de uma seqüência de pontos de A. Em particular, A é fechado se e somente se dada uma seqüência $(x_n)_{n\geq 1}$ em A que converge para um ponto $x \in M$ então $x \in A$.*

Demonstração. Se x é o limite de uma seqüência $(x_n)_{n\geq 1}$ de pontos de A então dado $r > 0$ temos $x_n \in \mathrm{B}(x;r)$ para n suficientemente grande e em particular para tais n temos $x_n \in \mathrm{B}(x;r) \cap A$, donde $\mathrm{B}(x;r) \cap A \neq \emptyset$. Reciprocamente, se x é um ponto de aderência de A então dado $n \geq 1$ temos $\mathrm{B}\big(x;\frac{1}{n}\big) \cap A \neq \emptyset$, donde existe $x_n \in A$ com $d(x_n,x) < \frac{1}{n}$. Daí $x_n \to x$ e $x_n \in A$ para todo $n \geq 1$. $\blacksquare$

Corolário. *A é denso em M se e somente se todo ponto de M é limite de uma seqüência de pontos de A.* $\blacksquare$

Teorema. *Sejam $A \subset M$ e $x \in M$. Então x é ponto de acumulação de A se e somente se x é o limite de uma seqüência $(x_n)_{n\geq 1}$ de pontos de A,* todos distintos de *x.*

Demonstração. Se x é o limite de uma seqüência $(x_n)_{n\geq 1}$ de pontos de A distintos de x então dado $r > 0$ temos $x_n \in \mathrm{B}(x;r)$ para todo n suficientemente grande e em particular existem pontos em $\mathrm{B}(x;r) \cap A$ distintos de x. Reciprocamente, se x é um ponto de acumulação de A então dado $n \geq 1$ existe $x_n \in \mathrm{B}\big(x;\frac{1}{n}\big) \cap A$ com $x_n \neq x$. Daí $x_n \to x$.
$\blacksquare$

(3) **Limites de funções e funções contínuas**.

No que segue, (M, d) e (N, d') denotam espaços métricos.

Definição. *Seja A um subconjunto de M e $f : A \subset M \to N$ uma função. Dado um ponto de acumulação $x_0 \in M$ de A então dizemos que $L \in N$ é um* limite *de $f(x)$ quando x tende a x_0 se dado $\varepsilon > 0$ existe $\delta > 0$ tal que $d'\big(f(x), L\big) < \varepsilon$ para todo $x \in A$ tal que $0 < d(x, x_0) < \delta$; escrevemos $\lim_{x\to x_0} f(x) = L$.*

Nos exercícios pedimos para voces mostrarem que o limite $\lim_{x\to x_0} f(x)$ é único (quando existir).

Definição. *Seja $f : M \to N$ uma função e $x_0 \in M$ um ponto. Dizemos que f é* contínua *no ponto x_0 se dado $\varepsilon > 0$ existe $\delta > 0$ tal que $d'\big(f(x), f(x_0)\big) < \varepsilon$ para todo $x \in M$ com $d(x, x_0) < \delta$.*

Para quem estiver usando um pouco de senso crítico enquanto lê essas notas, uma pergunta imediatamente surge à cabeça: "por que será que para definir limite ele considera f definida num subconjunto de M e para definir continuidade ele considera f definida em todo espaço M? E se eu precisasse falar de continuidade de uma função que está definida só num subconjunto de M?" Pois é, isso tudo tem uma explicação. Se eu precisar falar em continuidade de uma função f definida só num subconjunto A de M então eu posso simplesmente lembrar que A *também é um espaço métrico* com a métrica induzida de M: daí recaímos outra vez no caso de uma função definida num espaço métrico tomando valores em outro espaço métrico! No caso de limite ocorre um problema com esse raciocínio: é importante também definir limite de $f(x)$ quando x tende a pontos *fora do domínio de f*; para dar uma definição de limite que cobre essa situação precisamos pensar no domínio A de f como um subconjunto de um espaço métrico M possivelmente maior.

Nos exercícios pedimos para vocês provarem a relação familiar entre limite e continuidade.

Definição. *Uma função $f : M \to N$ é dita um* homeomorfismo *se f é bijetora e tanto f como f^{-1} são contínuas.*

Teorema. *Sejam $f : A \subset M \to N$, $g : N \to P$ funções e $x_0 \in M$ um ponto de acumulação de A; suponha que $\lim_{x\to x_0} f(x) = L \in N$ e que g seja contínua no ponto L. Então:*

$$\lim_{x\to x_0} g\big(f(x)\big) = g(L).$$

Demonstração. Dado $\varepsilon > 0$, existe $\eta > 0$ tal que $g\big(\mathrm{B}(L; \eta)\big) \subset \mathrm{B}\big(g(L); \varepsilon\big)$, pela continuidade de g. A partir de $\eta > 0$ encontramos $\delta > 0$ tal que $f(x) \in \mathrm{B}(L; \eta)$, sempre que $x \in \mathrm{B}(x_0; \delta) \cap A$ e $x \neq x_0$. Daí $g\big(f(x)\big) \in \mathrm{B}\big(g(L); \varepsilon\big)$ sempre que $x \in \mathrm{B}(x_0; \delta) \cap A$ e $x \neq x_0$. ■

Compare o teorema acima com o Exercício 29.

Teorema. *Se $f : M \to N$ é contínua em $x_0 \in M$ e $g : N \to P$ é contínua em $f(x_0) \in N$ então $g \circ f$ é contínua em x_0 ("a composta de funções contínuas é contínua").*

Demonstração. Dado $\varepsilon > 0$, existe $\eta > 0$ tal que $g\big(\mathrm{B}(f(x_0); \eta)\big) \subset \mathrm{B}\big(g(f(x_0)); \varepsilon\big)$, pela continuidade de g. Pela continuidade de f, achamos $\delta > 0$ com $f\big(\mathrm{B}(x_0; \delta)\big) \subset \mathrm{B}\big(f(x_0); \eta\big)$. Daí $(g \circ f)\big(\mathrm{B}(x_0; \delta)\big) \subset \mathrm{B}\big(g(f(x_0)); \varepsilon\big)$. ■

Teorema. *Sejam (N_i, d_i), $i = 1, \ldots, n$, espaços métricos e considere $N = \prod_{i=1}^n N_i$ munido da métrica produto d. Seja $f : A \subset M \to N$ uma função e seja $x_0 \in M$. Escreva $f = (f^1, \ldots, f^n)$ com cada $f^i : M \to N_i$. Então:*

(a) se x_0 é um ponto de acumulação de A temos $\lim_{x\to x_0} f(x) = L \in N$ se e somente se $\lim_{x\to x_0} f^i(x) = L^i \in N_i$ para todo $i = 1, \ldots, n$ ("limites são feitos coordenada por coordenada");

(b) se $x_0 \in A$ então f é contínua em x_0 se e somente se f^i é contínua em x_0 para $i = 1, \ldots, n$ ("continuidade é feita coordenada por coordenada").

Demonstração. Segue facilmente observando que, para todo $\varepsilon > 0$ e todo $x \in A$:

$$\left(\sum_{i=1}^n d_i\big(f^i(x), L^i\big)^2\right)^{\frac{1}{2}} = d\big(f(x), L\big) < \varepsilon \Longrightarrow d_i(f^i(x), L^i) < \varepsilon, \quad i = 1, \ldots, n,$$

$$d_i\big(f^i(x), L^i\big) < \frac{\varepsilon}{\sqrt{n}},\ i = 1, \ldots, n \Longrightarrow d\big(f(x), L\big) = \left(\sum_{i=1}^n d_i\big(f^i(x), L^i\big)^2\right)^{\frac{1}{2}} < \varepsilon. \blacksquare$$

Corolário. *(propriedades operatórias dos limites) Sejam $f : A \subset M \to I\!R^n$, $g : A \subset M \to I\!R^n$ e $c : A \subset M \to I\!R$ funções. Temos:*

(a) se $x_0 \in M$ é um ponto de acumulação de A e f, g, c admitem limite quando $x \to x_0$ então:

$$\lim_{x\to x_0} \big[f(x) + g(x)\big] = \lim_{x\to x_0} f(x) + \lim_{x\to x_0} g(x), \quad \lim_{x\to x_0} \big[c(x)f(x)\big] = \lim_{x\to x_0} c(x) \lim_{x\to x_0} f(x);$$

(b) se f, g e c são contínuas num ponto $x_0 \in A$ então $f + g$ e cf são contínuas no ponto x_0.

Demonstração. Segue dos teoremas anteriores, observando que as funções soma e produto:

$$I\!R^n \times I\!R^n \ni (v, w) \longmapsto v + w \in I\!R^n, \quad I\!R \times I\!R^n \ni (c, v) \longmapsto cv \in I\!R^n,$$

são contínuas (voltaremos a esse assunto depois). $\blacksquare$

Vamos agora relacionar limite e continuidade de funções com limites de seqüências.

Teorema. *Sejam $f : A \subset M \to N$ uma função e $x_0 \in M$ um ponto de acumulação de A. Temos $\lim_{x\to x_0} f(x) = L \in N$ se e somente se vale a seguinte propriedade: para toda seqüência $(x_n)_{n\ge1}$ em A com $x_n \to x_0$ e $x_n \ne x_0$ para todo n temos $f(x_n) \to L$.*

Demonstração. Suponha que $\lim_{x\to x_0} f(x) = L$ e seja $(x_n)_{n\ge1}$ uma seqüência em A com $x_n \to x_0$ e $x_n \ne x_0$ para todo n. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $d'\big(f(x), L\big) < \varepsilon$ para todo $x \in A$ com $0 < d(x, x_0) < \delta$; a partir de $\delta > 0$ achamos $n_0 \in I\!N$ com $d(x_n, x_0) < \delta$ para todo $n \ge n_0$. Concluímos que $d'\big(f(x_n), L\big) < \varepsilon$ para todo $n \ge n_0$.

Reciprocamente, se a propriedade envolvendo seqüências que aparece no enunciado é satisfeita, vamos mostrar que $\lim_{x\to x_0} f(x) = L$. Se não fosse, existiria um $\varepsilon > 0$ tal que para todo $n \ge 1$, tomando $\delta = \frac{1}{n}$, existe $x_n \in A$ com $0 < d(x_n, x_0) < \frac{1}{n}$ mas $d'\big(f(x_n), L\big) \ge \varepsilon$. Daí $(x_n)_{n\ge1}$ é uma seqüência em A com $x_n \to x_0$, $x_n \ne x_0$ para todo n, mas $f(x_n) \not\to L$. $\blacksquare$

Teorema. *Seja $f : M \to N$ uma função e $x_0 \in M$ um ponto. Temos que f é contínua no ponto x_0 se e somente se vale a seguinte propriedade: dada uma seqüência $(x_n)_{n\geq 1}$ em M com $x_n \to x_0$ então $f(x_n) \to f(x_0)$.*

Demonstração. Suponha que f é contínua no ponto x_0 e seja $(x_n)_{n\geq 1}$ uma seqüência em M com $x_n \to x_0$. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $d\big(f(x), f(x_0)\big) < \varepsilon$ sempre que $d(x, x_0) < \delta$; a partir de $\delta > 0$ encontramos $n_0 \in \mathbb{N}$ com $d(x_n, x_0) < \delta$ sempre que $n \geq n_0$. Daí $d\big(f(x_n), f(x_0)\big) < \varepsilon$ sempre que $n \geq n_0$.

Reciprocamente, se a propriedade envolvendo seqüências que aparece no enunciado é satisfeita, vamos mostrar que f é contínua no ponto x_0. Se não fosse, existiria $\varepsilon > 0$ tal que para todo $n \geq 1$, tomando $\delta = \frac{1}{n}$, existiria $x_n \in M$ com $d(x_n, x_0) < \frac{1}{n}$ mas $d'\big(f(x_n), f(x_0)\big) \geq \varepsilon$. Daí $(x_n)_{n\geq 1}$ é uma seqüência em M com $x_n \to x_0$ e $f(x_n) \not\to f(x_0)$. ■

Corolário. *Se $x_k \to x$, $y_k \to y$ em $\mathbb{R}^n$ e se $c_k \to c$ em $\mathbb{R}$ então $x_k + y_k \to x + y$ e $c_k x_k \to cx$.*

Demonstração. Segue do fato que as funções soma $\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto x + y \in \mathbb{R}^n$ e produto $\mathbb{R} \times \mathbb{R}^n \ni (c, x) \mapsto cx \in \mathbb{R}^n$ são contínuas (isso será revisto depois), do fato que $(x_k, y_k) \to (x, y)$ se e somente se $x_k \to x$ e $y_k \to y$ e do fato que $(c_k, x_k) \to (c, x)$ se e somente se $c_k \to c$ e $x_k \to x$. ■

Teorema. *As seguintes propriedades são equivalentes sobre uma função $f : M \to N$:*

(a) f é contínua;
(b) para todo $U \subset N$ aberto, $f^{-1}(U) \subset M$ é aberto;
(c) para todo $F \subset N$ fechado, $f^{-1}(F) \subset M$ é fechado.

Demonstração. A equivalência entre (b) e (c) é um exercício usando o fato que U é aberto se e somente se U^{c} é fechado, juntamente com o fato que $f^{-1}(U^{\mathrm{c}}) = f^{-1}(U)^{\mathrm{c}}$.

Provemos (a) $\Rightarrow$ (b). Suponha que f é contínua; dado $U \subset N$ aberto, mostremos que $f^{-1}(U) \subset M$ é aberto. Seja $x \in f^{-1}(U)$. Daí $f(x) \in U$ e como U é aberto, existe $\varepsilon > 0$ com $\mathrm{B}\big(f(x); \varepsilon\big) \subset U$. Como f é contínua, existe $\delta > 0$ com $f\big(\mathrm{B}(x; \delta)\big) \subset \mathrm{B}\big(f(x); \varepsilon\big) \subset U$. Daí $\mathrm{B}(x; \delta) \subset f^{-1}(U)$.

Provemos (b) $\Rightarrow$ (a). Suponha (b) e mostremos que f é contínua num certo ponto $x \in M$. Dado $\varepsilon > 0$, então $U = \mathrm{B}\big(f(x); \varepsilon\big)$ é aberto em N e portanto $f^{-1}(U)$ é aberto em M. Como $x \in U$, existe $\delta > 0$ com $\mathrm{B}(x; \delta) \subset f^{-1}(U)$. Daí $f\big(\mathrm{B}(x; \delta)\big) \subset \mathrm{B}\big(f(x); \varepsilon\big)$. ■

Corolário. *Uma função bijetora $f : M \to N$ é um homeomorfismo se e somente se valem uma das seguintes propriedades equivalentes:*

(i) para todo $U \subset M$, U é aberto em $M \Leftrightarrow f(U)$ é aberto em N;
(ii) para todo $F \subset M$, F é fechado em $M \Leftrightarrow f(F)$ é fechado em N. ■

Corolário. *Sejam $f, g : M \to \mathbb{R}$ funções contínuas. Os conjuntos:*

$$\big\{x \in M : f(x) < g(x)\big\}, \qquad \big\{x \in M : f(x) \neq g(x)\big\}$$

são abertos e os conjuntos:

$$\big\{x \in M : f(x) \leq g(x)\big\}, \qquad \big\{x \in M : f(x) = g(x)\big\}$$

são fechados.

Demonstração. A função $h = f - g : M \to I\!R$ é contínua. Os dois primeiros conjuntos são $h^{-1}\big((-\infty, 0)\big)$ e $h^{-1}\big(I\!R \setminus \{0\}\big)$ e os dois últimos são $h^{-1}\big((-\infty, 0]\big)$ e $h^{-1}(0)$. ■

Definição. *Uma função $f : M \to N$ é dita* uniformemente contínua *quando dado $\varepsilon > 0$, existe $\delta > 0$ tal que, para todos $x, y \in M$, se $d(x, y) < \delta$ então $d'\big(f(x), f(y)\big) < \varepsilon$.*

Obviamente toda função uniformemente contínua é contínua.

Teorema. *Se $f : (M, d) \to (N, d')$ e $g : (N, d') \to (P, d'')$ são uniformemente contínuas então $g \circ f$ é uniformemente contínua.*

Demonstração. Dado $\varepsilon > 0$, existe $\eta > 0$ tal que $d'(u, v) < \eta$ implica $d''\big(g(u), g(v)\big) < \varepsilon$; a partir de $\eta > 0$ encontramos $\delta > 0$ tal que $d(x, y) < \delta$ implica $d'\big(f(x), f(y)\big) < \eta$. Daí $d(x, y) < \delta$ implica $d''\big((g \circ f)(x), (g \circ f)(y)\big) < \varepsilon$. ■

Teorema. *Funções uniformemente contínuas levam seqüências de Cauchy em seqüências de Cauchy, i.e., se $f : M \to N$ é uniformemente contínua e $(x_n)_{n \geq 1}$ é uma seqüência de Cauchy em M então $\big(f(x_n)\big)_{n \geq 1}$ é uma seqüência de Cauchy em N.*

Demonstração. Dado $\varepsilon > 0$, existe $\delta > 0$ tal que $d(x, y) < \delta$ implica $d'\big(f(x), f(y)\big) < \varepsilon$; a partir de $\delta > 0$ encontramos $n_0 \in I\!N$ tal que $n, m \geq n_0$ implica $d(x_n, x_m) < \delta$. Daí $d'\big(f(x_n), f(y_n)\big) < \varepsilon$ sempre que $n, m \geq n_0$. ■

Definição. *Uma função $f : M \to N$ é dita* Lipschitziana *quando existe $k \geq 0$ tal que $d'\big(f(x), f(y)\big) \leq k\, d(x, y)$ para todos $x, y \in M$; dizemos então que k é uma* constante de Lipschitz *para f. Quando f é Lipschitziana com constante $k = 1$ dizemos que f é uma* contração fraca *e quando f for Lipschitziana com constante $k < 1$ dizemos que f é uma* contração.

Teorema. *Toda função Lipschitziana é uniformemente contínua.*

Demonstração. Se $k > 0$ é uma constante de Lipschitz para f então, dado $\varepsilon > 0$, tome $\delta = \frac{\varepsilon}{k}$ (complete!). ■

`Exemplo`. Uma função afim $f : I\!R \to I\!R$, $f(x) = ax + b$ é sempre uniformemente contínua, pois é Lipschitziana (com constante $|a|$). A função $f : I\!R \to I\!R$, $f(x) = x^2$ é contínua mas não é uniformemente contínua; de fato, dado $\delta > 0$ então $|f(x + \delta) - f(x)| = |2\delta x + \delta^2|$ e essa quantidade torna-se arbitrariamente grande quando $x \to +\infty$.

`Exemplo`. As projeções $\pi_j : \prod_{i=1}^n M_i \to M_j$, $j = 1, \ldots, n$, de um produto cartesiano são contrações fracas e portanto uniformemente contínuas.

Teorema. *Sejam (N_i, d_i), $i = 1, \ldots, n$, espaços métricos e considere $N = \prod_{i=1}^n N_i$ munido da métrica produto. Uma função $f : M \to N$ é uniformemente contínua se e somente se cada coordenada $f^i : M \to N_i$ é uniformemente contínua, $i = 1, \ldots, n$.*

Demonstração. É similar à demonstração de que f é contínua se e somente se cada f^i é contínua. ■

Definição. *Uma aplicação* $f : (M,d) \to (N,d')$ *tal que* $d'\big(f(x), f(y)\big) = d(x,y)$ *para todos* $x, y \in M$ *é chamada uma* imersão isométrica*; quando* f *é uma imersão isométrica bijetora dizemos então que* f *é uma* isometria.

`Observação`. É fácil ver que toda imersão isométrica é injetora, de modo que *uma isometria é o mesmo que uma imersão isométrica sobrejetora*.

`Exemplo`. Se $f : (M,d) \to (N,d')$ é uma isometria então f é um homeomorfismo; na verdade, f e f^{-1} são ambas *uniformemente contínuas*. De fato, se f é uma isometria então também f^{-1} é uma isometria e isometrias são Lipschitzianas.

Definição. *Uma função* $f : M \to N$ *é dita* limitada *quando sua imagem* $f(M)$ *é um subconjunto limitado de* N*, i.e., quando existe* $c \geq 0$ *tal que* $d\big(f(x), f(y)\big) \leq c$ *para todos* $x, y \in M$.

(4) **Métricas equivalentes**.

Definição. *Duas métricas* d_1 *e* d_2 *num conjunto* M *são ditas* equivalentes *quando induzem a mesma topologia em* M*, i.e., quando* U *é aberto em* (M, d_1) *se e somente se* U *é aberto em* (M, d_2).

Teorema. *Duas métricas* d_1 *e* d_2 *em* M *são equivalentes se e somente se a aplicação identidade* $\mathrm{Id} : (M, d_1) \to (M, d_2)$ *é um homeomorfismo.*

Demonstração. A afirmação "a identidade $\mathrm{Id} : (M, d_1) \to (M, d_2)$ é um homeomorfismo" é equivalente à afirmação "U é aberto em (M, d_1) se e somente se $\mathrm{Id}(U) = U$ é aberto em (M, d_2)". ■

`Exemplo`. Se $h : (M,d) \to (N,d')$ é um homeomorfismo então a métrica $\tilde{d}$ em M definida por $\tilde{d}(x,y) = d'\big(h(x), h(y)\big)$ é equivalente a d. De fato, temos um diagrama comutativo:

$$\begin{tikzcd}
& (N,d') \\
(M,d) \arrow[ur, "h"] \arrow[r, "\mathrm{Id}"'] & (M,\tilde{d}) \arrow[u, "h"']
\end{tikzcd}$$

onde a flecha inclinada é um homeomorfismo e a flecha vertical é uma isometria (veja o Exercício 31); segue que a flecha horizontal é um homeomorfismo.

Observe que trocando uma métrica por outra equivalente não prejudicamos noções topológicas do espaço, i.e., noções que podem ser definidas usando apenas a topologia. Temos então o seguinte:

Teorema. *Sejam* d_1, d_2 *métricas equivalentes em* M*. Então, para todo* $A \subset M$ *temos:*

(i) A *é aberto em* (M, d_1) *se e somente se* A *é aberto em* (M, d_2)*;*

(ii) A *é fechado em* (M, d_1) *se e somente se* A *é fechado em* (M, d_2)*;*

(iii) os pontos interiores, de fronteira, de aderência e de acumulação de A *no espaço* (M, d_1) *coincidem respectivamente com os pontos interiores, de fronteira, de aderência e de acumulação de* A *no espaço* (M, d_2)*;*

(iv) A é denso em (M, d_1) se e somente se A é denso em (M, d_2). ■

Limites de funções, continuidade de funções e limites de seqüência também são noções topológicas: afirmações como $\lim_{x\to x_0} f(x) = L$, $x_n \to x$ e "f é contínua no ponto x_0" continuam verdadeiras quando trocamos as métricas nos espaços envolvidos por métricas equivalentes. Noções como completude, seqüências de Cauchy, conjuntos limitados, continuidade uniforme *não* são preservadas quando trocamos uma métrica por outra equivalente (vide exemplo a seguir).

`Exemplo`. A *função arco–tangente* $\operatorname{arc\,tg} : \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ é um homeomorfismo. Segue então que a métrica d em $\mathbb{R}$ definida por $d(x, y) = |\operatorname{arc\,tg} x - \operatorname{arc\,tg} y|$ é equivalente à métrica Euclideana. Observe que $\mathbb{R}$ é limitado (com diâmetro π) com a métrica d, mas não é limitado com a métrica Euclideana. Além do mais, $\mathbb{R}$ é completo com a métrica Euclideana, mas não é completo com a métrica d; de fato, se $(x_n)_{n\geq1}$ é uma seqüência em $\mathbb{R}$ tal que $\operatorname{arc\,tg} x_n \to \frac{\pi}{2}$ então $(x_n)_{n\geq1}$ é de Cauchy (pois $(\operatorname{arc\,tg} x_n)_{n\geq1}$ é de Cauchy em $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$), mas $(x_n)_{n\geq1}$ não é convergente em $\mathbb{R}$.

A noção de equivalência mais forte introduzida a seguir garante a invariância de noções como completude e seqüências de Cauchy.

Definição. *Duas métricas d_1 e d_2 são ditas* uniformemente equivalentes *quando as aplicações identidade* $\mathrm{Id} : (M, d_1) \to (M, d_2)$ e $\mathrm{Id} : (M, d_2) \to (M, d_1)$ *são ambas uniformemente contínuas.*

`Observação`. Uma condição suficiente para que d_1 e d_2 sejam uniformemente equivalentes é que as aplicações identidade $\mathrm{Id} : (M, d_1) \to (M, d_2)$ e $\mathrm{Id} : (M, d_2) \to (M, d_1)$ sejam Lipschitzianas; isso equivale à existência de constantes $\alpha, \beta > 0$ tais que:

$$\alpha\, d_1(x, y) \leq d_2(x, y) \leq \beta\, d_1(x, y),$$

para todos $x, y \in M$.

Teorema. *Sejam d_1 e d_2 métricas uniformemente equivalentes em M. Então:*

(i) uma seqüência $(x_n)_{n\geq1}$ é de Cauchy em (M, d_1) se e somente se $(x_n)_{n\geq1}$ for de Cauchy em (M, d_2);

(ii) (M, d_1) é completo se e somente se (M, d_2) é completo.

Demonstração. Segue do fato que funções uniformemente contínuas levam seqüências de Cauchy em seqüências de Cauchy. ■

É fácil também mostrar que uma função uniformemente contínua mantêm-se uniformemente contínua quando trocamos as métricas do domínio ou do contra-domínio por métricas uniformemente equivalentes (isso segue do fato que a composta de funções uniformemente contínuas é uniformemente contínua).

(5) **Subespaços e topologia induzida**.

Como foi mencionado no início da aula, se (M, d) é um espaço métrico e $S \subset M$ é um subconjunto então $\left(S, d|_{S\times S}\right)$ também é um espaço métrico; podemos então de maneira natural pensar num subconjunto de um espaço métrico como sendo também um espaço métrico. O nosso objetivo agora é estabelecer algumas relações entre M e S (principalmente de natureza topológica).

Teorema. *Se $U \subset M$ é aberto em M então $U \cap S$ é aberto em S. Além do mais, dado $Z \subset S$ aberto em S então existe $U \subset M$ aberto em M com $Z = U \cap S$ ("os abertos de S são os abertos de M interceptados com S").*

Antes de provar o teorema acima, fazemos uma convenção: para $x \in S$ e $r > 0$ denotamos por $\mathrm{B}(x;r)$ e $\mathrm{B}[x;r]$ respectivamente a bola aberta e a bola fechada de centro x e raio r *no espaço métrico* (M,d); nesta seção, precisaremos também nos referir às bolas aberta e fechada de centro x e raio $r > 0$ *no espaço métrico* $\big(S, d|_{S\times S}\big)$ — as últimas serão denotadas por:

$$\begin{aligned} \mathrm{B}_S(x;r) &= \big\{y \in S : d(x,y) < r\big\}, \\ \mathrm{B}_S[x;r] &= \big\{y \in S : d(x,y) \leq r\big\}, \end{aligned}.$$

A seguinte observação é trivial:

$$\mathrm{B}_S(x;r) = \mathrm{B}(x;r) \cap S, \qquad \mathrm{B}_S[x;r] = \mathrm{B}[x;r] \cap S.$$

Demonstração do Teorema. Se $U \subset M$ é aberto, mostremos que $U \cap S$ é aberto. Seja $x \in U \cap S$; como U é aberto em M, existe $r > 0$ com $\mathrm{B}(x;r) \subset U$. Daí $\mathrm{B}_S(x;r) \subset U \cap S$. Reciprocamente, suponha que $Z \subset S$ é aberto em S. Para cada $x \in Z$ existe então $r_x > 0$ tal que $\mathrm{B}_S(x;r_x) = \mathrm{B}(x;r_x) \cap S \subset Z$. Daí $U = \bigcup_{x \in Z} \mathrm{B}(x;r_x)$ é aberto em M e $Z = U \cap S$. ∎

Corolário. *Se S é aberto em M então os abertos de S são os abertos de M que estão contidos em S.* ∎

Teorema. *Se $F \subset M$ é fechado em M então $F \cap S$ é fechado em S. Além do mais, dado $H \subset S$ fechado em S então existe $F \subset M$ fechado em M com $H = F \cap S$ ("os fechados de S são os fechados de M interceptados com S").*

Demonstração. Se $F \subset M$ é fechado em M então F^{c} é aberto em M e $F^{\mathrm{c}} \cap S$ é aberto em S pelo teorema anterior. Daí $S \setminus (F^{\mathrm{c}} \cap S) = F \cap S$ é fechado em S. Reciprocamente, suponha que $H \subset S$ é fechado em S. Daí $S \setminus H$ é aberto em S e portanto existe $U \subset M$ aberto em M com $S \setminus H = U \cap S$, pelo teorema anterior. Temos então que $F = U^{\mathrm{c}}$ é fechado em M e $H = F \cap S$. ∎

Corolário. *Se S é fechado em M então os fechados de S são os fechados de M que estão contidos em S.* ∎

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Espaços métricos (terminologia básica).`

1. Seja (M, d) um espaço métrico. Mostre que se $r_1, r_2 > 0$ e $x_1, x_2 \in M$ são tais que $r_1 + r_2 < d(x_1, x_2)$ então as bolas fechadas $\mathrm{B}[x_1; r_1]$ e $\mathrm{B}[x_2; r_2]$ são disjuntas. Se $r_1 + r_2 \leq d(x_1, x_2)$ então mostre que as bolas abertas $\mathrm{B}(x_1; r_1)$ e $\mathrm{B}(x_2; r_2)$ são disjuntas.
2. Para um espaço métrico arbitrário (M, d) e um subconjunto $A \subset M$ mostre que:
   (a) $M = \mathrm{int}(A) \cup \partial A \cup \mathrm{ext}(A)$ e essa união é disjunta;
   (b) $\mathrm{int}(A) \subset A$ e $\mathrm{ext}(A) \subset A^{\mathrm{c}}$;
   (c) $A \cap \partial A = A \setminus \mathrm{int}(A)$;
   (d) $\overline{A} = A \cup \partial A = A \cup A'$;
   (e) $\mathrm{ext}(A) = \overline{A}^{\mathrm{c}}$;
   (f) $\mathrm{int}(A)^{\mathrm{c}} = \overline{A^{\mathrm{c}}}$.
3. Mostre que as seguintes condições são equivalentes sobre um subconjunto A de um espaço métrico M:
   (a) A é aberto;
   (b) $A = \mathrm{int}(A)$;
   (c) para todo $x \in A$ existe $r > 0$ com $B(x; r) \subset A$;
   (d) para todo $x \in A$ existe $r > 0$ com $B[x; r] \subset A$.
4. Mostre que as seguintes condições são equivalentes sobre um subconjunto A de um espaço métrico M:
   (a) A é fechado;
   (b) $A = \overline{A}$;
   (c) para todo $x \in M$, se $x \notin A$ então existe $r > 0$ com $B(x; r) \cap A = \emptyset$;
   (d) para todo $x \in M$, se $x \notin A$ então existe $r > 0$ com $B[x; r] \cap A = \emptyset$;
   (e) o complementar de A é aberto (logo "ser fechado" é uma noção topológica);
   (f) $\partial A \subset A$;
   (g) $A' \subset A$;
5. Mostre que as seguintes condições são equivalentes sobre um subconjunto A de um espaço métrico M:
   (a) A é denso;
   (b) $\overline{A} = M$;
   (c) todo aberto não vazio de M intercepta A (logo "ser denso" é uma noção topológica);
   (d) toda bola aberta de M intercepta A;
   (e) toda bola fechada de M intercepta A;
   (f) dado $x \in M$ e $\varepsilon > 0$ existe $y \in A$ com $d(x, y) < \varepsilon$;
   (g) $\mathrm{int}(A^{\mathrm{c}}) = \emptyset$.

6. Sejam (M,d) um espaço métrico e $A \subset M$ um subconjunto. O *diâmetro* de A é definido por:

$$\operatorname{diam}(A) = \sup_{x,y \in A} d(x,y) \in [0,+\infty].$$

   Mostre que A é limitado se e somente se $\operatorname{diam}(A) < +\infty$.

7. Mostre que toda bola aberta é um conjunto aberto e que toda bola fechada é um conjunto fechado (ninguém acha que apenas os nomes "bola aberta" e "bola fechada" implicam automaticamente que esses conjuntos sejam abertos ou fechados, certo?).

8. Se A é um subconjunto de um espaço métrico M mostre que:

   (a) $\operatorname{int}(A)$ e $\operatorname{ext}(A)$ são abertos;

   (b) ∂A e $\overline{A}$ são fechados;

   (c) $\operatorname{int}\big(\operatorname{int}(A)\big) = \operatorname{int}(A)$ e $\overline{\overline{A}} = \overline{A}$.

9. Seja A um subconjunto de um espaço métrico M. Mostre que $\operatorname{int}(A)$ é o *maior aberto contido em* A, i.e., que $\operatorname{int}(A)$ é aberto e que se $U \subset M$ é um aberto contido em A então $U \subset \operatorname{int}(A)$.

10. Seja A um subconjunto de um espaço métrico M. Mostre que $\overline{A}$ é o *menor fechado que contém* A, i.e., que $\overline{A}$ é fechado e que se $F \subset M$ é um fechado que contém A então $\overline{A} \subset F$.

11. Dados subconjuntos A e B de um espaço métrico M, mostre que $A \subset B \Rightarrow \overline{A} \subset \overline{B}$.

12. Dado um espaço métrico M, um subconjunto $A \subset M$ e um aberto $U \subset M$ mostre que $U \cap A \neq \emptyset$ se e somente se $U \cap \overline{A} \neq \emptyset$.

13. Sobre um espaço métrico M, um subconjunto $A \subset M$ e um ponto $x \in M$ mostre que:

   (a) x é um ponto interior de A se e somente se x pertence a um aberto contido em A (logo "ponto interior" é uma noção topológica);

   (b) x é um ponto de aderência de A se e somente se todo aberto contendo x intercepta A (logo "ponto de aderência é uma noção topológica");

   (c) x é um ponto de fronteira de A se e somente se todo aberto que contém x possui pontos de A e pontos fora de A (logo "ponto de fronteira" é uma noção topológica);

   (d) x é um ponto de acumulação de A se e somente se todo aberto que contém x intercepta A num ponto diferente de x (logo "ponto de acumulação" é uma noção topológica).

14. Sobre um espaço métrico M, um subconjunto $A \subset M$ e um ponto $x \in M$ mostre que:
    (a) x é um ponto interior de A se e somente se x possui uma vizinhança contida em A;
    (b) x é um ponto de aderência de A se e somente se toda vizinhança de x intercepta A;
    (c) x é um ponto de fronteira de A se e somente se toda vizinhaça de x possui pontos de A e pontos fora de A;
    (d) x é um ponto de acumulação de A se e somente se toda vizinhança de x intercepta A num ponto diferente de x;
    (e) A é aberto se e somente se todo ponto de A possui uma vizinhança contida em A;
    (f) A é fechado se e somente se todo ponto fora de A possui uma vizinhança disjunta de A.

15. Sobre um espaço métrico M mostre que:
    (a) O conjunto vazio e o próprio espaço M são abertos e fechados;
    (b) A união de uma família arbitrária de abertos é um aberto;
    (c) A interseção de dois (ou, mais geralmente, de uma família *finita*) de abertos é aberto;
    (d) A interseção de uma família arbitrária (não vazia) de fechados é um fechado;
    (e) A união de dois (ou, mais geralmente, de uma família *finita*) de fechados é fechado.

16. Encontre em $I\!R^2$ uma família enumerável de abertos cuja interseção não é aberta e uma família enumerável de fechados cuja união não é fechada.

17. Sobre um espaço métrico M e subconjuntos $A, B \subset M$ mostre que:
    (a) $\overline{A \cup B} = \overline{A} \cup \overline{B}$;
    (b) $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ (dê um contra-exemplo para a igualdade com $M = I\!R^2$);
    (c) $\mathrm{int}(A \cap B) = \mathrm{int}(A) \cap \mathrm{int}(B)$;
    (d) $\mathrm{int}(A) \cup \mathrm{int}(B) \subset \mathrm{int}(A \cup B)$ (dê um contra-exemplo para a igualdade com $M = I\!R^2$).

18. Sejam (M_i, d_i), $i = 1, \ldots, n$, espaços métricos e seja $M = \prod_{i=1}^n M_i$. Defina uma aplicação $d : M \times M \to I\!R$ fazendo:

$$d(x, y) = \left(\sum_{i=1}^n d_i(x_i, y_i)^2 \right)^{\frac{1}{2}},$$

para $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n \in M$. Mostre que d satisfaz a desigualdade triangular (*dica*: para $x, y \in M$ denote por $\alpha_{xy} \in I\!R^n$ o vetor cuja i-ésima coordenada é $d_i(x_i, y_i)$; observe que $d(x, y) = \|\alpha_{xy}\|$, onde $\|\cdot\|$ denota a norma Euclideana em $I\!R^n$. Use a desigualdade triangular para a norma Euclideana).

19. Seja (M, d) um espaço métrico. Um *sistema fundamental de vizinhanças* para um ponto $x \in M$ é um conjunto $\mathcal{V}$ de vizinhanças de x tal que toda vizinhança de x contém uma vizinhança de x pertencente a $\mathcal{V}$. Quando todos os elementos de $\mathcal{V}$ são abertos, dizemos que $\mathcal{V}$ é um *sistema fundamental de vizinhanças abertas* de x. Mostre que:
    (a) $\mathcal{V} = \big\{\mathrm{B}[x; r] : r > 0\big\}$ é um sistema fundamental de vizinhanças de x e $\mathcal{V} = \big\{\mathrm{B}(x; r) : r > 0\big\}$ é um sistema fundamental de vizinhanças abertas de x;
    (b) $\mathcal{V} = \big\{\mathrm{B}\big[x; \frac{1}{n}\big] : n \in I\!N\big\}$ é um sistema fundamental de vizinhanças de x e $\mathcal{V} = \big\{\mathrm{B}\big(x; \frac{1}{n}\big) : n \in I\!N\big\}$ é um sistema fundamental de vizinhanças abertas de x;
    (c) para cada $x \in M$ seja $\mathcal{V}_x$ um sistema fundamental de vizinhanças de x. Reprove os itens (a)–(f) do Exercício 14 trocando o termo "vizinhança de x" por "vizinhança de x pertencente a $\mathcal{V}_x$".

20. Sejam (M_i, d_i), $i = 1, \ldots, n$, espaços métricos e considere o produto $M = \prod_{i=1}^n M_i$ munida da métrica produto d. Mostre que:
    (a) se U_i é aberto em M_i para cada $i = 1, \ldots, n$ então $\prod_{i=1}^n U_i$ é aberto em M ("o produto de abertos é aberto");
    (b) se F_i é fechado em M_i para cada $i = 1, \ldots, n$ então $\prod_{i=1}^n F_i$ é fechado em M ("o produto de fechados é fechado");
    (c) para $x = (x_i)_{i=1}^n \in M$, mostre que:
    $$\mathcal{V} = \left\{\prod_{i=1}^n \mathrm{B}(x_i; r_i) : r_1, \ldots, r_n \in (0, +\infty)\right\}$$
    é um sistema fundamental de vizinhanças para x em M;
    (d) o fecho do produto é o produto dos fechos, i.e.:
    $$\overline{A_1 \times \cdots \times A_n} = \overline{A_1} \times \cdots \times \overline{A_n},$$
    para quaisquer $A_i \subset M_i$, $i = 1, \ldots, n$.

`Espaços métricos (seqüências).`

21. Seja M um espaço métrico. Mostre que se y_1 e y_2 são ambos limites de uma seqüência $(x_n)_{n\geq 1}$ em M então $y_1 = y_2$.

22. Mostre que toda seqüência convergente num espaço métrico é uma seqüência de Cauchy.

23. Seja (M, d) um espaço métrico e $(x_n)_{n\geq 1}$ uma seqüência em M. Mostre que x_n converge para um certo $x \in M$ se e somente se $d(x_n, x) \to 0$ em $I\!R$.

24. Se $(x_n)_{n\geq 1}$ é uma seqüência e $\phi : I\!N \to I\!N$ é uma função estritamente crescente então dizemos que $(x_{\phi(k)})_{k\geq 1}$ é uma *subseqüência* de $(x_n)_{n\geq 1}$; tipicamente escrevemos $n_k = \phi(k)$ e dizemos que $(x_{n_k})_{k\geq 1}$ é uma subseqüência de $(x_n)_{n\geq 1}$. Mostre que se uma seqüência $(x_n)_{n\geq 1}$ converge para x num espaço métrico M então toda subseqüência $(x_{n_k})_{k\geq 1}$ converge para x.

25. (*o teorema do sanduíche*) Mostre que se $y_n \to L$ e $z_n \to L$ em $\mathbb{R}$ e se $y_n \le x_n \le z_n$ para todo $n \in \mathbb{N}$ então $x_n \to L$ em $\mathbb{R}$.

`Espaços métricos (limite e continuidade).`

26. Seja $f : A \subset M \to N$ uma função, onde (M, d) e (N, d') são espaços métricos. Dado um ponto de acumulação $x_0 \in M$ de A, mostre que se L e $L' \in N$ são ambos limites de $f(x)$ quando x tende a x_0 então $L = L'$.

`Observação.` Para quem gosta de pensar em coisas vaziamente satisfeitas: mostre que se $x_0 \in M$ não é um ponto de acumulação de A então seria verdade que *todo* $L \in N$ é um limite de $f(x)$ quando x tende a x_0.

27. Sejam (M, d), (N, d') espaços métricos e $f : M \to N$ uma função. Dado $x_0 \in M$ mostre que:

   (a) se x_0 é um ponto de acumulação de M então f é contínua no ponto x_0 se e somente se o limite $\lim_{x \to x_0} f(x)$ existe e é igual a $f(x_0)$.

   (b) se x_0 não é um ponto de acumulação de M (i.e., se x_0 é um *ponto isolado* de M) então f é *sempre* contínua no ponto x_0.

28. Sejam (M, d), (N, d') espaços métricos e $f : M \to N$ uma função. Mostre que f é contínua num ponto $x \in M$ se e somente se dado $\varepsilon > 0$, existe $\delta > 0$ com:

$$f\big(\mathrm{B}(x; \delta)\big) \subset \mathrm{B}\big(f(x); \varepsilon\big).$$

29. Sejam (M, d), (N, d') e (P, d'') espaços métricos, $A \subset M$, $B \subset N$ subconjuntos, $x_0 \in M$ um ponto de acumulação de A e $f : A \to N$, $g : B \to P$ funções com $f(A) \subset B$. Suponha que $\lim_{x \to x_0} f(x) = L \in N$ e que L *não* pertence à imagem de f. Mostre que:

   (a) L é um ponto de acumulação de B;

   (b) se $\lim_{y \to L} g(y) = K \in P$ então $\lim_{x \to x_0} g\big(f(x)\big) = K$;

   (c) ache um contra-exemplo para os itens (a) e (b) quando removemos a hipótese que L não pertence à imagem de f.

30. (*teorema do sanduíche para funções*) Sejam $f, g, h : A \subset M \to \mathbb{R}$ funções, onde M é um espaço métrico. Suponha que $x_0 \in M$ é um ponto de acumulação de A, que $\lim_{x \to x_0} f(x) = \lim_{x \to x_0} h(x) = L \in \mathbb{R}$ e que $f(x) \le g(x) \le h(x)$ para todos $x \in A$. Mostre que $\lim_{x \to x_0} g(x) = L$.

31. (*métricas induzidas por funções*) Seja $f : M \to N$ uma função bijetora, onde M é um conjunto e (N, d') é um espaço métrico. Mostre que $d(x, y) = d'\big(f(x), f(y)\big)$ é uma métrica em M; mostre também que d é a *única* métrica em M que torna f uma isometria.

`Espaços métricos (métricas equivalentes).`

32. Para $x \in I\!\!R^n$ denote por $\|x\|$ a norma Euclideana de x e defina:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|x\|_\infty = \max\big\{|x_1|, \ldots, |x_n|\big\}.$$

Mostre que:

$$\|x\|_\infty \le \|x\|, \quad \|x\| \le \sqrt{n}\,\|x\|_\infty, \quad \|x\|_\infty \le \|x\|_1, \quad \|x\|_1 \le n\|x\|_\infty,$$

e conclua que:

$$\|x\| \le \sqrt{n}\,\|x\|_1, \quad \|x\|_1 \le n\|x\|.$$

33. Sejam (M_i, d_i), $i = 1, \ldots, n$, espaços métricos e considere o produto $M = \prod_{i=1}^n M_i$. Mostre que as fórmulas:

$$d_1(x, y) = \sum_{i=1}^{n} d_i(x_i, y_i), \quad d_\infty(x, y) = \max\big\{d_1(x_1, y_1), \ldots, d_n(x_n, y_n)\big\},$$

definem métricas em M, onde $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n \in M$. Usando o Exercício anterior, mostre que as métricas d_1 e d_∞ são uniformemente equivalentes entre si e também são uniformemente equivalentes à métrica produto d.

34. Sejam d e d' métricas em M de modo que as aplicações identidade:

$$\mathrm{Id} : (M, d) \to (M, d'), \quad \mathrm{Id} : (M, d') \to (M, d)$$

sejam Lipschitzianas. Mostre que $A \subset M$ é limitado em (M, d) se e somente se A é limitado em (M, d').

`Espaços métricos (subespaços e topologia induzida).`

35. Seja (M, d) um espaço métrico e considere $S \subset M$ munido da métrica induzida por d. Mostre que:
    (a) uma seqüência $(x_n)_{n\ge1}$ em S é de Cauchy em S se e somente se $(x_n)_{n\ge1}$ for de Cauchy em M;
    (b) uma seqüência $(x_n)_{n\ge1}$ em S converge para $x \in S$ no espaço S se e somente se $(x_n)_{n\ge1}$ converge para $x \in S$ no espaço M;
    (c) para $A \subset S$, o fecho de A no espaço métrico S coincide com $\overline{A} \cap S$;
    (d) a função inclusão $i : S \to M$ é uniformemente contínua;

    se (N, d') é um outro espaço métrico então:
    (e) se uma função $f : M \to N$ é (uniformemente) contínua então sua restrição $f|_S : S \to N$ é (uniformemente) contínua;
    (f) se uma função $f : N \to M$ tem imagem contida em S então $f : N \to M$ é (uniformemente) contínua se e somente se $f : N \to S$ é (uniformemente) contínua.

36. Seja (M, d) um espaço métrico e considere $S \subset M$ munido da métrica induzida por d. Mostre que:
    (a) se S é um espaço métrico completo então S é fechado em M;
    (b) se M é um espaço métrico completo então S é completo se e somente se S é fechado em M.
37. Mostre que *continuidade é uma propriedade local*, i.e., que dada uma função $f : M \to N$ entre espaços métricos, um ponto $x \in M$ e uma vizinhança $V \subset M$ de x então f é contínua no ponto x se e somente se a restrição $f|_V$ é contínua no ponto x.
38. Sejam (M, d), (N, d') espaços métricos e suponha que $M = \bigcup_{i=1}^n F_i$ com cada F_i fechado em M. Seja $f : M \to N$ uma função tal que $f|_{F_i} : F_i \to N$ é contínua para todo $i = 1, \ldots, n$. Mostre que f é contínua (*dica*: use que f é contínua se e somente se $f^{-1}(F)$ é fechado em M para todo F fechado em N).
39. Sejam (M, d) um espaço métrico e $S \subset M$ um subconjunto. Um ponto $x \in S$ é dito *isolado* de S quando $x \notin S'$, i.e., quando existe $r > 0$ tal que $\mathrm{B}(x; r) \cap S = \{x\}$. Mostre que $x \in S$ é isolado de S se e somente se $\{x\}$ é um aberto do espaço métrico $\big(S, d|_{S \times S}\big)$.
40. Um espaço métrico (M, d) é dito *discreto* quando todo subconjunto de M é aberto. Mostre que:
    (a) M é discreto se e somente se $\{x\}$ é aberto em M para todo $x \in M$;
    (b) M é discreto se e somente se todo subconjunto de M é fechado em M;
    (c) um conjunto M munido da métrica zero-um é sempre um espaço métrico discreto;
    (d) considerando $S \subset M$ com a métrica induzida de M então S é um espaço discreto se e somente se todo ponto de S é isolado;
    (e) $\mathbb{Z}$ é discreto com a métrica induzida de $\mathbb{R}$.

Aula número 6 (22/03)

A aula número 6 cobriu o material das Seções (3) e (4) originalmente destinado à aula número 5.

Aula número 7 (27/03)

(1) **Espaços conexos**.

`Observação`. Esta aula deve iniciar cobrindo a seção (5) da aula número 5 (20/03) sobre o tema "Subespaços e topologia induzida"; para compreensão dos resultados sobre espaços conexos tal assunto é fundamental.

Seja (M, d) um espaço métrico. Se A, B são abertos disjuntos em M tais que $M = A \cup B$ então dizemos que $M = A \cup B$ é uma *cisão* de M; quando $A = \emptyset$, $B = M$ ou $A = M$, $B = \emptyset$ então dizemos que $M = A \cup B$ é uma *cisão trivial* de M.

Definição. *Dizemos que o espaço métrico M é* conexo *quando M só admite a cisão trivial, i.e., se dados abertos disjuntos A, $B \subset M$ com $M = A \cup B$ então $A = \emptyset$ ou $B = \emptyset$.*

Algumas maneiras equivalentes de definir espaço conexo estão listadas no Exercício 6. Particularmente importante é a observação que M é conexo se e somente se os únicos subconjuntos de M que são ao mesmo tempo abertos e fechados são $\emptyset$ e M.

`Observação`. Conexidade é uma *noção intrínseca*, i.e., não depende de espaço ambiente (ao contrário de noções como aberto, fechado e denso). Por exemplo, se (M, d) é um espaço métrico e $S \subset M$ é um subconjunto então afirmações como "S é aberto", "S é fechado", "S é denso" são *relativas a M*; se M for um subespaço métrico de um espaço métrico maior N então é bem possível que S seja aberto (ou fechado, ou denso) em M, mas *não* em N. Por outro lado, a definição de conexidade *não faz referência a um ambiente externo*; dizemos que um subconjunto S de um espaço métrico M é conexo quando o *espaço métrico S com a métrica $d|_{S\times S}$ induzida de M for conexo* no sentido da definição anterior. Note que se M é um subespaço métrico de N e $S \subset M$ então M e N *induzem a mesma métrica em S* (isso é meio óbvio!) e portanto S é conexo visto como subconjunto de M se e somente se S é conexo visto como subconjunto de N. Observamos que também as noções de *espaço limitado* e *espaço completo* mencionadas anteriormente são intrínsecas.

`Observação`. Conexidade é uma noção topológica, i.e., pode ser definida apenas em termos de abertos (sem referência direta à métrica). Segue que dois espaços homeomorfos são ambos conexos ou ambos desconexos.

`Exemplo bobo`. Se $M = \emptyset$ ou se M tem só um ponto então M é conexo.

Teorema. *Sejam (M, d), (N, d') espaços métricos e seja $f : M \to N$ uma função contínua. Se M é conexo então a imagem de f é conexa ("a imagem de um conexo por uma função contínua é conexa").*

Demonstração. Podemos substituir N por $f(M)$ (isso não afeta a continuidade de f) e portanto podemos supor sem perda de generalidade que f seja sobrejetora. Seja $A \subset N$ aberto e fechado. Daí $f^{-1}(A)$ é aberto e fechado em M e portanto $f^{-1}(A) = \emptyset$ ou $f^{-1}(A) = M$; no primeiro caso temos $A = \emptyset$ e no segundo $A = N$. ■

Teorema. *Seja (M, d) um espaço métrico e seja $(C_i)_{i \in I}$ uma família de subconjuntos conexos de M tal que existe $p \in \bigcap_{i \in I} C_i$. Então a união $C = \bigcup_{i \in I} C_i$ é conexa ("a união de conexos com ponto em comum é conexa").*

Demonstração. Seja $A \subset C$ aberto e fechado em C. Para cada $i \in I$ temos que $A \cap C_i$ é aberto e fechado em C_i e portanto $A \cap C_i = \emptyset$ ou $A \cap C_i = C_i$, pela conexidade de C_i. Se $p \in A$ então deve ser $A \cap C_i = C_i$ para todo $i \in I$ e portanto $A = C$; se $p \notin A$ então deve ser $A \cap C_i = \emptyset$ para todo $i \in I$ e portanto $A = \emptyset$. ■

Teorema. *Sejam (M, d) um espaço métrico e $X \subset M$ um subconjunto conexo. Se C é tal que $X \subset C \subset \overline{X}$ então C é conexo (em particular, o fecho de um conexo é conexo).*

Demonstração. Seja $A \subset C$ aberto e fechado em C. Temos que $A \cap X$ é aberto e fechado em X, donde $A \cap X = \emptyset$ ou $A \cap X = X$, já que X é conexo. Como X é denso em C e A é aberto em C temos que $A \cap X = \emptyset$ implica $A = \emptyset$. Se for $A \cap X = X$ (i.e., $X \subset A$) então $A = C$ pois A é fechado em C e X é denso em C. ■

Teorema. *A reta $I\!R$ (com a métrica Euclideana, lógico) é conexa.*

Demonstração. Suponha por absurdo que existam abertos disjuntos não vazios A, B em $I\!R$ com $I\!R = A \cup B$. Escolha $a \in A$ e $b \in B$; como A e B são disjuntos temos $a \neq b$ e podemos por exemplo supor que $a < b$ (senão é só trocar A por B). Considere o conjunto $X = \big\{x \in A : x < b\big\} \subset I\!R$. Temos que X é limitado superiormente (por b) e não vazio ($a \in X$), donde existe $c = \sup X \in I\!R$; obviamente $c \leq b$. Para todo $\varepsilon > 0$, existem elementos de X (e portanto de A) no intervalo $(c - \varepsilon, c]$, donde c é um ponto de aderência de A. Como A é fechado, temos $c \in A$; como $A \cap B = \emptyset$, temos $c \neq b$ (e portanto $c < b$). Como A é aberto, existe $\varepsilon > 0$ tal que $(c - \varepsilon, c + \varepsilon) \subset A$ e podemos supor também que $c + \varepsilon < b$; daí $[c, c + \varepsilon) \subset X$, contradizendo $c = \sup X$. ■

Corolário. *Todo intervalo de $I\!R$ é conexo.*

Demonstração. Intervalos abertos são homeomorfos a $I\!R$ (mostre!) e portanto conexos; se o intervalo I contém alguma de suas extremidades então existe um intervalo aberto J com $J \subset I \subset \overline{J}$, donde I é conexo. ■

Teorema. *Se $I \subset I\!R$ é conexo então ou $I = \emptyset$, ou I tem um único ponto ou I é um intervalo.*

Demonstração. Sejam $a, b \in I$ com $a < b$. Afirmamos que $[a, b] \subset I$; de fato, se existisse $c \in (a, b)$ com $c \notin I$ então $I = \big(I \cap (-\infty, c)\big) \cup \big(I \cap (c, +\infty)\big)$ seria uma cisão não trivial de I, contradizendo o fato que I é conexo. Suponha que I tem mais de um ponto; é fácil ver então que $(\inf I, \sup I) \subset I \subset [\inf I, \sup I]$, donde I é um intervalo (é possível que $\inf I = -\infty$ ou $\sup I = +\infty$). ■

Corolário. *(o teorema do valor intermediário) Se $f : I \subset \mathbb{R} \to \mathbb{R}$ é uma função contínua definida num intervalo $I \subset \mathbb{R}$ e se f assume os valores $x, y \in \mathbb{R}$ então f assume todos os valores entre x e y.*

Demonstração. Observe que $f(I) \subset \mathbb{R}$ é conexo e portanto é um intervalo (ou um único ponto). ■

Teorema. *O produto $M = \prod_{i=1}^{n} M_i$ de espaços métricos conexos (M_i, d_i) é conexo.*

Demonstração. É suficiente mostrar o caso $n = 2$, sendo que o caso geral segue por indução (observe que o produto cartesiano de espaços métricos é associativo). Vamos supor M_1 e M_2 não vazios, caso contrário o resultado é trivial. Fixe $x \in M_1$; para cada $y \in M_2$, os espaços $\{x\} \times M_2$ e $M_1 \times \{y\}$ são conexos (pois são homeomorfos respectivamente à M_2 e à M_1) e possuem o ponto (x, y) em comum. Segue que:

$$C_y = \big(\{x\} \times M_2\big) \cup \big(M_1 \times \{y\}\big)$$

é conexo para todo $y \in M_2$. Observe agora que $M_1 \times M_2 = \bigcup_{y \in M_2} C_y$, onde cada C_y é conexo e $\{x\} \times M_2 \subset \bigcap_{y \in M_2} C_y \neq \emptyset$. ■

Corolário. *$\mathbb{R}^n$ é conexo.* ■

Definição. *Um espaço métrico (M, d) é dito* conexo por caminhos *(dizemos também* conexo por arcos*) quando dados $x, y \in M$ existe uma aplicação contínua (dizemos também uma* curva *contínua) $\gamma : [0, 1] \to M$ tal que $\gamma(0) = x$ e $\gamma(1) = y$ (dizemos que γ é um* caminho ligando x a y*).*

Teorema. *Se M é conexo por caminhos então M é conexo.*

Demonstração. Escolha $x \in M$ (o caso $M = \emptyset$ é trivial). Para cada $y \in M$ podemos escolher uma curva contínua $\gamma_y : [0, 1] \to M$ com $\gamma_y(0) = x$ e $\gamma_y(1) = y$. Como γ_y é contínua e $[0, 1] \subset \mathbb{R}$ é conexo temos que a imagem de γ_y é conexa; daí $M = \bigcup_{y \in M} \mathrm{Im}(\gamma_y)$ e $x \in \mathrm{Im}(\gamma_y)$ para todo $y \in M$, i.e., M é união de conexos com ponto em comum. ■

`Exemplo`. Dados $x, y \in \mathbb{R}^n$ então o *segmento de reta* ligando x a y é definido por:

$$[x, y] = \big\{(1 - t)x + ty : t \in [0, 1]\big\};$$

um subconjunto $S \subset \mathbb{R}^n$ é dito *convexo* quando dados $x, y \in S$ então $[x, y] \subset S$. Todo subconjunto convexo S de $\mathbb{R}^n$ é conexo por caminhos ($[0, 1] \ni t \mapsto (1 - t)x + ty \in S$ é um caminho ligando $x, y \in S$) e portanto é conexo. Bolas abertas e bolas fechadas em $\mathbb{R}^n$ são convexas e portanto conexas por caminhos (e conexas); de fato, vamos verificar por exemplo que a bola aberta $\mathrm{B}(a; r)$, $a \in \mathbb{R}^n$, $r > 0$ é convexa. Dados $x, y \in \mathbb{R}^n$ com $\|x - a\| < r$, $\|y - a\| < r$ então para todo $t \in [0, 1]$:

$$\big\|(1 - t)x + ty - a\big\| = \big\|(1 - t)x - (1 - t)a + ty - ta\big\| \leq (1 - t)\|x - a\| + t\|y - a\| < r.$$

Teorema. *Seja M um espaço métrico conexo e suponha que todo ponto de M possui uma vizinhança conexa por caminhos. Então M é conexo por caminhos.*

Demonstração. Defina uma relação de equivalência $\sim$ em M fazendo:

$$x \sim y \quad \Longleftrightarrow \quad \text{existe um caminho em } M \text{ ligando } x \text{ a } y,$$

para todos $x, y \in M$ (veja o Exercício 8). Seja C uma classe de equivalência; vamos mostrar que C é aberta. Dado $x \in C$ então x possui uma vizinhança V em M que é conexa por caminhos, donde $y \sim x$ para todo $y \in V$; daí $V \subset C$ e x é um ponto interior de C. Mostramos então que C é aberta; mas C também é fechada, pois seu complementar é a união das outras classes de equivalência (que também são abertas). Como $C \neq \emptyset$ (classes de equivalência nunca são vazias, por definição; no caso trivial $M = \emptyset$ não temos nenhuma classe de equivalência), temos $C = M$ e portanto M é conexo por caminhos. ■

Corolário. *Se um aberto $U \subset \mathbb{R}^n$ é conexo então ele é conexo por caminhos.* ■

`Exemplo`. Existem espaços conexos que não são conexos por caminhos. Um exemplo famoso é a *senóide dos topólogos* definida da seguinte maneira: seja $S \subset \mathbb{R}^2$ o gráfico da função $(0, 1] \ni t \mapsto \operatorname{sen} \frac{1}{t}$, i.e.:

$$S = \big\{ \big(t, \operatorname{sen} \tfrac{1}{t}\big) : t \in (0, 1] \big\}.$$

Temos que S é conexo pois é a imagem do conexo $(0, 1]$ pela função contínua $t \mapsto \big(t, \operatorname{sen} \frac{1}{t}\big)$. A senóide dos topólogos é por definição o fecho de S, ou seja:

$$\overline{S} = S \cup \big(\{0\} \times [-1, 1]\big). \qquad \text{(verifique!)}$$

Temos que $\overline{S}$ é conexo pois é o fecho de um conexo. Ocorre que não existe um caminho contínuo em $\overline{S}$ ligando um ponto de S a um ponto do segmento vertical $\{0\} \times [-1, 1]$ (veja o Exercício 10).

(2) **Espaços compactos**.

Seja (M, d) um espaço métrico. Uma *cobertura* de M é uma família $(U_i)_{i \in I}$ de subconjuntos de M com $M = \bigcup_{i \in I} U_i$; dizemos que $(U_i)_{i \in I}$ é uma *cobertura aberta* de M se cada U_i for aberto em M. Uma *subcobertura* de $(U_i)_{i \in I}$ é uma família da forma $(U_i)_{i \in J}$ com $J \subset I$ tal que $M = \bigcup_{i \in J} U_i$.

Definição. *Um espaço métrico M é dito* compacto *quando toda cobertura aberta de M admite uma subcobertura finita, i.e., se $M = \bigcup_{i \in I} U_i$ com cada $U_i \subset M$ aberto então existem $i_1, \ldots, i_n \in I$ com $M = \bigcup_{k=1}^{n} U_{i_k}$.*

`Observação`. Assim como a noção de espaço conexo, a noção de espaço compacto é intrínseca, i.e., não faz referência a um ambiente onde o espaço está imerso – um subconjunto K de um espaço métrico M é dito *compacto* quando K é um espaço métrico compacto munido da métrica induzida de M. Assim, $K \subset M$ é compacto significa que toda cobertura aberta de K *por abertos relativos a K* admite uma subcobertura finita. No entanto, a noção intrínseca de compacto é *equivalente* à noção *extrínseca* de compacto dada no seguinte:

Teorema. *Seja M um espaço métrico e $K \subset M$ um subconjunto. Então K é um espaço compacto (com a métrica induzida de M) se e somente se toda cobertura aberta de K por abertos de M admite uma subcobertura finita, i.e., dada uma família $(U_i)_{i\in I}$ de abertos de M com $K \subset \bigcup_{i\in I} U_i$ então existem $i_1, \ldots, i_n \in I$ com $K \subset \bigcup_{k=1}^n U_{i_k}$.*

Demonstração. Suponha que K é compacto (na métrica induzida) e seja $K \subset \bigcup_{i\in I} U_i$ uma cobertura de K por abertos de M. Daí $K = \bigcup_{i\in I}(U_i \cap K)$ é uma cobertura aberta de K (por abertos de K) e portanto existem $i_1, \ldots, i_n$ com $K = \bigcup_{k=1}^n (U_{i_k} \cap K)$; daí $K \subset \bigcup_{k=1}^n U_{i_k}$. Reciprocamente, suponha que toda cobertura aberta de K por abertos de M admite uma subcobertura finita; seja $K = \bigcup_{i\in I} Z_i$ uma cobertura de K por abertos de K. Cada Z_i é da forma $U_i \cap K$ com U_i aberto em M; daí $K \subset \bigcup_{i\in I} U_i$ é uma cobertura de K por abertos de M e portanto existem $i_1, \ldots, i_n$ com $K \subset \bigcup_{k=1}^n U_{i_k}$. Daí $K = \bigcup_{k=1}^n (U_{i_k} \cap K) = \bigcup_{k=1}^n Z_{i_k}$ e portanto K é compacto. ■

`Observação`. Compacidade é uma noção topológica.

Teorema. *Sejam M, N espaços métricos. Se $f : M \to N$ é uma função contínua e M é compacto então a imagem de f é compacta ("a imagem de um compacto por uma função contínua é compacta").*

Demonstração. Trocando N por $f(M)$ (o que não afeta a continuidade de f) podemos supor sem perda de generalidade que f é sobrejetora. Seja $N = \bigcup_{i\in I} U_i$ uma cobertura aberta de N; daí $M = \bigcup_{i\in I} f^{-1}(U_i)$ é uma cobertura aberta de M, da qual extraímos uma subcobertura finita $M = \bigcup_{k=1}^n f^{-1}(U_{i_k})$. Daí $N = \bigcup_{k=1}^n U_{i_k}$. ■

Teorema. *Se M é compacto então M é limitado.*

Demonstração. Fixe $x \in M$ (o caso $M = \emptyset$ é trivial); temos que $M = \bigcup_{n\in \mathbb{N}} \mathrm{B}(x; n)$ é uma cobertura aberta de M e portanto existem $n_1, \ldots, n_k \in \mathbb{N}$ com $M = \bigcup_{i=1}^k \mathrm{B}(x; n_i)$. Tomando $n = \max\{n_1, \ldots, n_k\}$ então $M = \mathrm{B}(x; n)$. ■

Corolário. *Se $f : M \to N$ é contínua e M é compacto então f é limitada.* ■

Teorema. *Se (M, d) é um espaço métrico e $K \subset M$ é compacto então K é fechado em M.*

Demonstração. Seja $x \in M$ com $x \notin K$. Para cada $y \in K$, como $x \neq y$, existem abertos disjuntos $U_y, V_y \subset M$ com $y \in U_y$ e $x \in V_y$ (por exemplo, tome $U_y = \mathrm{B}(y; r)$ e $V_y = \mathrm{B}(x; r)$ com $r = \frac{1}{2}d(x, y) > 0$). Daí $K \subset \bigcup_{y\in K} U_y$ é uma cobertura aberta (por abertos de M!) e portanto podemos extrair uma subcobertura finita $K \subset \bigcup_{i=1}^n U_{y_i}$. Seja $V = \bigcap_{i=1}^n V_{y_i}$; obviamente V é uma vizinhança aberta de x. Além do mais, V é disjunto de K (na verdade V é disjunto até de $\bigcup_{i=1}^n U_{y_i} \supset K$). Daí x não é um ponto de aderência de K e K é fechado em M. ■

Corolário. *Se M é compacto não vazio e $f : M \to \mathbb{R}$ é contínua então f assume máximo e mínimo em M.*

Demonstração. Como $f(M) \subset \mathbb{R}$ é fechado e limitado temos $\sup f(M) \in f(M)$ e $\inf f(M) \in f(M)$. ■

Teorema. *Se (M,d) é compacto e $K \subset M$ é fechado então K é compacto.*

Demonstração. Seja $K \subset \bigcup_{i\in I} U_i$ uma cobertura aberta de K por abertos de M; daí $M = (M \setminus K) \cup \bigcup_{i\in I} U_i$ é uma cobertura aberta de M, já que K é fechado. Obtemos então uma subcobertura finita dessa última que pode ou não envolver $M \setminus K$; de qualquer modo, $M \setminus K$ não cobre nenhum ponto de K e portanto obtivemos na verdade uma cobertura finita de K por alguns U_i's. ■

Teorema. *(Heine–Borel) Um subconjunto de $\mathbb{R}$ é compacto se e somente se for fechado e limitado.*

Demonstração. Em vista dos teoremas anteriores, basta provar que um intervalo fechado $[a,b]$ é compacto. Seja $[a,b] \subset \bigcup_{i\in I} U_i$ uma cobertura de $[a,b]$ por abertos de $\mathbb{R}$; suponha por absurdo que ela não possui uma subcobertura finita. Divida o intervalo fechado $I_0 = [a,b]$ no meio; uma das duas metades não pode ser coberta por um número finito de U_i's (senão I_0 poderia) — denote essa metade por I_1. Agora divida I_1 no meio e denote por I_2 uma metade que não pode ser coberta por um número finito de U_i's. Seguindo com esse processo, obteremos uma seqüência decrescente $I_0 \supset I_1 \supset I_2 \supset I_3 \supset \cdots$ de intervalos fechados com $\mathrm{diam}(I_k) = \frac{1}{2^k}\mathrm{diam}(I_0)$, onde cada I_k não pode ser coberto por um número finito de U_i's. Pelo princípio dos intervalos encaixantes, existe $c \in \bigcap_{k\geq 1} I_k$; como $c \in I_0$, existe U_i com $c \in U_i$. Como $\mathrm{diam}(I_k) \to 0$, existe I_k com $c \in I_k \subset U_i$, contradizendo o fato que I_k não pode ser coberto por um número finito de U_i's. ■

Corolário. *(o teorema de Weierstrass) Uma função contínua $f : [a,b] \to \mathbb{R}$ é limitada e assume máximo e mínimo.* ■

A demonstração do teorema de Heine–Borel motiva a introdução da seguinte:

Definição. *Um espaço métrico M é dito* totalmente limitado *quando dado $\varepsilon > 0$ então M admite uma cobertura finita por subconjuntos de diâmetro menor que ε, i.e., existem $M_1, \ldots, M_n \subset M$ com cada $\mathrm{diam}(M_i) < \varepsilon$ e $M = \bigcup_{i=1}^n M_i$. Um subconjunto S de um espaço métrico é dito* totalmente limitado *quando S for um espaço métrico totalmente limitado munido da métrica induzida de M ("totalmente limitado" é uma noção intrínseca).*

`Exemplo`. Todo subconjunto limitado de $\mathbb{R}$ é totalmente limitado. De fato, todo subconjunto limitado de $\mathbb{R}$ está contido num intervalo $[a,b]$ e para todo $\varepsilon > 0$, o intervalo $[a,b]$ pode ser particionado em um número finito de intervalos de comprimento menor que ε.

`Observação`. Obviamente todo espaço totalmente limitado é limitado, de modo que espaços não limitados fornecem exemplos de espaços que não são totalmente limitados. Qual seria um exemplo de um espaço limitado que não é totalmente limitado? Bom, generalizando o exemplo acima não é difícil ver que todo subconjunto limitado de $\mathbb{R}^n$ é totalmente limitado. Para encontrar um exemplo de espaço limitado que não é totalmente limitado precisamos então procurar fora de $\mathbb{R}^n$ (o que na verdade foge um pouco do espírito do curso) — o exemplo é o seguinte: considere um conjunto infinito M munido da métrica zero-um (um outro exemplo seria uma bola de um espaço normado de dimensão infinita, mas isso é outra história).

Teorema. *Seja M um espaço métrico. As seguintes condições são equivalentes:*
(i) M é compacto;
(ii) toda seqüência em M possui uma subseqüência convergente (em M);
(iii) M é completo e totalmente limitado.

Demonstração.
(i)$\Rightarrow$(ii). Seja $(x_n)_{n\geq 1}$ uma seqüência em M. Suponha por absurdo que $(x_n)_{n\geq 1}$ não possui subseqüência convergente, i.e., que $(x_n)_{n\geq 1}$ não possui pontos aderentes (veja o Exercício 13). Daí, para todo $y \in M$ temos que y não é um ponto aderente de $(x_n)_{n\geq 1}$ e portanto (negue a definição!) existe um aberto U_y em M contendo y tal que $x_n \notin U_y$ para todo n suficientemente grande (i.e., o conjunto $\{n \in I\!N : x_n \in U_y\}$ é finito). A cobertura aberta $M = \bigcup_{y\in M} U_y$ admite uma subcobertura finita $M = \bigcup_{i=1}^{k} U_{y_i}$; daí:

$$I\!N = \{n \in I\!N : x_n \in M\} = \bigcup_{i=1}^{k} \{n \in I\!N : x_n \in U_{y_i}\}$$

é finito, uma contradição.
(ii)$\Rightarrow$(iii). É fácil ver que se uma seqüência de Cauchy possui uma subseqüência convergente então ela mesmo é convergente; portanto M é completo. Seja agora $\varepsilon > 0$ e vamos mostrar que M pode ser coberto por um número finito de subconjuntos de diâmetro menor que ε. Suponha que não. Daí $M \neq \emptyset$ e podemos escolher $x_1 \in M$; a bola $\mathrm{B}(x_1; \frac{\varepsilon}{3})$ tem diâmetro menor que ε e portanto não pode cobrir M – podemos então encontrar $x_2 \in M$ fora de $\mathrm{B}(x_1; \frac{\varepsilon}{3})$. Analogamente, as bolas $\mathrm{B}(x_1; \frac{\varepsilon}{3})$ e $\mathrm{B}(x_2; \frac{\varepsilon}{3})$ possuem diâmetro menor que ε e portanto não podem cobrir M; podemos encontrar então $x_3 \in M$ fora dessas bolas. Prosseguindo a construção indutivamente, supondo $x_1, \ldots, x_k$ construídos então como $\bigcup_{i=1}^{k} \mathrm{B}(x_i; \frac{\varepsilon}{3})$ não pode ser M, podemos encontrar x_{k+1} fora da união dessas k bolas. Obtemos assim uma seqüência $(x_n)_{n\geq 1}$ em M com $d(x_n, x_m) \geq \frac{\varepsilon}{3}$ para todo $n \neq m$ e tal seqüência não pode ter subseqüência convergente, uma contradição.
(iii)$\Rightarrow$(i). (essa parte imita a demonstração do teorema de Heine-Borel) Seja $M = \bigcup_{i\in I} U_i$ uma cobertura aberta de M. Suponha por absurdo que tal cobertura não admite uma subcobertura finita. Dado $\varepsilon = 1$ então M pode ser coberto por um número finito de subconjuntos de diâmetro menor que 1; podemos supor que tais subconjuntos são fechados (pois $\mathrm{diam}(S) = \mathrm{diam}(\overline{S})$; veja o Exercício 12). Se todos esses conjuntos pudessem ser cobertos por um número finito de U_i's então também M poderia — logo algum desses fechados, digamos F_1, não admite uma subcobertura finita (em particular $F_1 \neq \emptyset$). Temos que também F_1 é totalmente limitado, já que F_1 é um subespaço de M; logo, dado $\varepsilon = \frac{1}{2}$ podemos cobrir F_1 por um número finito de fechados de diâmetro menor que $\frac{1}{2}$. Novamente, um desses fechados, digamos $F_2 \subset F_1$, não pode ser coberto por um número finito de U_i's. Seguindo o raciocício indutivamente, construímos uma seqüência decrescente de fechados não vazios $F_1 \supset F_2 \supset F_3 \supset \cdots$ onde $\mathrm{diam}(F_k) < \frac{1}{k}$ para todo $k \geq 1$ e nenhum F_k pode ser coberto por um número finito de U_i's. Como M é completo (veja o Exercício 14) existe $x \in \bigcap_{k\in I\!N} F_k$; tal x pertence a algum U_i e como $\mathrm{diam}(F_k) \to 0$ devemos ter $F_k \subset U_i$ para algum k, contradizendo o fato que F_k não pode ser coberto por um número finito de U_i's.
■

Corolário. *O produto $M_1 \times \cdots \times M_n$ de espaços métricos compactos (M_i, d_i) é compacto.*

Demonstração. Podemos supor que $n = 2$; o caso geral segue indutivamente (já que o produto cartesiano de espaços métricos é associativo). Suponha que M_1 e M_2 são compactos; vamos mostrar que toda seqüência em $M_1 \times M_2$ admite uma subseqüência convergente. Seja $\big((x_m, y_m)\big)_{m\geq 1}$ uma seqüência com $x_m \in M_1$ e $y_m \in M_2$ para todo m. Como M_1 é compacto, podemos extrair uma subseqüência convergente $(x_{m_k})_{k\geq 1}$ de $(x_m)_{m\geq 1}$; da seqüência $(y_{m_k})_{k\geq 1}$ em M_2 podemos também extrair uma subseqüência convergente $(y_{m_{k_i}})_{i\geq 1}$, já que M_2 também é compacto. Logo, $\big((x_{m_{k_i}}, y_{m_{k_i}})\big)_{i\geq 1}$ é uma subseqüência convergente de $\big((x_m, y_m)\big)_{m\geq 1}$ em $M_1 \times M_2$. ■

Corolário. *Um subconjunto de $\mathbb{R}^n$ é compacto se e somente se é limitado e fechado.*

Demonstração. Basta ver que blocos retangulares fechados $\prod_{i=1}^n [a_i, b_i]$ são compactos; isso segue do Corolário anterior e do teorema de Heine-Borel. ■

`Observação`. Como já havíamos prometido, a teoria desta seção produz uma demonstração que $\mathbb{R}$ é completo. De fato, se $(x_n)_{n\geq 1}$ é uma seqüência de Cauchy em $\mathbb{R}$ então $(x_n)_{n\geq 1}$ fica em algum intervalo $[a, b]$, o qual é completo (pois é compacto).

`Observação`. Para espaços métricos quaisquer, nem de longe é verdade que limitado + fechado $\Rightarrow$ compacto. Por exemplo, se $M \subset \mathbb{R}^n$ é limitado mas não fechado então M é limitado e fechado em si próprio, mas não é compacto (para os que conhecem um pouco de Análise Funcional: num espaço de Banach de dimensão infinita, *nenhuma* bola fechada é compacta, apesar do espaço ser completo).

Teorema. *Sejam M, N espaços métricos e $f : M \to N$ uma função contínua. Se M é compacto então f é uniformemente contínua ("uma função contínua num compacto é uniformemente contínua").*

Demonstração. Se não fosse, existiria $\varepsilon > 0$ tal que para todo $\delta = \frac{1}{n}$, poderíamos encontrar $x_n, y_n \in M$ com $d(x_n, y_n) < \frac{1}{n}$ mas $d\big(f(x_n), f(y_n)\big) \geq \varepsilon$. Como M é compacto, a seqüência $(x_n)_{n\geq 1}$ possui uma subseqüência convergente $(x_{n_k})_{k\geq 1}$, digamos $x_{n_k} \to x$. Como $d(x_{n_k}, y_{n_k}) < \frac{1}{n_k} \to 0$, temos também $y_{n_k} \to x$. Daí, pela continuidade de f, $f(x_{n_k}) \to f(x)$ e $f(y_{n_k}) \to f(x)$, contradizendo $d\big(f(x_{n_k}), f(y_{n_k})\big) \geq \varepsilon$ para todo k. ■

Definição. *Sejam M um espaço métrico e $M = \bigcup_{i\in I} U_i$ uma cobertura de M (não necessariamente aberta). Um número positivo $\delta > 0$ é chamado um* número de Lebesgue *para a cobertura $M = \bigcup_{i\in I} U_i$ se todo subconjunto $S \subset M$ com diâmetro menor que δ está contido em algum U_i.*

A importância da definição acima está no seguinte:

Teorema. *Se M é um espaço métrico compacto então toda cobertura aberta $M = \bigcup_{i\in I} U_i$ $(I \neq \emptyset)$ admite um número de Lebesgue.*

Demonstração. Para cada $x \in M$ existe $i \in I$ com $x \in U_i$; como U_i é aberto podemos encontrar $r_x > 0$ com $\mathrm{B}(x; r_x) \subset U_i$. Da cobertura aberta $M = \bigcup_{x\in M} \mathrm{B}\big(x; \frac{r_x}{2}\big)$ podemos extrair uma subcobertura finita $M = \bigcup_{k=1}^n \mathrm{B}\big(x_k; \frac{r_{x_k}}{2}\big)$; tome $\delta = \frac{1}{2}\min\{r_{x_1}, \ldots, r_{x_n}\} > 0$. Vamos mostrar que δ é um número de Lebesgue para $M = \bigcup_{i\in I} U_i$. Seja $S \subset M$ com

$\text{diam}(S) < \delta$; escolha $x \in S$ (o caso $S = \emptyset$ é trivial). Temos $x \in \mathrm{B}\big(x_k; \frac{r_{x_k}}{2}\big)$ para algum $k = 1, \ldots, n$ e $\mathrm{B}(x_k; r_{x_k}) \subset U_i$ para algum $i \in I$. Daí, para todo $y \in S$ temos:

$$d(x_k, y) \le d(x_k, x) + d(x, y) < \frac{r_{x_k}}{2} + \delta \le r_{x_k},$$

donde $y \in \mathrm{B}(x_k; r_{x_k}) \subset U_i$. Daí $S \subset U_i$. ■

(3) **Base enumerável e separabilidade**.

`Observação`. Espero que todo mundo esteja familiarizado com o conceito de conjunto enumerável (um *conjunto enumerável* é um conjunto que é finito ou que admite uma bijeção com $\mathbb{N}$) — espero também que todo mundo saiba coisas como "a união enumerável de enumeráveis é enumerável", "o produto finito de enumeráveis é enumerável", "$\mathbb{Q}$ é enumerável" e "$\mathbb{R}$ não é enumerável".

Definição. *Seja M um espaço métrico. Uma* base de abertos *para M é uma coleção $\mathfrak{B}$ de abertos de M tal que todo aberto de M pode ser escrito como uma reunião de abertos de M pertencentes à coleção $\mathfrak{B}$. Quando M admite uma base de abertos enumerável então diz-se que M é um* espaço com base enumerável *ou um* espaço que satisfaz o segundo axioma da enumerabilidade *(em inglês é mais curto: diz-se* second countable space*).*

No Exercício 24 apresentamos uma maneira equivalente (e talvez mais simples) de definir base de abertos.

`Observação`. Como vocês podem imaginar, o nome "segundo axioma da enumerabilidade" vem do fato que existe um "primeiro axioma da enumerabilidade"; esse axioma só aparece em cursos de topologia geral, já que o mesmo é automaticamente satisfeito em todo espaço métrico (para os curiosos: o primeiro axioma da enumerabilidade diz que todo ponto possui um sistema fundamental de vizinhanças enumerável).

`Exemplo`. Os intervalos abertos $(a, b) \subset \mathbb{R}$ com a e b racionais formam uma base de abertos enumerável para o espaço métrico $\mathbb{R}$ (verifique!).

Teorema. *Sejam M um espaço métrico, $S \subset M$ um subconjunto e $\mathfrak{B}$ uma base de abertos para M. Então:*

$$\mathfrak{B}_S = \big\{B \cap S : B \in \mathfrak{B}\big\}$$

é uma base de abertos para S. Em particular, se M admite uma base enumerável então também S admite uma base enumerável (diz-se que a propriedade de existência de base enumerável é hereditária para subespaços*).*

Demonstração. É claro que $\mathfrak{B}_S$ é um conjunto de abertos relativos a S. Seja Z um aberto relativo a S; temos $Z = U \cap S$ com U aberto em M. Podemos escrever $U = \bigcup_{i \in I} B_i$ com cada $B_i \in \mathfrak{B}$; daí $Z = U \cap S = \bigcup_{i \in I}(B_i \cap S)$ e cada $B_i \cap S$ está em $\mathfrak{B}_S$. ■

Definição. *Um espaço métrico M é dito* separável *quando o mesmo admite um subconjunto enumerável denso (diz-se também que M satisfaz o* terceiro axioma da enumerabilidade*).*

`Observação`. Existência de base enumerável e separabilidade são propriedades intrínsecas e topológicas.

`Exemplo`. $\mathbb{R}$ é separável pois $\mathbb{Q}$ é um subconjunto enumerável denso de $\mathbb{R}$. Também $\mathbb{R}^n$ é separável pois $\mathbb{Q}^n$ é um subconjunto enumerável denso de $\mathbb{R}^n$.

Teorema. *Um espaço métrico M admite base enumerável de abertos se e somente se for separável.*

Demonstração. Suponha que M admite uma base enumerável $\mathfrak{B}$ de abertos, digamos $\mathfrak{B} = \{B_n\}_{n\geq 1}$ (para facilitar a notação vamos fazer de conta que $\mathfrak{B}$ é infinito, mas na realidade $\mathfrak{B}$ poderia ser finito, tanto faz). Podemos supor que todos os B_n's são não vazios (se algum B_n for vazio, jogue ele fora); podemos então escolher $x_n \in B_n$ para todo n. Daí $D = \{x_n\}_{n\geq 1}$ é um subconjunto enumerável denso de M; de fato, se $U \subset M$ é um aberto não vazio então existe $n \in I\!\!N$ com $B_n \subset U$ e daí $x_n \in D \cap U$. Logo M é separável.

Reciprocamente, suponha que M é separável. Seja $D = \{x_n\}_{n\geq 1}$ um subconjunto enumerável denso de M (de novo vamos fazer de conta que D é infinito para facilitar a notação, mas não tem problema se D for finito). Defina:

$$\mathfrak{B} = \big\{\mathrm{B}\big(x_n; \tfrac{1}{k}\big) : n, k \in I\!\!N\big\};$$

é claro que $\mathfrak{B}$ é um conjunto enumerável de abertos de M. Vamos mostrar que $\mathfrak{B}$ é uma base. Seja $U \subset M$ aberto e seja $x \in U$; devemos encontrar um elemento de $\mathfrak{B}$ que contém x e que esteja contido em U, i.e., devemos encontrar $n, k \in I\!\!N$ com $x \in \mathrm{B}\big(x_n; \frac{1}{k}\big) \subset U$. Como U é aberto, existe $r > 0$ com $\mathrm{B}(x; r) \subset U$; seja $k \in I\!\!N$ com $\frac{1}{k} \leq \frac{r}{2}$. Como D é denso, existe $n \in I\!\!N$ com $x_n \in \mathrm{B}\big(x; \frac{1}{k}\big)$; obviamente $x \in \mathrm{B}\big(x_n; \frac{1}{k}\big)$ (pois $d(x, x_n) < \frac{1}{k}$). Além do mais, $\mathrm{B}\big(x_n; \frac{1}{k}\big) \subset \mathrm{B}(x; r) \subset U$, pois dado $y \in \mathrm{B}\big(x_n; \frac{1}{k}\big)$, temos:

$$d(y, x) \leq d(y, x_n) + d(x_n, x) < \frac{1}{k} + \frac{1}{k} \leq \frac{r}{2} + \frac{r}{2} = r,$$

donde $y \in \mathrm{B}(x; r)$. ■

Corolário. *$I\!\!R^n$ (e todos os seus subespaços) admitem base enumerável de abertos.* ■

`Observação`. Na teoria geral dos espaços topológicos, não é verdade que separável implica existência de base enumerável; é por isso que os dois nomes diferentes existem. Como mostramos, no contexto de espaços métricos (que é o que nos interessa), existência de base enumerável é o mesmo que separabilidade.

Teorema de Lindelöf. *Seja M um espaço métrico separável (ou que admita base enumerável, tanto faz). Então toda cobertura aberta $M = \bigcup_{i\in I} U_i$ de M admite uma subcobertura enumerável, i.e., existe $J \subset I$ enumerável com $M = \bigcup_{j\in J} U_j$.*

Demonstração. Seja $\mathfrak{B}$ uma base enumerável de abertos, digamos $\mathfrak{B} = \{B_n\}_{n\geq 1}$; seja $M = \bigcup_{i\in I} U_i$ uma cobertura aberta de M. Para cada $n \in I\!\!N$ escolha, se existir, $i_n \in I$ com $B_n \subset U_{i_n}$; se tal i_n não existir, não escolha nada. Seja $J \subset I$ o conjunto dos i_n's que foram escolhidos nesse processo; obviamente J é enumerável. Vamos mostrar que $M = \bigcup_{j\in J} U_j$. Seja $x \in M$; sabemos que existe $i \in I$ com $x \in U_i$. Como $\mathfrak{B}$ é uma base de abertos e U_i é aberto, existe $n \in I\!\!N$ com $x \in B_n \subset U_i$. Daí esse n é tal que B_n está contido em algum dos U_i's; portanto, na parte inicial onde fizemos as escolhas dos i_n's, necessariamente escolhemos um i_n com $B_n \subset U_{i_n}$ (pode ser que $i_n = i$ ou não!). Em qualquer caso, temos $i_n \in J$ e $x \in B_n \subset U_{i_n}$; em particular $x \in \bigcup_{j\in J} U_j$. ■

`Observação`. De modo análogo ao que foi discutido na seção sobre espaços compactos, temos que para um subconjunto S de um espaço métrico M são equivalentes:
(i) toda cobertura de S por abertos de S admite uma subcobertura enumerável;
(ii) toda cobertura de S por abertos de M admite uma subcobertura enumerável.

Corolário. *Dada uma cobertura aberta de um subconjunto $S \subset \mathbb{R}^n$ (por abertos relativos a S ou por abertos de $\mathbb{R}^n$) então tal cobertura possui uma subcobertura enumerável.* ■

(4) **Espaços normados e espaços com produto interno**.

Nesta seção V denota um espaço vetorial real.

Definição. *Uma* norma *em V é uma aplicação $V \ni x \mapsto \|x\| \in \mathbb{R}$ satisfazendo as seguintes propriedades:*
(N1) $\|x\| \geq 0$ *para todo* $x \in V$ *e* $\|x\| = 0$ *se e somente se* $x = 0$;
(N2) $\|\lambda x\| = |\lambda|\|x\|$ *para todo* $x \in V$ *e todo* $\lambda \in \mathbb{R}$;
(N3) $\|x + y\| \leq \|x\| + \|y\|$ *para todos* $x, y \in V$ (desigualdade triangular).

É fácil ver que se $\|\cdot\|$ é uma norma em V então a fórmula:

$$d(x, y) = \|x - y\|, \quad x, y \in V,$$

define uma métrica em V; dizemos que d é a *métrica induzida* pela norma $\|\cdot\|$. Todo espaço vetorial normado será pensado sempre como um espaço métrico munido da métrica acima.

Definição. *Duas normas $\|\cdot\|$ e $\|\cdot\|'$ são ditas* equivalentes *quando as métricas correspondentes são equivalentes.*

Teorema. *Sejam V, W espaços normados e $T : V \to W$ uma aplicação linear. As seguintes afirmações são equivalentes:*
(i) T é contínua;
(ii) T é contínua na origem;
(iii) T é limitada na bola unitária $\mathrm{B}[0; 1]$ *de V;*
(iv) existe uma constante $c \geq 0$ tal que $\|T(x)\| \leq c\|x\|$ para todo $x \in V$;
(v) T é Lipschitziana.

Demonstração.
(i)⇒(ii). É óbvio.
(ii)⇒(iii). Dado $\varepsilon = 1$ existe $\delta > 0$ tal que $\|x\| < \delta$ implica $\|T(x)\| < 1$. Daí para todo $x \in \mathrm{B}[0; 1] \subset V$ temos $\left\|\frac{\delta}{2}x\right\| < \delta$ e portanto $\left\|T\left(\frac{\delta}{2}x\right)\right\| < 1$; daí $\|T(x)\| < \frac{2}{\delta}$ para todo $x \in \mathrm{B}[0; 1]$.
(iii)⇒(iv). Tome $c = \sup_{\|x\| \leq 1} \|T(x)\| < +\infty$. Daí $\left\|T\left(\frac{x}{\|x\|}\right)\right\| \leq c$ para todo $x \neq 0$ e portanto $\|T(x)\| \leq c\|x\|$ para todo $x \in V$ (o caso $x = 0$ é trivial).
(iv)⇒(v). $\|T(x) - T(y)\| = \|T(x - y)\| \leq c\|x - y\|$.
(v)⇒(i). É óbvio. ■

Corolário. *Duas normas* $\|\cdot\|$ *e* $\|\cdot\|'$ *são equivalentes se e somente se existem constantes* $\alpha, \beta > 0$ *tais que:*

$$\alpha\|x\| \le \|x\|' \le \beta\|x\|,$$

para todo $x \in V$*. Em particular, as métricas associadas a duas normas equivalentes são também uniformemente equivalentes (até mesmo "Lipschitz equivalentes").*

Demonstração. As aplicações identidade

$$\mathrm{Id} : (V, \|\cdot\|) \longrightarrow (V, \|\cdot\|'), \quad \mathrm{Id} : (V, \|\cdot\|') \longrightarrow (V, \|\cdot\|)$$

são lineares e portanto são contínuas se e somente se são Lipschitzianas. ■

`Exemplo`. A norma Euclideana $\|x\| = \big(\sum_{i=1}^n x_i^2\big)^{\frac{1}{2}}$ é uma norma em $I\!R^n$; a demonstração da desigualdade triangular será feita logo adiante. As fórmulas:

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_\infty = \max\big\{|x_1|, \ldots, |x_n|\big\},$$

definem normas em $I\!R^n$, ambas equivalentes à norma Euclideana (recorde Exercício 32 da aula número 5).

Teorema. *Seja* V *um espaço vetorial normado. A norma de* V *é uma contração fraca de* $(V, \|\cdot\|)$ *em* $I\!R$*. Em particular a norma de* V *é uniformemente contínua.*

Demonstração. É fácil ver que a desigualdade:

$$\big|\|x\| - \|y\|\big| \le \|x - y\| = d(x, y), \quad x, y \in V,$$

segue da desigualdade triangular para $\|\cdot\|$. ■

Teorema. *Todas as normas num espaço vetorial* V *de dimensão finita são equivalentes.*

Demonstração. Podemos supor sem perda de generalidade que $V = I\!R^n$ (veja também a observação a seguir). Seja $\|\cdot\|'$ uma norma arbitrária em $I\!R^n$; por transitividade, basta mostrar que $\|\cdot\|'$ é equivalente à norma $\|\cdot\|_1$ definida acima. Se $(e_1, \ldots, e_n)$ denota a base canônica de $I\!R^n$ então:

$$\|x\|' = \Big\|\textstyle\sum_{i=1}^n x_i e_i\Big\|' \le \displaystyle\sum_{i=1}^n |x_i|\,\|e_i\|' \le k\|x\|_1,$$

para todo $x \in I\!R^n$, onde $k = \max\big\{\|e_1\|', \ldots, \|e_n\|'\big\}$. Isso mostra que a aplicação identidade

$$\mathrm{Id} : (I\!R^n, \|\cdot\|_1) \longrightarrow (I\!R^n, \|\cdot\|')$$

é contínua; segue então também que a aplicação $(I\!R^n, \|\cdot\|_1) \ni x \mapsto \|x\|' \in I\!R$ é contínua (pois é uma composta de funções contínuas). Como a esfera unitária $\big\{x \in I\!R^n : \|x\|_1 = 1\big\}$ de $(I\!R^n, \|\cdot\|_1)$ é compacta em $(I\!R^n, \|\cdot\|_1)$, segue que $x \mapsto \|x\|'$ assume um valor mínimo $\alpha > 0$ em $\big\{x \in I\!R^n : \|x\|_1 = 1\big\}$. Daí, para todo $x \in I\!R^n$ não nulo temos:

$$\left\|\frac{x}{\|x\|_1}\right\|' \ge \alpha \Longrightarrow \|x\|_1 \le \frac{1}{\alpha}\|x\|',$$

donde $\mathrm{Id} : (V, \|\cdot\|') \to (V, \|\cdot\|_1)$ é contínua. ■

Corolário. *Dados espaços normados V e W, se V tem dimensão finita, então toda aplicação linear $T : V \to W$ é contínua.*

Demonstração. Trocando W por $T(V)$ (o que não afeta a continuidade de T), podemos supor sem perda de generalidade que T é sobrejetora; daí W também tem dimensão finita. Pelo Teorema acima, podemos supor também que $V = I\!R^n$, $W = I\!R^m$ e que ambos possuem a métrica Euclidena. Daí cada coordenada de T é uma combinação linear de projeções de $I\!R^n$ e portanto T é contínua. ∎

`Observação`. Vamos entender essa estória de "supor sem perda de generalidade". Quando queremos provar um teorema do tipo $h \Rightarrow t$ e dizemos que "é possível supor sem perda de generalidade que vale h'", isso significa que podemos provar o teorema desejado $h \Rightarrow t$ a partir do teorema mais fraco $h + h' \Rightarrow t$. Por exemplo, acima mostramos que duas normas quaisquer em $I\!R^n$ são equivalentes; como podemos de fato concluir a partir daí que duas normas num espaço vetorial de dimensão finita qualquer são equivalentes? Fazemos o seguinte: seja V um espaço vetorial de dimensão $n < +\infty$ e sejam $\|\cdot\|$ e $\|\cdot\|'$ normas em V. Sabemos da álgebra linear que existe um isomorfismo linear $\phi : I\!R^n \to V$; tal isomorfismo induz normas $\|\cdot\|_\phi$ e $\|\cdot\|'_\phi$ em $I\!R^n$ fazendo:

$$\|x\|_\phi = \|\phi(x)\|, \quad \|x\|'_\phi = \|\phi(x)\|', \qquad x \in I\!R^n,$$

(veja o Exercício 30). Temos então um diagrama comutativo:

$$\begin{tikzcd}
(V, \|\cdot\|) \arrow[r, "\mathrm{Id}"] & (V, \|\cdot\|') \\
(I\!R^n, \|\cdot\|_\phi) \arrow[u, "\phi"] \arrow[r, "\mathrm{Id}"'] & (I\!R^n, \|\cdot\|'_\phi) \arrow[u, "\phi"']
\end{tikzcd}$$

no diagrama acima as flechas verticais são isometrias e a flecha horizontal de baixo é um homeomorfismo, pois as normas $\|\cdot\|_\phi$ e $\|\cdot\|'_\phi$ em $I\!R^n$ são equivalentes. Segue então que a flecha horizontal de cima é um homeomorfismo e portanto $\|\cdot\|$ e $\|\cdot\|'$ são de fato equivalentes em V.

`Exemplo`. Se V, W são espaços vetoriais normados então o conjunto $\mathrm{Lin}(V, W)$ de todas as aplicações lineares contínuas $T : V \to W$ é um subespaço do espaço vetorial de todas as aplicações lineares de V em W; observe que quando V tem dimensão finita então $\mathrm{Lin}(V, W)$ denota simplesmente o espaço de todas as aplicações lineares de V em W. Para $T \in \mathrm{Lin}(V, W)$ a quantidade:

$$\|T\| = \sup_{\substack{x \in V \\ \|x\|=1}} \|T(x)\| = \sup_{\substack{x \in V \\ \|x\|\le 1}} \|T(x)\|$$

é finita. É um exercício simples de sup mostrar que $T \mapsto \|T\|$ é de fato uma norma no espaço vetorial $\mathrm{Lin}(V, W)$; tal norma é conhecida como a *norma de operadores*. As seguintes desigualdades são fáceis de verificar:

$$\|T(x)\| \le \|T\|\,\|x\|, \quad \|T \circ S\| \le \|T\|\,\|S\|,$$

para todos $x \in V$, $T \in \mathrm{Lin}(V,W)$, $S \in \mathrm{Lin}(Z,V)$ (onde Z é um outro espaço normado qualquer). Observe que em $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ temos a norma de operadores e a norma Euclideana bem definidas (podemos identificar $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ com $\mathbb{R}^{mn}$ da maneira óbvia) — tais normas *são diferentes*, mas são equivalentes, pelo que já mostramos.

`Observação`. Uma outra norma famosa em $\mathbb{R}^n$ é a *norma p*; para $p \in \mathbb{R}$, $p \geq 1$ definimos:

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \quad x \in \mathbb{R}^n.$$

Quando $p = 2$ obtemos a norma Euclideana. A desigualdade triangular para $\|\cdot\|_p$ é conhecida como a *desigualdade de Minkowski* e sua demonstração é um tanto chata. Não teremos nenhum uso para a norma p com $p \notin \{1, 2, \infty\}$.

Definição. *Um* produto interno *em V é uma aplicação bilinear*

$$V \times V \ni (x,y) \longmapsto \langle x, y\rangle \in \mathbb{R}$$

satisfazendo as seguintes propriedades:
(PI1) $\langle x, y\rangle = \langle y, x\rangle$, para todos $x, y \in V$ (simetria)*;*
(PI2) $\langle x, x\rangle > 0$ para todo $x \in V$ não nulo (positividade).

Se $\langle\cdot,\cdot\rangle$ é um produto interno num espaço vetorial V então definimos:

$$\|x\| = \langle x, x\rangle^{\frac{1}{2}};$$

a menos da desigualdade triangular, é fácil ver que $\|\cdot\|$ satisfaz todas as propriedades requeridas para uma norma. Veremos logo adiante que $\|\cdot\|$ de fato satisfaz a desigualdade triangular e portanto *é uma norma* em V; dizemos que $\|\cdot\|$ é a *norma induzida pelo produto interno* $\langle\cdot,\cdot\rangle$. Todo espaço vetorial com produto interno será sempre pensado como um espaço vetorial normado através da norma induzida pelo produto interno.

Teorema. *(desigualdade de Cauchy-Schwarz) Se V é um espaço vetorial com produto interno $\langle\cdot,\cdot\rangle$ então para todos $x, y \in V$ temos:*

$$\big|\langle x, y\rangle\big| \leq \|x\|\,\|y\|;$$

a igualdade vale se e somente se x e y são linearmente dependentes.

Demonstração. Seja $f : \mathbb{R} \to \mathbb{R}$ definida por $f(t) = \langle x + ty, x + ty\rangle$; temos:

$$f(t) = \|y\|^2 t^2 + 2\langle x, y\rangle t + \|x\|^2,$$

donde f é um polinômio de grau 2 em t (vamos excluir o caso $y = 0$ que é trivial). Como f é não negativa, temos que o discriminante Δ de f é menor ou igual a zero, ou seja:

$$\Delta = 4\langle x, y\rangle^2 - 4\|x\|^2\|y\|^2 \leq 0;$$

a igualdade vale se e somente se existe $t \in \mathbb{R}$ com $f(t) = 0$ e nesse caso $x + ty = 0$ e x, y são linearmente dependentes. ■

Corolário. *$\|x\| = \langle x, x\rangle^{\frac{1}{2}}$ satisfaz a desigualdade triangular e portanto realmente define uma norma em V.*

Demonstração. Temos:

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2. \blacksquare$$

`Exemplo`. Em $I\!R^n$, o produto interno canônico $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$ é de fato um produto interno; a norma correspondente é a norma Euclideana (e finalmente provamos a desigualdade triangular para essa norma!). *Para quem sabe mais álgebra linear*: se A é uma matriz real simétrica $n \times n$ que possui apenas autovalores positivos então $\langle x, y\rangle' = \sum_{i,j=1}^n A_{ij} x_i y_j$ é um produto interno em $I\!R^n$ (e na verdade todo produto interno em $I\!R^n$ é desse tipo).

`Observação`. O conceito de métrica e topologia faz sentido em conjuntos quaisquer; normas e produtos internos só fazem sentido em espaços vetoriais reais (e na verdade em complexos também, mas não nos interessamos por isso aqui). Num espaço vetorial real V, existem métricas que não vem de normas (como a métrica zero-um) e existem normas que não vem de produto interno (como a norma $\|\cdot\|_p$ com $p \neq 2$). Note que *métricas que não vem de normas podem não ser equivalentes à métricas que vem de norma*, mesmo quando V tem dimensão finita (a métrica zero-um não é equivalente à métrica Euclideana em $I\!R^n$).

`Observação`. Para cultura geral: um espaço normado completo chama-se um *espaço de Banach* e um espaço com produto interno completo chama-se um *espaço de Hilbert*. Esses são os objetos básicos de estudo da análise funcional abstrata. Nesse caso o interesse maior é pelos espaços de dimensão infinita; para tais espaços não é verdade que todas as normas são equivalentes.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Adendo aos exercícios da aula número 5.`

-1. No Exercício 2 da aula número 5 (20/03) acrescentar o seguinte item:
   (g) $\mathrm{ext}(A) = \mathrm{int}(A^c)$.

0. No Exercício 3 da aula número 5 (20/03) acrescentar o seguinte item:
   (e) $\partial A \cap A = \emptyset$.

`Mais espaços métricos.`

1. Mostre que todo subconjunto finito de um espaço métrico é fechado.
2. Seja (M, d) um espaço métrico. Dados $x, y, a \in M$ mostre que:

$$\big|d(x,a) - d(y,a)\big| \le d(x,y).$$

   (*dica*: mostre que $d(x,a) - d(y,a) \le d(x,y)$ e que $d(y,a) - d(x,a) \le d(x,y)$). Conclua que a função *distância até o ponto a* definida por $d_a : M \to I\!R$, $d_a(x) = d(x,a)$, é uma contração fraca e portanto é uniformemente contínua.
3. Seja (M, d) um espaço métrico.
   (a) Mostre que para todos $x, x', y, y' \in M$ vale a desigualdade:

$$\big|d(x,y) - d(x',y')\big| \le d(x,x') + d(y,y').$$

   (*dica*: $d(x,y) - d(x',y') = d(x,y) - d(x',y) + d(x',y) - d(x',y')$; use o Exercício anterior).

   (b) Considere o produto cartesiano $M \times M$ munido da seguinte métrica D:

$$D\big((x,y),(x',y')\big) = d(x,x') + d(y,y'), \quad (x,y),(x',y') \in M \times M;$$

   (recorde do Exercício 33 da aula número 5 que D é uniformemente equivalente à métrica produto em $M \times M$). Conclua do item (a) que a própria métrica de M:

$$d : M \times M \longrightarrow I\!R$$

   é uma contração fraca e portanto é uma função uniformemente contínua.

   (c) Mostre que se $f, g : N \to M$ são funções contínuas (onde N é um outro espaço métrico qualquer) então a função $h : N \to I\!R$ definida por $h(x) = d\big(f(x), g(x)\big)$, é contínua (*dica*: use o item (b)).

   (d) Refaça o item (c) trocando "função contínua" por "função uniformemente contínua" (i.e., suponha f, g uniformemente contínuas e conclua que h também o é).

   (e) Mostre que se $f, g : N \to M$ são funções contínuas então o conjunto

$$\big\{x \in N : f(x) = g(x)\big\}$$

dos pontos aonde elas coincidem é fechado em N — cuidado pois $h = f - g$ *não faz sentido* se M não é um espaço vetorial! (*dica*: mas $h(x) = d\big(f(x), g(x)\big)$ faz sentido).

(f) Mostre que se $(x_n)_{n\geq 1}$ e $(y_n)_{n\geq 1}$ são seqüências de Cauchy em M então a seqüência $\big(d(x_n, y_n)\big)_{n\geq 1}$ é convergente em $I\!R$ (*dica*: funções uniformemente contínuas levam seqüências de Cauchy em seqüências de Cauchy).

4. (*distância de um ponto a um conjunto*) Sejam (M, d) um espaço métrico, $x \in M$ um ponto e $A \subset M$ um subconjunto. A *distância de x até A* é definida por:

$$d(x, A) = \inf_{a\in A} d(x, a).$$

Mostre que:

(a) $d(x, A) = 0$ se e somente se $x \in \overline{A}$; conclua que se A é fechado e $x \notin A$ então $d(x, A) > 0$.

(b) $d(x, A) = d(x, \overline{A})$.

(c) (*desigualdade triangular generalizada*) para $x, y \in M$:

$$d(x, A) \leq d(x, y) + d(y, A).$$

(*dica*: se $\phi(a) \leq \psi(a)$ para todo a então $\inf_a \phi(a) \leq \inf_a \psi(a)$).

(d) Para $x, y \in M$:

$$\big|d(x, A) - d(y, A)\big| \leq d(x, y);$$

conclua que a função *distância até o conjunto A* definida por $d_A : M \to I\!R$, $d_A(x) = d(x, A)$, é uma contração fraca e portanto é uniformemente contínua.

(e) (*o Lema de Urisohn para espaços métricos*) Sejam $F, G \subset M$ fechados disjuntos. Mostre que a expressão:

$$f(x) = \frac{d(x, F)}{d(x, F) + d(x, G)}$$

define uma função contínua $f : M \to [0, 1]$ tal que $f|_F \equiv 0$ e $f|_G \equiv 1$.

(f) (*a propriedade T4 dos espaços métricos*) Conclua do item (d) que se $F, G \subset M$ são fechados disjuntos então existem abertos disjuntos $U, V \subset M$ com $F \subset U$ e $G \subset V$ (*dica*: considere $f^{-1}\big((-\infty, \frac{1}{2})\big)$ e $f^{-1}\big((\frac{1}{2}, +\infty)\big)$).

`Espaços conexos.`

5. Sejam M um espaço métrico e $A \subset M$ um subconjunto. Mostre que as seguintes afirmações são equivalentes:

(a) A é aberto e fechado em M;

(b) A e A^{c} são abertos em M;

(c) A e A^{c} são fechados em M;

(d) a fronteira de A é vazia.

6. Seja M um espaço métrico. Mostre que são equivalentes:
   (a) M é conexo;
   (b) se $M = A \cup B$ com A, B fechados disjuntos em M então $A = \emptyset$ ou $B = \emptyset$;
   (c) se A é aberto e fechado em M então $A = \emptyset$ ou $A = M$;
   (d) se $A \subset M$ tem fronteira vazia então $A = \emptyset$ ou $A = M$.
7. (*componentes conexas*) Seja M um espaço métrico. Defina uma relação $\sim$ em M fazendo:

$$x \sim y \quad \Longleftrightarrow \quad \text{existe } C \subset M \text{ conexo com } x, y \in C,$$

   para todos $x, y \in M$.
   (a) Mostre que $\sim$ é uma relação de equivalência (*dica*: se $C, C' \subset M$ são conexos tais que $x, y \in C$ e $y, z \in C'$, olhe para $C \cup C'$). As classes de equivalência correspondentes a $\sim$ são chamadas as *componentes conexas* de M (mostre que elas são de fato conexas).
   (b) Mostre que se $C \subset M$ é uma componente conexa, $A \subset M$ é conexo e $A \cap C \neq \emptyset$ então $A \subset C$; conclua que a componente conexa C de M que contém um certo ponto x é o maior subconjunto conexo de M que contém x, i.e., se $A \subset M$ é conexo e $x \in A$ então $A \subset C$.
   (c) Mostre que se $A \subset M$ é aberto, fechado, conexo e não vazio então A é uma componente conexa de M (*dica*: se $C \subset M$ é uma componente conexa então $A \cap C$ é aberto e fechado em C).
   (d) Mostre que as componentes conexas de um espaço métrico são fechadas (*dica*: o fecho de um conjunto conexo é conexo).
   (e) Determine as componentes conexas de $\mathbb{Q}$ e mostre que elas *não são abertas* em $\mathbb{Q}$.
   (f) Se $U \subset \mathbb{R}^n$ é um aberto (munido da métrica induzida da métrica Euclideana de $\mathbb{R}^n$), mostre que as componentes conexas de U são abertas em U — e portanto abertas em $\mathbb{R}^n$ (*dica*: uma bola aberta em $\mathbb{R}^n$ é conexa).
   (g) Vamos generalizar o item (f). Um espaço métrico M é dito *localmente conexo* quando todo ponto de M possui um sistema fundamental de vizinhanças conexas (recorde Exercício 19 da aula número 5), i.e., se toda vizinhança de um ponto $x \in M$ contém uma vizinhança conexa de x. Mostre que se M é localmente conexo então as componentes conexas de um aberto U de M são abertas em U — e portanto abertas em M.
8. (*componentes conexas por arcos*) Seja M um espaço métrico. Dadas curvas contínuas $\gamma, \mu : [0,1] \to M$ com $\gamma(1) = \mu(0)$ então a *concatenação* de γ e μ é a curva $\sigma = \gamma \cdot \mu$, $\sigma : [0,1] \to M$, definida por:

$$\sigma(t) = \begin{cases} \gamma(2t), & t \in \left[0, \frac{1}{2}\right], \\ \mu(2t-1), & t \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

   (intuitivamente, σ consiste de γ percorrida com o dobro da velocidade na primeira metade de $[0,1]$, seguida de μ percorrida com o dobro da velocidade na segunda metade de $[0,1]$).

(a) Mostre que $\sigma = \gamma \cdot \mu$ é contínua (*dica*: use o resultado do Exercício 38, aula número 5).

(b) Defina uma relação $\sim$ em M fazendo:

$$x \sim y \iff \text{existe } \gamma : [0,1] \to M \text{ contínua com } \gamma(0) = x \text{ e } \gamma(1) = y,$$

para todos $x, y \in M$. Mostre que $\sim$ é uma relação de equivalência; as classes de equivalência correspondentes a $\sim$ são chamadas as *componentes conexas por arcos* de M (mostre que elas são de fato conexas por arcos).

(c) Mostre que toda componente conexa de M é uma união de componentes conexas por arcos de M.

9. Um subconjunto $S \subset \mathbb{R}^n$ é dito *estrelado* com respeito a $x \in S$ se para todo $y \in S$ o segmento de reta $[x, y]$ está contido em S (observe que S é convexo se e somente se S é estrelado com respeito a todos os seus pontos). Mostre que todo subconjunto estrelado de $\mathbb{R}^n$ é conexo por caminhos (e portanto conexo).

10. Seja $S \subset \mathbb{R}^2$ o gráfico de $(0,1] \ni t \mapsto \operatorname{sen} \frac{1}{t}$, de modo que o fecho de S é a senóide dos topólogos. Vamos mostrar neste exercício que não existe uma curva contínua $\gamma : [0,1] \to \overline{S}$ com $\gamma(0) \in S$ e $\gamma(1) \in \{0\} \times [0,1]$. Suponha por absurdo que tal curva existe e escreva $\gamma(t) = \big(x(t), y(t)\big)$; considere o conjunto:

$$A = \big\{t \in [0,1] : x(t) > 0\big\}.$$

(a) Mostre que A é limitado superiormente e não vazio, de modo que existe $c = \sup A$.

(b) Mostre que $x(c) = 0$.

(c) Mostre que para todo $\varepsilon > 0$ a função $x : [0,1] \to \mathbb{R}$ assume valores da forma $\frac{1}{k\pi}$ e da forma $\frac{1}{\frac{\pi}{2} + k\pi}$ ($k \in \mathbb{N}$) no intervalo $[c - \varepsilon, c]$ (*dica*: use o teorema do valor intermediário).

(d) Conclua do item (c) que $\lim_{t \to c^-} y(t)$ não existe, contradizendo o fato que y é contínua.

11. Mostre que um intervalo fechado $[a, b] \subset \mathbb{R}$ não pode ser homeomorfo ao círculo $S^1 = \big\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\big\}$ (*dica*: é possível tirar um ponto de $[a, b]$ e torná-lo desconexo, mas tirando um ponto qualquer de S^1 ele continua conexo).

`Espaços compactos.`

12. Seja M um espaço métrico. Mostre que:

(a) se M é totalmente limitado então todo subconjunto $S \subset M$ é totalmente limitado;

(b) para $S \subset M$ temos $\operatorname{diam}(S) = \operatorname{diam}(\overline{S})$; conclua que S é totalmente limitado se e somente se $\overline{S}$ é totalmente limitado.

13. Seja M um espaço métrico e seja $(x_n)_{n \geq 1}$ uma seqüência em M. Dizemos que $x \in M$ é um *ponto aderente* (ou um *ponto de aderência*) à seqüência $(x_n)_{n \geq 1}$ se dado $\varepsilon > 0$ então temos $x_n \in \mathrm{B}(x; \varepsilon)$ para n *arbitrariamente grande*, i.e., se o conjunto

$$\big\{n \in \mathbb{N} : x_n \in \mathrm{B}(x; \varepsilon)\big\}$$

é infinito. Mostre que as seguintes condições são equivalentes:

(a) x é um ponto de aderência de $(x_n)_{n\geq 1}$;

(b) dado um aberto $U \subset M$ contendo x então $x_n \in U$ para n arbitrariamente grande (*"ponto aderente a uma seqüência" é uma noção topológica*);

(c) existe uma subseqüência $(x_{n_k})_{k\geq 1}$ de $(x_n)_{n\geq 1}$ convergindo para x (recorde que para que $(x_{n_k})_{k\geq 1}$ seja chamada uma *subseqüência* de $(x_n)_{n\geq 1}$ é necessário que $n_1 < n_2 < n_3 < \cdots$);

(d) x é um ponto de acumulação do conjunto $\big\{x_n : n \in I\!N\big\}$ *ou* a seqüência $(x_n)_{n\geq 1}$ possui infinitos termos iguais a x (i.e., o conjunto $\big\{n \in I\!N : x_n = x\big\}$ é infinito).

`Observação`. Não confundir "x é aderente à seqüência $(x_n)_{n\geq 1}$" com "x é aderente ao conjunto $\big\{x_n : n \in I\!N\big\}$". Por exemplo, $\frac{1}{n}$ é aderente ao conjunto $\big\{\frac{1}{n} : n \in I\!N\big\}$ para todo $n \in I\!N$ mas *nenhum* número da forma $\frac{1}{n}$ é aderente à seqüência $\big(\frac{1}{n}\big)_{n\geq 1}$.

`Observação`. "existem infinitos n's com $x_n \in \mathrm{B}(x;\varepsilon)$" e "existem infinitos x_n's em $\mathrm{B}(x;\varepsilon)$" são afirmações diferentes! O conjunto $\big\{n \in I\!N : x_n \in \mathrm{B}(x;\varepsilon)\big\}$ possui cardinalidade maior ou igual a do conjunto $\big\{x_n : n \in I\!N\big\} \cap \mathrm{B}(x;\varepsilon)$. Por exemplo, se $x_n = x$ para todo $n \in I\!N$ então existe um único x_n em $\mathrm{B}(x;\varepsilon)$ mas existem infinitos n's com $x_n \in \mathrm{B}(x;\varepsilon)$.

14. Seja M um espaço métrico completo. Seja $(F_n)_{n\geq 1}$ uma seqüência de subconjuntos fechados não vazios $F_n \subset M$ com $F_1 \supset F_2 \supset F_3 \supset \cdots$ e $\mathrm{diam}(F_n) \to 0$. Mostre que a interseção $\bigcap_{n\geq 1} F_n$ possui exatamente um ponto (*dica*: escolha $x_n \in F_n$ para cada n e mostre que $(x_n)_{n\geq 1}$ é uma seqüência de Cauchy).

`Observação`. Sem a hipótese $\mathrm{diam}(F_n) \to 0$ no Exercício 14 é possível que a interseção $\bigcap_{n\geq 1} F_n$ seja vazia! Por exemplo, tome $M = I\!R$ e $F_n = [n, +\infty)$.

15. Seja M um espaço métrico e seja $(K_n)_{n\geq 1}$ uma seqüência de compactos não vazios $K_n \subset M$ com $K_1 \supset K_2 \supset K_3 \supset \cdots$. Mostre que a interseção $\bigcap_{n\geq 1} K_n$ é não vazia (*dica*: supondo por absurdo que $\bigcap_{n\geq 1} K_n = \emptyset$, a família $(M \setminus K_n)_{n\geq 1}$ seria uma cobertura aberta de K_1).

16. Mostre que todo espaço métrico finito é compacto. Mostre também que a união finita de subespaços compactos de um espaço métrico ainda é compacta.

17. (*produto por um compacto*) Seja K um espaço métrico compacto e M um espaço métrico qualquer. Dado um ponto $p \in M$ e um aberto $Z \subset K \times M$ que contém $K \times \{p\}$, mostre que existe uma vizinhança aberta V de p em M tal que $K \times V \subset Z$ ("se um aberto contém uma linha compacta então ele contém uma faixa em torno dessa linha") — *dica*: para todo $x \in K$, existem abertos $U_x \subset K$ e $V_x \subset M$ com $x \in U_x$, $p \in V_x$ e $U_x \times V_x \subset Z$ (veja o Exercício 20(c) da aula número 5); adivinhe o que fazer com a cobertura aberta $K = \bigcup_{x\in K} U_x$!

18. (*mais uniformidades*) Sejam K um espaço métrico compacto e M, N espaços métricos quaisquer. Seja $f : K \times M \to N$ uma função contínua; mostre que a continuidade de f é *uniforme com respeito à variável em* K, i.e., que dados $y_0 \in M$ e $\varepsilon > 0$ então existe $\delta > 0$ tal que para todo $x \in K$ e todo $y \in \mathrm{B}(y_0;\delta) \subset M$ temos $d\big(f(x,y), f(x,y_0)\big) < \varepsilon$ ($\delta > 0$ não depende de $x \in K$) — *dica*: o conjunto

$$Z = \big\{(x,y) \in K \times M : d\big(f(x,y), f(x,y_0)\big) < \varepsilon\big\}$$

é aberto em $K \times M$ e contém a linha $K \times \{y_0\}$; use o Exercício anterior.

19. (*distância entre conjuntos*) Seja M um espaço métrico. Dados subconjuntos $A, B \subset M$ definimos a *distância entre A e B* por:

$$d(A,B) = \inf_{\substack{a\in A\\ b\in B}} d(a,b) = \inf_{a\in A} d(a,B) = \inf_{b\in B} d(A,b).$$

Mostre que:

(a) se $K \subset M$ é compacto e $F \subset M$ é fechado com $K \neq \emptyset$ então existe $x \in K$ com $d(x,F) = d(K,F)$ (*dica*: $K \ni x \mapsto d(x,F)$ é uma função contínua num compacto). Conclua que se $K \cap F = \emptyset$ então $d(K,F) > 0$.
(b) Se $K_1, K_2 \subset M$ são compactos não vazios então existem $x \in K_1$, $y \in K_2$ com $d(x,y) = d(K_1,K_2)$, i.e., *a distância entre compactos é efetivamente assumida* (*dica*: $d|_{K_1\times K_2}$ é uma função contínua num compacto).
(c) Se $K \subset \mathbb{R}^n$ é compacto e $F \subset \mathbb{R}^n$ é fechado, com K e F não vazios, então existem $x \in K$, $y \in F$ com $d(x,y) = d(K,F)$, i.e., *a distância entre um compacto e um fechado em $\mathbb{R}^n$ é efetivamente assumida* (*dica*: escolha $k > 0$ grande tal que $F_0 = \big\{y \in F : d(y,K) \le k\big\}$ é não vazio; daí F_0 é compacto e $d(K,F) = d(K,F_0)$).
(d) Ache subconjuntos fechados disjuntos $F_1, F_2 \subset \mathbb{R}^2$ tais que $d(F_1,F_2) = 0$.

`Observação`. O item (c) do Exercício 19 *não vale* se o espaço métrico ambiente não é o $\mathbb{R}^n$; por exemplo, se $M = \mathbb{R} \setminus \{0\}$, $K = \{-1\}$ e $F = (0,+\infty)$ então F é fechado em M, K é compacto, mas a distância entre K e F não é efetivamente assumida. É possível também obter contra-exemplos onde o espaço ambiente M é completo.

20. Mostre que se $(x_n)_{n\ge1}$ é uma seqüência convergente num espaço métrico M, digamos $x_n \to x \in M$, então o conjunto $\big\{x_n : n \ge 1\big\} \cup \{x\}$ é compacto.

21. Sejam M um espaço métrico e $S \subset M$ um subconjunto. Dizemos que S é *relativamente compacto* em M se o fecho de S em M é compacto ("relativamente compacto" *não* é uma noção intrínseca). Mostre que as seguintes condições são equivalentes:

(a) S é relativamente compacto em M;
(b) toda seqüência em S possui uma subseqüência convergente em M.

22. Seja M um espaço métrico discreto (recorde Exercício 40 da aula número 5). Mostre que M é compacto se e somente se M é finito.

23. Mostre que a esfera $S^2 = \big\{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\big\}$ não é homeomorfa a $\mathbb{R}^2$, mas que a esfera S^2 menos um ponto é homeomorfa a $\mathbb{R}^2$ (lembram da projeção estereográfica?).

`Base enumerável e separabilidade`.

24. Sejam M um espaço métrico e $\mathfrak{B}$ uma coleção de abertos de M. Mostre que as seguintes propriedades são equivalentes:

(i) $\mathfrak{B}$ é uma base de abertos para M;
(ii) dado $U \subset M$ aberto e $x \in U$ então existe $V \in \mathfrak{B}$ com $x \in V \subset U$.

25. (*para acompanhar este exercício, você deve recordar os Exercícios 19 e 20 da aula número 5*) Seja M um espaço métrico. Mostre que:
   (a) se $\mathfrak{B}$ é uma base de abertos para M então para todo $x \in M$ a coleção
   $$\mathcal{V}_x = \big\{ V \in \mathfrak{B} : x \in V \big\}$$
   é um sistema fundamental de vizinhanças abertas para $x \in M$.
   (b) Se para cada $x \in M$ escolhemos um sistema fundamental de vizinhanças abertas $\mathcal{V}_x$ de $x \in M$ então a coleção $\mathfrak{B} = \bigcup_{x \in M} \mathcal{V}_x$ é uma base de abertos para M.
   (c) Se (M_i, d_i), $i = 1, \ldots, n$, são espaços métricos então:
   $$\mathfrak{B} = \Big\{ \textstyle\prod_{i=1}^n U_i : U_i \subset M_i \text{ aberto}, \ i = 1, \ldots, n \Big\}$$
   é uma base de abertos para o produto $\prod_{i=1}^n M_i$. Mais geralmente, se $\mathfrak{B}_i$ é uma base de abertos para M_i então:
   $$\mathfrak{B} = \Big\{ \textstyle\prod_{i=1}^n U_i : U_i \in \mathfrak{B}_i, \ i = 1, \ldots, n \Big\}$$
   é uma base de abertos para $\prod_{i=1}^n M_i$.
   (d) Conclua do item (c) que o produto de espaços métricos com base enumerável ainda é um espaço métrico com base enumerável.
26. Mostre que todo espaço métrico totalmente limitado M é separável (*dica*: para todo $n \geq 1$ escolha $S_n \subset M$ finito tal que $M = \bigcup_{x \in S_n} \mathrm{B}\big(x; \frac{1}{n}\big)$ — defina $D = \bigcup_{n \geq 1} S_n$).

`Espaços normados e espaços com produto interno.`

Vamos generalizar alguns resultados da seção para aplicações bilineares e multi-lineares:

27. Sejam V, W e Z espaços vetoriais normados. Sobre uma aplicação bilinear $B : V \times W \to Z$ mostre que são equivalentes:
   (i) B é contínua;
   (ii) B é contínua em $(0, 0)$;
   (iii) existe $c \geq 0$ com $\|B(x, y)\| \leq c$ para todos $x \in V$, $y \in W$ com $\|x\| = \|y\| = 1$;
   (iv) existe $c \geq 0$ com $\|B(x, y)\| \leq c\|x\| \, \|y\|$ para todo $x \in V$, $y \in W$.
   Generalize o exercício para aplicações multi-lineares.

`Observação.` Segue do Exercício acima e da desigualdade de Cauchy-Schwarz que um produto interno é contínuo (com respeito à métrica definida por ele mesmo).

`Observação.` Aplicações bilineares em geral não são Lipschitzianas (e nem mesmo uniformemente contínuas).

28. Sejam V, W e Z espaços vetoriais normados. Mostre que se V e W tem dimensão finita então toda aplicação bilinear $B : V \times W \to Z$ é contínua (*dica:* troque Z por um subespaço de dimensão finita que contém a imagem de B; podemos supor então que B é uma aplicação bilinear $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ onde $\mathbb{R}^m$, $\mathbb{R}^n$ e $\mathbb{R}^p$ possuem a norma Euclideana). Generalize o resultado para aplicações multi-lineares.

29. Sejam V, W e Z espaços vetoriais normados. Denote por $\mathrm{Bil}(V,W;Z)$ o conjunto de todas as aplicações bilineares contínuas $B : V \times W \to Z$. Mostre que:

    (a) $\mathrm{Bil}(V,W;Z)$ é um subespaço do espaço vetorial de todas as aplicações bilineares $V \times W \to Z$;

    (b) a fórmula:

    $$\|B\| = \sup_{\substack{x\in V,\ \|x\|=1\\ y\in W,\ \|y\|=1}} \|B(x,y)\| = \sup_{\substack{x\in V,\ \|x\|\le 1\\ y\in W,\ \|y\|\le 1}} \|B(x,y)\|,$$

    define uma norma em $\mathrm{Bil}(V,W;Z)$;

    (c) generalize os itens (a) e (b) para o caso de aplicações multi-lineares.

30. (*norma induzida por um isomorfismo*) Sejam V, W espaços vetoriais, $\phi : V \to W$ um isomorfismo e $\|\cdot\|$ uma norma em W. Defina:

    $$\|x\|_\phi = \|\phi(x)\|,$$

    para todo $x \in V$. Mostre que $\|\cdot\|_\phi$ é uma norma em V e que $\phi : (V, \|\cdot\|_\phi) \to (W, \|\cdot\|)$ é uma isometria. Dizemos que $\|\cdot\|_\phi$ é a norma em V *induzida* pelo isomorfismo ϕ e pela norma de W.

31. (*norma num subespaço*) Sejam V um espaço vetorial e $S \subset V$ um subespaço. Mostre que se $\|\cdot\|$ é uma norma em V então $\|\cdot\|$ restringe-se a uma norma $\|\cdot\||_S$ em S. Mostre que a métrica induzida em S por $\|\cdot\||_S$ coincide com a restrição da métrica induzida por $\|\cdot\|$ em V.

32. (*produto interno induzido por uma aplicação*) Seja $\phi : V \to W$ um isomorfismo entre espaços vetoriais V, W e seja $\langle\cdot,\cdot\rangle$ um produto interno em W. Mostre que:

    $$\langle x, y\rangle_\phi = \langle \phi(x), \phi(y)\rangle, \quad x, y \in V,$$

    define um produto interno em V; dizemos que $\langle\cdot,\cdot\rangle_\phi$ é o produto interno *induzido* por ϕ e por $\langle\cdot,\cdot\rangle$. Mostre que $\phi : (V, \langle\cdot,\cdot\rangle_\phi) \to (W, \langle\cdot,\cdot\rangle)$ é uma isometria.

33. (*produto interno induzido num subespaço*) Sejam V um espaço vetorial e $S \subset V$ um subespaço. Seja $\langle\cdot,\cdot\rangle$ um produto interno em V. Mostre que $\langle\cdot,\cdot\rangle$ restringe-se a um produto interno $\langle\cdot,\cdot\rangle|_{S\times S}$ em S. Mostre que a norma induzida em S por $\langle\cdot,\cdot\rangle|_{S\times S}$ coincide com a restrição da norma induzida em V por $\langle\cdot,\cdot\rangle$.

`Observação`. É possível unificar os exercícios 30 e 31 (assim como os exercícios 32 e 33) trocando a hipótese que ϕ seja um isomorfismo pela hipótese que ϕ seja linear injetora (obviamente nesse caso vamos concluir apenas que ϕ é uma imersão isométrica em vez de concluir que ϕ é uma isometria).

34. (*generalizando os Exercícios 18 e 33 da aula 5*) Seja $\|\cdot\|$ uma norma em $I\!R^n$ satisfazendo a seguinte propriedade:

$$\text{se } x, y \in [0, +\infty)^n \text{ são tais que } x_i \leq y_i,\ i = 1, \ldots, n \text{ então } \|x\| \leq \|y\| \qquad (*)$$

(observe que a norma Euclideana e as normas $\|\cdot\|_p$ satisfazem a propriedade $(*)$).

(a) Sejam V_i, $i = 1, \ldots, n$ espaços vetoriais; o produto $V = \prod_{i=1}^n V_i$ também é um espaço vetorial com as operações definidas coordenada por coordenada — tal espaço é usualmente denotado por $\bigoplus_{i=1}^n V_i$ (ou $V_1 \oplus \cdots \oplus V_n$) e é chamado a *soma direta* ou a *soma direta externa* dos espaços V_i. Se $\|\cdot\|^i$ é uma norma em V_i, defina:

$$\|x\|^{\text{prod}} = \Big\| \big(\|x_1\|^1, \ldots, \|x_n\|^n \big) \Big\|,$$

para todo $x = (x_i)_{i=1}^n \in V$. Mostre que $\|\cdot\|^{\text{prod}}$ é de fato uma norma em V; mostre também que trocando $\|\cdot\|$ por uma outra norma $\|\cdot\|'$ em $I\!R^n$ (satisfazendo $(*)$) então obtemos uma norma equivalente a $\|\cdot\|^{\text{prod}}$ em V.

(b) Sejam (M_i, d_i), $i = 1, \ldots, n$ espaços métricos. Seja $M = \prod_{i=1}^n M_i$ e defina:

$$d^{\text{prod}}(x, y) = \Big\| \big(d_1(x_1, y_1), \ldots, d_n(x_n, y_n) \big) \Big\|,$$

para todos $x = (x_i)_{i=1}^n$, $y = (y_i)_{i=1}^n$ em M. Mostre que d^{prod} é de fato uma métrica em M; mostre também que d^{prod} é uniformemente equivalente à métrica produto usual em M (a equivalência é até mesmo "Lipschitz").

(c) Se d_i é a métrica induzida pela norma $\|\cdot\|^i$ no espaço V_i, mostre que d^{prod} é a métrica induzida pela norma $\|\cdot\|^{\text{prod}}$ em $V = \bigoplus_{i=1}^n V_i$.

Aula número 8 (29/03)

A aula número 8 cobriu o material das Seções (2) e (4) originalmente destinado à aula número 7 — os seguintes assuntos foram omitidos:

(i) a noção de número de Lebesgue de uma cobertura;
(ii) a seção (3) sobre base enumerável e separabilidade;
(iii) a noção de normas equivalentes e os resultados relativos à continuidade de aplicações lineares.

Aula número 9 (03/04)

A aula começa cobrindo a parte que faltou da seção (4) da aula número 7 (equivalência de normas, continuidade de aplicações lineares cujo domínio tem dimensão finita e norma de operadores).

(1) **A prova da regra da cadeia**.

Sejam $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$ abertos e $f : U \to \mathbb{R}^n$, $g : V \to \mathbb{R}^p$ funções com $f(U) \subset V$. Suponha que f é diferenciável num ponto $x \in U$ e que g é diferenciável no ponto $f(x) \in V$. Nós devemos mostrar que $g \circ f$ é diferenciável no ponto $x \in U$ e que sua diferencial é dada por:

$$\mathrm{d}(g \circ f)(x) = \mathrm{d}g\big(f(x)\big) \circ \mathrm{d}f(x).$$

Escrevemos:

$$f(x + h) = f(x) + \mathrm{d}f(x) \cdot h + \rho(h)\|h\|,$$
$$g\big(f(x) + k\big) = g\big(f(x)\big) + \mathrm{d}g\big(f(x)\big) \cdot k + \sigma(k)\|k\|,$$

com $\lim_{h\to 0} \rho(h) = 0$ e $\lim_{k\to 0} \sigma(k) = 0$. Aplicamos g dos dois lados da primeira identidade acima e utilizamos a segunda com $k = \mathrm{d}f(x) \cdot h + \rho(h)\|h\|$; obtemos:

$$(g \circ f)(x + h) = (g \circ f)(x) + \big[\mathrm{d}g\big(f(x)\big) \circ \mathrm{d}f(x)\big] \cdot h + \|h\|\big[\mathrm{d}g\big(f(x)\big) \cdot \rho(h)\big] + \sigma(k)\|k\|.$$

Como $\mathrm{d}g\big(f(x)\big) \circ \mathrm{d}f(x)$ é linear, a conclusão segue se mostrarmos que:

$$\lim_{h\to 0} \mathrm{d}g\big(f(x)\big) \cdot \rho(h) + \sigma(k)\frac{\|k\|}{\|h\|} = 0.$$

Sabemos que $\lim_{h\to 0} \rho(h) = 0$ e pela continuidade da transformação linear $\mathrm{d}g\big(f(x)\big)$ obtemos $\lim_{h\to 0} \mathrm{d}g\big(f(x)\big) \cdot \rho(h) = 0$. Além do mais, temos:

$$\|k\| \le \big[\|\mathrm{d}f(x)\| + \|\rho(h)\|\big]\|h\|,$$

donde segue que $\lim_{h\to 0} k = 0$ e que $\frac{\|k\|}{\|h\|}$ é uma quantidade limitada para h numa vizinhança da origem. A conclusão segue do fato que $\lim_{k\to 0} \sigma(k) = 0$. ■

(2) **A igualdade e a desigualdade do valor médio**.

Recordem do Cálculo 1 o seguinte:

Teorema. *(do valor médio) Se $f : [a, b] \to I\!R$ é uma função contínua, derivável no intervalo aberto (a, b), então existe $c \in (a, b)$ tal que:*

$$f(b) - f(a) = f'(c)(b - a).$$

Demonstração. Fazemos primeiro o caso $f(a) = f(b)$ (conhecido como o *Teorema de Rolle*). A função contínua f assume um máximo e um mínimo no compacto $[a, b]$; se ambos fossem assumidos nas extremidades do intervalo então f seria constante (pois $f(a) = f(b)$) — daí $f' \equiv 0$. Se f assume um máximo ou um mínimo num ponto $c \in (a, b)$ então (como todo mundo deve saber do Cálculo I) temos $f'(c) = 0$.

O caso geral segue diretamente aplicando o Teorema de Rolle para a função $g(t) = f(t) - \big(f(b) - f(a)\big)\frac{t-a}{b-a} - f(a)$. ■

É muito fácil generalizar o teorema acima para o caso de funções definidas em abertos de $I\!R^m$ tomando valores em $I\!R$; para isso, recorde que o *segmento de reta* ligando dois pontos $x, y \in I\!R^m$ é definido por:

$$[x, y] = \big\{(1 - t)x + ty : t \in [0, 1]\big\}.$$

Denotamos também por (x, y) o *segmento de reta aberto* definido por:

$$(x, y) = \big\{(1 - t)x + ty : t \in (0, 1)\big\}.$$

(espero que não ocorra confusão com o par ordenado (x, y)).

Teorema. *Sejam $f : U \subset I\!R^m \to I\!R$ uma função definida num aberto $U \subset I\!R^m$ e $[x, y] \subset U$ um segmento de reta* contido *em U. Suponha que f é contínua nos pontos de $[x, y]$ e diferenciável nos pontos de (x, y); então existe $z \in (x, y)$ tal que:*

$$f(y) - f(x) = \mathrm{d}f(z) \cdot (y - x).$$

Demonstração. Seja $\phi : [0, 1] \to I\!R$ a aplicação definida por $\phi(t) = f\big((1 - t)x + ty\big)$; note que f é bem definida pois $[x, y] \subset U$. Como f é contínua em $[x, y]$, temos que ϕ é contínua; como f é diferenciável em (x, y), segue da regra da cadeia que ϕ é diferenciável em $(0, 1)$ e que sua derivada é dada por:

$$\phi'(t) = \mathrm{d}f\big((1 - t)x + ty\big) \cdot (y - x),$$

para todo $t \in (0, 1)$. A conclusão segue diretamente aplicando o teorema do valor médio do Cálculo I para ϕ. ■

`Exemplo`. O teorema do valor médio não vale em geral para funções a valores vetoriais, i.e., para funções a valores em $I\!R^n$ com $n > 1$. Por exemplo, a curva diferenciável $\gamma : [0, 2\pi] \to I\!R^2$ definida por $\gamma(t) = (\cos t, \operatorname{sen} t)$ satisfaz $\gamma(0) = \gamma(2\pi)$ mas não existe $t \in [0, 2\pi]$ com $\gamma'(t) = 0$ (*para refletir*: por que a idéia óbvia de aplicar o teorema do valor médio em cada coordenada de γ não funciona?).

A generalização correta do teorema do valor médio para funções a valores vetoriais é dada no seguinte:

Teorema. *(desigualdade do valor médio) Sejam $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$ e $[x, y] \subset U$ um segmento de reta* contido *em U. Suponha que f é contínua nos pontos de $[x, y]$ e diferenciável nos pontos de (x, y); vale a desigualdade:*

$$\big\|f(y) - f(x)\big\| \le \sup_{z\in(x,y)} \big\|\mathrm{d}f(z)\big\| \, \|y - x\|,$$

onde utilizamos normas arbitrárias *em $\mathbb{R}^m$ e $\mathbb{R}^n$ e a norma usada em* $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ *(para fazer $\big\|\mathrm{d}f(z)\big\|$) é a* norma de operadores *correspondente.*

A demonstração da desigualdade do valor médio segue facilmente do teorema do valor médio se utilizarmos o seguinte:

Lema. *Seja V um espaço vetorial normado não nulo. Dado $v \in V$ então existe um funcional linear $\lambda : V \to \mathbb{R}$ com $\|\lambda\| = 1$ e $\lambda(v) = \|v\|$.*

Dizemos que λ é um funcional linear de norma 1 que *reproduz a norma* do vetor $v \in V$.

`Observação`. Para quem não sabe, um *funcional linear* num espaço vetorial real V é simplesmente uma aplicação linear $\lambda : V \to \mathbb{R}$ definida em V tomando valores em $\mathbb{R}$. O espaço $\mathrm{Lin}(V, \mathbb{R})$ de todos os funcionais lineares em V é chamado o *espaço dual* de V e é denotado por V^*. Mais adiante no curso precisaremos fazer uma revisão sobre a teoria do espaço dual — isso será importante para entender a teoria sobre integral de linha usando 1-formas.

Prova da desigualdade do valor médio. Pelo lema, existe um funcional linear $\lambda \in {\mathbb{R}^n}^*$ com $\|\lambda\| = 1$ e que reproduz a norma do vetor $f(y) - f(x) \in \mathbb{R}^n$, i.e.:

$$\lambda\big(f(y) - f(x)\big) = \big\|f(y) - f(x)\big\|.$$

Aplicando o teorema do valor médio para a função $\lambda \circ f : U \subset \mathbb{R}^m \longrightarrow \mathbb{R}$ obtemos que existe $z \in (x, y)$ com:

$$\lambda\big(f(y)\big) - \lambda\big(f(x)\big) = \mathrm{d}(\lambda \circ f)(z) \cdot (y - x);$$

usando a regra da cadeia e a linearidade de λ obtemos:

$$\lambda\big(f(y) - f(x)\big) = \lambda\big[\mathrm{d}f(z) \cdot (y - x)\big].$$

Finalmente, como $\|\lambda\| = 1$ e λ reproduz a norma de $f(y) - f(x)$ obtemos:

$$\begin{aligned}\big\|f(y) - f(x)\big\| &= \lambda\big(f(y) - f(x)\big) = \lambda\big[\mathrm{d}f(z) \cdot (y - x)\big] \le \Big|\lambda\big[\mathrm{d}f(z) \cdot (y - x)\big]\Big| \\ &\le \big\|\mathrm{d}f(z) \cdot (y - x)\big\| \le \big\|\mathrm{d}f(z)\big\| \, \|y - x\| \le \sup_{z\in(x,y)} \big\|\mathrm{d}f(z)\big\| \, \|y - x\|. \ \blacksquare\end{aligned}$$

O lema sobre a existência de um funcional $\lambda \in V^*$ que reproduz a norma de um vetor dado é bem fácil de ser provado quando a norma de V provém de um produto interno $\langle\cdot,\cdot\rangle$. Fazemos assim: tomamos $v \in V$ e consideramos o funcional $\lambda = \big\langle \frac{v}{\|v\|}, \cdot \big\rangle$ associado ao vetor unitário $\frac{v}{\|v\|}$; é fácil ver que $\|\lambda\| = 1$ e que $\lambda(v) = \|v\|$ (veja o Exercício 1). Essa demonstração não funciona para $v = 0$, mas esse caso é trivial (basta escolher *qualquer* $\lambda \in V^*$ com $\|\lambda\| = 1$).

A demonstração do lema para espaços normados quaisquer seguirá do seguinte:

Teorema de Hahn–Banach. *Sejam V um espaço vetorial normado, $S \subset V$ um subespaço e $\lambda : S \to \mathbb{R}$ um funcional linear. Então existe uma extensão linear $\mu : V \to \mathbb{R}$ de λ (i.e., $\lambda = \mu|_S$) tal que $\|\mu\| = \|\lambda\|$ (onde $\|\mu\|$ e $\|\lambda\|$ denotam as normas de operadores de $\mu \in V^*$ e $\lambda \in S^*$ respectivamente).*

Antes de provar o Teorema de Hahn–Banach, vamos mostrar como o lema sobre funcionais que reproduzem a norma de vetores segue diretamente de tal teorema. Seja V um espaço vetorial normado e seja $v \in V$ não nulo (o caso $v = 0$ é trivial). Seja $S = \mathbb{R}v \subset V$ o subespaço unidimensional gerado por v. Podemos definir $\lambda \in S^*$ fazendo $\lambda(v) = \|v\|$ (pois v é uma base para S) e é fácil ver que $\|\lambda\| = 1$ (a esfera unitária de S possui apenas os vetores $\pm\frac{v}{\|v\|}$). Pelo Teorema de Hahn–Banach podemos estender λ a um funcional linear em V que ainda possui norma 1 (e que ainda reproduz a norma de v, obviamente).

Prova do Teorema de Hahn–Banach. A parte central da demonstração do Teorema de Hahn–Banach é a seguinte: consideramos um vetor $v \in V$ fora de S e mostramos que é possível estender λ para um funcional linear μ no espaço $S \oplus \mathbb{R}v$ gerado por S e v, de modo que $\|\mu\| = \|\lambda\|$. Uma vez que esse resultado tenha sido demonstrado, o teorema de Hahn–Banach segue da seguinte maneira:

(i) se V tem dimensão finita (que é o caso que nos interessa) então basta repetir o processo descrito acima um número finito de vezes e eventualmente obteremos uma extensão de λ para V;

(ii) se V é arbitrário, a conclusão final segue facilmente do *Lema de Zorn*. O caso de espaços de dimensão infinita é totalmente irrelevante para esse curso, de modo que quem não tem familiaridade com o Lema de Zorn pode simplesmente ignorar esse caso.

Vamos então demonstrar o passo central, i.e., que podemos estender $\lambda \in S^*$ para $\mu \in (S \oplus \mathbb{R}v)^*$ de modo que $\|\lambda\| = \|\mu\|$. A extensão μ de λ fica totalmente definida quando escolhemos o valor de μ em v (veja o Exercício 3); esse valor será um número real $c \in \mathbb{R}$. Daí:

$$\mu(x + tv) = \lambda(x) + tc,$$

para todos $x \in S$ e $t \in \mathbb{R}$. Obviamente teremos $\|\mu\| \geq \|\lambda\|$ para qualquer escolha de $c \in \mathbb{R}$; o problema é mostrar que é possível escolher $c \in \mathbb{R}$ de modo que:

$$\big|\lambda(x) + tc\big| \leq \|\lambda\| \, \|x + tv\|,$$

para todos $x \in S$, $t \in \mathbb{R}$. A condição acima é trivialmente satisfeita para $t = 0$; se $t \neq 0$, podemos dividir os dois lados por $|t|$, obtendo a condição equivalente:

$$\big|\lambda\big(\tfrac{x}{t}\big) + c\big| \leq \|\lambda\| \, \big\|\tfrac{x}{t} + v\big\|,$$

para todos $x \in S$, $t \in \mathbb{R} \setminus \{0\}$. Observe que $\frac{x}{t}$ é também um vetor em S e portanto nosso trabalho ficará completo se mostrarmos que $c \in \mathbb{R}$ pode ser escolhido de modo que:

$$\big|\lambda(y) + c\big| \leq \|\lambda\| \, \|y + v\|,$$

para todo $y \in S$. A condição acima é equivalente a:

$$-\|\lambda\| \, \|y+v\| \le \lambda(y) + c \le \|\lambda\| \, \|y+v\|;$$

a última é por sua vez equivalente a:

$$-\|\lambda\| \, \|y+v\| - \lambda(y) \le c \le \|\lambda\| \, \|y+v\| - \lambda(y),$$

para todo $y \in S$. Para completar a demonstração, escolheremos $c \in I\!R$ tal que:

$$\sup_{y\in S} \big[-\|\lambda\| \, \|y+v\| - \lambda(y)\big] \le c \le \inf_{y\in S} \big[\|\lambda\| \, \|y+v\| - \lambda(y)\big];$$

para mostrar que isso é possível, devemos mostrar antes que:

$$\sup_{y\in S} \big[-\|\lambda\| \, \|y+v\| - \lambda(y)\big] \le \inf_{y\in S} \big[\|\lambda\| \, \|y+v\| - \lambda(y)\big].$$

A condição acima é equivalente a:

$$-\|\lambda\| \, \|y+v\| - \lambda(y) \le \|\lambda\| \, \|z+v\| - \lambda(z),$$

para todos $y, z \in S$. Temos agora:

$$\begin{aligned} -\|\lambda\| \, \|y+v\| - \lambda(y) &\le \|\lambda\| \, \|z+v\| - \lambda(z) \\ &\iff \lambda(z) - \lambda(y) \le \|\lambda\| \big[\|y+v\| + \|z+v\|\big] \\ &\iff \lambda(z-y) \le \|\lambda\| \big[\|y+v\| + \|z+v\|\big]; \end{aligned}$$

mas:

$$\begin{aligned} \lambda(z-y) \le \big|\lambda(z-y)\big| &\le \|\lambda\| \, \|z-y\| = \|\lambda\| \, \big\|(z+v) - (y+v)\big\| \\ &\le \|\lambda\| \big[\|z+v\| + \|y+v\|\big], \end{aligned}$$

para todos $y, z \in S$. Isso completa a demonstração do Teorema de Hahn–Banach. ■

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Álgebra Linear.`

1. Seja V um espaço vetorial real munido de um produto interno $\langle\cdot,\cdot\rangle$. A forma bilinear $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ corresponde à aplicação linear:

$$\phi : V \ni v \longmapsto \langle v, \cdot\rangle \in \mathrm{Lin}(V, \mathbb{R}) = V^*$$

onde $\langle v, \cdot\rangle$ denota o funcional linear $V \ni w \mapsto \langle v, w\rangle \in \mathbb{R}$.

   (a) Mostre que ϕ é injetor (*dica*: ϕ injetor $\Leftrightarrow$ $\mathrm{Ker}(\phi) = 0$; quanto vale $\phi(v)(v)$?). Conclua que se V tem dimensão finita então $\phi : V \to V^*$ é um isomorfismo.

   (b) Suponha que V tem dimensão finita e considere $\mathrm{Lin}(V, \mathbb{R}) = V^*$ munido da norma de operadores. Mostre que ϕ é uma isometria, i.e., que $\|\phi(v)\| = \|v\|$ para todo $v \in V$ (*dica*: para mostrar que $\|\phi(v)\| \le \|v\|$ use a desigualdade de Cauchy-Schwarz; para mostrar a igualdade avalie $\phi(v)$ em $\frac{v}{\|v\|}$).

2. (*somas diretas*) Sejam V um espaço vetorial e $S_1, S_2 \subset V$ subespaços. Denotamos por $S_1 + S_2$ a *soma* dos subespaços S_1 e S_2, ou seja:

$$S_1 + S_2 = \big\{v_1 + v_2 : v_1 \in V_1,\ v_2 \in V_2\big\};$$

$S_1 + S_2$ coincide com o *subespaço gerado* pela união $S_1 \cup S_2$, i.e., $S_1 + S_2$ é o menor subespaço de V que contém $S_1 \cup S_2$. Mostre que as seguintes condições são equivalentes:

   (a) $S_1 \cap S_2 = \{0\}$ e $V = S_1 + S_2$;

   (b) todo $v \in V$ se escreve de modo *único* na forma $v = v_1 + v_2$ com $v_1 \in S_1$ e $v_2 \in S_2$.

Quando uma das condições acima é satisfeita dizemos que V é a *soma direta* (ou a *soma direta interna*) dos subespaços S_1 e S_2 e escrevemos $V = S_1 \oplus S_2$; dizemos também que S_2 é um *subespaço complementar* a S_1 (*não* confundir com o *conjunto complementar* $V \setminus S_1 = \big\{v \in V : v \notin S_1\big\}$ que não é sequer um subespaço). Defina aplicações $P_1 : V \to S_1$ e $P_2 : V \to S_2$ da seguinte forma: dado $v \in V$ escreva $v = v_1 + v_2$ com $v_1 \in V_1$, $v_2 \in V_2$ (isso pode ser feito de modo único) e defina:

$$P_1(v) = v_1, \quad P_2(v) = v_2;$$

mostre que as aplicações P_1 e P_2 são lineares (elas são chamadas as *projeções* relativas à soma direta $V = S_1 \oplus S_2$).

3. Seja V um espaço vetorial e suponha que V se escreve como soma direta $V = V_1 \oplus V_2$ de dois subespaços V_1 e V_2. Dados um outro espaço vetorial W qualquer e aplicações lineares $T_1 : V_1 \to W$, $T_2 : V_2 \to W$, mostre que existe *uma única* aplicação linear $T : V \to W$ tal que $T|_{V_1} = T_1$ e $T|_{V_2} = T_2$ (*dica*: defina $T = T_1 \circ P_1 + T_2 \circ P_2$, onde P_1 e P_2 denotam as projeções relativas à soma direta $V = V_1 \oplus V_2$).

4. Sejam V um espaço vetorial e $S_1, S_2 \subset V$ subespaços de V. Mostre que:
   (a) se $V = S_1 \oplus S_2$ então a concatenação de uma base qualquer de S_1 com uma base qualquer de S_2 produz uma base de V;
   (b) se a concatenação de alguma base de S_1 com alguma base de S_2 produz uma base de V então $V = S_1 \oplus S_2$.

   Conclua que todo subespaço de V admite um subespaço complementar (*dica*: todo subconjunto linearmente independente estende-se a uma base).

5. (*somas diretas externas*) Se S_1 e S_2 são espaços vetoriais (não necessariamente subespaços de um espaço vetorial maior dado) então definimos no produto cartesiano $S_1 \times S_2$ as operações de *soma e produto por escalar coordenada por coordenada*:

$$(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2), \quad c(v_1, v_2) = (cv_1, cv_2).$$

   Mostre que $S_1 \times S_2$ munido das operações acima torna-se um espaço vetorial; dizemos que $S_1 \times S_2$ é a *soma direta* (ou a *soma direta externa*) dos espaços vetoriais S_1 e S_2. Usualmente escrevemos $S_1 \oplus S_2$ em vez de $S_1 \times S_2$.
   (a) Se $S_1 \oplus S_2$ denota a soma direta externa de S_1 e S_2, mostre que as aplicações:

$$S_1 \ni v_1 \longmapsto (v_1, 0) \in S_1 \oplus S_2, \quad S_2 \ni v_2 \longmapsto (0, v_2) \in S_1 \oplus S_2$$

   são lineares injetoras e que $S_1 \oplus S_2$ é a soma direta interna de suas imagens. As imagens das aplicações lineares injetoras acima são usualmente identificadas com os espaços S_1 e S_2.
   (b) Se um espaço vetorial V escreve-se como soma direta interna de subespaços S_1 e S_2, mostre que a aplicação:

$$\underbrace{S_1 \oplus S_2}_{\substack{\text{soma direta} \\ \text{externa}}} \ni (v_1, v_2) \longmapsto v_1 + v_2 \in \underbrace{S_1 \oplus S_2}_{\substack{\text{soma direta} \\ \text{interna}}} = V$$

   é um isomorfismo.

6. (*somas diretas arbitrárias*) Tente generalizar os Exercícios 2–5 acima para o caso de somas diretas de uma quantidade arbitrária (até infinita) de subespaços. Vamos dar apenas algumas idéias centrais:
   (i) se V é um espaço vetorial e $(S_i)_{i \in I}$ é uma família de subespaços $S_i \subset V$ então denota-se por $\sum_{i \in I} S_i$ a *soma* dos subespaços S_i que é o subespaço gerado pela união $\bigcup_{i \in I} S_i$; na prática, $\sum_{i \in I} S_i$ consiste de todos os elementos da forma $\sum_{i \in I} v_i$, onde cada $v_i \in S_i$ e $v_i \neq 0$ no *máximo para um número finito* de i's (*não se faz somas infinitas de vetores*);
   (ii) a condição $S_1 \cap S_2 = \{0\}$ no item (a) do Exercício 2 deve ser trocada por:

$$S_i \cap \sum_{\substack{j \in I \\ j \neq i}} S_j = \{0\},$$

para todo $i \in I$ (*não basta* $S_i \cap S_j = \{0\}$ para todos $i, j \in I$);

(iii) o item (b) do Exercício 2 deve ser trocado por "cada $v \in V$ se escreve de modo único na forma $v = \sum_{i \in I} v_i$, $v_i \in V_i$", onde a soma $\sum_{i \in I} v_i$ deve ser entendida no sentido que $v_i \neq 0$ no máximo para um número finito de i's (mais uma vez: *não se faz* somas infinitas de vetores);

(iv) a notação padrão para uma soma direta de uma família arbitrária (interna ou externa) é $\bigoplus_{i \in I} S_i$;

(v) no Exercício 5, quando define-se somas diretas externas, *não se deve usar* o produto cartesiano $\prod_{i \in I} S_i$, mas sim o subconjunto de $\prod_{i \in I} S_i$ formado pelas famílias $(v_i)_{i \in I}$ tais que o conjunto $\big\{i \in I : v_i \neq 0\big\}$ é finito (uma tal família é chamada *quase nula*).

`Diferenciação`.

7. Mostre que a noção de aplicação diferenciável não depende da escolha particular de uma norma, i.e., mostre que se $f : U \subset I\!R^m \to I\!R^n$ é diferenciável no ponto $x \in U$ quando usamos as normas Euclideanas então f continua diferenciável em $x \in U$ (com a mesma diferencial) se usarmos outras normas em $I\!R^m$ e em $I\!R^n$ (*observação*: a noção de limite é topológica e portanto não se altera quando trocamos uma métrica por outra equivalente; tome um cuidado porém: a norma do espaço aparece explicitamente na definição de diferenciabilidade).

Aula número 10 (05/04)

(1) **Um critério para diferenciabilidade em termos de derivadas parciais**.

Teorema. *Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$. Se f admite derivadas parciais $\frac{\partial f}{\partial x^i}$, $i = 1, \ldots, m$, em todo ponto de U e se essas derivadas parciais são funções contínuas $\frac{\partial f}{\partial x^i} : U \to \mathbb{R}^n$ num ponto $x \in U$ então f é diferenciável no ponto x e sua diferencial é dada por:*

$$\mathrm{d}f(x) \cdot h = \sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(x) h_i,$$

para todo $h \in \mathbb{R}^m$ (resumindo: "se uma função admite derivadas parciais contínuas num aberto então ela é diferenciável").

Demonstração. Uma vez que fique provado que f é diferenciável em x, a fórmula acima para $\mathrm{d}f(x)$ seguirá trivialmente do fato que $\mathrm{d}f(x)$ é linear e do fato que $\frac{\partial f}{\partial x^i}(x)$ é obtida aplicando $\mathrm{d}f(x)$ no i-ésimo vetor da base canônica de $\mathbb{R}^m$. Na verdade, como o lado direito da fórmula acima para $\mathrm{d}f(x)$ define de fato uma aplicação linear em h, podemos usar tal fórmula como "candidato" a $\mathrm{d}f(x)$ quando provamos que f é diferenciável em x; escrevemos então:

$$f(x + h) = f(x) + \sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(x) h_i + r(h),$$

e devemos mostrar que $\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$. Temos:

$$f(x + h) - f(x) = \sum_{i=1}^{m} \Delta_i(x, h),$$

onde:

$$\Delta_i(x, h) = f(x_1 + h_1, \ldots, x_i + h_i, x_{i+1}, \ldots, x_m) - f(x_1 + h_1, \ldots, x_{i-1} + h_{i-1}, x_i, \ldots, x_m),$$

para $i = 1, \ldots, m$. A idéia é estimar o valor de $\Delta_i(x, h)$ usando o teorema do valor médio. Seria possível agora terminar a demonstração usando a *desigualdade* do valor médio; por outro lado, como "f é diferenciável $\Leftrightarrow$ cada f^j é diferenciável", podemos supor sem perda de generalidade que $n = 1$ e daí podemos usar diretamente a *igualdade* do valor médio. Como $U \subset \mathbb{R}^m$ é aberto, podemos escolher $\varepsilon > 0$ tal que:

$$\|h\|_\infty \le \varepsilon \Longleftrightarrow |h_1| \le \varepsilon, \ldots, |h_m| \le \varepsilon \Longrightarrow x + h \in U;$$

a hipótese que f admite a i-ésima derivada parcial em U, implica que a função:

$$[x_i - \varepsilon, x_i + \varepsilon] \ni t \longmapsto f(x_1 + h_1, \ldots, x_{i-1} + h_{i-1}, t, x_{i+1}, \ldots, x_m) \in \mathbb{R}$$

é derivável e portanto o teorema do valor médio nos dá:

$$\Delta_i(x,h) = h_i \frac{\partial f}{\partial x^i}(x_1 + h_1, \ldots, x_{i-1} + h_{i-1}, c_i, x_{i+1}, \ldots, x_m),$$

para algum $c_i = c_i(x,h)$ entre x_i e $x_i + h_i$, desde que $\|h\|_\infty \leq \varepsilon$. Se $\|h\|_\infty \leq \varepsilon$ obtemos então:

$$r(h) = \sum_{i=1}^{m} \rho_i(x,h) h_i,$$

onde:

$$\rho_i(x,h) = \frac{\partial f}{\partial x^i}(x_1 + h_1, \ldots, x_{i-1} + h_{i-1}, c_i, x_{i+1}, \ldots, x_m) - \frac{\partial f}{\partial x^i}(x_1, \ldots, x_m).$$

A continuidade de $\frac{\partial f}{\partial x^i}$ em x implica que $\lim_{h \to 0} \rho_i(x,h) = 0$; além do mais:

$$\frac{\big|r(h)\big|}{\|h\|_\infty} \leq \sum_{i=1}^{m} \big|\rho_i(x,h)\big| \frac{|h_i|}{\|h\|_\infty} \leq \sum_{i=1}^{m} \big|\rho_i(x,h)\big|,$$

o que mostra que $\lim_{h \to 0} \frac{r(h)}{\|h\|_\infty} = 0$. ■

(2) **Diferenciação de ordem superior**.

Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$. Se f é diferenciável em todos os pontos de U então podemos considerar a função:

$$\mathrm{d}f : U \subset \mathbb{R}^m \longrightarrow \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n);$$

essa é novamente uma função definida num aberto de um espaço vetorial real de dimensão finita e tomando valores num espaço vetorial real de dimensão finita. Se a mesma for diferenciável em todos os pontos de U podemos considerar a função:

$$\mathrm{d}(\mathrm{d}f) = \mathrm{d}^2 f : U \subset \mathbb{R}^m \longrightarrow \mathrm{Lin}\big(\mathbb{R}^m, \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)\big);$$

iterando o raciocínio acima, vemos que se f pode ser diferenciada $k-1$ vezes numa vizinhança de $x \in U$ e se essa $(k-1)$-ésima diferencial pode ser diferenciada no ponto x então fica bem definida a *diferencial de k-ésima ordem* de f no ponto x, que é uma aplicação linear da forma:

$$\mathrm{d}^k f(x) : \mathbb{R}^m \longrightarrow \underbrace{\mathrm{Lin}\big(\mathbb{R}^m, \mathrm{Lin}\big(\mathbb{R}^m, \ldots, \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n) \ldots\big)\big)}_{k-1 \text{ Lin's}}.$$

Quando f pode ser diferenciada k vezes em todos os pontos de U obtemos uma aplicação:

$$\mathrm{d}^k f : U \subset \mathbb{R}^m \longrightarrow \underbrace{\mathrm{Lin}\big(\mathbb{R}^m, \mathrm{Lin}\big(\mathbb{R}^m, \ldots, \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n) \ldots\big)\big)}_{k \text{ Lin's}}.$$

Definição. *Se f pode ser diferenciada $k-1$ vezes numa vizinhança de x e se sua $(k-1)$-ésima diferencial pode ser diferenciada no ponto x então dizemos que f é k* vezes diferenciável *no ponto x; a aplicação linear* $\mathrm{d}^k f(x)$ *é chamada a* diferencial de k-ésima ordem *de f no ponto x. Quando f é k vezes diferenciável em todos os pontos de U, dizemos que f é k* vezes diferenciável em U. *Se f é k vezes diferenciável em U e se a aplicação $\mathrm{d}^k f$ é contínua em U então dizemos que f é uma* função de classe C^k *em U. Quando f é k vezes diferenciável para todo $k \geq 1$ (e portanto de classe C^k para todo $k \geq 1$) dizemos que f é uma* função de classe C^∞ *em U.*

`Observação`. Uma função contínua é às vezes chamada uma *função de classe* C^0.

A coisa por enquanto parece feia mas não é tanto assim. Como podemos pensar numa aplicação linear:

$$T \in \mathrm{Lin}\big(V_1, \mathrm{Lin}\big(V_2, \ldots, \mathrm{Lin}(V_k, W)\ldots\big)\big)$$

de maneira prática? Aplicamos T num vetor $v_1 \in V_1$ e obtemos uma aplicação linear:

$$T(v_1) \in \mathrm{Lin}\big(V_2, \mathrm{Lin}\big(V_3, \ldots, \mathrm{Lin}(V_k, W)\ldots\big)\big);$$

aplicamos agora $T(v_1)$ sucessivamente num vetor $v_2 \in V_2$, num vetor $v_3 \in V_3$ e assim por diante. Obtemos um vetor:

$$T(v_1)(v_2)\ldots(v_k) \in W.$$

Em vez de aplicar T sucessivamente nos vetores v_1, v_2, ..., v_k, poderíamos pensar diretamente na regra que associa a k-upla $(v_1, \ldots, v_k)$ a um vetor de W. Definimos:

$$B : V_1 \times V_2 \times \cdots \times V_k \longrightarrow W$$

fazendo $B(v_1, \ldots, v_k) = T(v_1)\ldots(v_k)$. Não é difícil ver que a linearidade de T é equivalente à *multi-linearidade* de B. Na verdade, temos um isomorfismo:

$$\psi_k : \mathrm{Lin}\big(V_1, \mathrm{Lin}\big(V_2, \ldots \mathrm{Lin}(V_k, W)\ldots\big)\big) \ni T \longrightarrow B \in \text{Mult-lin}_k(V_1, V_2, \ldots, V_k; W)$$

definido como acima; denotamos por $\text{Mult-lin}_k(V_1, V_2, \ldots, V_k; W)$ o espaço das aplicações k-lineares $B : V_1 \times \cdots \times V_k \to W$. A demonstração formal de que a aplicação ψ_k é de fato bem definida e é um isomorfismo de espaços vetoriais pode ser feita por indução em k com todos os detalhes — acreditamos que pouco se ganha com isso. Quem já trabalhou com o caso $k = 2$ e adquiriu alguma intuição sobre o assunto deve se convencer facilmente de que a mesma idéia funciona para k arbitrário.

`Notação`. Quando $V_1 = V_2 = \cdots = V_k = V$ o espaço $\text{Mult-lin}_k(V_1, \ldots, V_k; W)$ será denotado simplesmente por $\text{Mult-lin}_k(V; W)$.

`Observação`. Usaremos o nome ψ_k para o isomorfismo descrito acima sejam lá quais forem os espaços V_1, ..., V_k, W envolvidos; em princípio, isso introduz alguma ambigüidade de notação, mas essa ambigüidade dificilmente gerará confusão.

Utilizaremos agora o isomorfismo ψ_k para identificar a diferencial de k-ésima ordem $\mathrm{d}^k f(x)$ com um objeto mais amigável. Definimos:

$$\mathrm{d}^{(k)} f(x) = \psi_k\big(\mathrm{d}^k f(x)\big);$$

isso significa que $\mathrm{d}^{(k)}f(x) \in \text{Mult-lin}_k(I\!R^m; I\!R^n)$ e que:

$$\mathrm{d}^{(k)}f(x)(v_1, \ldots, v_k) = \mathrm{d}^k f(x)(v_1) \ldots (v_k),$$

para todos $v_1, \ldots, v_k \in I\!R^m$. Na prática, é comum denotar os objetos $\mathrm{d}^k f(x)$ e $\mathrm{d}^{(k)}f(x)$ pelo mesmo símbolo (tipicamente $\mathrm{d}^k f(x)$) — o uso do isomorfismo ψ_k deve ser automatizado! No início, porém, é mais saudável utilizar uma notação mais explícita (i.e., não omitir ψ_k). Observamos que as notações $\mathrm{d}^k f(x)$ e $\mathrm{d}^{(k)}f(x)$ *não são usuais*; estamos utilizando-as apenas neste texto — muitos autores identificam os objetos $\mathrm{d}^k f(x)$ e $\mathrm{d}^{(k)}f(x)$ desde o princípio, causando uma certa confusão para o leitor iniciante. Nós usaremos o nome "*diferencial de k-ésima ordem*" tanto para $\mathrm{d}^k f(x)$ como para $\mathrm{d}^{(k)}f(x)$ — esperamos que apenas a distinção a nível de notação simbólica seja suficiente para evitar a confusão.

Definição. *A aplicação bilinear* $\mathrm{d}^{(2)}f(x) : I\!R^m \times I\!R^m \to I\!R^n$ *associada à diferencial segunda de* f *no ponto* x *é conhecida como o* Hessiano *de* f *no ponto* x *e é denotada por* $\mathrm{Hess}(f)_x$.

`Observação`. Uma aplicação f é de classe C^k se e somente se f for k vezes diferenciável e a aplicação $\mathrm{d}^{(k)}f$ for contínua. De fato, o isomorfismo ψ_k é um homeomorfismo e portanto $\mathrm{d}^k f$ é contínua se e somente se $\mathrm{d}^{(k)}f = \psi_k \circ \mathrm{d}^k f$ é contínua.

`Observação`. Segue diretamente da definição da diferencial de k-ésima ordem que:

$$\mathrm{d}^{k+1}f(x) = \mathrm{d}\big(\mathrm{d}^k f\big)(x).$$

Diferenciando a igualdade $\mathrm{d}^{(k)}f = \psi_k \circ \mathrm{d}^k f$ e utilizando a regra da cadeia obtemos:

$$\mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x) = \psi_k \circ \mathrm{d}^{k+1}f(x) : I\!R^m \longrightarrow \text{Mult-lin}_k(I\!R^m; I\!R^n);$$

isso significa que:

$$\mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x)(v_1)(v_2, \ldots, v_{k+1}) = \mathrm{d}^{(k+1)}f(x)(v_1, \ldots, v_{k+1}), \qquad (*)$$

para todos $v_1, \ldots, v_{k+1} \in I\!R^m$.

`Exemplo`. Uma função constante é de classe C^∞ pois todas as suas diferenciais são nulas. Uma aplicação linear $T : I\!R^m \to I\!R^n$ também é de classe C^∞ pois sua diferencial

$$\mathrm{d}T : I\!R^m \longrightarrow \mathrm{Lin}(I\!R^m, I\!R^n)$$

é constante; de fato, temos $\mathrm{d}T(x) = T$ para todo $x \in I\!R^m$. Uma aplicação bilinear $B : I\!R^m \times I\!R^n \to I\!R^p$ também é de classe C^∞. De fato, a diferencial de B:

$$\mathrm{d}B : I\!R^m \times I\!R^n \to \mathrm{Lin}(I\!R^m \times I\!R^n, I\!R^p)$$

é dada por:

$$\mathrm{d}B(x, y) \cdot (h, k) = B(x, k) + B(h, y);$$

não é difícil ver que $\mathrm{d}B$ é linear, o que conclui a demonstração do fato que B é de classe C^∞. Segue também da linearidade de $\mathrm{d}B$ que a diferencial segunda de B (i.e., o Hessiano de B) é dada por:

$$\mathrm{d}^{(2)}B(x, y) \cdot \big((h_1, k_1), (h_2, k_2)\big) = \mathrm{d}B(h_1, k_1) \cdot (h_2, k_2) = B(h_1, k_2) + B(h_2, k_1).$$

É verdade também que aplicações multi-lineares quaisquer são de classe C^∞ (veja o Exercício 1).

Teorema. *Sejam $f, g : U \subset I\!R^m \to I\!R^n$ funções definidas num aberto $U \subset I\!R^m$. Se f e g são k vezes diferenciáveis num ponto $x \in U$ então $f + g$ é k vezes diferenciável no ponto x e valem as identidades:*

$$\mathrm{d}^k(f+g)(x) = \mathrm{d}^k f(x) + \mathrm{d}^k g(x), \quad \mathrm{d}^{(k)}(f+g)(x) = \mathrm{d}^{(k)} f(x) + \mathrm{d}^{(k)} g(x);$$

em particular, se f e g são de classe C^k então $f + g$ também é de classe C^k.

Demonstração. Usamos indução em k. O caso $k = 1$ já é conhecido. Suponha o resultado válido para um certo $k \geq 1$. Se f e g são $k+1$ vezes diferenciáveis em x então f e g também são k vezes diferenciáveis em x; pela hipótese de indução, $f + g$ é k vezes diferenciável em x e vale a identidade $\mathrm{d}^k(f + g) = \mathrm{d}^k f + \mathrm{d}^k g$. Como $\mathrm{d}^k f$ e $\mathrm{d}^k g$ são diferenciáveis no ponto x, segue que $\mathrm{d}^k(f + g)$ é diferenciável no ponto x, donde $f + g$ é $k + 1$ vezes diferenciável no ponto x. Calculamos agora:

$$\begin{aligned}\mathrm{d}^{k+1}(f+g)(x) &= \mathrm{d}\big[\mathrm{d}^k(f+g)\big](x) = \mathrm{d}\big[\mathrm{d}^k f + \mathrm{d}^k g\big](x) = \mathrm{d}\big(\mathrm{d}^k f\big)(x) + \mathrm{d}\big(\mathrm{d}^k g\big)(x)\\ &= \mathrm{d}^{k+1} f(x) + \mathrm{d}^{k+1} g(x);\end{aligned}$$

isso completa a demonstração do passo de indução. Observe agora que:

$$\begin{aligned}\mathrm{d}^{(k)}(f+g)(x) &= \psi_k\big(\mathrm{d}^k(f+g)(x)\big) = \psi_k\big(\mathrm{d}^k f(x) + \mathrm{d}^k g(x)\big) = \psi_k\big(\mathrm{d}^k f(x)\big) + \psi_k\big(\mathrm{d}^k g(x)\big)\\ &= \mathrm{d}^{(k)} f(x) + \mathrm{d}^{(k)} g(x).\end{aligned}$$

Quando f e g são de classe C^k então $\mathrm{d}^k f$ e $\mathrm{d}^k g$ são contínuas, donde a soma $\mathrm{d}^k(f + g) = \mathrm{d}^k f + \mathrm{d}^k g$ é contínua e $f + g$ também é de classe C^k. ■

Lema. *Seja $f : U \subset I\!R^m \to I\!R^n$ uma função definida num aberto $U \subset I\!R^m$ e seja $T : I\!R^n \to I\!R^p$ uma aplicação linear. Se f é k vezes diferenciável num ponto $x \in U$ então $T \circ f$ é k vezes diferenciável no ponto x e vale a identidade:*

$$\mathrm{d}^{(k)}(T \circ f)(x) = T \circ \mathrm{d}^{(k)} f(x) : \underbrace{I\!R^m \times \cdots \times I\!R^m}_{k \text{ vezes}} \longrightarrow I\!R^p;$$

em particular, se f é de classe C^k então $T \circ f$ é de classe C^k.

Demonstração. Usamos indução em k. O caso $k = 1$ já é conhecido. Suponha o resultado válido para um certo $k \geq 1$ e seja f uma função $k + 1$ vezes diferenciável no ponto $x \in U$. Em particular, f é k vezes diferenciável no ponto x, donde pela hipótese de indução a função $T \circ f$ é k vezes diferenciável no ponto x e vale a identidade $\mathrm{d}^{(k)}(T \circ f)(x) = T \circ \mathrm{d}^{(k)} f(x)$. Considere a aplicação linear:

$$\tau : \text{Mult-lin}_k(I\!R^m; I\!R^n) \ni B \longmapsto T \circ B \in \text{Mult-lin}_k(I\!R^m; I\!R^p);$$

sabemos que:

$$\mathrm{d}^{(k)}(T \circ f) = \tau \circ \mathrm{d}^{(k)} f.$$

Como $\mathrm{d}^k(T\circ f)=\psi_k^{-1}\circ \mathrm{d}^{(k)}(T\circ f)$ e como $\mathrm{d}^{(k)}f=\psi_k\circ \mathrm{d}^k f$ é diferenciável em x, segue da identidade acima que $\mathrm{d}^k(T\circ f)$ é diferenciável no ponto x, donde $T\circ f$ é $k+1$ vezes diferenciável no ponto x. Diferenciando a identidade acima e utilizando a regra da cadeia obtemos:

$$\mathrm{d}\big(\mathrm{d}^{(k)}(T\circ f)\big)(x)=\tau\circ \mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x):I\!R^m\longrightarrow \text{Mult-lin}_k(I\!R^m;I\!R^p);$$

aplicando os dois lados da igualdade acima em $v_1\in I\!R^m$ e depois na k-upla $(v_2,\ldots,v_{k+1})$ obtemos:

$$\begin{aligned}\mathrm{d}\big(\mathrm{d}^{(k)}(T\circ f)\big)(x)(v_1)(v_2,\ldots,v_{k+1})&=\tau\Big[\mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x)(v_1)\Big](v_2,\ldots,v_{k+1})\\&=\Big[T\circ \mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x)(v_1)\Big](v_2,\ldots,v_{k+1})\\&=T\Big[\mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x)(v_1)(v_2,\ldots,v_{k+1})\Big].\end{aligned}$$

Usando a relação observada anteriormente (veja fórmula $(*)$) entre $\mathrm{d}\big(\mathrm{d}^{(k)}f\big)$ e $\mathrm{d}^{(k+1)}f$ obtemos finalmente:

$$\mathrm{d}^{(k+1)}(T\circ f)(x)(v_1,\ldots,v_{k+1})=T\big[\mathrm{d}^{(k+1)}f(x)(v_1,\ldots,v_{k+1})\big],$$

o que completa a demonstração do passo de indução. Finalmente, se f é de classe C^k então $\mathrm{d}^{(k)}f$ é contínua e portanto $\mathrm{d}^{(k)}(T\circ f)=\tau\circ \mathrm{d}^{(k)}f$ é contínua, donde $T\circ f$ é de classe C^k.
■

Corolário. *Uma função $f:U\subset I\!R^m\to I\!R^n$ é k vezes diferenciável num ponto $x\in U$ se e somente se cada coordenada $f^i:U\to I\!R$ é k vezes diferenciável no ponto x; além do mais, a i-ésima coordenada de $\mathrm{d}^{(k)}f(x)\in \text{Mult-lin}_k(I\!R^m;I\!R^n)$ coincide com $\mathrm{d}^{(k)}f^i(x)\in \text{Mult-lin}_k(I\!R^m;I\!R)$. Analogamente, f é de classe C^k se e somente se cada f^i é de classe C^k.*

Demonstração. Seja $\pi_i:I\!R^n\to I\!R$ a i-ésima projeção. Temos $f^i=\pi_i\circ f$; como π_i é linear, o lema nos diz que se f é k vezes diferenciável no ponto $x\in U$ então cada f^i é k vezes diferenciável em x e que a i-ésima coordenada de $\mathrm{d}^{(k)}f(x)$ coincide com $\mathrm{d}^{(k)}f^i(x)$. Reciprocamente, suponha que cada f^i é k vezes diferenciável no ponto x; se $e_i\in I\!R^n$ denota o i-ésimo vetor da base canônica então a aplicação:

$$I\!R\ni c\longmapsto c\,e_i\in I\!R^n$$

é linear donde pelo lema a aplicação $U\ni y\mapsto f^i(y)e_i$ é k vezes diferenciável no ponto x. Como $f=\sum_{i=1}^n f^ie_i$, concluímos que f é k vezes diferenciável no ponto x. A demonstração de que f é de classe C^k se e somente se cada f^i é de classe C^k é feita de maneira análoga.
■

Teorema. *("regra da cadeia" C^k) Sejam $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$, $g : V \subset \mathbb{R}^n \to \mathbb{R}^p$ funções definidas em abertos $U \subset \mathbb{R}^m$, $V \subset \mathbb{R}^n$, de modo que $f(U) \subset V$. Se f é k vezes diferenciável num ponto $x \in U$ e g é k vezes diferenciável no ponto $f(x) \in V$ então a aplicação composta $g \circ f$ é k vezes diferenciável no ponto x. Além do mais, se f e g são de classe C^k então também $g \circ f$ é de classe C^k.*

Demonstração. Seja $\mathcal{C} : \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^p) \times \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n) \to \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^p)$ a *aplicação composição* definida por $\mathcal{C}(S, T) = S \circ T$. Sabemos que $\mathcal{C}$ é bilinear; da regra da cadeia segue que:

$$\mathrm{d}(g \circ f) = \mathcal{C} \circ (\mathrm{d}g \circ f, \mathrm{d}f).$$

Mostraremos agora o resultado desejado por indução em k. O caso $k = 1$ é conhecido. Supondo o resultado válido para um certo $k \geq 1$, suponha que f é $k+1$ vezes diferenciável em x e que g é $k+1$ vezes diferenciável em $f(x)$. Daí $g \circ f$ é diferenciável numa vizinhança de x e vale a fórmula acima para $\mathrm{d}(g \circ f)$. Sabemos que $\mathrm{d}f$ é k vezes diferenciável no ponto x, que f é k vezes diferenciável no ponto x, que $\mathrm{d}g$ é k vezes diferenciável no ponto $f(x)$ e que $\mathcal{C}$ é k vezes diferenciável em qualquer ponto (na verdade $\mathcal{C}$ é de classe C^∞) — segue da hipótese de indução que $\mathrm{d}g \circ f$ é k vezes diferenciável no ponto x (e portanto que $(\mathrm{d}g \circ f, \mathrm{d}f)$ é k vezes diferenciável no ponto x) e que $\mathrm{d}(g \circ f) = \mathcal{C} \circ (\mathrm{d}g \circ f, \mathrm{d}f)$ é k vezes diferenciável no ponto x. Daí $g \circ f$ é $k+1$ vezes diferenciável no ponto x. Supondo agora que f e g são de classe C^{k+1}, concluímos de modo análogo usando a identidade $\mathrm{d}(g \circ f) = \mathcal{C} \circ (\mathrm{d}g \circ f, \mathrm{d}f)$ e a hipótese de indução que $\mathrm{d}(g \circ f)$ é de classe C^k; daí $g \circ f$ é de classe C^{k+1}, o que completa a demonstração. ■

Corolário. *Se $f, g : U \subset \mathbb{R}^m \to \mathbb{R}$ são k vezes diferenciáveis no ponto $x \in U$ então o produto $fg : U \to \mathbb{R}$ é k vezes diferenciável em x. Além do mais, se f e g são de classe C^k então também fg é de classe C^k.*

Demonstração. Observe que $fg = B \circ (f, g)$, onde $B : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ denota a multiplicação de números reais. ■

`Observação`. Juntando todos os resultados mostrados até agora, já podemos concluir que se as coordenadas f^i de uma função $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ são dadas por *fórmulas* (i.e., funções elementares) que não envolvem raízes de expressões que podem se anular então f é de classe C^∞.

`Exemplo`. Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função diferenciável em U. Podemos identificar o espaço vetorial $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ com o espaço das matrizes reais $n \times m$, o qual pode ser identificado com $\mathbb{R}^{mn}$ da maneira óbvia. Temos que a função $\mathrm{d}f : U \to \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ é contínua se e somente se é contínua coordenada por coordenada quando identificamos $\mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$ com $\mathbb{R}^{mn}$. Recordando que a matriz que representa a aplicação linear $\mathrm{d}f(x)$ na base canônica (i.e., a matriz Jacobiana de f no ponto x) possui como entradas as derivadas parciais $\frac{\partial f^i}{\partial x^j}(x)$, vemos que f é de classe C^1 se e somente se f é diferenciável e possui derivadas parciais contínuas — mas já sabemos que uma função que possui derivadas parciais contínuas é diferenciável. Mostramos então que:

$$f \text{ é de classe } C^1 \Longleftrightarrow f \text{ possui todas as derivadas parciais } \frac{\partial f^i}{\partial x^j} \text{ contínuas.}$$

Obviamente uma função f é de classe C^k se e somente se f for diferenciável e sua diferencial $\mathrm{d}f$ for de classe C^{k-1}; mas $\mathrm{d}f$ é de classe C^{k-1} se e somente se suas coordenadas $\frac{\partial f^i}{\partial x^j}$ são de classe C^{k-1}. Mostramos então (por indução em k) que f é de classe C^k se e somente se f admite *derivadas parciais de ordem* k:

$$\frac{\partial}{\partial x^{j_1}}\frac{\partial}{\partial x^{j_2}}\cdots\frac{\partial f^i}{\partial x^{j_k}}(x)=\frac{\partial^k f^i}{\partial x^{j_1}\partial x^{j_2}\ldots\partial x^{j_k}}(x),$$

para todos $j_1,\ldots,j_k=1,\ldots,m$, $i=1,\ldots,n$, todas contínuas em U.

Teorema. *Suponha que $f:U\subset \mathbb{R}^m\to\mathbb{R}^n$ é k vezes diferenciável num ponto $x\in U$. Então dados vetores $e_{j_1},\ldots,e_{j_k}\in\mathbb{R}^m$ da base canônica de $\mathbb{R}^m$ temos:*

$$\mathrm{d}^{(k)}f(x)(e_{j_1},\ldots,e_{j_k})=\frac{\partial^k f}{\partial x^{j_1}\partial x^{j_2}\ldots\partial x^{j_k}}(x).$$

Demonstração. Fazemos indução em k. O caso $k=1$ é a definição de derivada parcial. Suponha o resultado válido para um certo $k\geq 1$ e suponha que f é $k+1$ vezes diferenciável no ponto x. Fixados $j_1,\ldots,j_{k+1}\in\{1,\ldots,m\}$, denote por λ a aplicação linear de *avaliação na k-upla* $(e_{j_2},\ldots,e_{j_{k+1}})$, i.e.:

$$\lambda:\text{Mult-lin}_k(\mathbb{R}^m;\mathbb{R}^n)\ni B\longmapsto B(e_{j_2},\ldots,e_{j_{k+1}})\in\mathbb{R}^n.$$

Pela hipótese de indução temos:

$$\lambda\big(\mathrm{d}^{(k)}f(x)\big)=\frac{\partial^k f}{\partial x^{j_2}\partial x^{j_3}\ldots\partial x^{j_{k+1}}}(x);$$

diferenciando a igualdade $\lambda\circ\mathrm{d}^{(k)}f=\frac{\partial^k f}{\partial x^{j_2}\ldots\partial x^{j_{k+1}}}$ no ponto x e aplicando o resultado em e_{j_1} obtemos:

$$\mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x)(e_{j_1})(e_{j_2},\ldots,e_{j_{k+1}})=\frac{\partial^{k+1} f}{\partial x^{j_1}\partial x^{j_2}\ldots\partial x^{j_{k+1}}}(x).$$

Usando a fórmula $(*)$ que relaciona $\mathrm{d}\big(\mathrm{d}^{(k)}f\big)(x)$ e $\mathrm{d}^{(k+1)}f(x)$ obtemos o resultado desejado. ■

`Exemplo`. (*o caso $m=1$*) Se $f:U\subset\mathbb{R}\to\mathbb{R}^n$ é uma curva k vezes diferenciável no instante $t\in U$ então pelo teorema acima temos:

$$\mathrm{d}^{(k)}f(t)(\underbrace{1,1,\ldots,1}_{k\text{ vezes}})=\frac{\mathrm{d}^k f}{\mathrm{d}t^k}(t);$$

além do mais, da multi-linearidade de $\mathrm{d}^{(k)}f(t)$ obtemos:

$$\mathrm{d}^{(k)}f(t)(h_1,h_2,\ldots,h_k)=\frac{\mathrm{d}^k f}{\mathrm{d}t^k}(t)h_1h_2\ldots h_k,$$

para todos $h_1,h_2,\ldots,h_k\in\mathbb{R}$.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Diferenciação.`

1. Seja $B : \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \cdots \times \mathbb{R}^{m_p} \to \mathbb{R}^n$ uma aplicação p-linear.

   (a) Mostre que B é diferenciável e que sua diferencial é dada por:

   $$\mathrm{d}B(x_1, \ldots, x_p) \cdot (h_1, \ldots, h_p) = \sum_{i=1}^{p} B(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_p),$$

   para todos $x_i \in \mathbb{R}^{m_i}$, $h_i \in \mathbb{R}^{m_i}$, $i = 1, \ldots, p$.

   (b) para $i = 1, \ldots, p$ considere a aplicação:

   $$B_i : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_{i-1}} \times \mathbb{R}^{m_{i+1}} \times \cdots \times \mathbb{R}^{m_p} \longrightarrow \mathrm{Lin}(\mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_p}, \mathbb{R}^n)$$

   definida por $B_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p)(h) = B(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_p)$, onde $h = (h_1, \ldots, h_p)$. Mostre que B_i é uma aplicação $(p-1)$-linear.

   (c) Seja $\mathfrak{p}_i : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_p} \to \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_{i-1}} \times \mathbb{R}^{m_{i+1}} \times \cdots \times \mathbb{R}^{m_p}$ a aplicação linear que *esquece a i-ésima coordenada*, i.e.:

   $$\mathfrak{p}_i(x_1, \ldots, x_p) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_p).$$

   Mostre que:

   $$\mathrm{d}B = \sum_{i=1}^{p} B_i \circ \mathfrak{p}_i.$$

   (d) Utilize indução em p e o resultado do item (c) para concluir que toda aplicação p-linear é de classe C^∞.

Adendo à Aula número 9

Aí vai uma aplicação importante da desigualdade do valor médio que ficou faltando na aula número 9.

Teorema. *Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função diferenciável com $\mathrm{d}f(x) = 0$ para todo $x \in U$, onde U é um aberto conexo de $\mathbb{R}^m$. Então f é constante.*

Demonstração. Se U é convexo então para todos $x, y \in U$ podemos aplicar a desigualdade do valor médio para f no segmento $[x, y]$ obtendo:

$$\big\|f(y) - f(x)\big\| \le \sup_{z \in [x,y]} \big\|\mathrm{d}f(z)\big\| \|y - x\| = 0,$$

donde $f(x) = f(y)$ e f é constante em U. Vamos agora mostrar o caso geral, i.e., o caso que U é um aberto conexo qualquer de $\mathbb{R}^m$. Seja $c \in \mathbb{R}^n$ um elemento qualquer da imagem de f (o caso $U = \emptyset$ é meio ridículo) — considere o conjunto:

$$A = \big\{x \in U : f(x) = c\big\} = f^{-1}(c).$$

Temos $A \neq \emptyset$ pois $c \in f(U)$; além do mais, A é fechado *em* U pois f é contínua e $\{c\}$ é fechado em $\mathbb{R}^n$. Se mostrarmos que A é aberto (em U ou em $\mathbb{R}^m$, dá no mesmo), poderemos concluir da conexidade de U que $A = U$, o que completará a demonstração. Observe então que se $x \in A$ então temos uma bola $\mathrm{B}(x; r) \subset U$ contida em U e como a bola é convexa, segue da primeira parte da demonstração que f é constante em $\mathrm{B}(x; r)$ e portanto $\mathrm{B}(x; r) \subset A$. Isso completa a demonstração. ■

Aula número 11 (17/04)

A aula número 11 cobriu o “adendo à aula número 9” e boa parte do material originalmente destinado à aula número 10 sobre “diferenciação de ordem superior”, faltando apenas a relação entre as diferenciais de ordem superior e as derivadas parciais de ordem superior.

Aula número 12 (19/04)

A aula começa cobrindo o material sobre a relação entre as diferenciais de ordem superior e as derivadas parciais de ordem superior, originalmente destinado à aula número 10.

(1) **O teorema de Schwarz**.

Teorema. *(de Schwarz) Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função duas vezes diferenciável num ponto $x \in U$, onde $U \subset \mathbb{R}^m$ é um aberto. Então:*

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(x) = \frac{\partial^2 f}{\partial x^j \partial x^i}(x),$$

para todos $i, j = 1, \ldots, m$.

Demonstração. O teorema pode obviamente ser demonstrado coordenada por coordenada e portanto podemos supor sem perda de generalidade que $n = 1$. Na verdade, podemos supor também que $m = 2$; o caso geral segue do caso $m = 2$ considerando a função:

$$(x'_i, x'_j) \longmapsto f(x_1, \ldots, x'_i, \ldots, x'_j, \ldots, x_m)$$

definida numa vizinhança aberta do ponto (x_i, x_j) em $\mathbb{R}^2$.

Seja então $f : U \to \mathbb{R}$ uma função definida num aberto U em $\mathbb{R}^2$, duas vezes diferenciável num ponto $(x, y) \in \mathbb{R}^2$ — vamos denotar a diferenciação parcial para funções definidas em abertos de $\mathbb{R}^2$ usando os símbolos $\frac{\partial}{\partial x}$ e $\frac{\partial}{\partial y}$. Devemos mostrar então que:

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y).$$

Considere a função:

$$\varphi(t) = f(x + t, y + t) - f(x + t, y) - f(x, y + t) + f(x, y),$$

definida para t numa vizinhança aberta da origem em $\mathbb{R}$. A estratégia da demonstração é mostrar que tanto $\frac{\partial^2 f}{\partial x \partial y}(x, y)$ como $\frac{\partial^2 f}{\partial y \partial x}(x, y)$ são iguais ao limite $\lim_{t \to 0} \frac{\varphi(t)}{t^2}$.

Como U é aberto, podemos escolher $\varepsilon > 0$ tal que $|x' - x| \le \varepsilon$ e $|y' - y| \le \varepsilon$ implicam $(x', y') \in U$. Fixado t com $|t| \le \varepsilon$, podemos considerar a função:

$$[y - \varepsilon, y + \varepsilon] \ni y' \longmapsto \xi(y') = f(x + t, y') - f(x, y') \in \mathbb{R}$$

e observar que $\varphi(t) = \xi(y+t) - \xi(y)$. Aplicando o teorema do valor médio para ξ concluímos que existe $\theta \in (0, 1)$ tal que:

$$\varphi(t) = \xi(y + t) - \xi(y) = t\xi'(y + \theta t) = t\left[\frac{\partial f}{\partial y}(x + t, y + \theta t) - \frac{\partial f}{\partial y}(x, y + \theta t)\right];$$

como $\mathrm{d}f$ é diferenciável em x, temos que $\frac{\partial f}{\partial y}$ é diferenciável em x (veja o Exercício 1) e portanto podemos escrever:

$$\frac{\partial f}{\partial y}(x + h, y + k) = \frac{\partial f}{\partial y}(x, y) + \frac{\partial^2 f}{\partial x \partial y}(x, y)h + \frac{\partial^2 f}{\partial y^2}(x, y)k + \rho(h, k)\big\|(h, k)\big\|,$$

com $\lim_{(h,k)\to 0} \rho(h, k) = 0$. Usando a fórmula acima com $(h, k) = (t, \theta t)$ e com $(h, k) = (0, \theta t)$ obtemos:

$$\begin{aligned}
\varphi(t) &= t\left[\frac{\partial f}{\partial y}(x, y) + \frac{\partial^2 f}{\partial x \partial y}(x, y)t + \frac{\partial^2 f}{\partial y^2}(x, y)\theta t + \rho(t, \theta t)\big\|(t, \theta t)\big\|\right] \\
&\quad - t\left[\frac{\partial f}{\partial y}(x, y) + \frac{\partial^2 f}{\partial y^2}(x, y)\theta t + \rho(0, \theta t)\big\|(0, \theta t)\big\|\right] \\
&= t^2 \frac{\partial^2 f}{\partial x \partial y}(x, y) + t\rho(t, \theta t)\big\|(t, \theta t)\big\| - t\rho(0, \theta t)\big\|(0, \theta t)\big\|,
\end{aligned}$$

e portanto

$$\lim_{t\to 0} \frac{\varphi(t)}{t^2} = \frac{\partial^2 f}{\partial x \partial y}(x, y).$$

(observe que θ não é uma constante quando fazemos $t \to 0$, mas sim uma *função* de t; no entanto, o valor de θ fica no intervalo $(0, 1)$ para todo t). Observe agora que, fixado $t \in [-\varepsilon, \varepsilon]$, temos $\varphi(t) = \eta(x + t) - \eta(x)$, onde η é a função definida por:

$$[x - \varepsilon, x + \varepsilon] \ni x' \longmapsto \eta(x') = f(x', y + t) - f(x', y).$$

Repetimos agora o que foi feito acima usando η no lugar de ξ e usando a diferenciabilidade de $\frac{\partial f}{\partial x}$; concluiremos que:

$$\lim_{t\to 0} \frac{\varphi(t)}{t^2} = \frac{\partial^2 f}{\partial y \partial x}(x, y),$$

o que completa a demonstração. ■

Uma aplicação bilinear $B : V \times V \to W$ é dita *simétrica* quando $B(v_1, v_2) = B(v_2, v_1)$ para todos $v_1, v_2 \in V$. Mais geralmente, uma aplicação multi-linear $B \in \text{Mult-lin}_k(V; W)$

é dita *simétrica* quando para todos $v_1, \ldots, v_k \in V$, o valor de $B(v_1, \ldots, v_k)$ é invariante (i.e., não se altera) por permutações dos vetores v_i, i.e., quando:

$$B(v_1, \ldots, v_k) = B(v_{\sigma(1)}, \ldots, v_{\sigma(k)}),$$

para toda bijeção $\sigma : \{1, \ldots, k\} \to \{1, \ldots, k\}$.

`Observação`. Na verdade, para que $B \in \text{Mult-lin}_k(V; W)$ seja simétrica é suficiente que para todos $v_1, \ldots, v_k \in V$, o valor de $B(v_1, \ldots, v_k)$ seja invariante por trocas de vetores consecutivos v_i, v_{i+1}, i.e., que:

$$B(v_1, \ldots, v_i, v_{i+1}, \ldots, v_k) = B(v_1, \ldots, v_{i+1}, v_i, \ldots, v_k),$$

para todo $i = 1, \ldots, k-1$. Isso segue do fato extremamente intuitivo de que toda permutação de k elementos pode ser realizada através de trocas de elementos consecutivos; a formalização correta desse fato occore quando se estuda em álgebra a estrutura do *grupo de permutações* S_k (para quem já estudou um pouco de teoria dos grupos, veja os Exercícios 6, 7 e 8).

Corolário. *Se $f : U \subset I\!R^m \to I\!R^n$ é duas vezes diferenciável num ponto x do aberto U então o Hessiano de f no ponto x é uma aplicação bilinear simétrica.*

Demonstração. Segue do Teorema de Schwarz e do Exercício 2.

Corolário. *Se $f : U \subset I\!R^m \to I\!R^n$ é k vezes diferenciável num ponto x do aberto U então a diferencial de k-ésima ordem $\mathrm{d}^{(k)} f(x)$ é uma aplicação k-linear simétrica.*

Demonstração. Se f é k vezes diferenciável em x então as derivadas parciais de f de ordem l com $l \leq k-2$ são duas vezes diferenciáveis em x (veja o Exercício 1) e portanto aplicando o Teorema de Schwarz indutivamente concluímos que o valor de uma derivada parcial de ordem k de f no ponto x não depende da ordem em que as k derivadas parciais são feitas. A conclusão segue do Exercício 3. ■

(2) **Aplicações multi-lineares em coordenadas**.

O objetivo desta seção é o de fazer uma preparação para o estudo da fórmula de Taylor para funções de várias variáveis.

Seja $B \in \text{Mult-lin}_k(I\!R^m; I\!R^n)$ uma aplicação k-linear. Vamos recordar um pouco o que foi deixado como exercício na aula número 3 (13/03) — dissemos lá que uma aplicação multi-linear fica univocamente determinada por seus valores em vetores de uma base. Vamos explorar isso mais a fundo: seja $(e_i)_{i=1}^m$ a base canônica de $I\!R^m$. A aplicação B deve ser calculada numa k-upla de vetores de $I\!R^m$; vamos então calcular B em k vetores da base canônica, digamos $e_{i_1}, \ldots, e_{i_k}$, onde os índices $i_1, \ldots, i_k$ variam em $\{1, \ldots, m\}$. Escrevemos então:

$$B(e_{i_1}, e_{i_2}, \ldots, e_{i_k}) = b_{i_1 i_2 \ldots i_k} \in I\!R^n.$$

Esse processo produz m^k vetores de $I\!R^n$. O fato importante para se ter em mente é que os vetores $b_{i_1 i_2 \ldots i_k}$, $i_1, \ldots, i_k = 1, \ldots, m$, *descrevem a aplicação k-linear B completamente.* Vamos fazer algumas contas. Digamos que os vetores $b_{i_1 \ldots i_k} \in I\!R^n$ sejam dados e digamos que queremos calcular o valor de B numa k-upla arbitrária $(v_1, \ldots, v_k)$ de vetores de $\mathbb{R}^m$, sendo que sabemos que $B(e_{i_1}, \ldots, e_{i_k}) = b_{i_1 \ldots i_k}$. Procedemos da seguinte forma: escrevemos cada v_α, $\alpha = 1, \ldots, k$, como combinação linear dos vetores da base canônica, i.e.:

$$v_\alpha = \sum_{i_\alpha = 1}^{m} v_\alpha^{i_\alpha} e_{i_\alpha}, \quad \alpha = 1, \ldots, k;$$

denotamos por i_α o índice usado na α-ésima somatória — em princípio, na fórmula acima, não é importante usar nomes diferentes para os índices das k somatórias, porém essa distinção será necessária a seguir. Calculamos:

$$B(v_1, \ldots, v_k) = B\Big(\sum\nolimits_{i_1=1}^{m} v_1^{i_1} e_{i_1}, \sum\nolimits_{i_2=1}^{m} v_2^{i_2} e_{i_2}, \ldots, \sum\nolimits_{i_k=1}^{m} v_k^{i_k} e_{i_k} \Big),$$

e pela multi-linearidade de B obtemos:

$$\begin{aligned} B(v_1, \ldots, v_k) &= \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \cdots \sum_{i_k=1}^{m} v_1^{i_1} v_2^{i_2} \cdots v_k^{i_k} B(e_{i_1}, e_{i_2}, \ldots, e_{i_k}) \\ &= \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} \cdots \sum_{i_k=1}^{m} v_1^{i_1} v_2^{i_2} \cdots v_k^{i_k} b_{i_1 i_2 \ldots i_k}. \end{aligned}$$

`Observação`. Para quem acha a quantidade de somatórias e índices acima uma coisa incompreensível, tente escrever sozinho(a) os casos $k = 1$ (que é bem fácil) e $k = 2$, ao menos — note que, quando k é pequeno, não surge a necessidade de usar índices i_α, $\alpha = 1, \ldots, k$. Por exemplo, se $k = 2$ você pode usar i e j em vez de i_1 e i_2.

O que foi mostrado acima pode ser resumido na seguinte afirmação: *a aplicação*

$$\mathrm{Mult\text{-}lin}_k(\mathbb{R}^m; \mathbb{R}^n) \ni B \longmapsto \big(B(e_{i_1}, \ldots, e_{i_k}) \big)_{i_1, \ldots, i_k = 1}^{m} \in (\mathbb{R}^n)^{(m^k)}$$

é um isomorfismo.

`Observação`. A *expansão em coordenadas* de B feita acima pode ser levada ainda um nível adiante: podemos descrever os vetores $b_{i_1 \ldots i_k} \in \mathbb{R}^n$ em termos de suas coordenadas na base canônica de $\mathbb{R}^n$, obtendo números reais $b^j_{i_1 \ldots i_k} \in \mathbb{R}$, $i_1, \ldots, i_k = 1, \ldots, m$, $j = 1, \ldots, n$. Podemos então pensar da seguinte maneira: *uma aplicação k-linear $B \in \mathrm{Mult\text{-}lin}_k(\mathbb{R}^m; \mathbb{R}^n)$ pode ser descrita por uma* hiper-matriz *real $n \times \underbrace{m \times \cdots \times m}_{k \text{ vezes}}$ com $k+1$ índices, i.e., por uma coleção de nm^k números reais.* Em particular, quando $k = 1$ obtemos novamente o fato familiar que aplicações lineares $T : \mathbb{R}^m \to \mathbb{R}^n$ podem ser descritas por matrizes reais $n \times m$; quando $k = 2$ e $n = 1$ obtemos o fato (familiar?) que formas bilineares $B : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ podem ser descritas por matrizes reais $m \times m$.

`Exemplo`. Se $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ é k vezes diferenciável no ponto $x \in U$ então a aplicação k-linear $\mathrm{d}^{(k)} f(x) \in \mathrm{Mult\text{-}lin}_k(\mathbb{R}^m; \mathbb{R}^n)$ é descrita pelos vetores:

$$\mathrm{d}^{(k)} f(x)(e_{i_1}, \ldots, e_{i_k}) = \frac{\partial^k f}{\partial x^{i_1} \ldots \partial x^{i_k}}(x) \in \mathbb{R}^n.$$

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Diferenciação.`

1. Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$. Mostre que se f é k vezes diferenciável num ponto $x \in U$ ($k \geq 1$) então as derivadas parciais $\frac{\partial f}{\partial x^i}$ de f são $k-1$ vezes diferenciáveis em x. Mostre também que se f é de classe C^k ($k \geq 1$) então as derivadas parciais $\frac{\partial f}{\partial x^i}$ são de classe C^{k-1} (*dica*: $\frac{\partial f}{\partial x_i} = \lambda \circ \mathrm{d}f$, onde $\lambda : T \mapsto T(e_i)$ é a aplicação de avaliação no i-ésimo vetor da base canônica de $\mathbb{R}^m$).

`Álgebra linear.`

2. Seja $B : V \times V \to W$ uma aplicação bilinear e seja $(e_i)_{i=1}^m$ uma base de V. Mostre que se $B(e_i, e_j) = B(e_j, e_i)$ para todos $i, j = 1, \ldots, m$ então B é simétrica (*dica*: defina uma aplicação bilinear $\widetilde{B} : V \times V \to W$ fazendo $\widetilde{B}(x, y) = B(y, x)$ para todos $x, y \in V$. Recorde dos Exercícios da Aula número 3 (13/03) que duas aplicações bilineares que coincidem sobre vetores de uma base são iguais).

3. Generalize o Exercício 2 para o caso de aplicações multi-lineares, i.e., dada uma aplicação k-linear $B \in \mathrm{Mult\text{-}lin}_k(V; W)$ e uma base $(e_i)_{i=1}^m$ de V, se a quantidade $B(e_{i_1}, e_{i_2}, \ldots, e_{i_k})$ é invariante por permutações dos índices $i_1, \ldots, i_k$, para todos $i_1, \ldots, i_k = 1, \ldots, m$ então B é simétrica (*dica*: se $\sigma : \{1, \ldots, k\} \to \{1, \ldots, k\}$ é uma bijeção, defina uma aplicação k-linear $B_\sigma \in \mathrm{Mult\text{-}lin}_k(V; W)$ fazendo:

$$B_\sigma(v_1, \ldots, v_k) = B(v_{\sigma(1)}, \ldots, v_{\sigma(k)}), \quad \forall\, v_1, \ldots, v_k \in V;$$

observe que B_σ e B coincidem sobre vetores da base e conclua que $B_\sigma = B$).

`Mais espaços métricos.`

4. Sejam M, N espaços métricos. Uma aplicação $f : M \to N$ é dita *fechada* quando f leva fechados em fechados, i.e., quando dado um fechado G em M então $f(G)$ é fechado em N. Mostre que:
   (a) se M é compacto então toda aplicação contínua $f : M \to N$ é fechada;
   (b) se $f : M \to N$ é uma bijeção contínua então f é fechada se e somente se f é um homeomorfismo;
   (c) se $f : M \to N$ é uma bijeção contínua e M é compacto então f é um homeomorfismo.

5. (*o teorema de alfândega*) Sejam M um espaço métrico, $C \subset M$ um subconjunto conexo e $S \subset M$ um subconjunto qualquer. Mostre que se C possui pontos de S e pontos do complementar de S então C possui pontos da fronteira de S, i.e.:

$$C \cap S \neq \emptyset,\ C \not\subset S \Longrightarrow C \cap \partial S \neq \emptyset.$$

(*dica*: se $C \cap \partial S = \emptyset$ então $C = \big(C \cap \mathrm{int}(S)\big) \cup \big(C \cap \mathrm{ext}(S)\big)$ é uma cisão de C).

`Álgebra (só para quem conhece um pouco de teoria dos grupos!).`

6. Seja S_k o grupo formado por todas as bijeções $\sigma : \{1, \ldots, k\} \to \{1, \ldots, k\}$ munido da operação de composição; o grupo S_k é conhecido como o *grupo das permutações de k elementos*. Dados $i, j \in \{1, \ldots, k\}$ com $i \neq j$ então a permutação $\sigma = (i\ j) \in S_k$ tal que $\sigma(i) = j$, $\sigma(j) = i$ e $\sigma(x) = x$ para $x \notin \{i, j\}$, é chamada a *transposição* dos elementos i e j. Mostre que as transposições $(i\ i+1)$ com $i = 1, \ldots, k-1$ geram o grupo S_k (*dica*: use indução em k).
7. Seja G um grupo e suponha que seja dada uma ação de G num conjunto X. Dada uma função $f : X \to Y$, um ponto $x \in X$ e um conjunto de geradores $S \subset G$, suponha que $f(g \cdot x) = f(x)$ para todo $g \in S$ e todo $x \in X$. Mostre que $f(g \cdot x) = f(x)$ para todo $g \in G$ e todo $x \in X$ (*dica*: o conjunto $H = \big\{g \in G : f(g \cdot x) = f(x),\ \forall x \in X\big\}$ é um subgrupo de G).
8. Seja $B \in \text{Mult-lin}_k(V; W)$ uma aplicação multi-linear. Seja X o conjunto das funções $\iota : \{1, \ldots, k\} \to V$, i.e., o conjunto V^k das k-uplas $(v_1, \ldots, v_k)$ de vetores de V. Para $\sigma \in S_k$, defina $\sigma \cdot \iota = \iota \circ \sigma^{-1}$, i.e., $\sigma \cdot (v_1, \ldots, v_k) = (v_{\sigma^{-1}(1)}, \ldots, v_{\sigma^{-1}(k)})$. Mostre que $(\sigma, \iota) \mapsto \sigma \cdot \iota$ define uma ação do grupo S_k no conjunto V^k. Conclua dos Exercícios 6 e 7 que B é simétrica se e somente se para todos $v_1, \ldots, v_k \in V$ o valor de $B(v_1, \ldots, v_k)$ é invariante por trocas de vetores consecutivos v_i, v_{i+1}, $i = 1, \ldots, k-1$ (*observação*: multi-linearidade não tem nada a ver com este exercício).

Aula número 13 (24/04)

(1) **Polinômios de várias variáveis**.

Um *monômio* em $\mathbb{R}^m$ é uma função da forma:

$$\mathbb{R}^m \ni x = (x_1, \ldots, x_m) \longmapsto x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_m^{\lambda_m} \in \mathbb{R},$$

onde $\lambda_1, \lambda_2, \ldots, \lambda_m$ são inteiros não negativos — por simplicidade de terminologia, é conveniente convencionar que $0^0 = 1$. O inteiro $\lambda_1 + \cdots + \lambda_m$ é chamado o *grau* do monômio acima. Uma *função polinomial* em $\mathbb{R}^m$ tomando valores em $\mathbb{R}^n$ é uma "combinação linear" de monômios em $\mathbb{R}^m$ com coeficientes em $\mathbb{R}^n$; de outro modo, uma função polinomial em $\mathbb{R}^m$ tomando valores em $\mathbb{R}^n$ é uma função da forma:

$$\mathbb{R}^m \ni x \longmapsto p(x) = \sum_{\lambda_1, \ldots, \lambda_m = 0}^{+\infty} a_{\lambda_1 \ldots \lambda_m} x_1^{\lambda_1} \cdots x_m^{\lambda_m} \in \mathbb{R}^n,$$

onde os coeficientes $a_{\lambda_1 \ldots \lambda_m}$ pertencem a $\mathbb{R}^n$ e *no máximo um número finito deles são não nulos*, i.e., a soma acima é *finita*. O *grau* de uma função polinomial p é o maior grau de um monômio aparecendo em p com coeficiente não nulo, i.e., o maior dos números $\lambda_1 + \cdots + \lambda_m$ tal que $a_{\lambda_1 \ldots \lambda_m} \neq 0$.

`Observação`. Uma função polinomial $p : \mathbb{R}^m \to \mathbb{R}^n$ de grau 1 é o mesmo que uma função afim não constante, i.e., p é da forma $p(x) = T(x) + c$ onde $T : \mathbb{R}^m \to \mathbb{R}^n$ é uma aplicação linear não nula e $c \in \mathbb{R}^n$ é um vetor. Uma função polinomial de grau zero é o mesmo que uma função constante não nula. Quanto à função nula, pode-se convencionar que a mesma *não possui grau definido*, ou *possui grau* -1 ou que *possui grau* $-\infty$, tanto faz.

Para trabalhar com polinômios de várias variáveis de maneira satisfatória é conveniente utilizar a *notação de multi-índice*. Isso significa o seguinte: denotamos uma m-upla $(\lambda_1, \ldots, \lambda_m)$ de inteiros não negativos pelo símbolo λ, o coeficiente $a_{\lambda_1 \ldots \lambda_m}$ é denotado por a_λ, o monômio $x_1^{\lambda_1} \cdots x_m^{\lambda_m}$ é denotado por x^λ e o grau $\lambda_1 + \cdots + \lambda_m$ desse monômio é denotado por $|\lambda|$. Uma função polinomial $p : \mathbb{R}^m \to \mathbb{R}^n$ de grau menor ou igual a k será denotada então por:

$$p(x) = \sum_{|\lambda| \leq k} a_\lambda x^\lambda, \quad x \in \mathbb{R}^m,$$

onde cada $a_\lambda \in \mathbb{R}^n$. Uma m-upla $\lambda = (\lambda_1, \ldots, \lambda_m)$ de inteiros não negativos é normalmente chamada um *multi-índice de dimensão* m e o número $|\lambda|$ é chamado o *grau* do multi-índice λ.

Definição. *Uma* função polinomial homogênea de grau k *é uma função polinomial que pode ser escrita usando apenas monômios de grau* k*, i.e., uma função polinomial* p *da forma:*

$$\mathbb{R}^m \ni x \longmapsto p(x) = \sum_{|\lambda| = k} a_\lambda x^\lambda \in \mathbb{R}^n,$$

com cada $a_\lambda \in \mathbb{R}^n$*.*

`Observação`. O leitor com senso crítico pode se sentir incomodado pelas definições acima, pelo seguinte motivo: em princípio é possível que uma função polinomial admita *mais de uma representação como combinação linear de monômios*, i.e., seria possível que:

$$\sum_{|\lambda|\le k} a_\lambda x^\lambda = \sum_{|\lambda|\le k} a'_\lambda x^\lambda,$$

para todo $x \in \mathbb{R}^m$, porém com $a_\lambda \ne a'_\lambda$ para algum λ. Isso tornaria muitas das definições acima ambíguas, pois o grau de uma função polinomial poderia em princípio depender da representação utilizada — além do mais, uma função polinomial homogênea poderia admitir também representações não homogêneas. Na verdade, *não há problema nenhum*! A representação de uma função polinomial como combinação de monômios *é de fato única*. Para os que tem uma boa base de álgebra, esse fato deve ser conhecido na seguinte forma: "*Se D é um domínio de integridade infinito (por exemplo, $D = \mathbb{R}$) então o anel de polinômios $D[X_1, \ldots, X_m]$ é canonicamente isomorfo ao anel de funções polinomiais de m variáveis em D.*"

Para os que não tem familiaridade com o fato algébrico acima, podemos usar um outro argumento: se $p : \mathbb{R}^m \to \mathbb{R}^n$ é dada por:

$$p(x) = \sum_{|\lambda|\le k} a_\lambda x^\lambda \in \mathbb{R}^n, \quad x \in \mathbb{R}^m,$$

então:

$$a_\lambda = \frac{1}{\lambda!}\frac{\partial^{|\lambda|}p}{\partial x^\lambda}(0),$$

onde $\lambda! = \lambda_1!\lambda_2!\cdots\lambda_m!$ e $\frac{\partial^{|\lambda|}p}{\partial x^\lambda} = \frac{\partial^{|\lambda|}p}{\partial x_1^{\lambda_1}\partial x_2^{\lambda_2}\cdots\partial x_m^{\lambda_m}}$. Isso mostra que os coeficientes a_λ são univocamente determinados pela função polinomial p.

`Observação`. Uma maneira agradável de lembrar os fatos acima usando a linguagem da álgebra linear é o seguinte: considere o espaço vetorial $\mathcal{F}(\mathbb{R}^m, \mathbb{R}^n)$ de *todas* as funções $f : \mathbb{R}^m \to \mathbb{R}^n$ — as operações de espaço vetorial em $\mathcal{F}(\mathbb{R}^m, \mathbb{R}^n)$ são definidas da maneira óbvia. As funções polinomiais homogêneas de grau k formam um subespaço de $\mathcal{F}(\mathbb{R}^m; \mathbb{R}^n)$ que será denotado por $\mathcal{P}_k(\mathbb{R}^m, \mathbb{R}^n)$. A soma dos subespaços $\mathcal{P}_k(\mathbb{R}^m, \mathbb{R}^n)$ de $\mathcal{F}(\mathbb{R}^m, \mathbb{R}^n)$ é *direta*, i.e., se para $i = 0, \ldots, k$, $p_i : \mathbb{R}^m \to \mathbb{R}^n$ é uma função polinomial homogênea de grau i então $p_0 + p_1 + \cdots + p_k = 0$ implica $p_i = 0$ para todo $i = 0, \ldots, k$; denotando por $\mathcal{P}(\mathbb{R}^m, \mathbb{R}^n)$ o subespaço de $\mathcal{F}(\mathbb{R}^m, \mathbb{R}^n)$ formado por todas as funções polinomiais obtemos então:

$$\mathcal{P}(\mathbb{R}^m, \mathbb{R}^n) = \bigoplus_{k=0}^{+\infty} \mathcal{P}_k(\mathbb{R}^m, \mathbb{R}^n).$$

A soma direta:

$$\mathcal{P}_{\le k}(\mathbb{R}^m, \mathbb{R}^n) = \bigoplus_{i=0}^{k} \mathcal{P}_i(\mathbb{R}^m, \mathbb{R}^n)$$

é igual ao subespaço $\mathcal{P}_{\leq k}(\mathbb{R}^m, \mathbb{R}^n)$ de $\mathcal{P}(\mathbb{R}^m, \mathbb{R}^n)$ formado pelas funções polinomiais de grau menor ou igual a k. Enquanto $\mathcal{P}(\mathbb{R}^m, \mathbb{R}^n)$ possui dimensão infinita (para $m, n \neq 0$), os espaços $\mathcal{P}_k(\mathbb{R}^m, \mathbb{R}^n)$ e $\mathcal{P}_{\leq k}(\mathbb{R}^m, \mathbb{R}^n)$ possuem dimensão finita (qual seria uma base para tais espaços? veja o Exercício 1!).

(2) **Relação entre aplicações multi-lineares e polinômios**.

Se $B \in \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n)$ é uma aplicação multi-linear então escrevemos:

$$B(x)^{(k)} = B(\underbrace{x, x, \ldots, x}_{k \text{ vezes}}), \quad x \in \mathbb{R}^m.$$

Considere a aplicação $\phi : \mathbb{R}^m \to \mathbb{R}^n$ definida por:

$$\phi(x) = B(x)^{(k)}, \quad x \in \mathbb{R}^m.$$

Escrevendo $x = \sum_{i=1}^m x_i e_i$ obtemos:

$$\phi(x) = \sum_{i_1, \ldots, i_k = 1}^m x_{i_1} x_{i_2} \cdots x_{i_k} b_{i_1 i_2 \ldots i_k},$$

onde $b_{i_1 \ldots i_k} = B(e_{i_1}, \ldots, e_{i_k}) \in \mathbb{R}^n$. Vocês agora podem se convencer que a expressão acima define uma *função polinomial homogênea de grau k*. De fato, dentre os índices $i_1, \ldots, i_k \in \{1, \ldots, m\}$, temos uma certa quantidade deles, digamos $\lambda_1 \geq 0$, que são iguais a 1; alguns outros, digamos $\lambda_2 \geq 0$, são iguais a 2. Em geral, podemos denotar por λ_j o número de índices $i_1, \ldots, i_k$ que é igual a j (com $j = 1, \ldots, m$). Obtemos então um multi-índice $\lambda = (\lambda_1, \ldots, \lambda_m)$ com $|\lambda| = k$; vamos denotá-lo por:

$$\lambda = \# i_1 \ldots i_k.$$

Observe que:

$$x_{i_1} x_{i_2} \cdots x_{i_k} = x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_m^{\lambda_m} \underset{\substack{\text{em notação de} \\ \text{multi-índice}}}{=} x^\lambda.$$

Concluímos então que:

$$\phi(x) = \sum_{|\lambda| = k} a_\lambda x^\lambda,$$

onde a_λ é dado por:

$$a_\lambda = \sum_{\substack{i_1, \ldots, i_k = 1 \\ \# i_1 \ldots i_k = \lambda}}^m b_{i_1 \ldots i_k} \in \mathbb{R}^n.$$

Um exercício simples de combinatória mostra que a somatória acima tem exatamente $\frac{k!}{\lambda!}$ termos (verifique!).

Vamos mostrar agora que toda função polinomial homogênea de grau k é da forma $x \mapsto B(x)^{(k)}$ para alguma aplicação k-linear $B \in \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n)$; na verdade, tal aplicação B pode ser escolhida de muitas formas. Por exemplo, dada uma função polinomial homogênea $p(x) = \sum_{|\lambda|=k} a_\lambda x^\lambda$ de grau k então podemos definir uma aplicação multi-linear $B \in \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n)$ fazendo:

$$B(e_{i_1}, \ldots, e_{i_k}) = b_{i_1 \ldots i_k} = \frac{a_\lambda}{k!} \lambda!,$$

para todos $i_1, \ldots, i_k = 1, \ldots, m$ com $\# i_1 \ldots i_k = \lambda$. Obtemos então uma aplicação k-linear *simétrica* $B \in \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n)$ com $p(x) = B(x)^{(k)}$ para todo $x \in \mathbb{R}^m$.

`Exemplo`. Suponha que $B : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ é uma aplicação bilinear e escreva $B(e_i, e_j) = b_{ij}$, para $i, j = 1, \ldots, m$. Temos:

$$B(x, x) = \sum_{i,j=1}^{m} b_{ij} x_i x_j,$$

para todo $x \in \mathbb{R}^m$. Podemos escrever também:

$$B(x, x) = \sum_{i=1}^{m} b_{ii} x_i^2 + \sum_{\substack{i,j=1 \\ i<j}}^{m} (b_{ij} + b_{ji}) x_i x_j.$$

Observe que se $\widetilde{B} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ é uma outra aplicação bilinear com:

$$\widetilde{B}(e_i, e_j) + \widetilde{B}(e_j, e_i) = B(e_i, e_j) + B(e_j, e_i),$$

para todos $i, j = 1, \ldots, m$ então B e $\widetilde{B}$ *definem a mesma função polinomial*, i.e., $B(x, x) = \widetilde{B}(x, x)$ para todo $x \in \mathbb{R}^m$. Em geral então é possível que *uma infinidade de aplicações multi-lineares definam a mesma função polinomial homogênea*! Veremos adiante, no entanto, que têm-se a unicidade da aplicação multi-linear associada a uma função polinomial homogênea quando restringe-se às *aplicações multi-lineares simétricas*!

`Observação`. Uma aplicação polinomial homogênea de grau 2 é também chamada uma *forma quadrática*. Recorde que se ϕ é uma forma quadrática em $\mathbb{R}^2$ então o conjunto solução de $\phi(x) = 1$ é chamado uma *cônica* e que se ϕ é uma forma quadrática em $\mathbb{R}^3$ então o conjunto solução de $\phi(x) = 1$ é chamado uma *quádrica*.

Teorema. *Seja* $B \in \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n)$ *uma aplicação k-linear* simétrica. *A aplicação* $\phi(x) = B(x)^{(k)}$ *é de classe* C^∞ *e sua r-ésima diferencial é dada por:*

$$\mathrm{d}^{(r)}\phi(x) \cdot (h_1, \ldots, h_r) = k(k-1) \cdots (k-r+1) B(\underbrace{x, \ldots, x}_{k-r \text{ vezes}}, h_1, \ldots, h_r),$$

para $r \leq k$ *e* $\mathrm{d}^{(r)}\phi(x) = 0$ *para* $r > k$. *Em particular:*

$$\mathrm{d}^{(k)}\phi(x) = k! B,$$

para todo $x \in \mathbb{R}^m$.

Demonstração. Usamos indução em r. O caso $r = 1$ segue observando que $\phi = B \circ \Delta$, onde $\Delta : \mathbb{R}^m \to (\mathbb{R}^m)^k$ é a *aplicação diagonal* dada por $\Delta(x) = (\underbrace{x, \ldots, x}_{k \text{ vezes}})$. Como Δ é linear e B é multi-linear, a regra da cadeia nos dá:

$$\mathrm{d}\phi(x) \cdot h = B(h, x, \ldots, x) + B(x, h, x, \ldots, x) + \cdots + B(x, x, \ldots, x, h) = kB(x, \ldots, x, h),$$

pois B é simétrica. Suponha agora o resultado verdadeiro para um certo $r \geq 1$. Considere a aplicação $(k - r)$-linear simétrica:

$$\widetilde{B} : \underbrace{\mathbb{R}^m \times \cdots \times \mathbb{R}^m}_{k-r \text{ vezes}} \longrightarrow \text{Mult-lin}_r(\mathbb{R}^m; \mathbb{R}^n)$$

definida por:

$$\widetilde{B}(x_1, \ldots, x_{k-r}) \cdot (h_1, \ldots, h_r) = B(x_1, \ldots, x_{k-r}, h_1, \ldots, h_r).$$

A hipótese de indução nos diz que:

$$\mathrm{d}^{(r)}\phi(x) = k(k-1)\cdots(k-r+1)\widetilde{B}(x)^{(k-r)};$$

diferenciando a igualdade acima dos dois lados obtemos:

$$\begin{aligned}
\mathrm{d}^{(r+1)}\phi(x) \cdot (h_1, \ldots, h_{r+1}) &= \mathrm{d}\big(\mathrm{d}^{(r)}\phi\big)(x) \cdot (h_1)(h_2, \ldots, h_{r+1}) \\
&= k(k-1)\cdots(k-r+1)(k-r)\widetilde{B}(x, \ldots, x, h_1) \cdot (h_2, \ldots, h_{r+1}) \\
&= k(k-1)\cdots(k-r+1)(k-r)B(x, \ldots, x, h_1, \ldots, h_{r+1}),
\end{aligned}$$

o que completa a demonstração. ■

Corolário. *Se* $\text{Mult-lin}_k^{\text{sim}}(\mathbb{R}^m; \mathbb{R}^n)$ *denota o espaço das aplicações* k*-lineares simétricas* $B : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}^n$ *então temos um isomorfismo:*

$$\text{Mult-lin}_k^{\text{sim}}(\mathbb{R}^m; \mathbb{R}^n) \ni B \longmapsto \phi_B \in \mathcal{P}_k(\mathbb{R}^m; \mathbb{R}^n)$$

onde $\phi_B(x) = B(x)^{(k)}$ *para todo* $x \in \mathbb{R}^m$.

Demonstração. A sobrejetividade já foi estabelecida quando observamos que toda função polinomial homogênea de grau k é da forma $x \mapsto B(x)^{(k)}$ para alguma aplicação k-linear simétrica $B \in \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n)$. A linearidade de $B \mapsto \phi_B$ é muito simples. Resta mostrar a injetividade. Suponha então que $B \in \text{Mult-lin}_k^{\text{sim}}(\mathbb{R}^m; \mathbb{R}^n)$ é tal que $\phi_B = 0$. Obtemos então:

$$\mathrm{d}^{(k)}\phi_B(x) = k!B = 0,$$

e portanto $B = 0$. ■

(3) **O polinômio de Taylor**.

Segue do que foi visto na seção anterior que toda aplicação polinomial $p : \mathbb{R}^m \to \mathbb{R}^n$ de grau menor ou igual a k pode ser escrita na forma:

$$p(x) = \sum_{i=0}^{k} B_i(x)^{(i)}, \quad x \in \mathbb{R}^m,$$

onde cada $B_i \in \text{Mult-lin}_i(\mathbb{R}^m; \mathbb{R}^n)$ é uma aplicação multi-linear simétrica (com exceção de B_0, que denota simplesmente um vetor fixado em $\mathbb{R}^n$). Além do mais, as aplicações B_i são *univocamente determinadas* por p.

Teorema. *Seja $f : U \subset I\!R^m \to I\!R^n$ uma função definida num aberto $U \subset I\!R^m$, k vezes diferenciável num ponto $x \in U$. Então:*

$$p(h) = \sum_{i=0}^{k} \frac{1}{i!}\,\mathrm{d}^{(i)} f(x) \cdot (h)^{(i)}, \quad h \in I\!R^m,$$

é a única função polinomial de grau menor ou igual a k tal que $p(0) = f(x)$ e $\mathrm{d}^{(i)}p(0) = \mathrm{d}^{(i)}f(x)$ para $i = 1, \ldots, k$.

Demonstração. Se $p(h) = \sum_{i=0}^{k} B_i(h)^{(i)}$ é uma aplicação polinomial de grau menor ou igual a k, onde cada B_i é uma aplicação i-linear simétrica $B_i \in \text{Mult-lin}_i(I\!R^m; I\!R^n)$ então é fácil ver que $p(0) = B_0$ e que $\mathrm{d}^{(i)}p(0) = i!B_i$ para $i = 1, \ldots, k$. A conclusão segue. ∎

Seja $f : U \subset I\!R^m \to I\!R^n$ uma função k vezes diferenciável num ponto $x \in U$, onde $U \subset I\!R^m$ é um aberto. A função polinomial $p(h) = \sum_{i=0}^{k} \frac{1}{i!}\,\mathrm{d}^{(i)}f(x) \cdot (h)^{(i)}$ é conhecida como o *k-ésimo polinômio de Taylor* de f em torno do ponto x. Podemos escrever:

$$f(x+h) = f(x) + \mathrm{d}f(x)\cdot h + \frac{1}{2!}\,\mathrm{d}^{(2)}f(x)\cdot(h)^{(2)} + \cdots + \frac{1}{k!}\,\mathrm{d}^{(k)}f(x)\cdot(h)^{(k)} + r(h),$$

onde r é uma função k vezes diferenciável no ponto 0 tal que $r(0) = 0$ e $\mathrm{d}^{(i)}r(0) = 0$ para $i = 1, \ldots, k$. Dizemos que r é o *resto do k-ésimo polinômio de Taylor de f em torno de x.*

`Observação`. Se r é o resto do k-ésimo polinômio de Taylor de f em torno do ponto x e se f é l vezes diferenciável em x para um certo $l \geq k$ então r também é l vezes diferenciável no ponto x e $\mathrm{d}^{(i)}r(x) = \mathrm{d}^{(i)}f(x)$ para todo $i = k+1, \ldots, l$.

O fato que r se anula no ponto 0 juntamente com suas k primeiras diferenciais implica em certo sentido que o resto r é "bem pequeno" em torno de 0. Essa idéia intuitiva de "bem pequeno" pode ser expressa precisamente em termos de três resultados que apresentamos agora.

Teorema. *(Taylor infinitesimal) Se f é k vezes diferenciável no ponto x ($k \geq 1$) e se r é o resto do seu k-ésimo polinômio de Taylor em torno de x então:*

$$\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0.$$

Demonstração. Vamos mostrar por indução em k que se r é uma função que se anula juntamente com suas k primeiras diferenciais no ponto 0 então $\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0$. O caso $k = 1$ segue trivialmente da definição de função diferenciável (quando se escreve a definição de função diferenciável para a função r no ponto 0 vê-se que r é o próprio resto). Supondo o resultado válido para um certo k, suponha que r se anula juntamente com suas $k+1$ primeiras diferenciais no ponto 0. Pela desigualdade do valor médio temos:

$$\|r(h)\| \leq \sup_{\theta\in(0,1)} \big\|\mathrm{d}r(\theta h)\big\|\,\|h\|,$$

para h suficientemente pequeno. Como $\mathrm{d}r$ se anula juntamente com suas k primeiras diferenciais em 0, segue da hipótese de indução que:

$$\lim_{h \to 0} \frac{\mathrm{d}r(h)}{\|h\|^k} = 0;$$

assim, dado $\varepsilon > 0$, podemos encontrar $\delta > 0$ tal que $\|h\| < \delta$, implica $\big\|\mathrm{d}r(h)\big\| \le \varepsilon \|h\|^k$. Daí $\|h\| < \delta$ implica:

$$\big\|r(h)\big\| \le \sup_{\theta \in (0,1)} \big\|\mathrm{d}r(\theta h)\big\| \, \|h\| \le \sup_{\theta \in (0,1)} \varepsilon \|\theta h\|^k \, \|h\| \le \varepsilon \|h\|^{k+1}.$$

Concluímos então que:

$$\lim_{h \to 0} \frac{r(h)}{\|h\|^{k+1}} = 0,$$

o que completa a demonstração. ■

Antes de provar nosso próximo teorema, precisamos do seguinte:

Lema. *(desigualdade do valor médio generalizada) Sejam $\phi : [a,b] \to I\!R^m$, $\psi : [a,b] \to I\!R$ funções contínuas, diferenciáveis no intervalo aberto (a,b), tais que $\big\|\phi'(t)\big\| \le \psi'(t)$ para todo $t \in (a,b)$, onde usamos uma* norma arbitrária $\|\cdot\|$ *em $I\!R^m$. Então:*

$$\big\|\phi(b) - \phi(a)\big\| \le \psi(b) - \psi(a).$$

Demonstração. Seja $\lambda : I\!R^m \to I\!R$ um funcional linear unitário (i.e., $\|\lambda\| = 1$) que reproduz a norma do vetor $\phi(b) - \phi(a)$, i.e., tal que $\lambda\big(\phi(b) - \phi(a)\big) = \big\|\phi(b) - \phi(a)\big\|$ (recorde Aula número 9 — 03/04). Aplicando o teorema do valor médio para a função $\psi - \lambda \circ \phi : [a,b] \to I\!R$ obtemos que existe $c \in (a,b)$ com:

$$(b-a)\big[\psi'(c) - \lambda\big(\phi'(c)\big)\big] = \psi(b) - \psi(a) - \lambda\big(\phi(b) - \phi(a)\big) = \psi(b) - \psi(a) - \big\|\phi(b) - \phi(a)\big\|;$$

observe agora que:

$$\lambda\big(\phi'(c)\big) \le \big|\lambda\big(\phi'(c)\big)\big| \le \big\|\phi'(c)\big\| \le \psi'(c),$$

donde $\psi'(c) - \lambda\big(\phi'(c)\big) \ge 0$. A conclusão segue. ■

Teorema. *(Taylor com resto de Lagrange) Suponha que o segmento de reta $[x, x+h]$ esteja contido no domínio U de f, que f seja de classe C^k em U $(k \ge 0)$ e $k+1$ vezes diferenciável nos pontos do segmento aberto $(x, x+h)$. Se r denota o resto do k-ésimo polinômio de Taylor de f em torno de x então:*

$$\|r(h)\| \le \frac{1}{(k+1)!} \sup_{z \in (x, x+h)} \big\|\mathrm{d}^{(k+1)} f(z)\big\| \, \|h\|^{k+1},$$

onde $\left\|\mathrm{d}^{(k+1)} f(z)\right\|$ *denota a norma da aplicação multi-linear* $\mathrm{d}^{(k+1)} f(z)$ *(veja o Exercício 29 da aula número 7 — 27/03).*

Demonstração. Considere a curva $\phi : [0,1] \to U \subset I\!R^m$ definida por:

$$\phi(t) = f(x+th) + (1-t)\,\mathrm{d}f(x+th)\cdot h + \frac{(1-t)^2}{2!}\,\mathrm{d}^{(2)} f(x+th)\cdot (h)^{(2)} + \cdots$$
$$\cdots + \frac{(1-t)^k}{k!}\,\mathrm{d}^{(k)} f(x+th)\cdot (h)^{(k)},$$

para todo $t \in [0,1]$. Temos que ϕ é contínua em $[0,1]$, derivável em $(0,1)$ e usando indução em k é fácil mostrar que (veja o Exercício 4):

$$\phi'(t) = \frac{(1-t)^k}{k!}\,\mathrm{d}^{(k+1)} f(x+th)\cdot (h)^{(k+1)},$$

para todo $t \in (0,1)$. Considere também a função $\psi : [0,1] \to I\!R$ definida por:

$$\psi(t) = -\frac{(1-t)^{k+1}}{(k+1)!} \sup_{z\in(x,x+h)} \left\|\mathrm{d}^{(k+1)} f(z)\right\| \|h\|^{k+1};$$

temos:

$$\left\|\phi'(t)\right\| \le \frac{(1-t)^k}{k!} \sup_{z\in(x,x+h)} \left\|\mathrm{d}^{(k+1)} f(z)\right\| \|h\|^{k+1} = \psi'(t), \quad t \in (0,1).$$

Como $\phi(1) - \phi(0) = r(h)$ e $\psi(1) - \psi(0) = \frac{1}{(k+1)!}\sup_{z\in(x,x+h)} \left\|\mathrm{d}^{(k+1)} f(z)\right\| \|h\|^{k+1}$, a conclusão segue do lema anterior. ■

Teorema. *(Taylor com resto integral) Se f é de classe C^{k+1} em U ($k \ge 0$) e se o segmento $[x, x+h]$ está contido em U então o resto r do k-ésimo polinômio de Taylor de f em torno de x é dado por:*

$$r(h) = \frac{1}{k!}\int_0^1 (1-t)^k \mathrm{d}^{(k+1)} f(x+th)\cdot (h)^{(k+1)}\,\mathrm{d}t.$$

Demonstração. Defina ϕ como na demonstração da fórmula de Taylor com resto de Lagrange. Observe que ϕ é de classe C^1 em $[0,1]$ e aplique o teorema fundamental do cálculo para ϕ, i.e., observe que:

$$r(h) = \phi(1) - \phi(0) = \int_0^1 \phi'(t)\,\mathrm{d}t.$$

A conclusão segue. ■

`Observação`. Uma boa maneira de lembrar os enunciados precisos das fórmulas de Taylor é o seguinte: a fórmula de Taylor com resto infinitesimal é uma generalização da definição de função diferenciável (recaímos na definição de função diferenciável quando $k = 1$). A fórmula de Taylor com resto de Lagrange é uma generalização da desigualdade do valor médio (reobtemos a desigualdade do valor médio com $k = 0$). Finalmente, a fórmula de Taylor com resto integral é uma generalização do Teorema Fundamental do Cálculo (recaímos no Teorema Fundamental do Cálculo para $k = 0$).

`Observação`. O lema que foi chamado acima de "desigualdade do valor médio generalizada", leva esse nome pois tal lema nos dá novamente a desigualdade do valor médio para a função $\phi : [a,b] \to I\!R^m$ se tomarmos $\psi(t) = t \sup_{s\in(a,b)} \left\|\phi'(s)\right\|$, $t \in [a,b]$.

`Exemplo`. Se $f : U \subset I\!R^m \to I\!R^n$ é k vezes diferenciável num ponto $x \in U$ então não é difícil mostrar que (em notação de multi-índice):

$$\mathrm{d}^{(k)} f(x) \cdot (h)^{(k)} = \sum_{|\lambda|=k} \frac{k!}{\lambda!} \frac{\partial^k f}{\partial x^\lambda}(x) h^\lambda,$$

para todo $h \in I\!R^m$. Segue que o polinômio de Taylor de ordem k de f em torno de x pode ser escrito de maneira explícita em termos de derivadas parciais na forma:

$$p(h) = \sum_{i=0}^{k} \sum_{|\lambda|=i} \frac{1}{\lambda!} \frac{\partial^i f}{\partial x^\lambda}(x) h^\lambda.$$

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Funções polinomiais.`

1. Denote por $(e_i)_{i=1}^n$ a base canônica de $\mathbb{R}^n$. Para todo multi-índice $\lambda = (\lambda_1, \dots, \lambda_m)$ de dimensão m, denote por $p_\lambda^i : \mathbb{R}^m \to \mathbb{R}^n$ a função polinomial dada por $p_\lambda^i(x) = e_i x^\lambda$, para todo $x \in \mathbb{R}^m$. Mostre que:
   (a) as funções p_λ^i com $i = 1, \dots, n$, $|\lambda| = k$ formam uma base do espaço $\mathcal{P}_k(\mathbb{R}^m, \mathbb{R}^n)$ de todas as funções polinomiais homogêneas $p : \mathbb{R}^m \to \mathbb{R}^n$ de grau k; conclua que $\mathcal{P}_k(\mathbb{R}^m, \mathbb{R}^n)$ tem dimensão $n\binom{m+k-1}{k}$;
   (b) as funções p_λ^i com $i = 1, \dots, n$, $|\lambda| \le k$ formam uma base do espaço $\mathcal{P}_{\le k}(\mathbb{R}^m, \mathbb{R}^n)$ de todas as funções polinomiais $p : \mathbb{R}^m \to \mathbb{R}^n$ de grau menor ou igual a k; conclua que $\mathcal{P}_{\le k}(\mathbb{R}^m, \mathbb{R}^n)$ tem dimensão $n\binom{m+k}{k}$.
2. Sejam V, W espaços vetoriais e seja $B \in \text{Mult-lin}_k(V; W)$ uma aplicação multi-linear. O *simetrizador* de B é definido por:

$$\text{Sim}(B)(v_1, \dots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} B(v_{\sigma(1)}, \dots, v_{\sigma(k)}),$$

   para todos $v_1, \dots, v_k \in V$, onde S_k denota o conjunto das bijeções $\sigma : \{1, \dots, k\} \to \{1, \dots, k\}$. Mostre que:
   (a) $\text{Sim}(B) \in \text{Mult-lin}_k(V; W)$, i.e., que o simetrizador de B é k-linear;
   (b) $\text{Sim}(B)$ é simétrica;
   (c) B é simétrica se e somente se $\text{Sim}(B) = B$;
   (d) $B(x)^{(k)} = \text{Sim}(B)(x)^{(k)}$ para todo $x \in V$, i.e., B e $\text{Sim}(B)$ definem a mesma função polinomial homogênea de grau k;
   (e) $\text{Mult-lin}_k(V; W) \ni B \mapsto \text{Sim}(B) \in \text{Mult-lin}_k(V; W)$ é um *operador de projeção* (i.e., um operador linear cujo quadrado é igual a si próprio) e sua imagem é $\text{Mult-lin}_k^{\text{sim}}(V; W)$.
3. Seja $B \in \text{Mult-lin}_k^{\text{sim}}(V; W)$ um aplicação multi-linear simétrica tal que $B(x)^{(k)} = 0$ para todo $x \in V$. O objetivo deste exercício é mostrar que $B = 0$ *sem usar derivadas*!
   (a) dados $x, y \in V$, mostre a *fórmula do binômio*:

$$B(x+y)^{(k)} = \sum_{i=0}^{k} \binom{k}{i} B(x)^{(i)}(y)^{(k-i)},$$

   onde $B(x)^{(i)}(y)^{(k-i)} = B(\underbrace{x, \dots, x}_{i \text{ vezes}}, \underbrace{y, \dots, y}_{k-i \text{ vezes}})$;
   (b) dados $x, y \in V$, $t \in \mathbb{R}$, use o fato que $B(x+ty)^{(k)} = 0$ juntamente com a fórmula do binômio para concluir que $B(x)^{(i)}(y)^{(k-i)} = 0$ para todos $x, y \in V$ (*dica*: vocês lembram do determinante de Vandermond?)

(c) defina uma aplicação $(k-1)$-linear simétrica $\widetilde{B} \in \text{Mult-lin}_{k-1}\big(V; \text{Lin}(V, W)\big)$ fazendo $\widetilde{B}(x_1, \ldots, x_{k-1})(y) = B(x_1, \ldots, x_{k-1}, y)$. Use o item (b) para concluir que $\widetilde{B}(x)^{(k-1)} = 0$ para todo $x \in V$ e obtenha a conclusão final usando indução em k.

`Diferenciação`.

4. Sejam $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ uma função (com $U \subset \mathbb{R}^m$ aberto), $[x, x+h]$ um segmento contido em U e $t \in (0,1)$. Se f é $k+1$ vezes diferenciável no ponto $x+th$, mostre que:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}^{(k)} f(x+th) \cdot (h)^{(k)} = \mathrm{d}^{(k+1)} f(x+th) \cdot (h)^{(k+1)}.$$

dica: considere a composta $\lambda \circ \mathrm{d}^{(k)} f \circ \iota$, onde $\iota : (0,1) \to \mathbb{R}^m$ é definida por $\iota(t) = x + th$ e $\lambda : \text{Mult-lin}_k(\mathbb{R}^m; \mathbb{R}^n) \to \mathbb{R}^n$ é definida por $\lambda(B) = B(h)^{(k)}$. Use também a equação $(*)$ da Aula número 10 (05/04).

Aula número 14 **(26/04)**

A aula começa cobrindo o material sobre as estimativas para o resto do polinômio de Taylor, originalmente destinado à aula número 13.

(1) **A recíproca da fórmula de Taylor**.

O objetivo desta seção é provar o seguinte:

Teorema. *Seja $f : U \subset I\!R^m \to I\!R^n$ uma função k vezes diferenciável num ponto x do aberto $U \subset I\!R^m$ $(k \geq 1)$ e suponha que $p \in \mathcal{P}_{\leq k}(I\!R^m, I\!R^n)$ é um polinômio de grau menor ou igual a k tal que, definindo r pela fórmula:*

$$f(x+h) = p(h) + r(h),$$

então $\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0$. Então p é o polinômio de Taylor de ordem k de f em torno de x.

O teorema acima nos diz que *o polinômio de Taylor é o único polinômio de grau menor ou igual a k que faz a "estimativa infinitesimal" do resto funcionar.*

`Observação`. A "estimativa infinitesimal" para o resto r do polinômio de Taylor de ordem k é a mais fraca das estimativas que consideramos, no sentido que tanto a "estimativa de Lagrange" para r como a fórmula integral para r implicam a "estimativa infinitesimal", i.e., que $\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0$ (supondo f de classe C^{k+1}, ao menos). Segue então do teorema acima que, se p é um polinômio de grau menor ou igual a k para o qual o resto r satisfaz a "estimativa de Lagrange" ou a fórmula integral então p é necessariamente o polinômio de Taylor de f de ordem k.

Antes de mostrar o teorema, mostraremos o seguinte:

Lema. *Seja $p \in \mathcal{P}_{\leq k}(I\!R^m, I\!R^n)$ um polinômio de grau menor ou igual a k e suponha que $\lim_{h\to 0} \frac{p(h)}{\|h\|^k} = 0$. Então $p = 0$.*

Demonstração. Fixado $v \in I\!R^m$, $v \neq 0$, fazemos $h = tv$ e $t \to 0$ obtendo:

$$\lim_{t\to 0} \frac{p(tv)}{|t|^k \|v\|^k} = \lim_{t\to 0} \frac{p(tv)}{t^k} = 0.$$

Escreva $p = \sum_{i=0}^k p_i$, onde cada $p_i \in \mathcal{P}_i(I\!R^m, I\!R^n)$ é um polinômio homogêneo de grau i (e $p_0 \in I\!R^n$ é um vetor fixado). Obtemos então:

$$\lim_{t\to 0} \frac{p_0 + p_1(v)t + p_2(v)t^2 + \cdots + p_k(v)t^k}{t^k} = 0.$$

Um exercício simples de Cálculo I mostra que a igualdade acima só é possível se $p_i(v) = 0$ para $i = 0, \ldots, k$. Como $v \in I\!R^m \setminus \{0\}$ é arbitrário (e obviamente $p_i(0) = 0$), temos $p_i = 0$ para todo i e portanto $p = 0$. ■

Vamos agora mostrar o resultado principal da seção. Suponha que $f : U \subset I\!R^m \to I\!R^n$ ($U \subset I\!R^m$ aberto) é k vezes diferenciável no ponto $x \in U$ $(k \geq 1)$, que $p \in \mathcal{P}_{\leq k}(I\!R^m, I\!R^n)$

é um polinômio de grau menor ou igual a k e que o resto $r(h) = f(x+h) - p(h)$ satisfaz $\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0$. Devemos mostrar que p é o polinômio de Taylor de ordem k de f em torno de x. Como f é k vezes diferenciável em x e p é de classe C^∞, temos que r é k vezes diferenciável em 0 e aplicando a fórmula de Taylor infinitesimal para r em torno de 0 obtemos:

$$r(h) = q(h) + \widetilde{r}(h),$$

onde $q(h) = r(0) + \mathrm{d}r(0)\cdot h + \cdots + \frac{1}{k!}\mathrm{d}^{(k)}r(0)\cdot(h)^{(k)}$ é o polinômio de Taylor de ordem k de r em torno de 0 e $\widetilde{r}$ satisfaz:

$$\lim_{h\to 0} \frac{\widetilde{r}(h)}{\|h\|^k} = 0;$$

como também $\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0$, obtemos $\lim_{h\to 0} \frac{q(h)}{\|h\|^k} = 0$ e segue do lema acima que $q = 0$. Concluímos então que $r = \widetilde{r}$ e portanto r se anula em 0 juntamente com suas k primeiras diferenciais (recorde que o resto de um polinômio de Taylor de ordem k *sempre* se anula juntamente com suas k primeiras diferenciais no ponto 0 — e $r = \widetilde{r}$). Daí $f(x) = p(0)$ e $\mathrm{d}^{(i)}f(x) = \mathrm{d}^{(i)}p(0)$ para $i = 1, \ldots, k$, donde segue do primeiro Teorema da Seção 3 da Aula número 13 (24/04) que p é o polinômio de Taylor de ordem k de f em torno de x. Isso completa o argumento.

Aula número 15 (03/05)

(1) **Aplicação da fórmula de Taylor: máximos e mínimos locais**.

Vocês provavelmente já sabem que se $f : U \to \mathbb{R}$ ($U \subset \mathbb{R}^m$ aberto) é diferenciável num ponto $x \in U$ e possui um máximo ou mínimo local em x então $\mathrm{d}f(x) = 0$ (uma maneira de ver isso é a seguinte: para todo $v \in \mathbb{R}^m$, a função de uma variável $t \mapsto f(x + tv)$ possui um máximo local em $t = 0$ e portanto sabemos do Cálculo I que sua derivada $\frac{\mathrm{d}}{\mathrm{d}t} f(x + tv)\big|_{t=0}$ em $t = 0$ se anula; isso significa que $\mathrm{d}f(x) \cdot v = 0$ e como $v \in \mathbb{R}^m$ é arbitrário temos $\mathrm{d}f(x) = 0$). Esse fato simples será redemonstrado nesta seção como conseqüência da fórmula de Taylor, e muito mais: obteremos uma série de critérios que relacionam os mínimos e máximos locais de f com suas diferenciais de ordem superior.

Definição. *Sejam $U \subset \mathbb{R}^m$ um aberto e $f : U \to \mathbb{R}$ uma função. Dizemos que $x \in U$ é um* ponto de máximo *(global) para f se $f(y) \le f(x)$ para todo $y \in U$; dizemos que $x \in U$ é um* ponto de máximo estrito *(global) para f se $f(y) < f(x)$ para todo $y \in U$ com $y \ne x$. Dizemos que $x \in U$ é um* ponto de máximo local *quando x possui uma vizinhança V em U (ou uma vizinhança V em $\mathbb{R}^m$ contida em U, tanto faz) tal que x é um ponto de máximo de $f|_V$; dizemos que x é um* ponto de máximo local estrito *se x possui uma vizinhança V em U tal que x é um ponto de máximo estrito de $f|_V$.*

Analogamente, define-se as noções de *ponto de mínimo* (global), *ponto de mínimo estrito* (global), *ponto de mínimo local* e de *ponto de mínimo local estrito*; é só trocar $f(y) \le f(x)$ por $f(y) \ge f(x)$ e $f(y) < f(x)$ por $f(y) > f(x)$ na definição acima.

`Observação`. Obviamente noções envolvendo máximos e mínimos só podem ser definidas para *funções a valores em $\mathbb{R}$*.

Seja $f : U \subset \mathbb{R}^m \to \mathbb{R}$ uma função k vezes diferenciável num ponto $x \in U$, onde $U \subset \mathbb{R}^m$ é um aberto. Suponha que as $k-1$ primeiras diferenciais de f em x sejam nulas, mas que a k-ésima não seja, i.e., $\mathrm{d}^{(i)}f(x) = 0$ para $i = 1, \ldots, k-1$ mas $\mathrm{d}^{(k)}f(x) \ne 0$. A fórmula de Taylor infinitesimal nos dá então:

$$f(x + h) = f(x) + \frac{1}{k!}\mathrm{d}^{(k)}f(x) \cdot (h)^{(k)} + r(h),$$

com $\lim_{h\to 0} \frac{r(h)}{\|h\|^k} = 0$.

`Observação`. Tipicamente, o que ocorre é o seguinte: temos uma função $f : U \subset \mathbb{R}^m \to \mathbb{R}$ de classe C^∞ num aberto $U \subset \mathbb{R}^m$ e escolhemos um ponto $x \in U$. Daí sempre podemos escolher o menor inteiro $k \ge 1$ tal que $\mathrm{d}^{(k)}f(x) \ne 0$ e portanto $\mathrm{d}^{(i)}f(x) = 0$ para $i = 1, \ldots, k-1$; o único caso em que nossa técnica falha é se *todas* as diferenciais de f no ponto x se anulam. Infelizmente, é de fato possível que todas as diferenciais de f se anulem em $x \in U$, sem que f seja constante numa vizinhança de x (veremos exemplos neste curso). De qualquer modo, a técnica explicada nesta seção ainda se aplica numa enorme classe de funções.

Voltando à situação acima, escrevemos:

$$f(x + h) = f(x) + \|h\|^k \left(\frac{1}{k!}\mathrm{d}^{(k)}f(x) \cdot \left(\tfrac{h}{\|h\|}\right)^{(k)} + \frac{r(h)}{\|h\|^k}\right), \qquad (*)$$

para todo $h \in I\!R^m$ suficientemente pequeno com $h \neq 0$. Nossa estratégia agora é a seguinte: como $\frac{r(h)}{\|h\|^k}$ é muito pequeno para h próximo da origem, vamos usar a equação $(*)$ para concluir que o sinal de $f(x+h)-f(x)$ é o mesmo de $\mathrm{d}^{(k)}f(x)\cdot\left(\frac{h}{\|h\|}\right)^{(k)}$ para h suficientemente próximo de 0, desde que $\mathrm{d}^{(k)}f(x)\cdot\left(\frac{h}{\|h\|}\right)^{(k)} \neq 0$. Note que, ao contrário de $\frac{r(h)}{\|h\|^k}$, o valor de $\mathrm{d}^{(k)}f(x)\cdot\left(\frac{h}{\|h\|}\right)^{(k)}$ *não diminui quando* $h \to 0$!

Passamos às considerações formais. Suponha que existam $h_1, h_2 \in I\!R^m$ com:

$$\mathrm{d}^{(k)}f(x)\cdot(h_1)^{(k)} > 0 \quad \text{e} \quad \mathrm{d}^{(k)}f(x)\cdot(h_2)^{(k)} < 0;$$

por exemplo, isso *sempre* ocorre se k é ímpar, pois nesse caso:

$$\mathrm{d}^{(k)}f(x)\cdot(-h)^{(k)} = -\mathrm{d}^{(k)}f(x)\cdot(h)^{(k)},$$

para todo $h \in I\!R^m$.

Fazendo $h = th_1$ com $t > 0$ na equação $(*)$ obtemos:

$$f(x+th_1) = f(x) + t^k\|h_1\|^k\left(\frac{1}{k!}\mathrm{d}^{(k)}f(x)\cdot\left(\frac{h_1}{\|h_1\|}\right)^{(k)} + \frac{r(th_1)}{\|th_1\|^k}\right);$$

como $\lim_{h\to 0}\frac{r(h)}{\|h\|^k} = 0$ temos $\lim_{t\to 0^+}\frac{r(th_1)}{\|th_1\|^k} = 0$ e portanto:

$$\left|\frac{r(th_1)}{\|th_1\|^k}\right| < \underbrace{\frac{1}{k!}\mathrm{d}^{(k)}f(x)\cdot\left(\frac{h_1}{\|h_1\|}\right)^{(k)}}_{>0},$$

para todo $t > 0$ suficientemente pequeno. Concluímos então que:

$$f(x+th_1) > f(x),$$

para todo $t > 0$ suficientemente pequeno e portanto x *não é um ponto de máximo local*. Repetindo o raciocínio acima com h_2 no lugar de h_1 concluiremos que:

$$f(x+th_2) < f(x),$$

para todo $t > 0$ suficientemente pequeno e portanto x também *não é um ponto de mínimo local*.

Mostramos acima que se f é k vezes diferenciável em $x \in U$ e $\mathrm{d}^{(i)}f(x) = 0$ para $i = 1, \ldots, k-1$ então *o valor de f cresce (estritamente) quando andamos a partir de x numa direção $h \in I\!R^m$ com $\mathrm{d}^{(k)}f(x)\cdot(h)^{(k)} > 0$ e o valor de f decresce (estritamente) quando andamos a partir de x numa direção $h \in I\!R^m$ com $\mathrm{d}^{(k)}f(x)\cdot(h)^{(k)} < 0$.*

Definição. *Seja* $B \in \text{Mult-lin}_k^{\text{sim}}(V;\mathbb{R})$ *uma aplicação* k*-linear simétrica definida num espaço vetorial real* V*. Dizemos que:*

(•) B *é* definida positiva *quando* $B(v)^{(k)} > 0$ *para todo* $v \in V$, $v \neq 0$*;*
(•) B *é* semi-definida positiva *quando* $B(v)^{(k)} \geq 0$ *para todo* $v \in V$*;*
(•) B *é* definida negativa *quando* $B(v)^{(k)} < 0$ *para todo* $v \in V$, $v \neq 0$*;*
(•) B *é* semi-definida negativa *quando* $B(v)^{(k)} \leq 0$ *para todo* $v \in V$*;*
(•) B *é* indefinida *quando existem* $v, w \in V$ *com* $B(v)^{(k)} > 0$ *e* $B(w)^{(k)} < 0$.

Observe que se k é ímpar então uma aplicação k-linear simétrica B é *sempre* indefinida, a menos que $B = 0$ (isso segue da igualdade $B(-v)^{(k)} = -B(v)^{(k)}$, como já havíamos comentado).

Suponha agora que $f : U \to \mathbb{R}$ é k vezes diferenciável num ponto $x \in U$ ($U \subset \mathbb{R}^m$ aberto), $\mathrm{d}^{(i)}f(x) = 0$ para $i = 1, \ldots, k-1$ e que $\mathrm{d}^{(k)}f(x)$ seja definida positiva. Pelo que mostramos acima, o valor de f cresce quando caminhamos a partir de x *em qualquer direção*. Dá para desconfiar então que x seja um mínimo local de f, mas ainda falta um argumento para que possamos concluir isso. De fato, sabemos que para todo $h \in \mathbb{R}^m$, $\|h\| = 1$, existe $\delta > 0$ tal que $f(x + th) > f(x)$ para $0 < t < \delta$ — o problema é que δ *pode em princípio depender de* h. Nosso objetivo agora é mostrar que *de fato é possível escolher* $\delta > 0$ *independente de* h.

Como aplicações multi-lineares são contínuas e a esfera unitária de $\mathbb{R}^m$ é compacta, temos:

$$\inf_{\|h\|=1} \frac{1}{k!}\mathrm{d}^{(k)}f(x) \cdot (h)^{(k)} = \varepsilon > 0,$$

pois uma função contínua num compacto assume seu valor máximo e seu valor mínimo. Como $\lim_{h \to 0} \frac{r(h)}{\|h\|^k} = 0$, podemos encontrar $\delta > 0$ tal que:

$$0 < \|h\| < \delta \Longrightarrow \left|\frac{r(h)}{\|h\|^k}\right| < \varepsilon \leq \frac{1}{k!}\mathrm{d}^{(k)}f(x) \cdot \left(\tfrac{h}{\|h\|}\right)^{(k)}.$$

A partir da equação $(*)$ obtemos então:

$$0 < \|h\| < \delta \Longrightarrow f(x + h) > f(x),$$

e portanto x *é de fato um mínimo local estrito de* f. Obviamente, um argumento análogo mostra que se $\mathrm{d}^{(k)}f(x)$ é definida negativa então x é um ponto de máximo local estrito de f.

Podemos resumir todas as informações obtidas nesta seção no seguinte:

Teorema. *Sejam* $U \subset \mathbb{R}^m$ *um aberto,* $x \in U$ *e* $f : U \to \mathbb{R}$ *uma função* k *vezes diferenciável no ponto* x ($k \geq 1$)*. Suponha que* $\mathrm{d}^{(i)}f(x) = 0$ *para* $i = 1, \ldots, k-1$ *e que* $\mathrm{d}^{(k)}f(x) \neq 0$*. Temos que:*

(a) *se* k *é ímpar então* x *não é nem um ponto de máximo local nem um ponto de mínimo local para* f*;*
(b) *se* k *é par e* $\mathrm{d}^{(k)}f(x)$ *é indefinida então* x *não é nem um ponto de máximo local nem um ponto de mínimo local para* f*;*

(c) *se k é par e* $\mathrm{d}^{(k)}f(x)$ *é definida positiva então x é um ponto de mínimo local estrito para f;*
(d) *se k é par e* $\mathrm{d}^{(k)}f(x)$ *é definida negativa então x é um ponto de máximo local estrito para f;*
(e) *se x é um ponto de máximo local para f então k é par e* $\mathrm{d}^{(k)}f(x)$ *é semi-definida negativa;*
(f) *se x é um ponto de mínimo local para f então k é par e* $\mathrm{d}^{(k)}f(x)$ *é semi-definida positiva.* ■

Corolário. *Seja $f : U \to \mathbb{R}$ uma função diferenciável num ponto x do aberto $U \subset \mathbb{R}^m$. Se x é um ponto de máximo ou mínimo local para f então* $\mathrm{d}f(x) = 0$.

Demonstração. Aplique o teorema acima com $k = 1$. ■

(2) **Seqüências de funções**.

Definição. *Sejam X um conjunto, (M, d) um espaço métrico e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$. Dizemos que $(f_n)_{n\geq 1}$* converge pontualmente *(ou* simplesmente*) para uma função $f : X \to M$ se para cada $x \in X$ a seqüência $\big(f_n(x)\big)_{n\geq 1}$ converge para $f(x)$ no espaço métrico M. Dito de outro modo, $(f_n)_{n\geq 1}$ converge pontualmente para $f : X \to M$ quando dados $x \in X$ e $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $d\big(f_n(x), f(x)\big) < \varepsilon$ para todo $n \geq n_0$. Dizemos que $(f_n)_{n\geq 1}$* converge uniformemente *para $f : X \to M$ quando dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $d\big(f_n(x), f(x)\big) < \varepsilon$ para todo $n \geq n_0$* e todo *$x \in X$.*

Usamos a seguinte notação:

$$f_n \xrightarrow{\text{pp}} f \quad \text{significa “}(f_n)_{n\geq 1} \text{ converge pontualmente para } f\text{”},$$
$$f_n \xrightarrow{\text{u}} f \quad \text{significa “}(f_n)_{n\geq 1} \text{ converge uniformemente para } f\text{”}.$$

A diferença entre convergência uniforme e pontual é apenas uma diferença no posicionamento de quantificadores; de fato, observe que:

$$f_n \xrightarrow{\text{pp}} f \iff \forall\, \varepsilon > 0,\ \forall\, x \in X,\ \exists\, n_0 \in \mathbb{N},\ \forall\, n \in \mathbb{N},\ n \geq n_0 \Longrightarrow d\big(f_n(x), f(x)\big) < \varepsilon,$$
$$f_n \xrightarrow{\text{u}} f \iff \forall\, \varepsilon > 0,\ \exists\, n_0 \in \mathbb{N},\ \forall\, x \in X,\ \forall\, n \in \mathbb{N},\ n \geq n_0 \Longrightarrow d\big(f_n(x), f(x)\big) < \varepsilon.$$

No primeiro caso, n_0 pode depender de x, e no segundo deve ser possível encontrar *um único* n_0 para *todo* $x \in X$. Essa “pequena” diferença faz uma diferença enorme, como veremos adiante. Observe que obviamente convergência uniforme implica convergência pontual, i.e., $f_n \xrightarrow{\text{u}} f$ implica $f_n \xrightarrow{\text{pp}} f$.

`Exemplo`. Para cada $n \in \mathbb{N}$, seja $f_n : \mathbb{R} \to \mathbb{R}$ a função dada por $f_n(x) = \frac{x}{n}$, para todo $x \in \mathbb{R}$. Obviamente $f_n(x) \to 0$ para todo $x \in \mathbb{R}$, donde $f_n \xrightarrow{\text{pp}} 0$. Por outro lado, fixado $n \in \mathbb{N}$ então $\lim_{x\to+\infty} \big|f_n(x)\big| = +\infty$ — isso implica que $(f_n)_{n\geq 1}$ *não converge uniformemente* para a função nula. Por outro lado, se $X \subset \mathbb{R}$ é um conjunto limitado então $f_n|_X \xrightarrow{\text{u}} f|_X$; de fato, se $M > 0$ é tal que $X \subset [-M, M]$ então $\big|f_n(x)\big| < \varepsilon$ para todo $x \in X$ e todo $n \geq n_0$, desde que $n_0 \in \mathbb{N}$ seja escolhido de modo que $\frac{M}{n_0} < \varepsilon$.

O exemplo acima poderia dar ao leitor uma falsa impressão de que convergência pontual implica convergência uniforme num espaço compacto (já que compactos normalmente estão associados à "uniformidades" — por exemplo, funções contínuas em compactos *são* uniformemente contínuas). Veremos no exemplo a seguir que uma seqüência pontualmente convergente num compacto *pode não ser uniformemente convergente, mesmo que todas as funções envolvidas sejam contínuas.*

`Exemplo`. Para cada $n \in I\!N$, seja $f_n : [0,1] \to I\!R$ a única função contínua que se anula fora de $\left[\frac{1}{n+1}, \frac{1}{n}\right]$, é afim em cada metade de $\left[\frac{1}{n+1}, \frac{1}{n}\right]$ e vale 1 no ponto médio de $\left[\frac{1}{n+1}, \frac{1}{n}\right]$; explicitamente:

$$f_n(x) = \begin{cases} 0, & x \in [0,1],\ x \notin \left[\frac{1}{n+1}, \frac{1}{n}\right], \\ 2n(n+1)\big(x - \frac{1}{n+1}\big), & \frac{1}{n+1} \le x \le \frac{2n+1}{2n(n+1)}, \\ 2n(n+1)\big(\frac{1}{n} - x\big), & \frac{2n+1}{2n(n+1)} \le x \le \frac{1}{n}. \end{cases}$$

Fixado $x \in [0,1]$, temos que $f_n(x) = 0$ para todo n suficientemente grande e portanto $f_n \xrightarrow{\text{pp}} 0$. Por outro lado, toda função f_n assume o valor 1 e portanto não pode ser $f_n \xrightarrow{\text{u}} 0$.

O próximo teorema é uma das principais motivações para a introdução do conceito de convergência uniforme.

Teorema. *Sejam M, N espaços métricos e $(f_n)_{n\ge1}$ uma seqüência de funções $f_n : M \to N$. Suponha que $(f_n)_{n\ge1}$ converge uniformemente para uma função $f : M \to N$ e que todas as funções f_n são contínuas num certo ponto $x \in M$. Então f também é contínua em x. Em particular, se cada f_n é contínua em M então f é contínua em M ("o limite uniforme de funções contínuas é contínuo").*

Demonstração. Seja dado $\varepsilon > 0$. Como $(f_n)_{n\ge1}$ converge uniformemente para f, existe $n \in I\!N$ tal que $d\big(f_n(y), f(y)\big) < \frac{\varepsilon}{3}$ para todo $y \in M$. Como f_n é contínua no ponto x, existe $\delta > 0$ tal que $d\big(f_n(y), f_n(x)\big) < \frac{\varepsilon}{3}$, para todo $y \in M$ com $d(y,x) < \delta$. Daí, para todo $y \in M$:

$$\begin{aligned} d(y,x) < \delta \Longrightarrow d\big(f(y), f(x)\big) &\le d\big(f(y), f_n(y)\big) + d\big(f_n(y), f_n(x)\big) + d\big(f_n(x), f(x)\big) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Logo f é contínua no ponto x. ■

`Exemplo`. Para cada $n \in I\!N$, seja $f_n : [0,1] \to I\!R$ a função definida por $f_n(x) = x^n$, para todo $x \in [0,1]$. Temos que $f_n \xrightarrow{\text{pp}} f$, onde $f(x) = 0$ para $x \in [0,1)$ e $f(1) = 1$. Observe que cada f_n é contínua, mas f não é (em particular, pelo teorema acima, não pode ser $f_n \xrightarrow{\text{u}} f$).

O seguinte conceito é útil para reconhecer seqüências uniformemente convergentes sem achar o limite explicitamente:

Definição. *Sejam X um conjunto, (M,d) um espaço métrico e $(f_n)_{n\ge1}$ uma seqüência de funções $f_n : X \to M$. Dizemos que $(f_n)_{n\ge1}$ é* uniformemente de Cauchy *quando dado $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que $d\big(f_n(x), f_m(x)\big) < \varepsilon$ para todo $x \in X$ e todos $n, m \ge n_0$ (n_0 pode depender de ε, mas* não *de x).*

Teorema. *Sejam X um conjunto, M um espaço métrico completo e $(f_n)_{n\ge1}$ uma seqüência uniformemente de Cauchy de funções $f_n : X \to M$. Então existe uma (única) função $f : X \to M$ tal que $f_n \xrightarrow{\text{u}} f$.*

Demonstração. Obviamente $\big(f_n(x)\big)_{n\ge1}$ é uma seqüência de Cauchy em M para todo $x \in M$ e portanto converge para um ponto de M; podemos definir então $f : X \to M$ fazendo $f(x) = \lim_{n\to+\infty} f_n(x)$. Obviamente $f_n \xrightarrow{\text{pp}} f$; devemos mostrar que $f_n \xrightarrow{\text{u}} f$. Seja dado $\varepsilon > 0$; como $(f_n)_{n\ge1}$ é uniformemente de Cauchy, existe $n_0 \in \mathbb{N}$ tal que $d\big(f_n(x), f_m(x)\big) < \varepsilon$ para todo $x \in X$ e todos $n, m \ge n_0$. Fixado $n \ge n_0$ e $x \in X$ temos:

$$d\big(f_n(x), f(x)\big) = \lim_{m\to+\infty} d\big(f_n(x), f_m(x)\big) \le \varepsilon,$$

onde a primeira igualdade acima é justificada pela continuidade da função $d : M \times M \to \mathbb{R}$ (recorde Exercício 3 da Aula número 7 — 27/03) e pelo fato que $\lim_{m\to+\infty} f_m(x) = f(x)$. Concluímos então que $d\big(f_n(x), f(x)\big) \le \varepsilon$ para todo $n \ge n_0$ e todo $x \in X$ (observe que n_0 foi escolhido antes de x!). Logo $f_n \xrightarrow{\text{u}} f$. ■

Nosso interesse sobre convergência de funções está em suas relações com a diferenciação. Seria razoável esperar resultados do tipo "$f_n \to f$ implica $\mathrm{d}f_n \to \mathrm{d}f$" *sob as hipóteses certas*. No exemplo a seguir, veremos que a possibilidade mais óbvia não funciona.

`Exemplo`. Para cada $n \in \mathbb{N}$, seja $f_n : \mathbb{R} \to \mathbb{R}$ a função definida por $f_n(x) = \frac{1}{n}\operatorname{sen}(nx)$, para todo $x \in \mathbb{R}$. Como $\big|f_n(x)\big| \le \frac{1}{n}$ para todo $x \in \mathbb{R}$, é claro que $f_n \xrightarrow{\text{u}} 0$. Por outro lado, $f_n'(x) = \cos(nx)$ e obviamente $(f_n')_{n\ge1}$ não converge para zero (nem pontualmente).

O que funciona é o seguinte:

Teorema. *Sejam $U \subset \mathbb{R}^m$ um aberto convexo limitado e $(f_k)_{k\ge1}$ uma seqüência de funções diferenciáveis $f_k : U \to \mathbb{R}^n$, de modo que $(\mathrm{d}f_k)_{k\ge1}$ converge uniformemente para uma função $g : U \to \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$. Suponha que existe algum ponto $x_0 \in U$ para o qual a seqüencia $\big(f_k(x_0)\big)_{k\ge1}$ converge em $\mathbb{R}^n$. Então $(f_k)_{k\ge1}$ converge uniformemente para alguma função $f : U \to \mathbb{R}^n$, f é diferenciável e $\mathrm{d}f = g$.*

Demonstração. Sejam $x \in U$, $k, l \in \mathbb{N}$; como U é convexo, podemos aplicar a desigualdade do valor médio para a função $f_k - f_l$ no segmento $[x_0, x]$ obtendo:

$$\Big\|\big(f_k(x) - f_l(x)\big) - \big(f_k(x_0) - f_l(x_0)\big)\Big\| \le \sup_{z\in(x_0,x)} \big\|\mathrm{d}f_k(z) - \mathrm{d}f_l(z)\big\|\,\|x - x_0\|.$$

Como U é limitado, existe $M > 0$ com $\|x - x_0\| \le M$ para todo $x \in U$. Como $(\mathrm{d}f_k)_{k\ge1}$ é uniformemente convergente, temos que $(\mathrm{d}f_k)_{k\ge1}$ é uniformemente de Cauchy (veja Exercício 5) e portanto, dado $\varepsilon > 0$, podemos encontrar $k_0 \in \mathbb{N}$ tal que

$\big\|\mathrm{d}f_k(z) - \mathrm{d}f_l(z)\big\| < \frac{\varepsilon}{2M}$ para todo $z \in U$ e todos $k, l \geq k_0$. Além do mais, a seqüência $\big(f_k(x_0)\big)_{k\geq 1}$ é convergente em $I\!R^n$ e portanto de Cauchy; podemos então aumentar $k_0 \in I\!N$ de modo que $\big\|f_k(x_0) - f_l(x_0)\big\| < \frac{\varepsilon}{2}$ para todos $k, l \geq k_0$. Obtemos então:

$$\big\|f_k(x) - f_l(x)\big\| \leq \big\|\big(f_k(x) - f_l(x)\big) - \big(f_k(x_0) - f_l(x_0)\big)\big\| + \big\|f_k(x_0) - f_l(x_0)\big\| < \frac{\varepsilon}{2M}M + \frac{\varepsilon}{2} = \varepsilon,$$

para todo $x \in U$ e todos $k, l \geq k_0$ (note que k_0 não depende de x!). Mostramos então que $(f_k)_{k\geq 1}$ é uniformemente de Cauchy e portanto existe uma função $f : U \to I\!R^n$ tal que $(f_k)_{k\geq 1}$ é uniformemente convergente para f. Falta mostrar que f é diferenciável e que $\mathrm{d}f = g$. Fixe então $x \in U$ e vamos mostrar que f é diferenciável em x com $\mathrm{d}f(x) = g(x)$; para isso escrevemos:

$$f(x + h) = f(x) + g(x) \cdot h + r(h),$$

e tentamos mostrar que $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$. Como f_k é diferenciável em x, podemos escrever:

$$f_k(x + h) = f_k(x) + \mathrm{d}f_k(x) \cdot h + r_k(h),$$

com $\lim_{h\to 0} \frac{r_k(h)}{\|h\|} = 0$ para todo $k \in I\!N$. Como $f_k \xrightarrow{\mathrm{u}} f$ e $\mathrm{d}f_k \xrightarrow{\mathrm{u}} g$ temos:

$$\lim_{k\to+\infty} r_k(h) = r(h),$$

para todo $h \in I\!R^m$ com $x + h \in U$ (note que nessa passagem só usamos que $f_k \xrightarrow{\mathrm{pp}} f$ e $\mathrm{d}f_k \xrightarrow{\mathrm{pp}} g$). Fixados $k, l \in I\!N$, considere a função $\phi : U \to I\!R^n$ definida por:

$$\phi(z) = f_k(z) - f_l(z) - \mathrm{d}f_k(x) \cdot z + \mathrm{d}f_l(x) \cdot z, \quad z \in U \subset I\!R^m;$$

para todo $h \in I\!R^m$ com $x + h \in U$ temos $\phi(x + h) - \phi(x) = r_k(h) - r_l(h)$ e aplicando a desigualdade do valor médio para ϕ no segmento $[x, x + h] \subset U$ obtemos:

$$\begin{aligned}\big\|r_k(h) - r_l(h)\big\| &\leq \sup_{z\in(x,x+h)} \big\|\mathrm{d}\phi(z)\big\| \, \|h\| \\ &= \sup_{z\in(x,x+h)} \big\|\big(\mathrm{d}f_k(z) - \mathrm{d}f_l(z)\big) - \big(\mathrm{d}f_k(x) - \mathrm{d}f_l(x)\big)\big\| \, \|h\|,\end{aligned}$$

para todo $h \in I\!R^m$ com $x + h \in U$. Dado $\varepsilon > 0$, como $(\mathrm{d}f_k)_{k\geq 1}$ é uniformemente de Cauchy, podemos encontrar $k_0 \in I\!N$ tal que $\big\|\mathrm{d}f_k(z) - \mathrm{d}f_l(z)\big\| < \frac{\varepsilon}{4}$ para todo $z \in U$ e todos $k, l \geq k_0$. Daí:

$$\big\|\big(\mathrm{d}f_k(z) - \mathrm{d}f_l(z)\big) - \big(\mathrm{d}f_k(x) - \mathrm{d}f_l(x)\big)\big\| \leq \big\|\mathrm{d}f_k(z) - \mathrm{d}f_l(z)\big\| + \big\|\mathrm{d}f_k(x) - \mathrm{d}f_l(x)\big\| < \frac{\varepsilon}{2},$$

para todo $z \in U$ e portanto:

$$\big\|r_k(h) - r_l(h)\big\| \leq \frac{\varepsilon}{2}\|h\|,$$

para todo $h \in \mathbb{R}^m$ com $x + h \in U$ e todos $k, l \geq k_0$. Fixando $h \in \mathbb{R}^m$, $k \geq k_0$ e fazendo $l \to +\infty$ obtemos (já que $r_l(h) \to r(h)$):

$$\big\|r_k(h) - r(h)\big\| \leq \frac{\varepsilon}{2}\|h\|,$$

para todo $k \geq k_0$ e todo $h \in \mathbb{R}^m$ com $x+h \in U$. Fixe agora $k = k_0$; como $\lim_{h\to 0} \frac{r_k(h)}{\|h\|} = 0$, vemos que existe $\delta > 0$ tal que $\|h\| < \delta$ implica $\|r_k(h)\| \leq \frac{\varepsilon}{2}\|h\|$. Daí:

$$\|h\| < \delta \Longrightarrow \big\|r(h)\big\| \leq \big\|r(h) - r_k(h)\big\| + \big\|r_k(h)\big\| \leq \frac{\varepsilon}{2}\|h\| + \frac{\varepsilon}{2}\|h\| = \varepsilon\|h\|.$$

Isso mostra que $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$ e completa a demonstração. ■

A hipótese que U seja convexo e limitado no teorema acima é bastante desagradável; na verdade, ela só é necessária se quisermos concluir que $(f_k)_{k\geq 1}$ *converge uniformemente* para f. Se admitirmos um tipo de convergência mais fraco como conclusão, será possível supor apenas que U é aberto! Esse é nosso objetivo agora:

Definição. *Sejam M, N espaços métricos e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : M \to N$. Dada uma função $f : M \to N$ então dizemos que:*

- (•) $(f_n)_{n\geq 1}$ *converge* local-uniformemente *para f se todo ponto de M possui uma vizinhança onde f_n converge uniformemente para f, i.e., se dado $x \in M$ então existe $V \subset M$ aberto com $x \in V$ e $f_n|_V \xrightarrow{\text{u}} f|_V$;*
- (•) $(f_n)_{n\geq 1}$ *converge* uniformemente sobre compactos *para f se para todo subespaço compacto $K \subset M$ temos $f_n|_K \xrightarrow{\text{u}} f|_K$.*

É óbvio que tanto convergência local uniforme como convergência uniforme sobre compactos implicam convergência pontual. É óbvio também que convergência uniforme implica tanto convergência local-uniforme como convergência uniforme sobre compactos. Para seqüências de funções definidas em abertos de $\mathbb{R}^m$ é fácil ver que *convergência local-uniforme é equivalente à convergência uniforme sobre compactos*! (veja os Exercícios 9, 10(a–d) e 11).

Teorema. *Sejam $U \subset \mathbb{R}^m$ um aberto, $(f_k)_{k\geq 1}$ uma seqüência de funções diferenciáveis $f_k : U \to \mathbb{R}^n$ que converge pontualmente para uma função $f : U \to \mathbb{R}^n$. Suponha que $(\mathrm{d}f_k)_{k\geq 1}$ converge local-uniformemente (ou uniformemente sobre compactos, tanto faz) para uma função $g : U \to \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$. Então $(f_k)_{k\geq 1}$ converge local-uniformemente (ou uniformemente sobre compactos) para f, f é diferenciável* e $\mathrm{d}f = g$.

Demonstração. Dado $x_0 \in U$, seja $V = \mathrm{B}(x_0; r)$ uma bola aberta contida em U e tal que $(\mathrm{d}f_k)_{k\geq 1}$ converge uniformemente para g em V. Aplicando o teorema anterior para as funções $f_k|_V$ e para $g|_V$, concluímos que $f_k|_V \xrightarrow{\text{u}} f|_V$, $f|_V$ é diferenciável e que $\mathrm{d}\big(f|_V\big) = g|_V$. Como $x_0 \in U$ é arbitrário, a conclusão segue. ■

(3) **Funções real-analíticas**.

Este assunto não será explorado a fundo no curso. Faremos apenas uma rápida exposição para cultura geral. Vamos começar recordando alguns fatos simples sobre séries.

Definição. *Seja $(x_k)_{k\geq 1}$ uma seqüência em $I\!R^n$. Dizemos que a* série $\sum_{k=1}^{+\infty} x_k$ converge *para um vetor $x \in I\!R^n$ se a seqüência $(s_r)_{r\geq 1}$ em $I\!R^n$ definida por $s_r = \sum_{k=1}^{r} x_k$ converge para x (dizemos que $(s_r)_{r\geq 1}$* é a seqüência das somas parciais *da série $\sum_{k=1}^{+\infty} x_k$). Dizemos que a série $\sum_{k=1}^{+\infty} x_k$* é normalmente convergente *quando a série $\sum_{k=1}^{+\infty} \|x_k\|$ de números reais é convergente (no caso $n = 1$ é mais usual o termo* absolutamente convergente *em vez de normalmente convergente). Dizemos que a série $\sum_{k=1}^{+\infty} x_k$* é comutativamente convergente *em $I\!R^n$ se essa série converge para o mesmo vetor $x \in I\!R^n$ "independentemente da ordem em que os termos são somados", i.e., se para toda bijeção $\phi : I\!N \to I\!N$ temos que $\sum_{k=1}^{+\infty} x_{\phi(k)} = x$.*

Definição. *Sejam X um conjunto e $(f_k)_{k\geq 1}$ uma seqüência de funções $f_k : X \to I\!R^n$. Dizemos que a* série de funções $\sum_{k=1}^{+\infty} f_k$ converge pontualmente *para uma função $f : X \to I\!R^n$ se a seqüência $(s_r)_{r\geq 1}$ formada pelas funções $s_r : X \to I\!R^n$ definidas por $s_r = \sum_{k=1}^{r} f_k$ converge pontualmente para a função f. Dizemos que a série $\sum_{k=1}^{+\infty} f_k$* converge uniformemente *para f se a seqüência $(s_r)_{r\geq 1}$ converge uniformemente para f (a seqüência $(s_r)_{r\geq 1}$ é chamada a* seqüência das somas parciais *da série $\sum_{k=1}^{+\infty} f_k$).*

Para séries de funções $\sum_{k=1}^{+\infty} f_k$ (cada $f_k : X \to I\!R^n$) é comum também usar os termos "normalmente convergente" e "comutativamente convergente" significando respectivamente que para todo $x \in X$ a série de vetores $\sum_{k=1}^{+\infty} f_k(x)$ é normalmente convergente ou comutativamente convergente.

Seja $f : U \subset I\!R^m \to I\!R^n$ uma função de classe C^∞, onde $U \subset I\!R^m$ é um aberto. A *série de Taylor* de f em torno de um ponto $x \in U$ é a série (na variável $h \in I\!R^m$):

$$\sum_{|\lambda|=0}^{+\infty} \frac{1}{\lambda!} \frac{\partial^{|\lambda|} f}{\partial x^\lambda}(x) h^\lambda;$$

o índice λ usado na somatória acima percorre todos os multi-índices m-dimensionais, i.e., todas as m-uplas $(\lambda_1, \ldots, \lambda_m)$ de inteiros não negativos. A soma acima não é uma *série* no sentido estrito do termo, pois seus termos *não estão indexados nos números naturais*; observe porém, que o conjunto de todos os multi-índices λ é *enumerável* e escolhendo uma enumeração específica para os λ's (i.e., uma bijeção entre $I\!N$ e o conjunto dos multi-índices m-dimensionais) obtemos uma série no sentido estrito. Quando dissermos que a série de Taylor converge, estaremos significando que ela "converge comutativamente", i.e., que o resultado da soma não depende de como os λ's foram enumerados. Essa não é uma hipótese dura, pois sabe-se que séries normalmente convergentes são de fato comutativamente convergentes (veja o Exercício 18 — veja também o Exercício 26).

Observe que o polinômio de Taylor de ordem k de f em torno de x é exatamente a porção da série de Taylor correspondente a $|\lambda| \leq k$. Embora o resto do k-ésimo polinômio de Taylor seja "pequeno" em vários sentidos, *não é verdade em geral que a série de Taylor de uma função C^∞ seja convergente*! É possível também que a série de Taylor de f seja convergente, mas *não para f*! Exemplos desse fenômeno serão vistos mais adiante no curso.

Definição. *Uma função $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ ($U \subset \mathbb{R}^m$ aberto) é chamada* real-analítica *quando for de classe C^∞ e para todo $x \in U$ tivermos:*

$$f(x+h) = \sum_{|\lambda|=0}^{+\infty} \frac{1}{\lambda!} \frac{\partial^{|\lambda|} f}{\partial x^\lambda}(x) h^\lambda,$$

para todo h em alguma vizinhança da origem em $\mathbb{R}^m$ (a série acima deve convergir comutativamente*). De outro modo, f é real-analítica quando for igual à soma de sua série de Taylor em torno de cada ponto (numa vizinhança suficientemente pequena desse ponto).*

`Observação`. É fácil ver que funções polinomiais são real-analíticas, já que suas respectivas séries de Taylor são finitas (i.e., possuem apenas um número finito de termos não nulos). Para algumas funções elementares (como seno, cosseno, exponencial) é possível mostrar "manualmente" a condição de real-analiticidade, fazendo estimativas sobre o resto do polinômio de Taylor. Na verdade, é possível mostrar que *muitas* funções são real-analíticas, no seguinte sentido: a soma, produto e a composição de funções real-analíticas ainda é real-analítica. Em particular, toda "fórmula" que não envolva o cálculo de raízes n-ésimas em torno de zero é real-analítica (na verdade, vale muito mais: as soluções de muitos tipos importantes de equações diferenciais são real-analíticas, se os coeficientes da equação original o forem). A demonstração "manual" desses resultados (usando séries) é bastante dolorosa: a maneira inteligente de trabalhar com funções real-analíticas é observar a relação que existe entre as mesmas e as *funções analíticas complexas*. Não entraremos em detalhes neste curso.

O seguinte teorema da uma idéia de que na verdade existem "poucas" funções real-analíticas:

Teorema. *Seja $U \subset \mathbb{R}^m$ um aberto conexo e sejam $f, g : U \to \mathbb{R}^n$ funções real-analíticas. Se f e g coincidem em algum subconjunto aberto não vazio de U então $f = g$.*

Demonstração. Seja $A \subset U$ o conjunto dos pontos onde f e g coincidem juntamente com todas as suas diferenciais, i.e.:

$$A = \big\{ x \in U : f(x) = g(x),\ \mathrm{d}^{(k)} f(x) = \mathrm{d}^{(k)} g(x),\ k = 1, 2, \ldots \big\}.$$

Como o conjunto dos pontos onde duas funções contínuas são iguais é fechado e como a interseção de conjuntos fechados é fechada, segue que A é fechado em U. Além do mais, se $x \in A$ então as séries de Taylor de f e g em torno de x são iguais termo a termo e portanto a condição de real-analiticidade implica que $f(x+h) = g(x+h)$ para $\|h\| < r$ e algum $r > 0$ suficientemente pequeno. Daí a bola $\mathrm{B}(x; r)$ está contida em A, o que mostra que A é aberto em $\mathbb{R}^m$ (e em U). Como f e g são iguais em algum aberto não vazio, segue que A é não vazio e, como U é conexo, $A = U$. ■

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Seqüências de funções.`

1. Sejam X um conjunto, M um espaço métrico, $S \subset X$ um subconjunto e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$. Dada uma função $f : X \to M$, mostre que:
   (a) se $f_n \xrightarrow{\text{pp}} f$ então $f_n|_S \xrightarrow{\text{pp}} f|_S$;
   (b) se $f_n \xrightarrow{\text{u}} f$ então $f_n|_S \xrightarrow{\text{u}} f|_S$.
2. Sejam X um conjunto, M um espaço métrico e $(f_n)_{n\geq 1}$ uma seqüência de funções. Dada uma função $f : X \to M$, mostre que:
   (a) dada uma família $(X_i)_{i\in I}$ de subconjuntos de X com $X = \bigcup_{i\in I} X_i$, se para todo $i \in I$, $f_n|_{X_i} \xrightarrow{\text{pp}} f|_{X_i}$ então $f_n \xrightarrow{\text{pp}} f$;
   (b) dada uma família *finita* $(X_i)_{i=1}^n$ de subconjuntos de X com $X = \bigcup_{i=1}^n X_i$, se $f_n|_{X_i} \xrightarrow{\text{u}} f|_{X_i}$ para todo $i = 1, \ldots, n$ então $f_n \xrightarrow{\text{u}} f$;
   (c) dê um contra-exemplo para o item (b) no caso que a família de subconjuntos X_i de X não é finita.
3. Sejam X um conjunto, (M, d) um espaço métrico e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$. Dada uma função $f : X \to M$, mostre que as seguintes afirmações são equivalentes:
   - $f_n \xrightarrow{\text{u}} f$;
   - $s_n = \sup_{x\in X} d\big(f_n(x), f(x)\big)$ é finito para n suficientemente grande e a seqüência $(s_n)_{n\geq 1}$ converge para zero em $I\!R$.
4. Sejam X um conjunto, (M, d) um espaço métrico e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$ que converge uniformemente para uma função $f : X \to M$. Mostre que se f é limitada então f_n é limitada para todo n suficientemente grande. Mostre que essa limitação é *uniforme para n grande*, i.e., existem $n_0 \in I\!N$ e $K > 0$ de modo que $d\big(f_n(x), f_n(y)\big) \leq K$ para todos $x, y \in X$ e todo $n \geq n_0$.
5. Sejam X um conjunto, M um espaço métrico e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$ que converge uniformemente para uma função $f : X \to M$. Mostre que $(f_n)_{n\geq 1}$ é uniformemente de Cauchy.
6. Sejam X um conjunto, (M, d) um espaço métrico e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$. Dada uma função $f : X \to M$, mostre que:
   (a) se $f_n \xrightarrow{\text{pp}} f$ e se $(f_n)_{n\geq 1}$ é *uniformemente limitada*, i.e., se existe $K > 0$ tal que $d\big(f_n(x), f_n(y)\big) \leq K$ para todos $x, y \in X$ e todo $n \in I\!N$, então f é limitada; na verdade, vale que $d\big(f(x), f(y)\big) \leq K$ para todos $x, y \in X$ (*dica*: a função $d : M \times M \to I\!R$ é contínua — veja o Exercício 3 da Aula número 7 – 27/03);
   (b) se $(f_n)_{n\geq 1}$ é uniformemente de Cauchy e cada f_n é limitada então $(f_n)_{n\geq 1}$ é uniformemente limitada.
   (c) se $(f_n)_{n\geq 1}$ converge uniformemente para uma função $f : X \to M$ e se cada f_n é limitada então f é limitada.

7. Sejam X um conjunto, M, N espaços métricos, $\phi : M \to N$ uma função uniformemente contínua e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : X \to M$. Mostre que:

   (a) se $(f_n)_{n\geq 1}$ é uniformemente de Cauchy então $(\phi \circ f_n)_{n\geq 1}$ é uniformemente de Cauchy;

   (b) se $(f_n)_{n\geq 1}$ converge uniformemente para alguma função $f : X \to M$ então a seqüência $(\phi \circ f_n)_{n\geq 1}$ converge uniformemente para $\phi \circ f$;

   (c) conclua do item (a) que se d e d' são métricas uniformemente equivalentes em M então uma seqüência de funções $f_n : X \to M$ é uniformemente de Cauchy com respeito a d se e somente se o for com respeito a d';

   (d) conclua do item (b) que se d e d' são métricas uniformemente equivalentes em M então uma seqüência de funções $f_n : X \to M$ converge uniformemente para $f : X \to M$ com respeito a d se e somente $(f_n)_{n\geq 1}$ converge uniformemente para f com respeito a d'.

8. Sejam X um conjunto e (M_i, d_i), $i = 1, \ldots, n$ espaços métricos. Considere o conjunto $M = \prod_{i=1}^n M_i$ munido da métrica $d(x, y) = \sum_{i=1}^n d_i(x_i, y_i)$ (ou de qualquer outra métrica usual para o produto que seja uniformemente equivalente a d — recorde os Exercícios 18 e 33 da aula número 5 (20/03) e o Exercício 34 da aula número 7 (27/03)). Dada uma seqüência de funções $(f_k)_{k\geq 1}$, $f_k : X \to M$, escreva $f_k = (f_k^1, \ldots, f_k^n)$ com $f_k^i : X \to M_i$. Mostre que:

   (a) $(f_k)_{k\geq 1}$ é uniformemente de Cauchy se e somente se $(f_k^i)_{k\geq 1}$ é uniformemente de Cauchy para todo $i = 1, \ldots, n$;

   (b) $(f_k)_{k\geq 1}$ converge uniformemente para uma função $f = (f^1, \ldots, f^n) : X \to M$ se e somente se $(f_k^i)_{k\geq 1}$ converge uniformemente para f^i para todo $i = 1, \ldots, n$.

9. Sejam M, N espaços métricos e $(f_n)_{n\geq 1}$ uma seqüência de funções $f_n : M \to N$. Mostre que se $(f_n)_{n\geq 1}$ converge local uniformemente para uma função $f : M \to N$ então $(f_n)_{n\geq 1}$ converge para f uniformemente sobre compactos (*dica*: use o item (b) do Exercício 2 acima).

10. Um espaço métrico M é dito *localmente compacto* quando todo ponto de M possui uma vizinhança compacta. Mostre que:

   (a) M é localmente compacto se e somente se todo ponto de M pertence a um aberto cujo fecho é compacto;

   (b) se M é localmente compacto então todo aberto $U \subset M$ é localmente compacto;

   (c) se M é localmente compacto então todo fechado $F \subset M$ é localmente compacto;

   (d) $\mathbb{R}^n$ é localmente compacto;

   (e) se M é localmente compacto, $U \subset M$ é aberto e $F \subset M$ é fechado então $F \cap U$ é localmente compacto (*dica*: use os itens (b) e (c));

   (f) se $S \subset M$ é um subespaço denso e localmente compacto então S é aberto em M (*dica*: dado $x \in S$, use o item (a) para concluir que existe uma vizinhança aberta U de x em M tal que $\overline{U \cap S} \cap S$ é compacto — mostre que $U \subset S$);

   (g) se $S \subset M$ é um subespaço localmente compacto então S é a interseção de um aberto U de M com um fechado F de M (*dica*: $F = \overline{S}$; use o item (f)).

11. Se M é um espaço métrico localmente compacto, N é um espaço métrico qualquer e $(f_n)_{n\geq 1}$ é uma seqüência de funções $f_n : M \to N$, mostre que $(f_n)_{n\geq 1}$ converge local uniformemente para $f : M \to N$ se e somente se $(f_n)_{n\geq 1}$ converge uniformemente sobre compactos para f.

`Séries`.

Muitos dos exercícios de séries que aparecem aqui não são fáceis! Na verdade, eles são apenas um incentivo para aqueles que querem recordar a teoria das séries (um livro legal para o assunto é o Curso de Análise vol. I do Elon).

12. Mostre que se $\sum_{k=1}^{+\infty} x_k$ é uma série convergente em $\mathbb{R}^n$ então a seqüência $(x_k)_{k\geq 1}$ converge para zero (*dica*: se $s_r = \sum_{k=1}^{r} x_k$ denota a r-ésima soma parcial da série então $x_k = s_k - s_{k-1}$).
13. Mostre que se $\sum_{k=1}^{+\infty} x_k$ é uma série convergente em $\mathbb{R}^n$ então para todo $r \geq 1$ a série $\sum_{k=r}^{+\infty} x_k$ é convergente e, denotando sua soma por S_r, temos que a seqüência $(S_r)_{r\geq 1}$ converge para zero (*dica*: S_r é a diferença entre a $(r-1)$-ésima soma parcial de $\sum_{k=1}^{+\infty} x_k$ e a soma total da série $\sum_{k=1}^{+\infty} x_k$).
14. Mostre que toda série normalmente convergente $\sum_{k=1}^{+\infty} x_k$ em $\mathbb{R}^n$ é convergente (*dica*: se $s_r = \sum_{k=1}^{r} x_k$ denota a r-ésima soma parcial da série, mostre que $\|s_r - s_u\| \leq \sum_{k=u+1}^{r} \|x_k\|$ para $u < r$ e conclua usando o Exercício 13 que $(s_r)_{r\geq 1}$ é uma seqüência de Cauchy).
15. Mostre que a afirmação "a série $\sum_{k=1}^{+\infty} x_k$ é normalmente convergente em $\mathbb{R}^n$" não depende da norma usada em $\mathbb{R}^n$.
16. (*desigualdade triangular para séries*) Mostre que se $\sum_{k=1}^{+\infty} x_k$ é uma série normalmente convergente em $\mathbb{R}^n$ então vale a desigualdade:

$$\left\| \sum_{k=1}^{+\infty} x_k \right\| \leq \sum_{k=1}^{+\infty} \|x_k\|.$$

17. Considere uma série $\sum_{k=1}^{+\infty} x_k$ em $\mathbb{R}^n$ e denote por x_k^i a i-ésima coordenada do vetor $x_k \in \mathbb{R}^n$. Mostre que:
    (a) $\sum_{k=1}^{+\infty} x_k$ converge em $\mathbb{R}^n$ se e somente se $\sum_{k=1}^{+\infty} x_k^i$ converge em $\mathbb{R}$ para todo $i = 1, \ldots, n$;
    (b) $\sum_{k=1}^{+\infty} x_k$ é normalmente convergente em $\mathbb{R}^n$ se e somente se $\sum_{k=1}^{\infty} x_k^i$ é absolutamente convergente em $\mathbb{R}$ para todo $i = 1, \ldots, n$.
18. Mostre que se uma série $\sum_{k=1}^{+\infty} x_k$ é normalmente convergente em $\mathbb{R}^n$ então essa série também é comutativamente convergente (*dica*: se $\phi : \mathbb{N} \to \mathbb{N}$ é uma bijeção então, dado $u \in \mathbb{N}$, temos que para $r \in \mathbb{N}$ suficientemente grande a soma parcial $\sum_{k=1}^{r} x_{\phi(k)}$ inclui ao menos os u primeiros termos de $\sum_{k=1}^{+\infty} x_k$ — agora use o Exercício 13 para concluir que $\sum_{k=u+1}^{+\infty} \|x_k\|$ tende a zero quando $u \to +\infty$).

`Observação`. Na verdade, vale a recíproca do resultado enunciado no Exercício 18, i.e., toda série comutativamente convergente em $\mathbb{R}^n$ também é normalmente convergente. Esse fato é mostrado da seguinte maneira: primeiro reduz-se ao caso $n = 1$ com um argumento do tipo “coordenada por coordenada”. Para $n = 1$ é possível provar o seguinte fato mais geral: se uma série é convergente em $I\!R$ mas não é absolutamente convergente (diz-se então que a série é *condicionalmente convergente*) então é possível reordenar os termos dessa série de modo a obter *qualquer número real como resultado da soma da série*. A idéia básica da prova desse fato é usar a parte positiva e a parte negativa da série de maneira balanceada, de modo a fazer com que as somas parciais oscilem em torno do resultado desejado. Uma prova completa desse fato interessante pode ser encontrada por exemplo no *Curso de Análise, vol. I* de *Elon Lages Lima*, pg. 120.

19. Mostre que uma série $\sum_{k=1}^{+\infty} x_k$ de números reais não negativos é convergente se e somente se a seqüência $s_r = \sum_{k=1}^{r} x_k$ de suas somas parciais é limitada — em caso afirmativo, mostre que $\sum_{k=1}^{+\infty} = \sup_{r\ge 1} s_r$.

20. Sejam $\sum_{k=1}^{+\infty} x_k$, $\sum_{k=1}^{+\infty} y_k$ duas séries convergentes de números reais não negativos. Mostre que a “série” $\sum_{k,l=1}^{+\infty} x_k y_l$ converge para o produto $\left(\sum_{k=1}^{+\infty} x_k\right)\left(\sum_{l=1}^{+\infty} y_l\right)$, onde para somar os produtos $x_k y_l$ podemos escolher *qualquer* enumeração dos pares $(k, l) \in I\!N \times I\!N$ (*dica*: é só mostrar que toda soma parcial de $\sum_{k,l=1}^{+\infty} x_k y_l$ é menor ou igual a algum produto da forma $(\sum_{k=1}^{r} x_k)(\sum_{l=1}^{r} y_l)$ e que, reciprocamente, qualquer produto dessa forma é menor ou igual a alguma soma parcial de $\sum_{k,l=1}^{+\infty} x_k y_l$).

21. Generalize o Exercício 20 para o caso de séries absolutamente convergentes, i.e., obtenha a conclusão do Exercício 20 supondo apenas que $\sum_{k=1}^{+\infty} x_k$ e $\sum_{k=1}^{+\infty} y_k$ sejam séries absolutamente convergentes de números reais (*dica*: usando o Exercício 20, você já sabe que $\sum_{k,l=1}^{+\infty} x_k y_l$ é absolutamente convergente e portanto comutativamente convergente; para obter a conclusão faça $r \to +\infty$ na identidade:

$$\sum_{k,l=1}^{r} x_k y_l = \left(\sum_{k=1}^{r} x_k\right)\left(\sum_{l=1}^{r} y_l\right).$$

Note que é possível ordenar os termos de $\sum_{k,l=1}^{+\infty} x_k y_l$ de modo que o lado esquerdo da igualdade acima seja uma soma parcial de $\sum_{k,l=1}^{+\infty} x_k y_l$).

22. (*a série geométrica*) Mostre que se $x \in I\!R$, $|x| < 1$ então $\sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$ (*dica*: quanto vale $\sum_{k=1}^{r} x^k$?). Mostre que a série $\sum_{k=1}^{+\infty} x^k$ é absolutamente convergente.

23. (*a série geométrica em várias variáveis*) Dado $x = (x_1, \ldots, x_n) \in I\!R^n$ com $|x_i| < 1$ para $i = 1, \ldots, n$, mostre que:

$$\frac{1}{(1 - x_1)(1 - x_2)\cdots(1 - x_n)} = \sum_{|\lambda|=0}^{+\infty} x^{\lambda},$$

onde na série do lado direito λ percorre os multi-índices de dimensão n — a série $\sum_{|\lambda|=0}^{+\infty} x^{\lambda}$ é absolutamente convergente (*dica: use os Exercícios 21 e 22*).

24. (*o teste M de Weierstrass*) Seja X um conjunto e seja $\sum_{k=1}^{+\infty} f_k$ uma série de funções $f_k : X \to I\!R^n$. Suponha que existem números reais não negativos M_k tais que $\|f_k(x)\| \le M_k$ para todo $x \in X$ e todo $k \in I\!N$; suponha também que $\sum_{k=1}^{+\infty} M_k$ converge em $I\!R$. Mostre que $\sum_{k=1}^{+\infty} f_k$ é uniformemente convergente para uma função $f : X \to I\!R^n$ (*dica*: se $s_r = \sum_{k=1}^{r} f_k$ denota a r-ésima soma parcial de $\sum_{k=1}^{+\infty} f_k$, mostre que $\big\|s_r(x) - s_u(x)\big\| \le \sum_{k=u+1}^{r} M_k$ para $u < r$ e todo $x \in X$; conclua usando o Exercício 13 que $(s_r)_{r\ge1}$ é uma seqüência uniformemente de Cauchy).

25. Sejam X um conjunto e $\sum_{k=1}^{+\infty} f_k$ uma série uniformemente convergente de funções $f_k : X \to I\!R^n$. Mostre que $f_k \xrightarrow{\mathrm{u}} 0$ (*dica*: se $s_r = \sum_{k=1}^{r} f_k$ denota a r-ésima soma parcial da série então por definição existe $s : X \to I\!R^n$ com $s_r \xrightarrow{\mathrm{u}} s$; observe que $f_k = s_k - s_{k-1}$).

`Funções real-analíticas.`

Como já foi mencionado, não nos aprofundaremos na teoria de funções real-analíticas neste curso. Os exercícios a seguir são apenas uma curiosidade.

26. Seja dado $m \in I\!N$. Para cada multi-índice m-dimensional λ, seja $a_\lambda \in I\!R^n$ um vetor fixado. Para cada $k \ge 0$, denote por $p_k : I\!R^m \to I\!R^n$ a função polinomial homogênea $p_k(x) = \sum_{|\lambda|=k} a_\lambda x^\lambda$ e por $B_k \in \text{Mult-lin}(I\!R^m; I\!R^n)$ a única aplicação k-linear simétrica com $B_k(x)^{(k)} = p_k(x)$ para todo $x \in I\!R^m$. Suponha que existe uma vizinhança $U \subset I\!R^m$ da origem tal que a série de potências $\sum_{|\lambda|=0}^{+\infty} a_\lambda x^\lambda$ converge para todo $x \in U$ *em alguma ordem*. Mostre que:
    (a) existem $M > 0$, $R > 0$ de modo que $\|a_\lambda\| \le \frac{M}{R^{|\lambda|}}$ para todo multi-índice λ (*dica*: tome $R > 0$ tal que o vetor $x = (\underbrace{R, \ldots, R}_{m \text{ vezes}})$ pertence a U; observe que a família $a_\lambda x^\lambda = a_\lambda R^{|\lambda|}$ em $I\!R^n$ é limitada).
    (b) se R é definido como no item (a) e $0 < R' < R$, mostre que a série $\sum_{|\lambda|=0}^{+\infty} a_\lambda x^\lambda$ converge normalmente e uniformemente no bloco retangular

$$[-R', R']^m = \big\{x \in I\!R^m : \|x\|_\infty \le R'\big\}$$

    (*dica*: use a majoração para $\|a_\lambda\|$ obtida no item (a), o teste M de Weierstrass e o Exercício 23).

27. (*unicidade da série de Taylor*) Seja $f : U \subset I\!R^m \to I\!R^n$ uma função definida num aberto $U \subset I\!R^m$ e seja $x \in U$ fixado. Suponha que $f(x+h) = \sum_{|\lambda|=0}^{+\infty} a_\lambda h^\lambda$ para todo $h \in I\!R^m$ numa vizinhança da origem, para alguma família de vetores $a_\lambda \in I\!R^n$, onde λ percorre o conjunto dos multi-índices de dimensão m. Mostre que f é de classe C^∞ numa vizinhança aberta de x em U e que $a_\lambda = \frac{1}{\lambda!}\frac{\partial^{|\lambda|} f}{\partial x^\lambda}(x)$ para todo λ (*dica*: use o Exercício 26 para justificar o fato que uma série de potências pode ser diferenciada termo a termo — calcule $\frac{\partial^{|\lambda|} f}{\partial x^\lambda}(x)$ diferenciando a série de potências $\sum_{|\lambda|=0}^{+\infty} a_\lambda h^\lambda$ com respeito a h no ponto $h = 0$ da maneira adequada).

Aula número 16 (08/05)

A aula começa cobrindo os tópicos: "seqüências de funções uniformemente de Cauchy", "diferenciação termo a termo" e "funções real-analíticas", originalmente destinados à aula número 15.

(1) **Uma função de classe $\mathbf{C}^\infty$ que não é real-analítica.**

Vamos nesta seção estudar um exemplo muito interessante: uma função $\xi : \mathbb{R} \to \mathbb{R}$ de classe C^∞ que se anula em $(-\infty, 0]$ e que não é identicamente nula. Em particular, esta função não pode ser real-analítica (pois duas funções real-analíticas num aberto conexo que coincidem num aberto não vazio são iguais). A existência dessa função não é apenas uma curiosidade: ela será muito importante quando estudarmos o conceito de *partição da unidade*.

Para mostrar que a função ξ que definiremos adiante é mesmo de classe C^∞, precisaremos de um teorema que apresentamos a seguir. Esse teorema é na verdade interessante em si mesmo, independentemente do seu uso nesta seção — também a técnica usada na sua demonstração é curiosa.

Teorema. *Sejam $U \subset \mathbb{R}^m$ um aberto, $x \in U$ um ponto, $f : U \to \mathbb{R}^n$ uma função contínua, diferenciável no aberto $U \setminus \{x\}$. Se o limite $\lim_{y\to x} \mathrm{d}f(y)$ existe (igual a $T \in \mathrm{Lin}(\mathbb{R}^m, \mathbb{R}^n)$, digamos) então f é diferenciável no ponto x e $\mathrm{d}f(x) = T$.*

Demonstração. Para $h \in \mathbb{R}^m$ suficientemente próximo da origem, temos que o segmento $[x, x+h]$ está contido em U e portanto podemos aplicar a desigualdade do valor médio para a função $f - T$ em tal segmento obtendo:

$$\big\|f(x+h) - f(x) - T(h)\big\| \le \sup_{y\in(x,x+h)} \big\|\mathrm{d}f(y) - T\big\|\, \|h\|;$$

o uso da desigualdade do valor médio é justificado pelo fato que $f-T$ é contínua em $[x, x+h]$ e diferenciável em $(x, x+h)$ (não sabemos ainda que $f - T$ é diferenciável em x, mas isso não é um problema). Definindo então r pela igualdade $f(x+h) = f(x)+T(h)+r(h)$ então, como $\lim_{y\to x} \mathrm{d}f(y) = T$, dado $\varepsilon > 0$, podemos encontrar $\delta > 0$ tal que $0 < \|y - x\| < \delta$ implica $\big\|\mathrm{d}f(y) - T\big\| < \varepsilon$ e portanto:

$$\left\|\frac{r(h)}{\|h\|}\right\| \le \varepsilon,$$

sempre que $\|h\| < \delta$. Logo $\lim_{h\to 0} \frac{r(h)}{\|h\|} = 0$, o que completa a demonstração. ■

`Observação`. Duas coisas são curiosas na demonstração acima: em primeiro lugar, usamos a desigualdade do valor médio com toda a sua força, i.e., levamos em conta o fato que a hipótese da desigualdade do valor médio exige diferenciabilidade da função *apenas no segmento aberto*. É interessante também a técnica de aplicar a desigualdade do valor médio para $f - T$. Em geral, quando f é uma função diferenciável numa vizinhança de um ponto

x, pode ser útil aplicar a desigualdade do valor médio para a função $y \mapsto f(y) - \mathrm{d}f(x) \cdot y$ obtendo a estimativa:

$$\big\|f(x+h) - f(x) - \mathrm{d}f(x) \cdot h\big\| \leq \sup_{y \in (x, x+h)} \big\|\mathrm{d}f(y) - \mathrm{d}f(x)\big\| \, \|h\|,$$

que é mais poderosa que a estimativa $\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$, dada pela definição de função diferenciável.

`Observação`. Apesar de termos demonstrado o teorema acima para funções de várias variáveis, só usaremo-lo para funções de uma variável.

Definimos agora nossa função $\xi : \mathbb{R} \to \mathbb{R}$ fazendo:

$$\xi(x) = \begin{cases} e^{-\frac{1}{x}}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Obviamente ξ é de classe C^∞ no aberto $\mathbb{R} \setminus \{0\}$; o problema é mostrar que ξ é infinitamente diferenciável no ponto 0. Em primeiro lugar, observe que $\lim_{x \to 0^+} e^{-\frac{1}{x}} = 0$, de modo que ξ é contínua em 0. Além do mais, $\xi'(x) = 0$ para $x < 0$ e:

$$\xi'(x) = \frac{1}{x^2} \, e^{-\frac{1}{x}},$$

para todo $x > 0$. Daí $\lim_{x \to 0} \xi'(x) = 0$ e pelo teorema provado acima, temos que ξ é diferenciável em $x = 0$ e $\xi'(0) = 0$. Como ξ' é contínua e diferenciável em $\mathbb{R} \setminus \{0\}$, podemos agora proceder de modo análogo para mostrar que ξ é duas vezes diferenciável em $x = 0$ e que $\xi''(0) = 0$. Para provar que ξ é de classe C^∞, usamos indução. Suponha que ξ é de classe C^k para algum $k \geq 1$ e que a k-ésima derivada $\xi^{(k)}$ de ξ seja da forma:

$$\xi^{(k)}(x) = \begin{cases} \frac{p(x)}{q(x)} \, e^{-\frac{1}{x}}, & x > 0, \\ 0, & x \leq 0, \end{cases}$$

onde $p, q : \mathbb{R} \to \mathbb{R}$ são polinômios. Daí $\xi^{(k)}$ é uma função contínua, diferenciável em $\mathbb{R} \setminus \{0\}$, $\xi^{(k+1)}(x) = 0$ para $x < 0$ e:

$$\xi^{(k+1)}(x) = \frac{p'(x)q(x) - p(x)q'(x)}{q(x)^2} \, e^{-\frac{1}{x}} + \frac{p(x)}{x^2 q(x)} \, e^{-\frac{1}{x}} = \frac{\widetilde{p}(x)}{\widetilde{q}(x)} \, e^{-\frac{1}{x}}, \quad x > 0,$$

onde $\widetilde{p}$ e $\widetilde{q}$ são polinômios. Por um argumento elementar de Cálculo I, segue que:

$$\lim_{x \to 0} \xi^{(k+1)}(x) = 0;$$

pelo teorema acima, ξ é $k+1$ vezes diferenciável em $\mathbb{R}$ e $\xi^{(k+1)}(0) = 0$ (e logo ξ é de classe C^{k+1}). Por indução em k, concluímos que ξ é de classe C^∞.

`Observação`. Note que *todos os coeficientes da série de Taylor de ξ em torno de 0 são nulos*, enquanto ξ não é nula em nenhuma vizinhança de 0, i.e., a série de Taylor de ξ em torno de 0 converge, mas não para o valor de ξ.

Curvas

Antes de continuar nosso estudo de cálculo diferencial, vamos estudar um pouco de curvas em $I\!R^n$. Na seção 2 a seguir faremos uma exposição da noção de comprimento de arco. Em seguida, relacionaremos o comprimento de arco com a integral da norma do vetor tangente, no caso de curvas de classe C^1. Para isso, precisaremos fazer uma rápida recordação da definição de integral de Riemann na seção 3 — não faremos um desenvolvimento completo da teoria de integração aqui, já que o mesmo pertence a outra parte do curso. Vamos apelar para o fato que estudantes neste estágio já devem ter um conhecimento básico de integral de Riemann (em uma variável, ao menos).

O próximo passo do curso será o estudo da noção de integral de linha. Primeiro, usaremos o *approach* clássico usando a linguagem de campos vetoriais. Depois passaremos ao *approach* moderno, usando a linguagem de formas diferenciais. Terminamos o estudo de curvas com a teoria dos campos conservativos (usando a linguagem de formas diferenciais, onde a expressão “campo conservativo” será trocada por “1-forma diferencial exata”).

(2) Curvas em $\mathbf{R^n}$ — o comprimento de arco.

Por uma *curva* em $I\!R^n$ significaremos uma aplicação arbitrária $\gamma : [a, b] \to I\!R^n$ definida num intervalo fechado $[a, b]$ (apesar do fato que na maior parte do tempo estaremos interessados em curvas que são ao menos contínuas). Por uma *partição* do intervalo $[a, b]$ significamos um subconjunto *finito* $P \subset [a, b]$ que contém os extremos a e b. Tipicamente escreveremos $P = \{t_0, \dots, t_k\}$, deixando subentendido que $k \geq 1$ e que:

$$a = t_0 < t_1 < \cdots < t_{k-1} < t_k = b.$$

A *variação* de uma curva $\gamma : [a, b] \to I\!R^n$ com respeito à partição $P = \{t_0, \dots, t_k\}$ é o número real não negativo $\mathrm{V}(\gamma; P)$ definido por:

$$\mathrm{V}(\gamma; P) = \sum_{i=0}^{k-1} \big\|\gamma(t_{i+1}) - \gamma(t_i)\big\|.$$

Definição. *O* comprimento de arco *(ou* variação total*) de uma curva $\gamma : [a, b] \to I\!R^n$ é definido por:*

$$\mathrm{V}(\gamma) = \sup_P \mathrm{V}(\gamma; P) \in [0, +\infty],$$

onde P percorre o conjunto de todas *as possíveis partições de $[a, b]$. Quando $\mathrm{V}(\gamma) < +\infty$ (i.e., quando existe $M > 0$ tal que $\mathrm{V}(\gamma; P) \leq M$ para toda partição P de $[a, b]$) então dizemos que a curva γ é* retificável *(ou* de variação limitada*).*

`Observação`. O nome “variação total” (no lugar de comprimento de arco) e “de variação limitada” (no lugar de retificável) é mais usado quando $n = 1$; isso provavelmente se deve ao fato que no caso $n = 1$ a imagem intuitiva que temos de uma “curva” não se encaixa bem.

`Observação`. É mais usual definir o comprimento de uma curva $\gamma : [a,b] \to I\!R^n$ apenas usando a norma Euclideana $\|\cdot\|$, apesar do fato que para o desenvolvimento da teoria que faremos, $\|\cdot\|$ pode em princípio denotar qualquer norma. É fácil mostrar que a condição de retificabilidade *não depende* da norma escolhida, apesar do fato que o comprimento de arco depende, obviamente (veja os Exercícios 3 e 4).

Se uma partição P' do intervalo $[a,b]$ contém uma partição P de $[a,b]$ então dizemos que P' *refina* (ou que é um *refinamento*) de P. É fácil ver usando a desigualdade triangular que para toda curva $\gamma : [a,b] \to I\!R^n$ temos:

$$\mathrm{V}(\gamma; P') \geq \mathrm{V}(\gamma; P);$$

de fato, é suficiente mostrar a desigualdade acima quando P' tem só um ponto a mais que P (o caso geral segue facilmente por indução). Se $P = \{t_0, \ldots, t_k\}$ e $P' = P \cup \{t\}$ com $t_i < t < t_{i+1}$ então a parcela $\big\|\gamma(t_{i+1}) - \gamma(t_i)\big\|$ em $\mathrm{V}(\gamma; P)$ é substituída pela soma $\big\|\gamma(t_{i+1}) - \gamma(t)\big\| + \big\|\gamma(t) - \gamma(t_i)\big\|$ em $\mathrm{V}(\gamma; P')$ e obviamente tal soma é maior ou igual à parcela original $\big\|\gamma(t_{i+1}) - \gamma(t_i)\big\|$.

A observação acima mostra que para verificar a retificabilidade de γ e para calcular o seu comprimento é suficiente considerar partições *suficientemente finas*, i.e., se P_0 é uma partição fixada de $[a,b]$ então:

$$\mathrm{V}(\gamma) = \sup_{P \supset P_0} \mathrm{V}(\gamma; P).$$

`Observação`. *A reta é o caminho mais curto entre dois pontos*, i.e., se $\gamma : [a,b] \to I\!R^n$ é uma curva então $\big\|\gamma(b) - \gamma(a)\big\| \leq \mathrm{V}(\gamma)$. Isso segue trivialmente da observação que $\big\|\gamma(b) - \gamma(a)\big\| = \mathrm{V}(\gamma; P)$, onde $P = \{a, b\}$.

Vamos provar algumas propriedades fáceis do comprimento de arco.

Teorema. *Se $\gamma : [a,b] \to I\!R^n$ é retificável e se $[c,d] \subset [a,b]$ é um subintervalo então $\gamma|_{[c,d]}$ também é retificável e $\mathrm{V}(\gamma|_{[c,d]}) \leq \mathrm{V}(\gamma)$.*

Demonstração. Se P é uma partição de $[c,d]$ então $P' = P \cup \{a, b\}$ é uma partição de $[a,b]$ e obviamente:

$$\mathrm{V}(\gamma|_{[c,d]}; P) \leq \big\|\gamma(c) - \gamma(a)\big\| + \mathrm{V}(\gamma|_{[c,d]}; P) + \big\|\gamma(b) - \gamma(d)\big\| = \mathrm{V}(\gamma; P') \leq \mathrm{V}(\gamma);$$

a conclusão segue tomando o supremo com respeito a P na desigualdade $\mathrm{V}(\gamma|_{[c,d]}; P) \leq \mathrm{V}(\gamma)$. ■

Teorema. *(aditividade por concatenação do comprimento de arco) Seja $\gamma : [a,b] \to I\!R^n$ uma curva e $c \in (a,b)$ um ponto qualquer. Se $\gamma|_{[a,c]}$ e $\gamma|_{[c,b]}$ são retificáveis então γ também é retificável e:*

$$\mathrm{V}(\gamma) = \mathrm{V}(\gamma|_{[a,c]}) + \mathrm{V}(\gamma|_{[c,b]}).$$

Demonstração. Para calcular o comprimento de γ é suficiente usar partições que contém c, ou seja:

$$\mathrm{V}(\gamma) = \sup_{c \in P} \mathrm{V}(\gamma; P);$$

mas toda partição P de $[a,b]$ contendo c é da forma $P = P' \cup P''$, onde P' e P'' são respectivamente partições de $[a,c]$ e de $[c,b]$. Reciprocamente, se P' é uma partição de $[a,c]$ e P'' é uma partição de $[c,b]$ então $P' \cup P''$ é uma partição de $[a,b]$ contendo c e:

$$\mathrm{V}(\gamma; P' \cup P'') = \mathrm{V}(\gamma|_{[a,c]}; P') + \mathrm{V}(\gamma|_{[c,b]}; P'');$$

tomando supremos em P' e P'' na igualdade acima obtemos a conclusão desejada. ■

Recordamos que uma aplicação $\sigma : [c,d] \to I\!R$ é dita *monótona* quando for crescente ou decrescente (*crescente*, aqui, significa apenas $x < y \Rightarrow \sigma(x) \le \sigma(y)$, i.e., crescente *não* significa estritamente crescente; comentário análogo vale para o termo decrescente).

Teorema. *(invariância do comprimento de arco por reparametrização) Seja $\gamma : [a,b] \to I\!R^n$ uma curva e $\sigma : [c,d] \to [a,b]$ uma aplicação monótona e sobrejetora. Então γ é retificável se e somente se $\gamma \circ \sigma$ o for e vale a igualdade:*

$$\mathrm{V}(\gamma) = \mathrm{V}(\gamma \circ \sigma).$$

Demonstração. Vamos supor para fixar as idéias que σ seja crescente; o caso em que σ é decrescente é análogo. Como σ é crescente e sobrejetora, obviamente vale $\sigma(c) = a$ e $\sigma(d) = b$. Daí, se P é uma partição de $[c,d]$ então $\sigma(P)$ é uma partição de $[a,b]$ e, reciprocamente (como σ é sobrejetora), toda partição P' de $[a,b]$ é da forma $\sigma(P)$ para alguma partição P de $[c,d]$. Além do mais, é fácil ver que se P é uma partição de $[c,d]$ e $P' = \sigma(P)$ então $\mathrm{V}(\gamma \circ \sigma; P) = \mathrm{V}(\gamma; P')$; de fato, escrevendo $P = \{t_0, \ldots, t_k\}$ com $c = t_0 < \cdots < t_k = d$ então $P' = \{\sigma(t_0), \ldots, \sigma(t_k)\}$ com $a = \sigma(t_0) \le \cdots \le \sigma(t_k) = b$ e portanto:

$$\mathrm{V}(\gamma \circ \sigma; P) = \sum_{i=0}^{k-1} \big\|(\gamma \circ \sigma)(t_{i+1}) - (\gamma \circ \sigma)(t_i)\big\| = \sum_{i=0}^{k-1} \big\|\gamma\big(\sigma(t_{i+1})\big) - \gamma\big(\sigma(t_i)\big)\big\| = \mathrm{V}(\gamma; P');$$

a última igualdade na fórmula acima é de fato correta apesar de termos possivelmente $\sigma(t_i) = \sigma(t_{i+1})$ para alguns i's (i.e., na somatória antes da última igualdade acima temos alguns termos nulos a mais que na definição padrão de $\mathrm{V}(\gamma; P')$). A conclusão segue então tomando o supremo em P na igualdade $\mathrm{V}(\gamma \circ \sigma; P) = \mathrm{V}\big(\gamma; \sigma(P)\big)$. ■

`Observação`. Se $\sigma : [c,d] \to [a,b]$ é monótona, contínua e se $\sigma\big(\{c,d\}\big) = \{a,b\}$ então σ é sobrejetora, pelo teorema do valor intermediário, de modo que as hipóteses do teorema acima são satisfeitas. Na verdade, se $\sigma : [c,d] \to [a,b]$ é monótona e sobrejetora então σ é automaticamente contínua (veja *Curso de Análise, vol. I, Elon Lages Lima*, pg. 182), de modo que não perdemos nada em supor σ contínua desde o começo no teorema acima.

`Exemplo`. Considere a função $f : [0,1] \to I\!R$ dada por:

$$f(t) = \begin{cases} t \cos \frac{1}{t}, & t \in (0,1], \\ 0, & t = 0. \end{cases}$$

Temos que f é contínua. Vamos mostrar que f não é de variação limitada. De fato, dado um inteiro $r \ge 1$, considere a partição P de $[0,1]$ definida por:

$$P = \{0,1\} \cup \bigcup_{k=1}^{r} \left\{ \frac{1}{2k\pi}, \frac{1}{2k\pi + \frac{\pi}{2}} \right\}.$$

Temos:

$$\mathrm{V}(f;P) \geq \sum_{k=1}^{r} \left| f\left(\frac{1}{2k\pi}\right) - f\left(\frac{1}{2k\pi + \frac{\pi}{2}}\right) \right| = \sum_{k=1}^{r} \frac{1}{2k\pi}.$$

Como a série $\sum_{k=1}^{+\infty} \frac{1}{2k\pi}$ é divergente, segue que $\mathrm{V}(f) = \sup_P \mathrm{V}(f;P) = +\infty$ e portanto f não é de variação limitada.

Definição. *Uma curva $\gamma : [a,b] \to I\!R^n$ é dita* parametrizada por comprimento de arco *quando for retificável e para todo $t \in (a,b]$ tivermos* $\mathrm{V}(\gamma|_{[a,t]}) = t - a$.

Usando a aditividade por concatenação do comprimento de arco, é fácil ver que uma curva $\gamma : [a,b] \to I\!R^n$ é parametrizada por comprimento de arco se e somente se for retificável e para todos $t_1, t_2 \in [a,b]$ com $t_1 < t_2$ tivermos $\mathrm{V}(\gamma|_{[t_1,t_2]}) = t_2 - t_1$. Observe também que se $\gamma : [a,b] \to I\!R^n$ é parametrizada por comprimento de arco então vale a desigualdade:

$$\big\|\gamma(t_1) - \gamma(t_2)\big\| \leq \mathrm{V}(\gamma|_{[t_1,t_2]}) = t_2 - t_1,$$

para todos $t_1, t_2 \in [a,b]$ com $t_1 < t_2$. Segue em particular que γ é Lipschitziana e portanto *contínua.*

(3) **O comprimento de arco para curvas de classe $\mathbf{C^1}$ usando integrais**.

Vamos começar com uma rápida recordação da definição de integral de Riemann. Não faremos um desenvolvimento completo da teoria de integração básica nessa parte do curso.

Seja $P = \{t_0, \ldots, t_k\}$ uma partição do intervalo $[a,b]$ (como sempre, convencionamos $a = t_0 < \cdots < t_k = b$). Uma *pontilhação* para a partição P é um conjunto $\tau = \{\tau_0, \ldots, \tau_{k-1}\}$, onde $t_i \leq \tau_i \leq t_{i+1}$ para $i = 0, \ldots, k-1$. O par (P, τ) será chamado uma *partição pontilhada* de $[a,b]$. Se $f : [a,b] \to I\!R$ é uma função então a *soma de Riemann* de f com respeito a (P,τ) é definida por:

$$S(f;P,\tau) = \sum_{i=0}^{k-1} f(\tau_i)(t_{i+1} - t_i).$$

A *norma* $\|P\|$ da partição P é o maior dos comprimentos dos intervalos $[t_i, t_{i+1}]$, i.e., $\|P\| = \max_{i=0}^{k-1}(t_{i+1} - t_i)$.

Definição. *Seja $f : [a,b] \to I\!R$ uma função limitada. Dizemos que f é* Riemann integrável *quando existe o "limite" $\lim_{\|P\|\to 0} S(f;P,\tau)$, i.e., quando existe um número real $I \in I\!R$ tal que para todo $\varepsilon > 0$, existe $\delta > 0$, de modo que $\big|S(f;P,\tau) - I\big| < \varepsilon$, para toda partição P de $[a,b]$ com $\|P\| < \delta$ e toda pontilhação τ de P. O número I (que, como é fácil ver, é único, quando existe) é chamado a* integral de Riemann *de f e é denotado por:*

$$I = \int_a^b f = \int_a^b f(x)\,\mathrm{d}x.$$

`Observação`. Apesar de termos considerado apenas funções a valores reais na definição acima, não haveria mal nenhum em considerar $f : [a,b] \to I\!R^n$ (na verdade, quando estudamos o resto integral da Fórmula de Taylor, usamos a integral de uma função a valores em $\mathbb{R}^n$). Definindo $\int_a^b f$ para $f : [a,b] \to \mathbb{R}^n$ de modo idêntico ao feito acima, é fácil ver que $f = (f^1, \ldots, f^n)$ é Riemann integrável se e somente se cada f^i, $i = 1, \ldots, n$, é Riemann integrável; nesse caso, a i-ésima coordenada de $\int_a^b f \in \mathbb{R}^n$ é $\int_a^b f^i$ para $i = 1, \ldots, n$.

`Observação`. A integral de Riemann só é definida da maneira acima *para funções limitadas*. A integral de Riemann para funções ilimitadas (chamada *integral imprópria de Riemann*) deve ser definida separadamente, como um limite de integrais próprias. Não teremos nenhum uso aqui para integrais impróprias de Riemann (uma das *muitas* vantagens da integral de Lebesgue, é a não necessidade de distinguir integrais próprias de impróprias).

Agora vamos ao resultado central desta seção.

Teorema. *Seja $\gamma : [a,b] \to \mathbb{R}^n$ uma curva de classe C^1. Então γ é retificável e seu comprimento é dado por:*

$$\mathrm{V}(\gamma) = \int_a^b \left\|\gamma'(t)\right\| \mathrm{d}t.$$

Demonstração. Como $\gamma' : [a,b] \to \mathbb{R}^n$ é contínua no compacto $[a,b]$, existe $M > 0$ tal que $\left\|\gamma'(t)\right\| \le M$ para todo $t \in [a,b]$. Daí, se $P = \{t_0, \ldots, t_k\}$ é uma partição de $[a,b]$, a desigualdade do valor médio nos dá:

$$\mathrm{V}(\gamma; P) = \sum_{i=0}^{k-1} \left\|\gamma(t_{i+1}) - \gamma(t_i)\right\| \le \sum_{i=0}^{k-1} M(t_{i+1} - t_i) = M(b-a).$$

Daí γ é retificável e $\mathrm{V}(\gamma) \le M(b-a)$. Falta mostrar agora que $\int_a^b \left\|\gamma'(t)\right\| \mathrm{d}t = \mathrm{V}(\gamma)$. Seja dado então $\varepsilon > 0$. Como γ é retificável, existe uma partição $P_0 = \{t_0, \ldots, t_k\}$ de $[a,b]$ tal que:

$$\mathrm{V}(\gamma) - \frac{\varepsilon}{3} < \mathrm{V}(\gamma; P_0) \le \mathrm{V}(\gamma).$$

Lembrando que uma função contínua num compacto é uniformemente contínua, podemos escolher $\delta > 0$ satisfazendo as três seguintes condições:

- $\delta < \min_{i=0}^{k-1}(t_{i+1} - t_i)$;
- para todos $t, s \in [a,b]$, $|t-s| < \delta \Rightarrow \left\|\gamma(t) - \gamma(s)\right\| < \frac{\varepsilon}{9k}$;
- para todos $t, s \in [a,b]$, $|t-s| < \delta \Rightarrow \left\|\gamma'(t) - \gamma'(s)\right\| < \frac{\varepsilon}{3(b-a)}$.

Seja agora $P = \{s_0, \ldots, s_l\}$ uma partição de $[a,b]$ com $\|P\| < \delta$ e seja $\tau = \{\tau_0, \ldots, \tau_{l-1}\}$ uma pontilhação para P. Para completar a demonstração, vamos mostrar que:

$$\left|S(f; P, \tau) - \mathrm{V}(\gamma)\right| < \varepsilon,$$

onde $f : [a,b] \to \mathbb{R}$ é definida por $f(t) = \left\|\gamma'(t)\right\|$. Dividimos o restante da demonstração em três partes.

(i) $\left|S(f;P,\tau) - \mathrm{V}(\gamma;P)\right| \le \frac{\varepsilon}{3}$;

fixado $i \in \{0,\ldots,l-1\}$, aplicamos a desigualdade do valor médio para a função $t \mapsto \gamma(t) - \gamma'(\tau_i)t$ no intervalo $[s_i, s_{i+1}]$ obtendo:

$$\begin{aligned}\big\|\gamma(s_{i+1}) - \gamma(s_i) - \gamma'(\tau_i)(s_{i+1}-s_i)\big\| &\le \sup_{t\in(s_i,s_{i+1})} \big\|\gamma'(t) - \gamma'(\tau_i)\big\|\,(s_{i+1}-s_i)\\ &\le \frac{\varepsilon}{3(b-a)}(s_{i+1}-s_i),\end{aligned}$$

já que $|t-\tau_i| < \delta$ para todo $t \in (s_i, s_{i+1})$. Concluímos então que:

$$\begin{aligned}\big|\mathrm{V}(\gamma;P) - S(f;P,\tau)\big| &\le \sum_{i=0}^{l-1}\Big|\big\|\gamma(s_{i+1}) - \gamma(s_i)\big\| - \big\|\gamma'(\tau_i)\big\|(s_{i+1}-s_i)\Big|\\ &\le \sum_{i=0}^{l-1}\big\|\gamma(s_{i+1}) - \gamma(s_i) - \gamma'(\tau_i)(s_{i+1}-s_i)\big\| \le \frac{\varepsilon}{3(b-a)}\sum_{i=0}^{l-1}(s_{i+1}-s_i) = \frac{\varepsilon}{3}.\end{aligned}$$

Isso completa a demonstração da parte (i).

(ii) $\left|\mathrm{V}(\gamma;P) - \mathrm{V}(\gamma;P\cup P_0)\right| < \frac{\varepsilon}{3}$;

como $\|P\| < \delta < \min_{i=0}^{k-1}(t_{i+1}-t_i)$, segue que cada intervalo da partição P contém no máximo um ponto da partição P_0 em seu interior. Definimos então:

$$\begin{aligned}I &= \Big\{i \in \{0,\ldots,l-1\} : (s_i,s_{i+1}) \cap P_0 = \emptyset\Big\},\\ I' &= \Big\{i \in \{0,\ldots,l-1\} : (s_i,s_{i+1}) \cap P_0 \ne \emptyset\Big\},\end{aligned}$$

e para cada $i \in I'$ denotamos por $t_{\rho(i)}$ ($\rho(i) \in \{1,\ldots,k-1\}$) o único ponto da partição P_0 que pertence ao intervalo aberto (s_i, s_{i+1}). Note que I' tem no máximo $k-1$ elementos. Temos:

$$\begin{aligned}\mathrm{V}(\gamma;P\cup P_0) &= \sum_{i\in I}\big\|\gamma(s_{i+1}) - \gamma(s_i)\big\| + \sum_{i\in I'}\big\|\gamma(s_{i+1}) - \gamma(t_{\rho(i)})\big\| + \big\|\gamma(t_{\rho(i)}) - \gamma(s_i)\big\|,\\ \mathrm{V}(\gamma;P) &= \sum_{i\in I}\big\|\gamma(s_{i+1}) - \gamma(s_i)\big\| + \sum_{i\in I'}\big\|\gamma(s_{i+1}) - \gamma(s_i)\big\|,\end{aligned}$$

e portanto:

$$\mathrm{V}(\gamma;P\cup P_0) - \mathrm{V}(\gamma;P) = \sum_{i\in I'}\big\|\gamma(s_{i+1}) - \gamma(t_{\rho(i)})\big\| + \big\|\gamma(t_{\rho(i)}) - \gamma(s_i)\big\| - \big\|\gamma(s_{i+1}) - \gamma(s_i)\big\|.$$

Como $|s_{i+1} - t_{\rho(i)}| < \delta$, $|t_{\rho(i)} - s_i| < \delta$ e $|s_{i+1} - s_i| < \delta$, o termo geral da somatória acima tem módulo menor que $\frac{\varepsilon}{3k}$; como a somatória tem no máximo k termos, concluímos que:

$$\mathrm{V}(\gamma;P) \le \mathrm{V}(\gamma;P\cup P_0) < \mathrm{V}(\gamma;P) + \frac{\varepsilon}{3},$$

o que completa a demonstração da parte (ii).

(iii) $\big|\mathrm{V}(\gamma; P \cup P_0) - \mathrm{V}(\gamma)\big| < \frac{\varepsilon}{3}$;

Basta observar que:

$$\mathrm{V}(\gamma) - \frac{\varepsilon}{3} < \mathrm{V}(\gamma; P_0) \le \mathrm{V}(\gamma; P \cup P_0) \le \mathrm{V}(\gamma).$$

Isso completa a demonstração da parte (iii) e a demonstração do teorema. ■

Definição. *Uma curva $\gamma : [a, b] \to I\!R^n$ é dita* de classe C^p por partes *($1 \le p \le \infty$) quando existe uma partição $P = \{t_0, \ldots, t_k\}$ do intervalo $[a, b]$ tal que $\gamma|_{[t_i, t_{i+1}]}$ é de classe C^p para todo $i = 0, \ldots, k-1$.*

Observe que se γ é de classe C^p por partes então γ é contínua em $[a, b]$ e (como $p \ge 1$) admite derivadas laterais em cada ponto t_i da partição. Nos pontos $t \in [a, b]$ que não estão em P (e nas extremidades $t = a$ e $t = b$) a curva γ é diferenciável.

Corolário. *Se $\gamma : [a, b] \to I\!R^n$ é de classe C^1 por partes então γ é retificável e seu comprimento é dado por:*

$$V(\gamma) = \int_a^b \big\|\gamma'(t)\big\| \, \mathrm{d}t.$$

O integrando acima não é bem definido num número finito de pontos, mas isso não é um problema: sabe-se que a integral de uma função não se altera quando alteramos a função num número finito de pontos e portanto podemos atribuir um valor arbitrário ao integrando em questão quando o mesmo não estiver bem definido.

Demonstração. Seja $P = \{t_0, \ldots, t_k\}$ uma partição de $[a, b]$ tal que $\gamma|_{[t_i, t_{i+1}]}$ é de classe C^1 para $i = 0, \ldots, k-1$. Pelo teorema anterior e pela aditividade do comprimento de arco (e usando uma propriedade elementar da integral de Riemann):

$$\mathrm{V}(\gamma) = \sum_{i=0}^{k-1} \mathrm{V}(\gamma|_{[t_i, t_{i+1}]}) = \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \big\|\gamma'(t)\big\| \, \mathrm{d}t = \int_a^b \big\|\gamma'(t)\big\| \, \mathrm{d}t. \blacksquare$$

Corolário. *Se $\gamma : [a, b] \to I\!R^n$ é uma curva de classe C^1 por partes então γ é parametrizada por comprimento de arco se e somente se $\big\|\gamma'(t)\big\| = 1$ para todo $t \in [a, b]$ tal que γ é diferenciável em t.*

Demonstração. Segue do Teorema Fundamental do Cálculo e da igualdade:

$$\mathrm{V}(\gamma|_{[a,t]}) = \int_a^t \big\|\gamma'(s)\big\| \, \mathrm{d}s,$$

para todo $t \in (a, b]$. ■

`Observação`. Para cultura geral: a fórmula $\mathrm{V}(\gamma) = \int_a^b \big\|\gamma'(t)\big\| \, \mathrm{d}t$ vale numa classe de curvas maior do que aquela das curvas C^1 por partes. Essa é a classe das curvas chamadas *absolutamente contínuas* (veja o último capítulo do livro de Teoria da Medida do Pedro

Jesus Fernandez, se tiver curiosidade); nesse contexto, a integral $\int_a^b \|\gamma'(t)\| \, \mathrm{d}t$ deve ser entendida no sentido de Lebesgue. A classe de curvas absolutamente contínuas inclui até mesmo a classe das curvas Lipschitzianas. Mostra-se então que uma curva absolutamente contínua $\gamma : [a,b] \to \mathbb{R}^n$ é parametrizada por comprimento de arco se e somente se a igualdade $\|\gamma'(t)\| = 1$ vale *para quase todo* $t \in [a,b]$, i.e., quando tal igualdade é falsa no máximo num subconjunto de $[a,b]$ que tem *medida de Lebesgue zero*.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Comprimento de arco.`

1. Seja $\gamma : [a,b] \to I\!R^n$ uma curva. Mostre que $\mathrm{V}(\gamma) \geq 0$ e que $\mathrm{V}(\gamma) = 0$ se e somente se γ é constante.
2. Seja $f : [a,b] \to I\!R$ uma função monótona. Mostre que f é de variação limitada e que $\mathrm{V}(f) = \big|f(b) - f(a)\big|$.
3. Mostre que a condição "$\gamma : [a,b] \to I\!R^n$ é retificável" não depende da norma usada em $I\!R^n$.
4. Seja $\|\cdot\|_1$ a norma em $I\!R^n$ definida por $\|x\|_1 = \sum_{i=1}^n |x_i|$. Dada uma curva $\gamma : [a,b] \to I\!R^n$ e uma partição P de $[a,b]$, denote por $\mathrm{V}_1(\gamma; P)$ a variação de γ com respeito a P definida usando a norma $\|\cdot\|_1$ e por $\mathrm{V}_1(\gamma) = \sup_P \mathrm{V}_1(\gamma; P)$ o comprimento de arco de γ definido usando a norma $\|\cdot\|_1$. Mostre que:
   (a) se $\gamma = (\gamma^1, \ldots, \gamma^n)$ com cada $\gamma^i : [a,b] \to I\!R$ então $\mathrm{V}_1(\gamma; P) = \sum_{i=1}^n \mathrm{V}_1(\gamma^i; P)$ (onde a norma usada em $I\!R$ é o módulo usual);
   (b) conclua do item (a) que γ é retificável (com respeito a $\|\cdot\|_1$ ou a qualquer outra norma — vide Exercício 3) se e somente se cada γ^i, $i = 1, \ldots, n$, é de variação limitada;
   (c) se γ é retificável mostre que $\mathrm{V}_1(\gamma) = \sum_{i=1}^n \mathrm{V}_1(\gamma^i)$ (*dica*: você deve precisar do seguinte fato: se $P_1, \ldots, P_n$ são partições de $[a,b]$ então $P = \bigcup_{i=1}^n P_i$ é uma partição de $[a,b]$ que refina todas as P_i's simultaneamente);
   (d) se $\mathrm{V}(\gamma)$ denota o comprimento de arco de γ com respeito à norma Euclideana, dê um exemplo para mostrar que $\mathrm{V}(\gamma)$ *não pode ser calculado a partir dos números* $\mathrm{V}(\gamma^i)$, $i = 1, \ldots, n$, i.e., procure curvas retificáveis $\gamma, \mu : [a,b] \to I\!R^n$ com $\mathrm{V}(\gamma^i) = \mathrm{V}(\mu^i)$ para todo $i = 1, \ldots, n$, mas com $\mathrm{V}(\gamma) \neq \mathrm{V}(\mu)$ (*dica*: é fácil achar um exemplo com $n = 2$ e com as aplicações γ^i, μ^i, $i = 1, 2$ todas monótonas).
5. Seja $\gamma : [a,b] \to I\!R^n$ uma curva retificável que é contínua no instante $t = a$. O objetivo deste exercício é mostrar que $\lim_{t \to a} \mathrm{V}(\gamma|_{[a,t]}) = 0$.
   (a) Seja $c = \inf_{t \in (a,b]} \mathrm{V}(\gamma|_{[a,t]}) \geq 0$ e suponha por absurdo que $c > 0$. Dado $\varepsilon > 0$ arbitrário, escolha $t \in (a,b]$ com $\mathrm{V}(\gamma|_{[a,t]}) < c + \varepsilon$. Mostre que para todo $s \in (a,t)$ temos $\mathrm{V}(\gamma|_{[s,t]}) < \varepsilon$.
   (b) Escolha agora $t \in (a,b]$ com $\mathrm{V}(\gamma|_{[a,t]}) < c + \varepsilon$ e tal que $\big\|\gamma(s) - \gamma(a)\big\| < \varepsilon$ para todo $s \in [a,t]$. Mostre que para toda partição P de $[a,t]$ temos $\mathrm{V}(\gamma|_{[a,t]}; P) < 2\varepsilon$ (*dica*: trate separadamente o primeiro termo de $\mathrm{V}(\gamma|_{[a,t]}; P)$).
   (c) Conclua do item (b) que $\mathrm{V}(\gamma|_{[a,t]}) \leq 2\varepsilon$ e obtenha uma contradição (de modo que $c = 0$).
   (d) Observando que a aplicação $(a,b] \ni t \mapsto \mathrm{V}(\gamma|_{[a,t]}) \in I\!R$ é crescente, conclua que $\lim_{t \to a} \mathrm{V}(\gamma|_{[a,t]}) = 0$.

6. Seja $\gamma : [a,b] \to I\!R^n$ uma curva retificável e defina $V : [a,b] \to I\!R$ fazendo $V(t) = \mathrm{V}(\gamma|_{[a,t]})$ para $t \in (a,b]$ e $V(a) = 0$. Mostre que os pontos de continuidade de γ coincidem com os pontos de continuidade de V (*dica*: se V é contínua em $t \in [a,b]$, use a desigualdade $\big\|\gamma(t_1) - \gamma(t_2)\big\| \le \big|V(t_1) - V(t_2)\big|$ para mostrar que γ também é contínua em t. Se γ é contínua em t, aplique o resultado do Exercício 5 para a curva $\gamma|_{[t,b]}$ para concluir que V é contínua à direita em t; repita o raciocínio para a *reparametrização reversa* $[-b,-a] \ni s \mapsto \gamma(-s) \in I\!R^n$ de γ para concluir que V é contínua à esquerda em t).
7. (*toda curva contínua é reparametrização de uma curva parametrizada por comprimento de arco*) Seja $\gamma : [a,b] \to I\!R^n$ uma curva contínua e retificável. Mostre que existe um único par $(\sigma, \tilde\gamma)$, onde $\sigma : [a,b] \to [0,L]$ e $\tilde\gamma : [0,L] \to I\!R^n$ são aplicações satisfazendo as seguintes condições:
   - $\sigma(a) = 0$ e $\sigma(b) = L$;
   - σ é contínua e monótona;
   - $\gamma = \tilde\gamma \circ \sigma$;
   - $\tilde\gamma$ é parametrizada por comprimento de arco.

(*dica*: para mostrar a unicidade de σ, considere as igualdades:

$$\sigma(t) = \mathrm{V}(\tilde\gamma|_{[0,\sigma(t)]}) = \mathrm{V}(\tilde\gamma \circ \sigma|_{[a,t]}) = \mathrm{V}(\gamma|_{[a,t]}),$$

que mostram que σ deve coincidir com a função V definida no Exercício 6. Para a existência, defina $\sigma = V$ e mostre que:

(i) *$\sigma : [a,b] \to [0,L]$ é sobrejetora, onde $L = \mathrm{V}(\gamma)$;*
(ii) *para $s,t \in [a,b]$, $\sigma(s) = \sigma(t) \Rightarrow \gamma(s) = \gamma(t)$.*

Conclua que existe uma única aplicação $\tilde\gamma : [0,L] \to I\!R^n$ tal que $\tilde\gamma \circ \sigma = \gamma$, i.e., $\tilde\gamma$ deve preencher o lugar da flecha pontilhada no diagrama comutativo:

$$\begin{tikzcd} {[a,b]} \arrow[r, "\gamma"] \arrow[d, "\sigma"'] & I\!R^n \\ {[0,L]} \arrow[ur, dotted, "\tilde\gamma"'] & \end{tikzcd}$$

Use novamente a invariância do comprimento de arco por reparametrizações para concluir que $\tilde\gamma$ é de fato parametrizada por comprimento de arco).

8. Dadas curvas retificáveis $\gamma, \mu : [a,b] \to I\!R^n$, mostre que:
   (a) a soma $\gamma+\mu$ é retificável e $\mathrm{V}(\gamma+\mu) \le \mathrm{V}(\gamma)+\mathrm{V}(\mu)$ (*dica*: mostre que $\mathrm{V}(\gamma+\mu;P) \le \mathrm{V}(\gamma;P) + \mathrm{V}(\mu;P)$ para toda partição P de $[a,b]$);
   (b) se $c \in I\!R$ então a curva $c\gamma$ (definida por $(c\gamma)(t) = c\gamma(t)$) é retificável e $\mathrm{V}(c\gamma) = |c|\mathrm{V}(\gamma)$ (*dica*: mostre que $\mathrm{V}(c\gamma;P) = |c|\mathrm{V}(\gamma;P)$ para toda partição P de $[a,b]$);
   (c) o conjunto $\mathrm{Ret}([a,b], I\!R^n)$ formado por todas as curvas retificáveis $\gamma : [a,b] \to I\!R^n$ é um subespaço do espaço vetorial real de todas as curvas $\gamma : [a,b] \to I\!R^n$ e a aplicação $\gamma \mapsto \big\|\gamma(a)\big\| + \mathrm{V}(\gamma)$ define uma norma em $\mathrm{Ret}([a,b], I\!R^n)$;
   (d) se $\sigma : [a,b] \to [a,b]$ é crescente e sobrejetora então a aplicação

$$\mathrm{Ret}([a,b], I\!R^n) \ni \gamma \longmapsto \gamma \circ \sigma \in \mathrm{Ret}([a,b], I\!R^n)$$

   é uma isometria linear.

Aula número 17 (10/05)

A aula começa cobrindo as seções 2 e 3 da aula número 16 (08/05), sobre comprimento de arco e sua relação com a integral da norma do vetor tangente (para curvas de classe C^1 por partes).

(1) **Integral de um campo vetorial ao longo de uma curva**.

Definição. *Um* campo vetorial *em $\mathbb{R}^n$ (definido num subconjunto $S \subset \mathbb{R}^n$) é uma função $X : S \to \mathbb{R}^n$. Se $X : S \to \mathbb{R}^n$ é um campo vetorial contínuo (i.e., a aplicação $X : S \to \mathbb{R}^n$ é contínua) e se $\gamma : [a, b] \to \mathbb{R}^n$ é uma curva de classe C^1 por partes cuja imagem está contida em S então definimos* a integral de X ao longo de γ *através da integral:*

$$\int_\gamma X = \int_a^b \big\langle X\big(\gamma(t)\big), \gamma'(t)\big\rangle \,\mathrm{d}t,$$

onde $\langle \cdot, \cdot \rangle$ denota o produto interno canônico de $\mathbb{R}^n$.

Se $X = (X^1, \ldots, X^n)$ e $\gamma = (\gamma^1, \ldots, \gamma^n)$ com cada $X^i : S \to \mathbb{R}$ e cada $\gamma^i : [a, b] \to \mathbb{R}$ então a integral de X ao longo de γ pode ser escrita mais explicitamente na forma:

$$\int_\gamma X = \sum_{i=1}^n \int_a^b X^i\big(\gamma^1(t), \ldots, \gamma^n(t)\big)(\gamma^i)'(t) \,\mathrm{d}t.$$

`Observação`. Como γ é apenas de classe C^1 por partes, o integrando $\big\langle X\big(\gamma(t)\big), \gamma'(t)\big\rangle$ não está bem definido num número finito de instantes $t \in [a, b]$ — isso não é um problema, pois a integral de uma função não se altera quando alteramos o valor da função num número finito de pontos. Podemos então atribuir qualquer valor ao integrando em questão nos pontos t onde γ não é diferenciável.

Vamos começar mostrando a invariância da integral de linha por reparametrizações.

Teorema. *Sejam $X : S \subset \mathbb{R}^n \to \mathbb{R}^n$ um campo vetorial contínuo e $\gamma : [a, b] \to \mathbb{R}^n$ uma curva de classe C^1 por partes com $\gamma([a, b]) \subset S$. Dada uma função injetora $\sigma : [c, d] \to [a, b]$ de classe C^1 por partes, então:*

$$\int_{\gamma\circ\sigma} X = \int_\gamma X, \qquad \text{se } \sigma(c) = a \text{ e } \sigma(d) = b,$$

$$\int_{\gamma\circ\sigma} X = -\int_\gamma X, \quad \text{se } \sigma(c) = b \text{ e } \sigma(d) = a.$$

Demonstração. Seja $P = \{t_0, \ldots, t_k\}$ uma partição de $[a, b]$ tal que $\gamma|_{[t_i, t_{i+1}]}$ é de classe C^1 para $i = 0, \ldots, k-1$ e seja $Q = \{s_0, \ldots, s_l\}$ uma partição de $[c, d]$ tal que $\sigma|_{[s_i, s_{i+1}]}$ é de classe C^1 para $i = 0, \ldots, l-1$. Temos que $Q \cup \sigma^{-1}(P) = \{u_0, \ldots, u_m\}$ é uma partição de $[a, b]$ (usamos aqui que σ é injetora!) e é fácil ver que $(\gamma \circ \sigma)|_{[u_i, u_{i+1}]}$ é de classe C^1

para $i = 0, \ldots, m-1$; daí $\gamma \circ \sigma$ também é de classe C^1 por partes e faz realmente sentido considerar $\int_{\gamma\circ\sigma} X$. Temos:

$$\int_{\gamma\circ\sigma} X = \int_c^d \big\langle X\big(\gamma(\sigma(t))\big), (\gamma\circ\sigma)'(t)\big\rangle \, \mathrm{d}t;$$

exceto para um número finitos de instantes $t \in [c, d]$ (mais precisamente, exceto para $t \in \{u_0, \ldots, u_m\}$) temos $(\gamma\circ\sigma)'(t) = \gamma'\big(\sigma(t)\big)\sigma'(t)$ e portanto:

$$\int_{\gamma\circ\sigma} X = \int_c^d \big\langle X\big(\gamma(\sigma(t))\big), \gamma'\big(\sigma(t)\big)\big\rangle \, \sigma'(t) \, \mathrm{d}t = \sum_{i=0}^{m-1} \int_{u_i}^{u_{i+1}} \big\langle X\big(\gamma(\sigma(t))\big), \gamma'\big(\sigma(t)\big)\big\rangle \, \sigma'(t) \, \mathrm{d}t.$$

Fazendo a mudança de variável $\sigma(t) = x$ nas integrais acima obtemos:

$$\int_{\gamma\circ\sigma} X = \sum_{i=0}^{m-1} \int_{\sigma(u_i)}^{\sigma(u_{i+1})} \big\langle X\big(\gamma(x)\big), \gamma'(x)\big\rangle \, \mathrm{d}x = \int_{\sigma(c)}^{\sigma(d)} \big\langle X\big(\gamma(x)\big), \gamma'(x)\big\rangle \, \mathrm{d}x,$$

e a última integral vale $\pm \int_\gamma X$, dependendo do valor de σ nas extremidades c e d. ■

`Observação`. A hipótese que $\sigma : [c, d] \to [a, b]$ é injetora no teorema acima é usada só para concluir que $\sigma^{-1}(P)$ é finito. Poderíamos alternativamente supor apenas que para todo $t \in [a, b]$ no qual γ não é diferenciável temos que o conjunto $\sigma^{-1}(t) \subset [c, d]$ é finito. Observe que sem a hipótese de injetividade para σ, é bem possível que σ não seja monótona; isso corresponde à idéia intuitiva que a reparametrização $\gamma \circ \sigma$ de γ pode fazer "um pouco de zigue-zague" que a curva γ original não fazia. Desde que $\sigma(c) = a$ e $\sigma(d) = b$, i.e., desde que $\gamma \circ \sigma$ "percorra todo o percurso de $\gamma(a)$ a $\gamma(b)$", a integral de linha $\int_\gamma X$ não se altera; na prática, o que ocorre é que trechos de γ percorridos a primeira vez num certo sentido e uma segunda vez em outro sentido se cancelam. Note bem que *esse fenômeno não ocorre no caso do comprimento de arco*! No caso do comprimento de arco os "zigue-zagues" *aumentam o comprimento total*, de modo que a hipótese de monotonicidade de σ é essencial.

`Observação`. Muitos livros elementares de Cálculo e de Física–Matemática usam notações como $\int_\gamma X \cdot \mathrm{d}\vec{r}$ em vez de $\int_\gamma X$. A motivação para essa notação é a seguinte: pensa-se em $\vec{r}$ como o "vetor posição da partícula que viaja ao longo de γ" e daí $\mathrm{d}\vec{r}$ é uma espécie de "vetor infinitesimal de variação de posição". Escreve-se então $X \cdot \mathrm{d}\vec{r}$ para denotar o produto escalar (i.e., o produto interno) do campo X pela "variação infinitesimal" $\mathrm{d}\vec{r}$.

`Observação`. A interpretação física para a integral de linha definida acima (quando $n \le 3$) é a seguinte: pensamos em X como um *campo de forças*, i.e., em cada ponto $x \in I\!R^n$ temos um "vetor força" $X(x)$. A curva γ corresponde à trajetória de uma partícula (o parâmetro t de γ é o tempo). A integral de linha $\int_\gamma X$ corresponde então ao *trabalho* da força X sobre a partícula que descreve a trajetória γ. O trabalho de uma força constante $\vec{F} \in I\!R^3$ sobre uma partícula que descreve um movimento retilínio do ponto $p \in I\!R^3$ ao ponto $q \in I\!R^3$ é dada por $\langle \vec{F}, q - p\rangle$ (pois a *componente da força $\vec{F}$ normal ao movimento não realiza trabalho*). A integral de linha corresponde a uma passagem ao limite dessa idéia: escolhemos uma partição $P = \{t_0, \dots, t_k\}$ do intervalo $[a, b]$ com norma pequena e instantes $\tau_i \in [t_i, t_{i+1}]$, $i = 0, \dots, k-1$. Num "intervalo pequeno" $[t_i, t_{i+1}]$ aproximamos o movimento da partícula por um segmento de reta e a força dada pelo campo X por uma força constante; o trabalho realizado por X do instante t_i ao instante t_{i+1} é aproximado então pelo produto $\big\langle X\big(\gamma(\tau_i)\big), \gamma(t_{i+1}) - \gamma(t_i)\big\rangle$. Agora aproximamos o vetor $\gamma(t_{i+1}) - \gamma(t_i)$ por $\gamma'(\tau_i)(t_{i+1} - t_i)$ e somamos em $i = 0, \dots, k-1$. Fazendo o limite quando $\|P\| \to 0$, obtemos a integral de linha $\int_\gamma X$ (veja a observação a seguir).

`Observação`. É muito comum definir a integral de linha $\int_\gamma X$ usando limites de somas da forma $\sum_{i=0}^{k-1} \big\langle X\big(\gamma(\tau_i)\big), \gamma(t_{i+1}) - \gamma(t_i)\big\rangle$ quando a norma da partição $P = \{t_0, \dots, t_k\}$ de $[a, b]$ tende a zero e $\{\tau_0, \dots, \tau_{k-1}\}$ é uma pontilhação de P. Mostra-se então que esse limite de somas coincide com a integral de Riemann $\int_a^b \big\langle X\big(\gamma(t)\big), \gamma'(t)\big\rangle \, \mathrm{d}t$ no caso que X é contínuo e γ é de classe C^1 por partes. Essa definição da integral de linha usando limites de somas é similar à definição da *integral de Riemann–Stieltjes* que não abordaremos nesse curso.
[para quem conhece a definição de integral de Riemann–Stieltjes, observamos que se $X = (X^1, \dots, X^n)$ e $\gamma = (\gamma^1, \dots, \gamma^n)$ então a integral de linha $\int_\gamma X$ coincide com a soma $\sum_{i=1}^n \int_a^b X^i\big(\gamma(t)\big) \, \mathrm{d}\gamma^i(t)$, onde $\int_a^b X^i\big(\gamma(t)\big) \, \mathrm{d}\gamma^i(t)$ denota a integral de Riemann–Stieltjes de $f = X^i \circ \gamma : [a, b] \to I\!R$ com respeito à função $\alpha = \gamma^i : [a, b] \to I\!R$. Os curiosos podem consultar o *Curso de Análise vol. II* do Elon Lages Lima, pgs. 193–204.]
Nós decidimos usar um caminho mais curto para simplificar a exposição: definir $\int_\gamma X$ *apenas* quando X é contínuo e γ é de classe C^1 por partes, usando diretamente a integral de Riemann. Na prática, raramente se usa a integral $\int_\gamma X$ em condições mais gerais do que a que consideramos. Nos Exercícios 10–15 apresentamos um roteiro para os interessados na definição de $\int_\gamma X$ como limite de somas.

`Observação`. Como mostramos acima, a integral de linha $\int_\gamma X$ é invariante por reparametrizações, mas apenas por reparametrizações que "preservam o sentido de percurso de γ" — se a reparametrização "inverte o sentido do percurso", o valor de $\int_\gamma X$ troca de sinal. Por esse motivo diz-se às vezes que $\int_\gamma X$ é uma *integral orientada*. Observe que, por outro lado, o comprimento de arco $\mathrm{V}(\gamma)$ é invariante *tanto por reparametrizações crescentes como por reparametrizações decrescentes* (não há mudança de sinal). O comprimento de arco é muitas vezes denotado em textos elementares de Cálculo e em textos de Física–Matemática por $\int_\gamma \mathrm{d}s$ ou por $\int_\gamma \|\mathrm{d}\vec{r}\|$; essa notação é motivada pela idéia que $\mathrm{d}s$ (ou $\|\mathrm{d}\vec{r}\|$) denotam o comprimento de uma "porção infinitesimal" de γ. Considera-se também "integrais" da forma $\int_a^b f \, \mathrm{d}s$ (denotadas também por $\int_a^b f \, \|\mathrm{d}\vec{r}\|$), onde $f : [a, b] \to I\!R$ é uma função escalar. Tais integrais podem ser definidas como limites de somas da forma $\sum_{i=0}^{k-1} f(\tau_i)\big\|\gamma(t_{i+1}) - \gamma(t_i)\big\|$ quando a norma da partição $P = \{t_0, \dots, t_k\}$ de $[a, b]$ tende a zero e $\{\tau_0, \dots, \tau_{k-1}\}$ é uma pontilhação de P. Mostra-se então que, se f é contínua e γ é de classe C^1 por partes, então $\int_\gamma f \, \mathrm{d}s$ coincide com a integral de Riemann $\int_a^b f(t)\big\|\gamma'(t)\big\| \, \mathrm{d}t$ (ou, para simplificar a exposição, alguns autores definem $\int_\gamma f \, \mathrm{d}s$ apenas quando f é contínua e γ é de classe C^1 por partes, usando diretamente a integral de Riemann $\int_a^b f(t)\big\|\gamma'(t)\big\| \, \mathrm{d}t$). Essa integral $\int_\gamma f \, \mathrm{d}s$ corresponde fisicamente, por exemplo, à *massa do "fio" γ cuja densidade linear de* *massa é expressa pela função* f (o comprimento de arco é reobtido quando $f \equiv 1$). A integral $\int_\gamma f\,\mathrm{d}s$, como o comprimento de arco, é invariante tanto por reparametrizações crescentes como por reparametrizações decrescentes, não havendo mudança de sinal; diz-se então que tais integrais são *não-orientadas*. Alguns autores usam também a terminologia *integral de linha de primeiro tipo* para integrais da forma $\int_\gamma f\,\mathrm{d}s$ e *integral de linha de segundo tipo* para integrais da forma $\int_\gamma X$.

Para finalizar, vamos enunciar algumas propriedades elementares da integral de linha cuja demonstração é imediata a partir de propriedades correspondentes da integral de Riemann.

Teorema. *Sejam* $X, Y : S \subset I\!R^n \to I\!R^n$ *campos vetoriais contínuos e* $\gamma : [a,b] \to I\!R^n$ *uma curva de classe* C^1 *por partes. Então:*

(a) $\int_\gamma X + Y = \int_\gamma X + \int_\gamma Y$;

(b) $\int_\gamma kX = k\int_\gamma X$, *para toda constante* $k \in I\!R$;

(c) $\int_\gamma X = \int_{\gamma|_{[a,c]}} X + \int_{\gamma|_{[c,b]}} X$, *para todo* c *no intervalo aberto* (a,b);

(d) $\left|\int_\gamma X\right| \le \sup_{t\in[a,b]} \big\|X\big(\gamma(t)\big)\big\|\,\mathrm{V}(\gamma)$.

Observe que a demonstração do item (d) do teorema acima segue trivialmente usando a desigualdade de Cauchy–Schwarz e a fórmula $\mathrm{V}(\gamma) = \int_a^b \big\|\gamma'(t)\big\|\,\mathrm{d}t$ para o comprimento de arco $\mathrm{V}(\gamma)$.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Funções de variação limitada`

0. Seja $f : [a, b] \to \mathbb{R}$ uma função de variação limitada. Defina $V : [a, b] \to \mathbb{R}$ fazendo $V(t) = \mathrm{V}(f|_{[a,t]})$ para $t \in (a, b]$ e $V(a) = 0$. Mostre que V e $V - f$ são funções crescentes (*dica*: $\big|f(x) - f(y)\big| \le \big|V(x) - V(y)\big|$). Conclua que uma função $g : [a, b] \to \mathbb{R}$ é de variação limitada se e somente se pode ser escrita como diferença de duas funções crescentes (*dica*: use os Exercícios 2 e 8(a) da Aula número 16 — 08/05).

`Recordação das propriedades elementares da integral de Riemann`.

Os exercícios que aparecem aqui são para aqueles que querem apreender a demonstrar algumas propriedades básicas da integral de Riemann definida em termos de limites de somas de Riemann. Alguns exercícios não são tão fáceis, mas todos esses resultados podem ser encontrados no *Curso de Análise, vol. I*, do Elon Lages Lima (capítulo sobre integração). Avisamos no entanto que o Elon utiliza uma definição diferente de integral de Riemann (em termos de integrais inferiores e superiores), mas depois ele mostra que essa definição diferente é de fato equivalente àquela adotada aqui.

1. Sejam $f, g : [a, b] \to \mathbb{R}$ funções limitadas, $k \in \mathbb{R}$ um número real e (P, τ) uma partição pontilhada de $[a, b]$. Mostre que:

$$S(f + g; P, \tau) = S(f; P, \tau) + S(g; P, \tau), \quad S(kf; P, \tau) = kS(f; P, \tau),$$

e que se $f(t) \le g(t)$ para todo $t \in [a, b]$ então $S(f; P, \tau) \le S(g; P, \tau)$. Conclua que se $f, g : [a, b] \to \mathbb{R}$ são Riemann integráveis então $f + g$ e kf também o são e:

$$\int_a^b f + g = \int_a^b f + \int_a^b g, \quad \int_a^b kf = k \int_a^b f;$$

além do mais, se $f(t) \le g(t)$ para todo $t \in [a, b]$, mostre que $\int_a^b f \le \int_a^b g$. Conclua que se $|f|$ também é Riemann integrável então $\left|\int_a^b f\right| \le \int_a^b |f|$ (*dica*: $-|f| \le f \le |f|$).

`Observação`. Na verdade, se f é integrável então $|f|$ é automaticamente integrável (veja observação mais adiante sobre a condição necessária e suficiente para a Riemann integrabilidade em termos do conjunto dos pontos de descontinuidade da função).

2. (*critério de Cauchy para integrabilidade*) Mostre que uma função limitada $f : [a, b] \to \mathbb{R}$ é Riemann integrável se e somente se vale a seguinte propriedade: dado $\varepsilon > 0$, existe $\delta > 0$ tal que dadas partições pontilhadas arbitrárias (P, τ) e (P', τ') de $[a, b]$ com $\|P\| < \delta$ e $\|P'\| < \delta$ então $\big|S(f; P, \tau) - S(f; P', \tau')\big| < \varepsilon$ (*dica*: para mostrar que f é Riemann integrável, escolha uma seqüência arbitrária (P_n, τ_n) de partições pontilhadas com $\|P_n\| \to 0$ e mostre que a seqüência $S(f; P_n, \tau_n)$ é de Cauchy em $\mathbb{R}$; mostre que o limite $I \in \mathbb{R}$ dessa seqüência é igual à integral $\int_a^b f$).

3. Se $f : [a,b] \to I\!R$ é Riemann integrável e $[c,d] \subset [a,b]$ é um subintervalo, mostre que $f|_{[c,d]}$ é Riemann integrável (*dica*: se (P,τ) é uma partição pontilhada de $[c,d]$ então $\widetilde{P} = P \cup \{a,b\}$ e $\widetilde{\tau} = \tau \cup \{a,b\}$ definem uma partição pontilhada $(\widetilde{P},\widetilde{\tau})$ de $[a,b]$; além do mais, se (P,τ) e (P',τ') são partições pontilhadas de $[c,d]$, $\widetilde{P} = P \cup \{a,b\}$, $\widetilde{P}' = P' \cup \{a,b\}$, $\widetilde{\tau} = \tau \cup \{a,b\}$ e $\widetilde{\tau}' = \tau' \cup \{a,b\}$ então:
$$S(f;\widetilde{P},\widetilde{\tau}) - S(f;\widetilde{P}',\widetilde{\tau}') = S(f|_{[c,d]};P,\tau) - S(f|_{[c,d]};P',\tau').$$
Use o critério de Cauchy explicado no Exercício 2).
4. Sejam $f : [a,b] \to I\!R$ uma função e c um ponto do intervalo aberto (a,b), de modo que $f|_{[a,c]}$ e $f|_{[c,b]}$ sejam Riemann integráveis. Mostre que f também é Riemann integrável e que:
$$\int_a^b f = \int_a^c f + \int_c^b f.$$
[*dica*: se (P,τ) é uma partição pontilhada de $[a,b]$ com $c \in P$ então podemos de maneira óbvia definir partições pontilhadas (P_1,τ_1) e (P_2,τ_2) de $[a,c]$ e de $[c,b]$ respectivamente (com $P_1 = P \cap [a,c]$ e $P_2 = P \cap [c,b]$) de modo que $S(f;P,\tau) = S(f|_{[a,c]};P_1,\tau_1) + S(f|_{[c,b]};P_2,\tau_2)$. Se P é uma partição de $[a,b]$ que não contém c e se $\|P\| < \delta$ então $P \cup \{c\}$ é uma partição de $[a,b]$ que contém c e $\big|S(f;P,\tau) - S(f;P\cup\{c\},\tau')\big| < 2M\delta$, onde $M = \sup_{t\in[a,b]} \big|f(t)\big|$ e τ' é uma pontilhação adequada (definida de maneira óbvia) de $P \cup \{c\}$].
5. Mostre que se $f : [a,b] \to I\!R$ é nula fora de um subconjunto finito de $[a,b]$ então f é Riemann integrável e $\int_a^b f = 0$ (*dica*: se o conjunto dos pontos de $[a,b]$ onde f não é nula possui k elementos e (P,τ) é uma partição pontilhada de $[a,b]$ com $\|P\| < \delta$ então $\big|S(f;P,\tau)\big| \le 2kM\delta$, onde $M = \sup_{t\in[a,b]} \big|f(t)\big|$). Conclua que se $f : [a,b] \to I\!R$ é Riemann integrável e $g : [a,b] \to I\!R$ difere de f só num conjunto finito então g também é Riemann integrável e $\int_a^b g = \int_a^b f$.
6. Mostre que toda função contínua $f : [a,b] \to I\!R$ é Riemann integrável (*dica*: use o Critério de Cauchy da seguinte forma: se P e P' são partições de $[a,b]$ então $P \cup P'$ é uma partição que refina P e P' simultaneamente. É suficiente então para aplicar o critério de Cauchy estimar o valor de $\big|S(f;P,\tau) - S(f;P',\tau')\big|$ com $\|P\| < \delta$, onde (P,τ), (P',τ') são partições pontilhadas de $[a,b]$ e P' refina P, i.e., $P' \supset P$. Use o fato que f é uniformemente contínua e limitada para escolher $\delta > 0$).
7. (*o Teorema Fundamental do Cálculo*) Mostre que se $f : [a,b] \to I\!R$ é uma função Riemann integrável e se $F : [a,b] \to I\!R$ é definida por $F(t) = \int_a^t f$ (com $F(a) = 0$) então F é contínua e, para todo $t \in [a,b]$ onde f é contínua, temos que F é derivável em t e $F'(t) = f(t)$. Conclua que se f é contínua então existe uma função $G : [a,b] \to I\!R$ de classe C^1 com $G' = f$ e, dada uma tal função, temos $\int_a^b f = G(b) - G(a)$.
8. Mostre que se $F : [a,b] \to I\!R$ é uma função de classe C^1 por partes então $\int_a^b F' = F(b) - F(a)$, onde atribuímos um valor arbitrário ao integrando no número finito de instantes onde F não é derivável (*dica*: use os Exercícios 4, 5 e 7). Usando a convenção

$\int_p^p f = 0$ e $\int_p^q f = -\int_q^p f$ para $p > q$, mostre que $\int_p^q F' = F(q) - F(p)$ para todos $p, q \in [a, b]$.

9. (*mudança de variáveis na integral de Riemann*) Se $f : [a, b] \to I\!R$ é uma função contínua e $g : [c, d] \to I\!R$ é uma função de classe C^1 por partes cuja imagem está contida em $[a, b]$ então:

$$\int_{g(c)}^{g(d)} f(t)\,\mathrm{d}t = \int_c^d f\big(g(x)\big)g'(x)\,\mathrm{d}x,$$

onde, como sempre, substituímos $g'(x)$ por um valor arbitrário no número finito de pontos $x \in [c, d]$ onde g não é derivável (*dica*: seja $F : [a, b] \to I\!R$ uma função tal que $F' = f$ e aplique o teorema fundamental do cálculo para a função composta $F \circ g : [c, d] \to I\!R$).

`Observação`. Sabe-se que uma função limitada $f : [a, b] \to I\!R$ é Riemann integrável se e somente se *o conjunto dos pontos onde f é descontínua possui medida de Lebesgue zero*. Para os interessados em mais detalhes, consultem o *Curso de Análise, vol. I* do Elon Lages Lima (capítulo sobre integração).

`A integral de linha como limite de somas.`

Esta série de Exercícios é para aqueles que desejam estudar a noção de integral de linha num contexto mais geral (i.e., para curvas que podem não ser de classe C^1 por partes e campos que podem não ser contínuos). Mais detalhes podem ser encontrados no *Curso de Análise, vol. II* do Elon Lages Lima (capítulo sobre integrais curvilíneas). Antes de começar a série de Exercícios, apresentamos a definição geral de integrais de linha da forma $\int_\gamma X$.

Sejam $X : S \subset I\!R^n \to I\!R^n$ um campo vetorial arbitrário e $\gamma : [a, b] \to I\!R^n$ uma curva arbitrária cuja imagem está contida em S. Se $P = \{t_0, \ldots, t_k\}$ é uma partição de $[a, b]$ e $\tau = \{\tau_0, \ldots, \tau_{k-1}\}$ é uma pontilhação para P então definimos a *soma de Riemann* $S(X, \gamma; P, \tau)$ fazendo:

$$S(X, \gamma; P, \tau) = \sum_{i=0}^{k-1} \big\langle X\big(\gamma(\tau_i)\big), \gamma(t_{i+1}) - \gamma(t_i)\big\rangle.$$

A *integral de X ao longo de γ* é definida como sendo o número real $I \in I\!R$ (se existir) tal que dado $\varepsilon > 0$, existe $\delta > 0$ tal que $\big|S(X, \gamma; P, \tau) - I\big| < \varepsilon$, para toda partição pontilhada (P, τ) de $[a, b]$ com $\|P\| < \delta$. É fácil ver que I é único se existir; escrevemos $I = \int_\gamma X$.

10. Mostre que se $X, Y : S \subset I\!R^n \to I\!R^n$ são campos vetoriais, $\gamma : [a, b] \to I\!R^n$ é uma curva com imagem contida em S e se $k \in I\!R$ é um número real então, se as integrais $\int_\gamma X$ e $\int_\gamma Y$ existem então também as integrais $\int_\gamma X + Y$ e $\int_\gamma kX$ existem e valem as identidades:

$$\int_\gamma X + Y = \int_\gamma X + \int_\gamma Y, \quad \int_\gamma kX = k \int_\gamma X.$$

(*dica*: a solução é quase igual a do Exercício 1).

11. (*critério de Cauchy para integral de linha*) Sejam $X : S \subset \mathbb{R}^n \to \mathbb{R}^n$ um campo vetorial e $\gamma : [a, b] \to \mathbb{R}^n$ uma curva cuja imagem está contida em S. Mostre que a integral $\int_\gamma X$ existe se e somente se vale a seguinte propriedade: dado $\varepsilon > 0$, existe $\delta > 0$ tal que dadas partições pontilhadas (P, τ), (P', τ') arbitrárias de $[a, b]$ com $\|P\| < \delta$, $\|P'\| < \delta$ então $\big|S(X, \gamma; P, \tau) - S(X, \gamma; P', \tau')\big| < \varepsilon$ (*dica*: a solução é quase igual a do Exercício 2).

12. Se $X : S \subset \mathbb{R}^n \to \mathbb{R}^n$ é um campo vetorial e $\gamma : [a, b] \to \mathbb{R}^n$ é uma curva cuja imagem está contida em S então, se a integral $\int_\gamma X$ existe então a integral $\int_{\gamma|_{[c,d]}} X$ também existe para todo subintervalo $[c, d] \subset [a, b]$ (*dica*: a solução é quase igual a do Exercício 3). Mostre também que se c é um ponto do intervalo aberto (a, b) então vale a identidade:
$$\int_\gamma X = \int_{\gamma|_{[a,c]}} X + \int_{\gamma|_{[c,b]}} X.$$
(*dica*: a solução é quase igual a do Exercício 4, mas como já estamos supondo que $\int_\gamma X$ existe, é suficiente considerar partições P de $[a, b]$ que contêm c. Observamos que em geral não é possível concluir que $\int_\gamma X$ existe a partir da existência de $\int_{\gamma|_{[a,c]}} X$ e de $\int_{\gamma|_{[c,b]}} X$; veja o *Curso de Análise vol. II* do Elon Lages Lima, pg. 236, Exercício 2.3 para um contra-exemplo no contexto da integral de Riemann-Stieltjes).

13. (*invariância por reparametrização*) Sejam $X : S \subset \mathbb{R}^n \to \mathbb{R}^n$ um campo vetorial e $\gamma : [a, b] \to \mathbb{R}^n$ uma curva arbitrária com imagem contida em S. Mostre que se a integral $\int_\gamma X$ existe e $\sigma : [c, d] \to [a, b]$ é uma função monótona e sobrejetora então a integral $\int_{\gamma\circ\sigma} X$ também existe e vale a igualdade:
$$\int_{\gamma\circ\sigma} X = \int_\gamma X, \qquad \text{se } \sigma \text{ é crescente},$$
$$\int_{\gamma\circ\sigma} X = -\int_\gamma X, \quad \text{se } \sigma \text{ é decrescente}.$$
[*dica*: use o fato que uma função monótona num intervalo é contínua se sua imagem também for um intervalo (veja *Curso de Análise, vol. I*, Elon Lages Lima, pg. 182). Daí σ é contínua e portanto uniformemente contínua. Mostre que se (P, τ) é uma partição pontilhada de $[c, d]$ então $P' = \sigma(P)$ é uma partição de $[a, b]$ e podemos escolher uma pontilhação τ' de P' (de maneira óbvia) de modo que $S(X, \gamma\circ\sigma; P, \tau) = \pm S(X, \gamma; P', \tau')$, sendo que o sinal $+$ aparece para σ crescente e o sinal $-$ aparece para σ decrescente (o argumento é similar ao usado na demonstração da invariância do comprimento de arco por reparametrizações, mas aqui aparece um sinal de $-$ no caso decrescente). Para concluir, use a continuidade uniforme de σ para mostrar que $\|P\| \to 0$ implica $\|P'\| \to 0$, i.e., que dado $\delta' > 0$, existe $\delta > 0$ de modo que $\|P\| < \delta$ implica $\big\|\sigma(P)\big\| < \delta'$].

14. Mostre que se $X : S \subset \mathbb{R}^n \to \mathbb{R}^n$ é um campo vetorial contínuo e $\gamma : [a,b] \to \mathbb{R}^n$ é uma curva contínua e retificável com imagem contida em S então a integral $\int_\gamma X$ existe e vale a desigualdade:

$$\left| \int_\gamma X \right| \le \sup_{t\in[a,b]} \big\| X\big(\gamma(t)\big) \big\| \, \mathrm{V}(\gamma).$$

(*dica*: a idéia da prova é similar a do Exercício 6).

15. Mostre que se $X : S \subset \mathbb{R}^n \to \mathbb{R}^n$ é um campo vetorial contínuo e $\gamma : [a,b] \to \mathbb{R}^n$ é uma curva de classe C^1 por partes cuja imagem está contida em S então:

$$\int_\gamma X = \int_a^b \big\langle X\big(\gamma(t)\big), \gamma'(t) \big\rangle \, \mathrm{d}t.$$

[*dica*: primeiro reduza o caso geral ao caso que γ é de classe C^1. Lembre que você *já sabe* que as integrais $\int_\gamma X$ e $\int_a^b \big\langle X\big(\gamma(t)\big), \gamma'(t)\big\rangle \, \mathrm{d}t$ existem, de modo que você não precisa se preocupar em usar partições pontilhadas arbitrárias para mostrar a existência das integrais. Se $P = \{t_0, \ldots, t_k\}$ é uma partição de $[a,b]$ de norma pequena e se $\tau = \{\tau_0, \ldots, \tau_{k-1}\}$ é uma pontilhação de P, use (para cada $i = 0, \ldots, k-1$) a desigualdade do valor médio para a função $t \mapsto \gamma(t) - \gamma'(\tau_i)t$ no intervalo $[t_i, t_{i+1}]$ para obter uma estimativa para o erro ao aproximar $\gamma(t_{i+1}) - \gamma(t_i)$ por $\gamma'(\tau_i)(t_{i+1} - t_i)$. Use o fato que $X \circ \gamma : [a,b] \to \mathbb{R}^n$ é limitado e o fato que $\gamma' : [a,b] \to \mathbb{R}^n$ é uniformemente contínua para concluir a demonstração].

Aula número 18 (15/05)

A aula começa cobrindo parte da demonstração da fórmula "$\mathrm{V}(\gamma) = \int_a^b \|\gamma'(t)\| \, \mathrm{d}t$" (para γ de classe C^1 por partes) e todo o material originalmente planejado para a aula número 17 (10/05).

(1) **Mais álgebra linear: o espaço dual**.

Definição. *Seja V um espaço vetorial real. O* espaço dual *de V é o espaço $V^* = \mathrm{Lin}(V, I\!R)$ formado por todos os funcionais lineares em V, i.e., por todas as aplicações lineares $T : V \to I\!R$.*

Os elementos de V^* são às vezes chamados também de *co-vetores*.

A teoria do espaço dual é extremamente simplificada quando V tem dimensão finita. *Suporemos então a partir de agora que todos os espaços vetoriais considerados são reais e têm dimensão finita.*

Se $\mathfrak{B} = (b_i)_{i=1}^n$ é uma base para V então para cada $i = 1, \ldots, n$, denotamos por $b_i^* : V \to I\!R$ a aplicação que associa a cada vetor $v \in V$ a sua i-ésima coordenada na base $\mathfrak{B}$, ou seja:

$$v = \sum_{i=1}^n b_i^*(v) b_i,$$

para todo $v \in V$. É fácil ver que b_i^* é linear e portanto define um elemento do espaço dual V^*. Além do mais, temos:

$$b_i^*(b_j) = \begin{cases} 0, & i \neq j, \\ 1, & i = j, \end{cases}$$

para todos $i, j = 1, \ldots, n$. Temos o seguinte:

Teorema. *Se $\mathfrak{B} = (b_i)_{i=1}^n$ é uma base de V então a família $\mathfrak{B}^* = (b_i^*)_{i=1}^n$ é uma base de V^*.*

Demonstração. Se $\sum_{i=1}^n k_i b_i^* = 0$ com cada $k_i \in I\!R$ então, avaliando os dois lados dessa igualdade em b_j obtemos $k_j = 0$ e portanto $\mathfrak{B}^*$ é linearmente independente. Dado agora $\alpha \in V^*$ temos:

$$\alpha(v) = \alpha\left(\sum_{i=1}^n b_i^*(v) b_i\right) = \left(\sum_{i=1}^n \alpha(b_i) b_i^*\right)(v),$$

para todo $v \in V$ e portanto $\alpha = \sum_{i=1}^n \alpha(b_i) b_i^*$. A conclusão segue. ■

Definição. *A base $\mathfrak{B}^*$ de V^* definida acima é chamada a* base dual *correspondente a $\mathfrak{B}$.*

`Observação`. A demonstração acima nos dá uma informação adicional importante: *a i-ésima coordenada de um funcional $\alpha \in V^*$ na base dual $\mathfrak{B}^*$ é exatamente o valor de α no i-ésimo vetor da base original $\mathfrak{B}$.*

Passamos agora a estudar o *espaço bidual* de V, i.e., o dual do dual de V; vamos denotá-lo por V^{**}. A cada $v \in V$, podemos associar canonicamente um funcional linear em V^* que é a *aplicação de avaliação em* v; vamos denotá-lo temporariamente por $\hat{v}$. Temos:

$$\hat{v}(\alpha) = \alpha(v),$$

para todo $\alpha \in V^*$. É fácil ver que $\hat{v}$ é realmente linear e portanto define um elemento de V^{**}. Além do mais, a aplicação:

$$V \ni v \longmapsto \hat{v} \in V^{**}$$

é linear. Vamos mostrar que tal aplicação é um isomorfismo. Em primeiro lugar, se $\hat{v} = 0$ então $\hat{v}(\alpha) = \alpha(v) = 0$ para todo $\alpha \in V^*$ e portanto $v = 0$ (se $v \neq 0$, sempre podemos encontrar $\alpha \in V^*$ com $\alpha(v) \neq 0$ — por exemplo, $b_i^*(v) \neq 0$ para algum $i = 1, \ldots, n$). Mostramos então que $v \mapsto \hat{v}$ é injetora. Como $\dim(V^{**}) = \dim(V^*) = \dim(V)$, demonstramos o seguinte:

Teorema. *A aplicação $V \ni v \mapsto \hat{v} \in V^{**}$ é um isomorfismo.* ■

É muito comum usar o isomorfismo $v \mapsto \hat{v}$ para *identificar* V com seu bidual V^{**}, i.e., denota-se o funcional de avaliação em v com o próprio símbolo v, em vez de $\hat{v}$. Essa identificação corresponde intuitivamente à seguinte idéia: se $v \in V$ e $\alpha \in V^*$ então "não distinguimos entre aplicar α em v ou aplicar v em α", i.e., escrevemos $\alpha(v) = v(\alpha)$; o vetor v identifica-se com a operação $\hat{v}$ que leva α em $\alpha(v)$.

`Observação`. A existência de um isomorfismo canônico $V \cong V^{**}$ num certo sentido explica a terminologia "espaço dual". Em geral, diz-se que "dois tipos de objetos são duais" quando "obtem-se um escalar ao juntar os dois". A idéia de identificar V com V^{**} corresponde em pensar na operação de juntar v e α como uma operação na qual v e α tem um papel de caráter similar, i.e., não se distingue o vetor da função.

`Observação`. Se $\mathfrak{B} = (b_i)_{i=1}^n$ é uma base de V então, como já observamos, as coordenadas de um funcional $\alpha \in V^*$ na base dual $\mathfrak{B}^* = (b_i^*)_{i=1}^n$ são exatamente as avaliações de α nos vetores b_i de $\mathfrak{B}$. Uma releitura desse fato nos diz o seguinte: a *base bidual* associada a $\mathfrak{B}$, i.e., a base $\mathfrak{B}^{**} = (b_i^{**})_{i=1}^n$ dual de $\mathfrak{B}^*$ coincide com a base original $\mathfrak{B}$ de V, se identificarmos V com V^{**} da maneira descrita acima.

Assim como o espaço bidual V^{**}, é claro que o espaço dual V^* também é isomorfo a V, já que $\dim(V) = \dim(V^*)$; por exemplo, se $\mathfrak{B} = (b_i)_{i=1}^n$ é uma base de V então podemos definir um isomorfismo de V sobre V^* que leva b_i em b_i^* para todo $i = 1, \ldots, n$. Ocorre no entanto uma diferença importante entre as afirmações "V é isomorfo a V^*" e "V é isomorfo a V^{**}": o isomorfismo que construímos entre V e V^* *não é canônico*, i.e., *depende da base $\mathfrak{B}$ escolhida para defini-lo*, enquanto que a definição do isomorfismo $V \cong V^{**}$ não depende de escolha alguma.

`Observação`. A noção de *isomorfismo canônico* pode ser formalizada dentro do contexto da *teoria das categorias*; na terminologia técnica de tal teoria diria-se que $V \mapsto V$ e $V \mapsto V^{**}$ são *funtores naturalmente isomorfos*. O diagrama comutativo que aparece no Exercício 7 fornece justamente a definição técnica de naturalidade de transformações entre funtores na teoria das categorias!

Para finalizar, observamos que *a escolha de um produto interno em V determina um isomorfismo de V sobre V^**. Em geral, se $B : V \times V \to I\!R$ é uma forma bilinear, sabemos que a aplicação $T : V \to \mathrm{Lin}(V, I\!R) = V^*$ definida por $T(v)(w) = B(v, w)$ é linear. Denotaremos o funcional linear $T(v)$ por $B(v, \cdot)$. Temos o seguinte:

Teorema. *Se $\langle\cdot,\cdot\rangle$ é um produto interno em V então a aplicação linear:*

$$V \ni v \longmapsto \langle v, \cdot\rangle \in V^*$$

correspondente à forma bilinear $\langle\cdot,\cdot\rangle$ é um isomorfismo.

Demonstração. Como $\dim(V) = \dim(V^*)$, basta mostrar que a aplicação linear em questão é injetora. Se $v \in V$ é tal que o funcional linear $\langle v, \cdot\rangle$ é nulo então em particular $\langle v, v\rangle = 0$ e portanto $v = 0$ (já que $\langle\cdot,\cdot\rangle$ é definida positiva). Isso mostra que $v \mapsto \langle v, \cdot\rangle$ tem núcleo zero e completa a demonstração. ■

Recorde que uma base $(b_i)_{i=1}^n$ de V é dita *ortonormal* com respeito a um produto interno $\langle\cdot,\cdot\rangle$ quando $\langle b_i, b_j\rangle = 0$ para $i \neq j$ e $\langle b_i, b_i\rangle = 1$ para todos $i, j = 1, \ldots, n$. O seguinte teorema nos permite identificar o isomorfismo $v \mapsto \langle v, \cdot\rangle$ mais concretamente em termos de bases ortonormais de V.

Teorema. *Seja $\langle\cdot,\cdot\rangle$ um produto interno em V e $\mathfrak{B} = (b_i)_{i=1}^n$ uma base ortonormal. Então o isomorfismo $V \to V^*$ determinado por $\langle\cdot,\cdot\rangle$ é o único isomorfismo que leva b_i sobre b_i^* para $i = 1, \ldots, n$. Em particular, as coordenadas de um vetor $v \in V$ com respeito à base $\mathfrak{B}$ coincidem com as coordenadas de $\alpha = \langle v, \cdot\rangle \in V^*$ com respeito à base dual $\mathfrak{B}^*$.*

Demonstração. Devemos mostrar que $\langle b_i, \cdot\rangle = b_i^*$ para todo $i = 1, \ldots, n$; basta mostrar então que esses funcionais coincidem sobre uma base de $\mathfrak{B}$, i.e., basta ver que $\langle b_i, b_j\rangle = b_i^*(b_j)$ para $j = 1, \ldots, n$. Isso segue diretamente da definição de base ortonormal e da definição de base dual. Quanto à útlima afirmação, se o isomorfismo $v \mapsto \langle v, \cdot\rangle$ leva a base $\mathfrak{B}$ sobre a base $\mathfrak{B}^*$ então as coordenadas de v na base $\mathfrak{B}$ coincidem com as coordenadas de $\langle v, \cdot\rangle$ na base $\mathfrak{B}^*$. ■

`Exemplo`. Considerando o $I\!R^n$ munido de seu produto interno canônico então a base canônica é ortonormal e portanto o isomorfismo $I\!R^n \ni v \mapsto \langle v, \cdot\rangle \in {I\!R^n}^*$ induzido pelo produto interno canônico leva o vetor $v = (v_1, \ldots, v_n) \in I\!R^n$ no funcional linear $\alpha \in {I\!R^n}^*$ que possue coordenadas $(v_1, \ldots, v_n)$ na base dual da base canônica. Esse funcional é dado por:

$$\alpha(x) = \sum_{i=1}^n v_i x_i = \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

para todo $x \in I\!R^n$ e portanto ele é representado pela matriz linha $\begin{pmatrix} v_1 \; v_2 \; \cdots \; v_n \end{pmatrix}$. Observe que o i-ésimo vetor da base dual da base canônica de $I\!R^n$ é simplesmente a i-ésimo aplicação de projeção $(x_1, \ldots, x_n) \mapsto x_i$.

(2) **1-formas em $\mathbb{R}^n$.**

Definição. *Uma* 1-forma *em $I\!R^n$ (definida num subconjunto $S \subset I\!R^n$) é uma aplicação $\omega : S \to I\!R^{n*}$ que a cada ponto $x \in S$ associa um funcional linear $\omega(x) : I\!R^n \to I\!R$.*

Vocês devem manter em mente a analogia entre a definição acima e a definição de campo vetorial: um campo vetorial associa a cada ponto de $S \subset I\!R^n$ um vetor, i.e., um elemento de $I\!R^n$; uma 1-forma associa a cada ponto de $S \subset I\!R^n$ um co-vetor, i.e., um elemento de $I\!R^{n*}$. Vamos agora entender a definição acima de maneira mais concreta.

Denotaremos, como sempre, por $(e_i)_{i=1}^n$ a base canônica de $I\!R^n$. De acordo com a notação da seção anterior, deveríamos denotar a base dual da base canônica de $I\!R^n$ por $(e_i^*)_{i=1}^n$; como observamos no último exemplo daquela seção, $e_i^* : I\!R^n \to I\!R$ é simplesmente a i-ésima aplicação de projeção $I\!R^n \ni x \mapsto x_i \in I\!R$. No contexto do cálculo com 1-formas, no entanto, notações como $(e_i^*)_{i=1}^n$ para a base dual da base canônica *não são usuais*; em vez, o i-ésimo vetor da base dual da base canônica de $I\!R^n$ é usualmente denotado por $\mathrm{d}x_i$. Embora em princípio não seja necessário justificar a introdução de uma notação, existe uma "explicação" padrão para o uso da notação $(\mathrm{d}x_i)_{i=1}^n$ para a base dual de $(e_i)_{i=1}^n$. A explicação é a seguinte: se denotamos por $x = (x_1, \ldots, x_n)$ um vetor de $I\!R^n$, então é perfeitamente natural denotar a i-ésima aplicação de projeção $x \mapsto x_i$ (i.e., o i-ésimo vetor da base dual da base canônica) simplesmente pelo símbolo x_i. Como tal aplicação de projeção é linear, sua diferencial em qualquer ponto é igual a ela própria; é razoável então denotar também por $\mathrm{d}x_i$ a própria aplicação de projeção $x \mapsto x_i$ (embora $\mathrm{d}x_i(x)$ seria possivelmente uma notação mais correta).

Se $\omega : S \subset I\!R^n \to I\!R^{n*}$ é uma 1-forma então para cada $x \in S$ o co-vetor $\omega(x) \in I\!R^{n*}$ pode ser escrito de modo único como combinação linear dos vetores da base dual da base canônica de $I\!R^n$. Temos então que a 1-forma ω pode ser escrita de modo único sob a forma:

$$\omega(x) = \sum_{i=1}^n a_i(x)\mathrm{d}x_i,$$

onde $a_i : S \to I\!R$, $i = 1, \ldots, n$, são funções a valores reais.

`Exemplo`. Se $f : U \subset I\!R^n \to I\!R$ é uma função diferenciável num aberto $U \subset I\!R^n$ então o diferencial de f é uma 1-forma $\mathrm{d}f : U \to I\!R^{n*}$. A i-ésima coordenada de $\mathrm{d}f(x)$ na base $(\mathrm{d}x_i)_{i=1}^n$ é $\mathrm{d}f(x) \cdot e_i = \frac{\partial f}{\partial x_i}(x)$; temos então:

$$\mathrm{d}f(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)\, \mathrm{d}x_i,$$

para todo $x \in U$.

Como já observamos na seção anterior, o produto interno canônico $\langle \cdot, \cdot \rangle$ de $I\!R^n$ induz um isomorfismo $v \mapsto \langle v, \cdot \rangle$ de $I\!R^n$ sobre $I\!R^{n*}$. Como a base canônica de $I\!R^n$ é ortonormal, sabemos que esse isomorfismo leva o vetor $v = (v_1, \ldots, v_n) \in I\!R^n$ sobre o co-vetor $\sum_{i=1}^n v_i \mathrm{d}x_i$. Esse isomorfismo entre $I\!R^n$ e $I\!R^{n*}$ fornece uma correspondência biunívoca entre campos vetoriais $X : S \subset I\!R^n \to I\!R^n$ e 1-formas $\omega : S \subset I\!R^n \to I\!R^{n*}$ de modo que $\omega(x) = \big\langle X(x), \cdot \big\rangle$ para todo $x \in S$. Em termos das bases canônicas temos:

$$X(x) = \sum_{i=1}^n X_i(x) e_i, \quad \omega(x) = \sum_{i=1}^n X_i(x) \mathrm{d}x_i,$$

para todo $x \in S$.

`Exemplo`. Se $f : U \subset I\!R^n \to I\!R$ é uma função diferenciável no aberto $U \subset I\!R^n$ então o campo vetorial associado à 1-forma $\mathrm{d}f : U \to I\!R^{n*}$ é conhecido como o *gradiente* de f e é denotado por ∇f. Temos:

$$\nabla f(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) e_i,$$

para todo $x \in U$. Além do mais, obtemos também a familiar fórmula:

$$\frac{\partial f}{\partial v}(x) = \mathrm{d}f(x) \cdot v = \big\langle \nabla f(x), v \big\rangle, \quad x \in U, \ v \in I\!R^n,$$

para as derivadas direcionais $\frac{\partial f}{\partial v}(x)$ de f. Se $\nabla f(x) \neq 0$, uma aplicação direta da desigualdade de Cauchy–Schwarz mostra que a direção $v = \frac{\nabla f(x)}{\|\nabla f(x)\|}$ maximiza o valor da derivada direcional $\frac{\partial f}{\partial v}(x)$ para $\|v\| = 1$.

(3) **Integração de 1-formas**.

Definição. *Seja $\omega : S \subset I\!R^n \to I\!R^{n*}$ uma 1-forma contínua definida num subconjunto $S \subset I\!R^n$. Se $\gamma : [a,b] \to I\!R^n$ é uma curva de classe C^1 por partes cuja imagem está contida em S então a* integral de ω ao longo de γ *é definida por:*

$$\int_\gamma \omega = \int_a^b \omega\big(\gamma(t)\big) \cdot \gamma'(t) \, \mathrm{d}t.$$

Obviamente, se $\omega(x) = \sum_{i=1}^n a_i(x) \mathrm{d}x_i$ e $X(x) = \sum_{i=1}^n a_i(x) e_i$ é o campo vetorial $X : S \to I\!R^n$ associado a ω pelo isomorfismo canônico de $I\!R^n$ sobre $I\!R^{n*}$ então:

$$\int_\gamma \omega = \sum_{i=1}^n \int_a^b a_i\big(\gamma(t)\big)(\gamma^i)'(t) \, \mathrm{d}t = \int_\gamma X.$$

Na verdade, não existe diferença nenhuma entre a teoria de integração de campos vetoriais sobre curvas em $I\!R^n$ e a teoria de integração de 1-formas sobre curvas em $I\!R^n$ — é só uma questão de trocar $\sum_{i=1}^n a_i(x) \mathrm{d}x_i$ por $\sum_{i=1}^n a_i(x) e_i$. Em particular, todos os resultados que mostramos sobre integração de campos vetoriais em curvas (invariância por reparametrização, aditividade por concatenação, etc) admitem versões análogas na teoria de integração de 1-formas em curvas.

`Observação`. Em vista do comentário acima, surge uma pergunta: “para que estudar integral de linha na linguagem de 1-formas?” — em primeiro lugar, observe que a definição da integral $\int_\gamma \omega$ é "mais natural" que a definição da integral $\int_\gamma X$ pois a integração de 1-formas não faz referência ao produto interno de $\mathbb{R}^n$. Na verdade, a grande diferença só é sentida quando se estuda integral de linha para curvas em *variedades diferenciáveis*: lá o produto interno em geral não está disponível e *apenas a integração de* 1*-formas faz sentido*!

`Observação`. Muitos livros elementares de Cálculo usam notações como $\int_\gamma P\mathrm{d}x + Q\mathrm{d}y$ significando apenas a integral do campo vetorial (P, Q) ao longo da curva γ. Nesse caso a notação $P\mathrm{d}x + Q\mathrm{d}y$ é usada sem a introdução formal da noção de forma diferencial. É interessante observar que com a escolha da notação $\mathrm{d}x_i$ para os vetores da base dual da base canônica, a notação clássica $\int_\gamma P\mathrm{d}x + Q\mathrm{d}y$ dos livros de Cálculo elementares é recuperada agora com um significado matemático formal.

`Observação`. Assim como a integral de campos vetoriais, a integração de 1-formas também pode ser definida em termos de limites de somas, caso seja desejável generalizar o significado de $\int_\gamma \omega$ para 1-formas ω que não são contínuas e curvas γ que não são de classe C^1 por partes (veja os Exercícios 9–12).

`Observação`. Para quem está curioso com o nome "1-forma", mencionamos agora (mas só estudaremos em detalhes depois) que uma k-forma em $\mathbb{R}^n$ é uma função que a cada ponto de um subconjunto $S \subset \mathbb{R}^n$ associa uma *aplicação k-linear anti-simétrica em $\mathbb{R}^n$* (note que no caso $k = 1$ a anti-simetria é redundante). A teoria de k-formas aparecerá naturalmente quando estudarmos integral de superfície.

(4) **Formas fechadas e exatas**.

Definição. *Uma* 1*-forma contínua* $\omega : U \to \mathbb{R}^n$ *definida num aberto* $U \subset \mathbb{R}^n$ *é dita* exata *quando existe uma função* $f : U \to \mathbb{R}$ *de classe* C^1 *tal que* $\mathrm{d}f = \omega$.

`Observação`. Se $X : U \to \mathbb{R}^n$ é o campo vetorial associado a uma 1-forma $\omega : U \to \mathbb{R}^{n*}$ então uma função $f : U \to \mathbb{R}$ de classe C^1 com $\mathrm{d}f = \omega$ é chamada um *potencial* para X (observe que $\nabla f = X$). Na terminologia dos cursos elementares de Cálculo, um campo vetorial que admite um potencial (i.e., um campo vetorial que corresponde a uma 1-forma exata) é chamado um *campo conservativo*.

Teorema. *Seja* $f : U \to \mathbb{R}$ *uma função de classe* C^1 *definida num aberto* $U \subset \mathbb{R}^n$ *e seja* $\gamma : [a, b] \to \mathbb{R}^n$ *uma curva de classe* C^1 *por partes cuja imagem está contida em* U. *Então:*

$$\int_\gamma \mathrm{d}f = f\big(\gamma(b)\big) - f\big(\gamma(a)\big).$$

Demonstração. É só calcular:

$$\int_\gamma \mathrm{d}f = \int_a^b \mathrm{d}f\big(\gamma(t)\big) \cdot \gamma'(t)\,\mathrm{d}t = \int_a^b (f \circ \gamma)'(t)\,\mathrm{d}t;$$

a conclusão segue do Teorema Fundamental do Cálculo. ■

Corolário. *Se $\omega : U \to I\!R^{n*}$ é uma 1-forma exata então a integral $\int_\gamma \omega$ não depende da curva γ, mas* somente de suas extremidades*; mais precisamente, se $\gamma : [a,b] \to I\!R^n$, $\mu : [a',b'] \to I\!R^n$ são curvas de classe C^1 por partes com imagem em U tais que $\gamma(a) = \mu(a')$ e $\gamma(b) = \mu(b')$ então:*

$$\int_\gamma \omega = \int_\mu \omega. \blacksquare$$

Vamos agora estudar o problema clássico da dependência do caminho nas integrais de linha, i.e., o problema de determinar se uma integral da forma $\int_\gamma \omega$ depende da curva γ ou só de suas extremidades $\gamma(a)$ e $\gamma(b)$. Para começar, temos o seguinte resultado muito simples:

Teorema. *Seja $\omega : U \to I\!R^{n*}$ uma 1-forma contínua definida num aberto $U \subset I\!R^n$. As seguinte condições são equivalentes:*

(a) *dadas curvas arbitrárias $\gamma : [a,b] \to I\!R^n$, $\mu : [a',b'] \to I\!R^n$ de classe C^k por partes ($1 \le k \le \infty$) com imagens contidas em U e $\gamma(a) = \mu(a')$, $\gamma(b) = \mu(b')$ então $\int_\gamma \omega = \int_\mu \omega$;*

(b) *dada uma curva $\gamma : [a,b] \to I\!R^n$ de classe C^k por partes com imagem contida em U então, se γ é* fechada *(i.e., se $\gamma(a) = \gamma(b)$), então $\int_\gamma \omega = 0$.*

Demonstração. Suponha (a). Se $\gamma : [a,b] \to I\!R^n$ é uma curva fechada de classe C^k por partes com imagem contida em U então temos $\int_\gamma \omega = \int_\mu \omega$, onde $\mu : [a,b] \to I\!R^n$ é a curva constante e igual a $\gamma(a)$ (observe que γ e μ tem as mesmas extremidades). Como obviamente $\int_\mu \omega = 0$, (b) segue. Suponha agora (b). Se $\gamma : [a,b] \to I\!R^n$ e $\mu : [a',b'] \to I\!R^n$ são curvas de classe C^k por partes com imagens contidas em U e as mesmas extremidades, então podemos definir uma curva de classe C^k por partes $\rho : [0,1] \to I\!R^n$ tal que $\rho|_{[0,\frac{1}{2}]}$ é uma reparametrização crescente de γ e $\rho|_{[\frac{1}{2},1]}$ é uma reparametrização decrescente de μ. Como ρ é uma curva fechada, temos:

$$0 = \int_\rho \omega = \int_\gamma \omega - \int_\mu \omega;$$

isso prova (a) e completa a demonstração. $\blacksquare$

Já estabelecemos que para 1-formas exatas ω a integral de linha $\int_\gamma \omega$ não depende da curva γ, mas só de suas extremidades. Mostramos agora a recíproca:

Teorema. *Seja $\omega : U \to I\!R^{n*}$ uma 1-forma contínua definida num aberto $U \subset I\!R^n$. Suponha que para quaisquer curvas $\gamma : [a,b] \to I\!R^n$, $\mu : [a',b'] \to I\!R^n$ de classe C^∞ por partes, com imagens contidas em U e $\gamma(a) = \mu(a')$, $\gamma(b) = \mu(b')$ temos $\int_\gamma \omega = \int_\mu \omega$. Então ω é exata.*

Demonstração. Não há perda de generalidade em supor que U é conexo. De fato, se U não é conexo então escrevemos $U = \bigcup_{i \in I} U_i$ como união de suas *componentes conexas* (recorde Exercício 7, Aula número 7 — 27/03). Daí cada U_i é um aberto conexo e a união $U = \bigcup_{i \in I} U_i$ é disjunta. Se pudermos para cada $i \in I$ definir uma função $f_i : U_i \to I\!R$

de classe C^1 com $\mathrm{d}f_i = \omega|_{U_i}$ então a conclusão será obtida definindo $f : U \to I\!R$ fazendo $f|_{U_i} = f_i$ para todo $i \in I$.

Suponha então que U é um aberto conexo e vamos construir $f : U \to I\!R$ de classe C^1 com $\mathrm{d}f = \omega$. Fixe arbitrariamente $x_0 \in U$ e defina:

$$f(x) = \int_\gamma \omega,$$

onde $\gamma : [a, b] \to I\!R^n$ é uma curva arbitrária de classe C^∞ por partes com imagem em U e $\gamma(a) = x_0$, $\gamma(b) = x$. Observe que por hipótese a integral $\int_\gamma \omega$ de fato não depende da escolha de γ. Obviamente devemos nos preocupar com a *existência* de ao menos uma tal curva γ, mas tal existência é demonstrada com um argumento rotineiro de conexidade (veja Exercício 8). Escreva $\omega(x) = \sum_{i=1}^n a_i(x)\mathrm{d}x_i$; fixado $x \in U$ e $i = 1, \ldots, n$, vamos mostrar que a derivada parcial $\frac{\partial f}{\partial x_i}(x)$ existe e é igual a $\omega(x)\cdot e_i = a_i(x)$. Isso basta para concluir que f é de classe C^1 e que $\mathrm{d}f = \omega$ (pois concluiremos que f tem derivadas parciais contínuas!). Escolha $\varepsilon > 0$ tal que $x + te_i \in U$ para todo $t \in (-\varepsilon, \varepsilon)$. Fixe uma curva qualquer γ de classe C^∞ por partes ligando x_0 a x e considere a curva $\mu : [0, 1] \to I\!R^n$ definida por $\mu(s) = x + ste_i$ para todo $s \in [0, 1]$. Podemos escolher uma curva $\rho : [0, 1] \to I\!R^n$ de classe C^∞ por partes que é obtida concatenando reparametrizações crescentes de γ e μ, i.e., ρ é tal que $\rho|_{\left[0,\frac{1}{2}\right]}$ é uma reparametrização crescente de γ e $\rho|_{\left[\frac{1}{2},1\right]}$ é uma reparametrização crescente de μ. Daí ρ é uma curva de classe C^∞ por partes com imagem contida em U tal que $\rho(0) = x_0$, $\rho(1) = x + te_i$; logo:

$$\begin{aligned} f(x + te_i) &= \int_\rho \omega = \int_\gamma \omega + \int_\mu \omega = \int_\gamma \omega + \int_0^1 \omega(x + ste_i) \cdot (te_i) \,\mathrm{d}s \\ &= \int_\gamma \omega + \int_0^1 a_i(x + ste_i) t \,\mathrm{d}s, \end{aligned}$$

para todo $t \in (-\varepsilon, \varepsilon)$. Fazendo a mudança de variável $st = u$ na última integral obtemos:

$$f(x + te_i) = \int_\gamma \omega + \int_0^t a_i(x + ue_i) \,\mathrm{d}u,$$

para todo $t \in (-\varepsilon, \varepsilon)$. Derivando a igualdade acima em $t = 0$ e usando o Teorema Fundamental do Cálculo (note que o termo $\int_\gamma \omega$ não depende de t!) obtemos:

$$\frac{\partial f}{\partial x_i}(x) = \frac{\mathrm{d}}{\mathrm{d}t} f(x + te_i)\Big|_{t=0} = a_i(x),$$

o que completa a demonstração. ■

Resumimos os resultados mostrados até agora omitindo os detalhes técnicos do enunciado para uma visualização melhor dos fatos. *Temos que as seguintes afirmações são equivalentes sobre uma* 1*-forma contínua* ω *num aberto de* $I\!R^n$:

(i) *ω é exata*;

(ii) *a integral ω sobre curvas não depende do caminho*;

(iii) *a integral de ω sobre curvas fechadas é zero.*

Precisamos agora de métodos mais diretos para determinar se uma 1-forma é ou não exata. Observe que se uma 1-forma $\omega(x) = \sum_{i=1}^{n} a_i(x)\mathrm{d}x_i$ de classe C^1 for exata então existe uma função f de classe C^2 tal que $\frac{\partial f}{\partial x_i}(x) = a_i(x)$ para todo x. Aplicando o Teorema de Schwarz obtemos:

$$\frac{\partial a_i}{\partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial a_j}{\partial x_i}(x),$$

para todo x. Isso motiva a seguinte:

Definição. *Uma 1-forma $\omega : U \to \mathbb{R}^{n*}$ de classe C^1 num aberto $U \subset \mathbb{R}^n$ é dita* fechada *se:*

$$\frac{\partial a_i}{\partial x_j}(x) = \frac{\partial a_j}{\partial x_i}(x),$$

para todo $x \in U$, e todos $i, j = 1, \ldots, n$, onde $\omega(x) = \sum_{i=1}^{n} a_i(x)\mathrm{d}x_i$.

Do Teorema de Schwarz segue trivialmente o seguinte:

Teorema. *Toda 1-forma exata de classe C^1 é fechada.* ■

A recíproca do teorema acima é falsa, mas não totalmente. Isso fica pra próxima...

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Espaço dual.`

Nessa série de exercícios, *todos os espaços vetoriais são reais e de dimensão finita.*

1. Sejam V, W espaços vetoriais. Se $T : V \to W$ é uma aplicação linear então a sua *aplicação transposta* $T^* : W^* \to V^*$ é definida por $T^*(\alpha) = \alpha \circ T$ para todo $\alpha \in W^*$. Mostre que:
   (a) T^* é bem definida e linear;
   (b) se $\mathfrak{B}$ é uma base de V, $\mathfrak{C}$ é uma base de W e $\mathfrak{B}^*$, $\mathfrak{C}^*$ são as bases duais a $\mathfrak{B}$ e a $\mathfrak{C}$ respectivamente então a matriz $[T]_{\mathfrak{B}\mathfrak{C}}$ que representa T com respeito às bases $\mathfrak{B}$ e $\mathfrak{C}$ é a *transposta* da matriz $[T^*]_{\mathfrak{C}^*\mathfrak{B}^*}$ que representa T^* com respeito às bases duais $\mathfrak{C}^*$ e $\mathfrak{B}^*$;
   (c) $(T_1 \circ T_2)^* = T_2^* \circ T_1^*$, se $T_1 : V \to W$, $T_2 : U \to V$ são aplicações lineares;
   (d) se Id denota a identidade de V então Id^* é a identidade de V^*;
   (e) se $T : V \to W$ é um isomorfismo então T^* também o é e $(T^*)^{-1} = (T^{-1})^*$.
2. Se S é um subespaço do espaço vetorial V então o *anulador* de S em V é definido por:
$$S^{\mathrm{o}} = \big\{\alpha \in V^* : \alpha|_S = 0\big\}.$$
   Mostre que:
   (a) a *aplicação restrição* definida por $V^* \ni \alpha \mapsto \alpha|_S \in S^*$ é linear, sobrejetora e seu núcleo é S^{o} (em particular, S^{o} é um subespaço de V^*);
   (b) conclua do item (a) que $\dim(S) + \dim(S^{\mathrm{o}}) = \dim(V)$;
   (c) se A é um *subconjunto* arbitrário de V, então o conjunto $\big\{\alpha \in V^* : \alpha|_A = 0\big\}$ coincide com o anulador do subespaço S gerado por A em V (e portanto é também um subespaço de V^*).
3. Se V é um espaço vetorial e Z é um subespaço do dual V^* então o *espaço anulado* por Z é definido por:
$$Z_{\mathrm{o}} = \big\{v \in V : \alpha(v) = 0 \text{ para todo } \alpha \in Z\big\}.$$
   Mostre que:
   (a) Z_{o} é um subespaço de V;
   (b) se S é um subespaço de V então $(S^{\mathrm{o}})_{\mathrm{o}} = S$, i.e., o espaço anulado pelo anulador de S é S (*dica*: se $v \in V$, $v \notin S$, construa $\alpha \in V^*$ com $\alpha|_S = 0$ e $\alpha(v) = 1$);
   (c) conclua do item (b) que se $S_1, S_2 \subset V$ são subespaços com $S_1^{\mathrm{o}} = S_2^{\mathrm{o}}$ então $S_1 = S_2$;
   (d) se $Z^{\mathrm{o}} \subset V^{**}$ denota o anulador de Z então, identificando V com V^{**} da maneira usual, Z^{o} identifica-se com Z_{o};
   (e) se S é um subespaço de V então, identificando V e V^{**} da maneira usual, temos que o espaço anulado por $S \subset V^{**}$ (que é um subespaço de V^*) é precisamente o anulador de S;

(f) conclua dos itens (b), (d) e (e) que $(Z_{\mathrm{o}})^{\mathrm{o}} = Z$ (*dica*: interprete a igualdade $(Z^{\mathrm{o}})_{\mathrm{o}} = Z$);

(g) conclua do item (f) que $\dim(Z) + \dim(Z_{\mathrm{o}}) = \dim(V)$;

(h) conclua do item (f) que se $Z_1, Z_2 \subset V^*$ são subespaços com $(Z_1)_{\mathrm{o}} = (Z_2)_{\mathrm{o}}$ então $Z_1 = Z_2$;

(i) se B é um *subconjunto* arbitrário de V^* então o conjunto:

$$\big\{v \in V : \alpha(v) = 0, \text{ para todo } \alpha \in B\big\}$$

coincide com o espaço anulado pelo subespaço Z gerado por B em V^* (e portanto é também um subespaço de V).

4. Se $\alpha_1, \ldots, \alpha_k \in V^*$ são funcionais lineares e se $S = \bigcap_{i=1}^k \mathrm{Ker}(\alpha_i)$ então todo $\alpha \in V^*$ pertencente ao anulador de S é uma combinação linear dos α_i's (*dica*: se $Z \subset V^*$ é o subespaço gerado pelos α_i's então $S = Z_{\mathrm{o}}$; use o item (f) do Exercício 3).

5. Se V, W são espaços vetoriais e $T : V \to W$ é uma aplicação linear, mostre que:

(a) $\mathrm{Ker}(T^*) = \big[\mathrm{Im}(T)\big]^{\mathrm{o}}$;

(b) $\mathrm{Im}(T^*) = \big[\mathrm{Ker}(T)\big]^{\mathrm{o}}$ (*dica*: se $\alpha \in \big[\mathrm{Ker}(T)\big]^{\mathrm{o}}$, defina $\beta : \mathrm{Im}(T) \to \mathbb{R}$ *por passagem ao quociente*, i.e., de modo a completar a flecha pontilhada no diagrama comutativo:

$$\begin{tikzcd}
V \arrow[d, "T"'] \arrow[dr, "\alpha"] & \\
\mathrm{Im}(T) \arrow[r, dotted, "\beta"'] & \mathbb{R}
\end{tikzcd}$$

escolha uma extensão linear arbitrária $\tilde{\beta} : W \to \mathbb{R}$ de β e mostre que $T^*\big(\tilde{\beta}\big) = \alpha$);

(c) conclua do item (a) que T é sobrejetora se e somente se T^* é injetora;

(d) conclua do item (b) que T é injetora se e somente se T^* é sobrejetora;

(e) conclua do item (b) que $\dim\big(\mathrm{Im}(T)\big) = \dim\big(\mathrm{Im}(T^*)\big)$;

(f) conclua do item (e) acima e do item (b) do Exercício 1 que o "posto coluna" de uma matriz (i.e., a dimensão do subespaço gerado por suas colunas) e o "posto linha" dessa matriz (i.e., a dimensão do subespaço gerado por suas linhas) coincidem (*dica*: a dimensão do subespaço gerado pelas colunas de uma matriz coincide com a dimensão da imagem da aplicação linear associada a essa matriz).

6. Sejam V um espaço vetorial e $T : V \to \mathbb{R}^n$ uma aplicação linear. Escreva $T = (T_1, \ldots, T_n)$ com cada $T_i : V \to \mathbb{R}$ (cada T_i é linear, obviamente). Mostre que T é sobrejetora se e somente se os funcionais $T_1, \ldots, T_n$ são linearmente independentes em V^* (*dica*: se $(e_i^*)_{i=1}^n$ denota a base dual da base canônica de $\mathbb{R}^n$ então $T_i = T^*(e_i^*)$; use o item (c) do Exercício 5). Use o item (b) do Exercício 5 para obter uma outra demonstração (além daquela dada no Exercício 4) para a igualdade:

$$\left[\bigcap_{i=1}^n \mathrm{Ker}(T_i)\right]^{\mathrm{o}} = \text{subespaço gerado por } \{T_1, \ldots, T_n\} \subset V^*.$$

7. Se V, W são espaços vetoriais, $\phi_V : V \to V^{**}$, $\phi_W : W \to W^{**}$ denotam os isomorfismos canônicos e $T : V \to W$ é uma aplicação linear, mostre que o diagrama:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\phi_V"', "\cong"] & W \arrow[d, "\phi_W", "\cong"'] \\ V^{**} \arrow[r, "T^{**}=(T^*)^*"'] & W^{**} \end{tikzcd}$$

comuta (isso significa que, fazendo as identificações usuais $V \cong V^{**}$, $W \cong W^{**}$ então a *aplicação bi-transposta* $T^{**} = (T^*)^*$ de T identifica-se com T). Conclua do item (e) do Exercício 1 que T é um isomorfismo se e somente se T^* o for.

`Abertos conexos de` $I\!R^n$.

8. Seja $U \subset I\!R^n$ um aberto conexo. Mostre que dados $x, y \in U$ existe uma curva $\gamma : [0,1] \to I\!R^n$ de classe C^∞ por partes com imagem contida em U e $\gamma(0) = x$, $\gamma(1) = y$ (*dica*: defina uma relação $\sim$ em U fazendo $x \sim y$ se e somente se existe uma curva de classe C^∞ por partes em U ligando x a y; mostre que $\sim$ é uma relação de equivalência em U e que as classes de equivalência são abertas).

`A integral de linha como limite de somas.`

Os Exercícios 9, 10 e 11 abaixo são preparações para o Exercício 12.

9. Sejam $U \subset I\!R^n$ um aberto e $\gamma : [a,b] \to I\!R^n$ uma curva contínua com imagem contida em U. Mostre que existe $\delta > 0$ tal que para todos $t, s \in [a,b]$ com $|t-s| \le \delta$ temos que o segmento de reta $\big[\gamma(t), \gamma(s)\big]$ está contido em U (*dica*: use a continuidade uniforme de γ e o fato que existe $\varepsilon > 0$ tal que a bola $\mathrm{B}\big(\gamma(t);\varepsilon\big)$ está contida em U para todo $t \in [a,b]$).

10. Sejam $\gamma : [a,b] \to I\!R^n$ uma curva contínua e $U \subset I\!R^n$ um aberto que contém a imagem de γ. Suponha que $\delta > 0$ é tal que o segmento $\big[\gamma(t), \gamma(s)\big]$ está contido em U para todos $t, s \in [a,b]$ com $|t-s| \le \delta$. Mostre que existe um compacto $K \subset U$ tal que $\big[\gamma(t), \gamma(s)\big] \subset K$ para todos $t, s \in [a,b]$ com $|t-s| \le \delta$ (*dica*: tome $K = \phi(A)$ onde A é definido por:

$$A = \big\{(t,s,u) \in [a,b] \times [a,b] \times [0,1] : |t-s| \le \delta\big\},$$

e ϕ é definida por $\phi(t,s,u) = (1-u)\gamma(t) + u\gamma(s) \in U \subset I\!R^n$).

Nos exercícios a seguir, você deve considerar a integral de linha definida em termos de limites de somas (como nos Exercícios 10–15 da aula número 17 — 10/05). No caso de integrais de 1-formas $\int_\gamma \omega$ (em vez de integrais de campos vetoriais $\int_\gamma X$) a integral é definida como limite de somas de Riemann da forma $\sum_{i=0}^{k-1} \omega\big(\gamma(\tau_i)\big) \cdot \big[\gamma(t_{i+1}) - \gamma(t_i)\big]$.

11. Sejam $U \subset \mathbb{R}^n$ um aberto, $\omega : U \to \mathbb{R}^{n*}$ uma 1-forma contínua e $\gamma : [a,b] \to \mathbb{R}^n$ uma curva contínua e retificável com imagem contida em U. Mostre que dado $\varepsilon > 0$, existe uma curva $\mu : [a,b] \to \mathbb{R}^n$ de classe C^∞ por partes tal que a imagem de μ está contida em U, $\mu(a) = \gamma(a)$, $\mu(b) = \gamma(b)$ e $\left\| \int_\gamma \omega - \int_\gamma \mu \right\| < \varepsilon$ (*dica*: tome $\delta > 0$ como no Exercício 9 e K como no Exercício 10. Você já sabe que a integral $\int_\gamma \omega$ existe e portanto é possível escolher uma partição $P = \{t_0, \ldots, t_k\}$ de $[a,b]$ com $\|P\| < \delta$ e tal que:
$$\left\| \int_\gamma \omega - \sum_{i=0}^{k-1} \omega\big(\gamma(t_i)\big) \cdot \big[\gamma(t_{i+1}) - \gamma(t_i)\big] \right\| < \varepsilon',$$
onde $\varepsilon' > 0$ é escolhido a partir de ε de maneira adequada. Defina $\mu : [a,b] \to \mathbb{R}^n$ fazendo:
$$\mu(t) = \frac{(t - t_i)\gamma(t_{i+1}) + (t_{i+1} - t)\gamma(t_i)}{t_{i+1} - t_i},$$
para todo $t \in [t_i, t_{i+1}]$ e todo $i = 0, \ldots, k-1$. Compare o termo:
$$\omega\big(\gamma(t_i)\big) \cdot \big[\gamma(t_{i+1}) - \gamma(t_i)\big] = \int_{t_i}^{t_{i+1}} \omega\big(\gamma(t_i)\big) \cdot \mu'(t) \, \mathrm{d}t$$
com a integral $\int_{\mu|_{[t_i, t_{i+1}]}} \omega$. Use a continuidade uniforme de ω em K e a continuidade uniforme de γ para refinar P de maneira adequada).

12. Suponha que $f : U \to \mathbb{R}$ é uma função de classe C^1 definida num aberto $U \subset \mathbb{R}^n$ e $\gamma : [a,b] \to \mathbb{R}^n$ é uma curva contínua e retificável com imagem contida em U. Mostre que:
$$\int_\gamma \mathrm{d}f = f\big(\gamma(b)\big) - f\big(\gamma(a)\big).$$
(*dica*: use o Exercício 11).

Aula número 19 (17/05)

A aula começa cobrindo o material das seções 2, 3 e 4 da aula número 18 (15/05) sobre 1-formas, integração de 1-formas e sobre formas fechadas e exatas.

(1) **Uma 1-forma fechada que não é exata**.

Para $(x, y) \in I\!R^2$, $(x, y) \neq 0$ definimos:

$$\omega(x, y) = -\frac{y}{x^2 + y^2}\,\mathrm{d}x + \frac{x}{x^2 + y^2}\,\mathrm{d}y,$$

onde $\mathrm{d}x$ e $\mathrm{d}y$ denotam os vetores da base canônica de $I\!R^{2^*}$. Temos que ω é uma 1-forma de classe C^∞ no aberto $I\!R^2 \setminus \{0\}$ em $I\!R^2$. Um cálculo direto mostra que:

$$\frac{\partial}{\partial y}\left(-\frac{y}{x^2 + y^2}\right) = \frac{y^2 - x^2}{(x^2 + y^2)^2} = \frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right),$$

e portanto ω é fechada.

Considere a curva fechada $\gamma : [0, 2\pi] \to I\!R^2$ de classe C^∞ definida por $\gamma(t) = (\cos t, \operatorname{sen} t)$. Observe que γ é simplesmente uma parametrização para o círculo unitário $x^2 + y^2 = 1$ no sentido anti-horário. Temos:

$$\int_\gamma \omega = \int_0^{2\pi} \operatorname{sen}^2 t + \cos^2 t\,\mathrm{d}t = 2\pi \neq 0,$$

e portanto ω não pode ser exata.

Vamos agora analisar a 1-forma ω de uma maneira mais geométrica. Começamos com a seguinte:

Definição. *Se $(x, y) \in I\!R^2$ é um ponto diferente da origem então um* ângulo *para (x, y) é um escalar $\theta \in I\!R$ tal que:*

$$x = r\cos\theta, \quad y = r\operatorname{sen}\theta,$$

onde $r = \sqrt{x^2 + y^2}$ denota a norma Euclideana de (x, y).

Observe que um ângulo para (x, y) é simplesmente um valor qualquer que pode ser atribuído à coordenada polar θ do ponto (x, y). Observe também que se θ é um ângulo para (x, y) então $\{\theta + 2k\pi : k \in \mathbb{Z}\}$ é o conjunto de todos os outros possíveis ângulos para (x, y).

Dado $\theta_0 \in I\!R$, denote por A_{θ_0} o subconjunto do plano obtido removendo a semi-reta fechada que contém os pontos com ângulo θ_0; explicitamente:

$$A_{\theta_0} = I\!R^2 \setminus \big\{(t\cos\theta_0, t\operatorname{sen}\theta_0) : t \geq 0\big\}.$$

Temos que A_{θ_0} é um aberto de $I\!R^2$ e que para todo $(x, y) \in A_{\theta_0}$ existe um único $\theta \in I\!R$ pertencente ao intervalo aberto $(\theta_0, \theta_0 + 2\pi)$ que seja um ângulo para (x, y); obtemos então uma função $\theta : A_{\theta_0} \to I\!R$ que a cada ponto $(x, y) \in A_{\theta_0}$ associa seu único ângulo no intervalo $(\theta_0, \theta_0 + 2\pi)$. Não é difícil ver que θ é uma função de classe C^∞ (veja o Exercício 2).

Definição. *Se $U \subset \mathbb{R}^2$ é um aberto que não contém a origem então uma* função ângulo *em U é uma função* contínua *$\theta : U \to \mathbb{R}$ tal que, para todo $(x, y) \in U$, $\theta(x, y) \in \mathbb{R}$ é um ângulo para o ponto (x, y).*

Não é difícil mostrar que toda função ângulo num aberto $U \subset \mathbb{R}^2 \setminus \{0\}$ é automaticamente de classe C^∞ (veja o Exercício 3). Nós mostramos acima que para todo $\theta_0 \in \mathbb{R}$, existe uma função ângulo no aberto A_{θ_0}. Veremos adiante que *não pode existir uma função ângulo definida em $\mathbb{R}^2 \setminus \{0\}$*.

Seja agora $U \subset \mathbb{R}^2 \setminus \{0\}$ um aberto onde esteja definida uma função ângulo $\theta : U \to \mathbb{R}$. Vamos calcular o diferencial de θ. Considere as funções $f : U \to \mathbb{R}^2$ e $g : \mathbb{R}^2 \to \mathbb{R}^2$ de classe C^∞ definidas por:

$$f(x, y) = \Big(\sqrt{x^2 + y^2}, \theta(x, y)\Big), \quad g(r, \vartheta) = (r \cos \vartheta, r \operatorname{sen} \vartheta),$$

para todos $(x, y) \in U$, $(r, \vartheta) \in \mathbb{R}^2$. O fato que θ é uma função ângulo em U nos diz que $g\big(f(x, y)\big) = (x, y)$ para todos $(x, y) \in U$, i.e., que $g \circ f = \mathrm{Id}|_U$. Diferenciando essa igualdade num ponto $(x, y) \in U$ e usando a regra da cadeia obtemos:

$$\mathrm{d}g\big(f(x, y)\big) \circ \mathrm{d}f(x, y) = \mathrm{Id} \Longrightarrow \mathrm{d}f(x, y) = \mathrm{d}g\big(f(x, y)\big)^{-1};$$

a matriz Jacobiana de g no ponto $f(x, y) = (r, \vartheta)$ é dada por:

$$\mathrm{J}g(r, \vartheta) = \begin{pmatrix} \cos \vartheta & -r \operatorname{sen} \vartheta \\ \operatorname{sen} \vartheta & r \cos \vartheta \end{pmatrix},$$

e portanto:

$$\mathrm{J}f(x, y) = \big[\mathrm{J}g(r, \vartheta)\big]^{-1} = \frac{1}{r} \begin{pmatrix} r \cos \vartheta & r \operatorname{sen} \vartheta \\ -\operatorname{sen} \vartheta & \cos \vartheta \end{pmatrix},$$

para todos $(x, y) \in U$, sendo $r = \sqrt{x^2 + y^2}$ e $\vartheta = \theta(x, y)$. Como $x = r \cos \vartheta$ e $y = r \operatorname{sen} \vartheta$, obtemos:

$$\mathrm{J}f(x, y) = \begin{pmatrix} \dfrac{x}{r} & \dfrac{y}{r} \\ -\dfrac{y}{r^2} & \dfrac{x}{r^2} \end{pmatrix},$$

para todo $(x, y) \in U$. Mas as derivadas parciais de θ aparecem na segunda linha de $\mathrm{J}f(x, y)$; isso significa que:

$$\mathrm{d}\theta(x, y) = \frac{\partial \theta}{\partial x}(x, y)\mathrm{d}x + \frac{\partial \theta}{\partial y}(x, y)\mathrm{d}y = -\frac{y}{x^2 + y^2}\,\mathrm{d}x + \frac{x}{x^2 + y^2}\,\mathrm{d}y = \omega(x, y),$$

i.e., $\mathrm{d}\theta = \omega$ em U!

Os cálculos acima mostraram que a 1-forma ω, apesar de não ser exata em $\mathbb{R}^2 \setminus \{0\}$, *torna-se exata quando restrita a um aberto $U \subset \mathbb{R}^2 \setminus \{0\}$ onde é possível definir uma função ângulo*! O fato que ω não é exata em $\mathbb{R}^2 \setminus \{0\}$ implica em particular o seguinte:

Teorema. *Não existe uma função ângulo em* $I\!R^2 \setminus \{0\}$. ■

`Observação`. Apesar do fato que o teorema acima apereceu como um corolário elegante da nossa teoria, não seria difícil demonstrá-lo diretamente.

`Observação`. Em função da igualdade $\mathrm{d}\theta = \omega$, a 1-forma ω é às vezes chamada de *forma elemento de ângulo no plano*.

(2) **Uma condição suficiente para que uma 1-forma fechada seja exata**.

Como vimos na seção anterior, é possível que uma 1-forma fechada não seja exata. No entanto, no exemplo estudado, foi possível tornar exata a 1-forma fechada em questão por uma simples restrição do domínio. Essa é na verdade a situação geral. Vamos agora estudar condições suficientes para o domínio de uma 1-forma de modo que a implicação "fechada $\Rightarrow$ exata" seja verdadeira.

Definição. *Um subconjunto* $S \subset I\!R^n$ *é dito* estrelado *em torno de um ponto* $x_0 \in S$ *se para todo* $x \in S$ *o segmento de reta* $[x_0, x]$ *está contido em* S.

Observe que um subconjunto $S \subset I\!R^n$ é convexo se e somente se for estrelado em torno de qualquer um de seus pontos.

Teorema. *Seja* $U \subset I\!R^n$ *um subconjunto aberto, estrelado em torno de algum* $x_0 \in U$. *Então toda 1-forma fechada em* U *(de classe* C^1*) é exata.*

Demonstração. Seja $\omega : U \to I\!R^{n*}$ uma 1-forma fechada de classe C^1 e escreva $\omega(x) = \sum_{i=1}^n a_i(x)\mathrm{d}x_i$, de modo que cada $a_i : U \to I\!R$ é uma função de classe C^1 e $\frac{\partial a_i}{\partial x_j} = \frac{\partial a_j}{\partial x_i}$ para todos $i, j = 1, \ldots, n$. Defina uma função $f : U \to I\!R$ pela fórmula:

$$f(x) = \int_0^1 \omega\big(x_0 + t(x - x_0)\big) \cdot (x - x_0)\,\mathrm{d}t = \sum_{i=1}^n \int_0^1 a_i\big(x_0 + t(x - x_0)\big)\big(x_i - (x_0)_i\big)\,\mathrm{d}t,$$

para todo $x \in U$. Observe que a fórmula acima só faz sentido pois $x_0 + t(x - x_0) \in U$ para todo $t \in [0, 1]$!

Dado agora $j = 1, \ldots, n$, calculamos a derivada parcial $\frac{\partial f}{\partial x_j}(x)$ usando a regra usual para derivação sob o sinal de integral (veja Exercício 5):

$$\begin{aligned}\frac{\partial f}{\partial x_j}(x) &= \int_0^1 a_j\big(x_0 + t(x - x_0)\big)\,\mathrm{d}t + \sum_{i=1}^n \int_0^1 t\,\frac{\partial a_i}{\partial x_j}\big(x_0 + t(x - x_0)\big)\big(x_i - (x_0)_i\big)\,\mathrm{d}t \\ &= \int_0^1 a_j\big(x_0 + t(x - x_0)\big)\,\mathrm{d}t + \sum_{i=1}^n \int_0^1 t\,\frac{\partial a_j}{\partial x_i}\big(x_0 + t(x - x_0)\big)\big(x_i - (x_0)_i\big)\,\mathrm{d}t.\end{aligned}$$

Um cálculo direto mostra agora que:

$$\frac{\partial f}{\partial x_j}(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\Big[t a_j\big(x_0 + t(x - x_0)\big)\Big]\,\mathrm{d}t;$$

aplicando o Teorema Fundamental do Cálculo obtemos então $\frac{\partial f}{\partial x_j}(x) = a_j(x)$, para todo $j = 1, \ldots, n$ e portanto $\mathrm{d}f = \omega$. ■

`Observação`. **Será necessário cobrir agora o material sobre o número de Lebesgue de uma cobertura, originalmente destinado à Aula número 7 (27/03), seção 2**.

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Funções ângulo.`

0. Seja M um espaço métrico conexo e N um espaço métrico discreto (recorde que um espaço métrico é dito *discreto* quando todos os seus pontos são abertos). Mostre que toda função contínua $f : M \to N$ é constante (*dica*: se $y \in N$ então $f^{-1}(y) \subset M$ é aberto e fechado).
1. Seja $U \subset I\!\!R^2$ um aberto conexo que não contém a origem. Mostre que se $\theta_1, \theta_2 : U \to I\!\!R$ são funções ângulo em U então existe $k \in \mathbb{Z}$ com $\theta_1(x) = \theta_2(x) + 2k\pi$ para todo $x \in U$ (*dica*: a função contínua $U \ni x \mapsto \frac{1}{2\pi}(\theta_1 - \theta_2)(x) \in I\!\!R$ toma valores em $\mathbb{Z}$; use o Exercício 0).
2. Seja $\theta_0 \in I\!\!R$ e defina $\theta : A_{\theta_0} \to I\!\!R$ de modo que $\theta(x, y)$ é o único ângulo para o ponto (x, y) que pertence ao intervalo aberto $(\theta_0, \theta_0 + 2\pi)$. Mostre que θ é de classe C^∞ (*dica*: use o fato que as funções trigonométricas inversas $\operatorname{arc\,cos} : (-1, 1) \to (0, \pi)$ e $\operatorname{arc\,sen} : (-1, 1) \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ são de classe C^∞; faça uma análise exaustiva de casos para mostrar que A_{θ_0} pode ser coberto por setores abertos onde $\theta(x, y)$ é da forma $2k\pi \pm \operatorname{arc\,cos}\left(\frac{x}{\sqrt{x^2+y^2}}\right)$ ou $k\pi \pm \operatorname{arc\,sen}\left(\frac{y}{\sqrt{x^2+y^2}}\right)$, $k \in \mathbb{Z}$).
3. Seja $U \subset I\!\!R^2$ um aberto que não contém a origem. Mostre que toda função ângulo $\theta : U \to I\!\!R$ é de classe C^∞ (*dica*: dado $x \in U$, escolha $\theta_0 \in I\!\!R$ com $x \in A_{\theta_0}$ e seja V uma vizinhança aberta e conexa de x contida em $U \cap A_{\theta_0}$; considere a única função ângulo $\widehat{\theta} : A_{\theta_0} \to I\!\!R$ que toma valores em $(\theta_0, \theta_0 + 2\pi)$ e estude a diferença $\theta - \widehat{\theta}$ em V, tendo em mente os Exercícios 1 e 2).

`Propriedades básicas da integral de Riemann.`

4. Sejam M um espaço métrico e $f : [a, b] \times M \to I\!\!R$ uma função contínua. Defina $F : M \to I\!\!R$ pela integral:
$$F(x) = \int_a^b F(t, x)\,\mathrm{d}t.$$
Mostre que F é contínua (*dica*: estime $\big|F(t, x) - F(t, x_0)\big|$ usando o fato que *a continuidade de f é uniforme com respeito à variável t*, i.e., use o Exercício 18 da Aula número 7 — 27/03).
5. (*derivação sob o sinal de integral*) Sejam $I \subset I\!\!R$ um intervalo e
$$[a, b] \times I \ni (t, s) \longmapsto f(t, s) \in I\!\!R$$
uma função derivável com respeito à variável $s \in I$ e tal que a derivada parcial $\frac{\partial f}{\partial s} : [a, b] \times I \to I\!\!R$ seja contínua. Suponha que para todo $s \in I$ a função $t \mapsto f(t, s)$ seja Riemann integrável em $[a, b]$ e defina $F : I \to I\!\!R$ pela integral:
$$F(s) = \int_a^b f(t, s)\,\mathrm{d}t.$$

Mostre que F é de classe C^1 e que $F'(s) = \int_a^b \frac{\partial f}{\partial s}(t,s)\,\mathrm{d}t$ (*dica*: estime o valor de:

$$\left|\frac{F(s+h)-F(s)}{h} - \int_a^b \frac{\partial f}{\partial s}(t,s)\,\mathrm{d}t\right|,$$

usando o teorema do valor médio para concluir que:

$$f(t,s+h) - f(t,s) = h\frac{\partial f}{\partial s}\big(t, s+\theta(t)h\big),$$

com $0 < \theta(t) < 1$. Como no Exercício 4 acima, use o fato que a continuidade de $\frac{\partial f}{\partial s}$ é uniforme com respeito à variável $t \in [a,b]$).

Aula número 20 (22/05)

A aula começa cobrindo todo o material originalmente destinado à aula número 19 (17/05).

`Observação`. Antes de mais nada, vamos fazer uma comparação entre a teoria de campos vetoriais e a teoria de 1-formas em abertos de $I\!R^n$. Sejam então $U \subset I\!R^n$ um aberto, $\omega(x) = \sum_{i=1}^n a_i(x)\mathrm{d}x_i$ uma 1-forma contínua e $X(x) = \sum_{i=1}^n a_i(x)e_i$ o campo vetorial (contínuo) correspondente. Como já observamos, ω é exata se e somente se o campo X é *conservativo* — uma função $f : U \to I\!R$ de classe C^1 satisfaz $\mathrm{d}f = \omega$ se e somente se $\nabla f = X$, i.e., se e somente se f é um *potencial* para X. Suponha agora ω (e portanto X) de classe C^1. A qual condição para X corresponde a condição "ω é fechada"? No caso $n = 3$ (ou $n = 2$), define-se nos cursos elementares de Cálculo o *rotacional* de um campo vetorial — é imediato então verificar que ω é fechada se e somente se X possui *rotacional zero*. Para quem estudou a teoria de campos conservativos em $I\!R^3$, o teorema "ω exata $\Rightarrow$ ω fechada" é mais familiar sob a forma "X conservativo $\Rightarrow$ $\mathrm{rot}\, X = 0$". A recíproca dessa última implicação não vale (como vimos na aula número 19); a obstrução para a validade dessa recíproca é enunciada informalmente às vezes como a "presença de buracos" no domínio do campo (ou, equivalentemente, da 1-forma associada). Nosso objetivo na seção que segue é estudar a condição de "ausência de buracos" de maneira matemática formal — isso é feito através do conceito de *homotopia*. Mostraremos que para "domínios sem buracos" (chamados de *simplesmente conexos*) a recíproca "ω fechada $\Rightarrow$ ω exata" é verdadeira. A demonstração dessa recíproca já foi feita anteriormente para abertos convexos (e, mais geralmente, para abertos estrelados); tal resultado foi apenas uma preparação para o que será feito a seguir, que é muito mais geral. Para terminar esta (longa) observação, mencionamos que o conceito de rotacional para campos vetoriais é típico de $I\!R^3$ e não faz sentido em $I\!R^n$, para $n > 3$; em tal caso, a condição "ω é fechada" significa simplesmente $\frac{\partial a_i}{\partial x_j} = \frac{\partial a_j}{\partial x_i}$, nada mais. Adiante, quando estudarmos *k-formas diferenciais*, definiremos a noção de *diferencial exterior de uma forma diferencial*. Esse conceito generalizará os operadores rotacional, gradiente e divergente, estudados em cursos de Cálculo Vetorial; daí uma forma fechada será definida como uma *forma com diferencial exterior zero*.

(1) **A noção de homotopia**.

A partir de agora *todas as curvas consideradas serão assumidas contínuas*; mais explicitamente, uma *curva* num espaço métrico M significará uma *aplicação contínua* $\gamma : [a, b] \to M$ definida em algum intervalo fechado $[a, b]$.

Definição. *Sejam M um espaço métrico e $\gamma, \mu : [a, b] \to M$ curvas em M. Uma* homotopia *de γ para μ é uma aplicação contínua $H : [a, b] \times [0, 1] \to M$ tal que:*

$$H(t, 0) = \gamma(t) \quad e \quad H(t, 1) = \mu(t),$$

para todo $t \in [a, b]$. Quando vale também a condição:

$$H(a, s) = H(a, 0) \quad e \quad H(b, s) = H(b, 0),$$

para todo $s \in [0,1]$ então dizemos que H é uma homotopia com extremos fixos *de γ para μ. Se existe uma homotopia de γ para μ, dizemos que γ e μ são* homotópicas*; quando existe uma homotopia com extremos fixos de γ para μ, dizemos que γ e μ são* homotópicas com extremos fixos.

`Observação`. A definição acima, assim como toda a teoria de homotopia que apresentaremos, normalmente é estudada no contexto de espaços topológicos — a presença de uma métrica é totalmente irrelevante. Neste curso, nós nos restringimos ao caso de espaços métricos para facilitar a vida daqueles que não tem familiaridade com topologia.

Se $H : [a,b] \times [0,1] \to M$ é uma aplicação contínua então, dado $s \in [0,1]$, é comum denotar por $H_s : [a,b] \to M$ a aplicação $H(\cdot, s)$, i.e., $H_s(t) = H(t,s)$ para todo $t \in [a,b]$. Se H é uma homotopia de γ para μ então, para cada $s \in [0,1]$, H_s é uma curva em M, $H_0 = \gamma$ e $H_1 = \mu$. Pensa-se então no parâmetro s como sendo o "tempo"; em cada instante $s \in [0,1]$ temos uma curva H_s em M que "varia continuamente" quando s cresce de 0 a 1. No instante inicial $s = 0$, temos a curva γ; no instante final $s = 1$, temos a curva μ. Uma homotopia de γ para μ significa então uma "deformação contínua" de γ até μ dentro do espaço M. Quando dizemos que a *homotopia tem extremos fixos*, queremos dizer justamente isso: durante a deformação as extremidades da curva sendo deformada permanecem fixas. Obviamente, para que γ e μ possam ser homotópicas com extremos fixos, o pré-requisito básico é que γ e μ *tenham os mesmos extremos*, i.e., que $\gamma(a) = \mu(a)$ e $\gamma(b) = \mu(b)$.

A seguinte variação da noção de homotopia também é importante.

Definição. *Um* laço *no espaço métrico M é uma curva fechada em M, i.e., uma curva $\gamma : [a,b] \to M$ com $\gamma(a) = \gamma(b)$. Se $\gamma, \mu : [a,b] \to M$ são laços então uma homotopia $H : [a,b] \times [0,1] \to M$ de γ para μ é chamada uma* homotopia de laços *se a curva H_s é um laço em M para todo $s \in [0,1]$, i.e., se $H(a,s) = H(b,s)$ para todo $s \in [0,1]$. Quando existe uma homotopia de laços de γ para μ dizemos que γ e μ são* homotópicas como laços.

Obviamente para que curvas γ e μ sejam homotópicas como laços, o pré-requisito básico é que as mesmas *sejam laços*. Observe que uma homotopia com extremos fixos entre dois laços é uma homotopia de laços, mas uma homotopia de laços não é em geral uma homotopia com extremos fixos (os Exercícios 7 e 8 relacionam as duas noções de homotopia para laços).

`Exemplo`. Seja M um subconjunto convexo de $\mathbb{R}^n$ (munido da métrica induzida de $\mathbb{R}^n$). Se $\gamma, \mu : [a,b] \to M$ são curvas então a expressão:

$$H(t,s) = (1-s)\gamma(t) + s\mu(t), \quad t \in [a,b], \ s \in [0,1],$$

define uma homotopia $H : [a,b] \times [0,1] \to M$ de γ para μ. Se γ e μ têm os mesmos extremos então H é uma homotopia com extremos fixos; se γ e μ são laços então H é uma homotopia de laços. Observe que a homotopia H deforma γ para μ de modo que o ponto $\gamma(t)$ é levado para o ponto $\mu(t)$ ao longo do segmento de reta $[\gamma(t), \mu(t)] \subset M$.

Definição. *Um espaço métrico M é dito* simplesmente conexo *quando qualquer laço $\gamma : [a, b] \to M$ é homotópico como laço a uma aplicação constante, i.e., se existe uma aplicação contínua $H : [a, b] \times [0, 1] \to M$ com $H(t, 0) = \gamma(t)$, $H(a, s) = H(b, s)$ e $H(t, 1) = H(a, 1)$ para todos $t \in [a, b]$, $s \in [0, 1]$.*

`Exemplo`. O exemplo acima mostrou que todo subconjunto convexo de $\mathbb{R}^n$ é simplesmente conexo. Se $M \subset \mathbb{R}^n$ é estrelado em torno de algum ponto $x_0 \in M$ então M é simplesmente conexo também: de fato, se $\gamma : [a, b] \to M$ é um laço então $H(t, s) = (1 - s)\gamma(t) + sx_0$, $s \in [0, 1]$, $t \in [a, b]$, define uma homotopia de laços de γ até o laço constante e igual a x_0.

`Exemplo`. É fácil ver que todo espaço métrico homeomorfo a um espaço métrico simplesmente conexo é ainda simplesmente conexo (veja Exercício 6). Em particular, se um espaço métrico M é homeomorfo a um subconjunto estrelado de $\mathbb{R}^n$ então M é simplesmente conexo.

`Exemplo`. Se $M = \mathbb{R}^2 \setminus \{0\}$ é o "plano furado" então M não é simplesmente conexo: o laço $[0, 2\pi] \ni t \mapsto \gamma(t) = (\cos t, \operatorname{sen} t) \in M$ não é homotópico como laço a uma aplicação constante. Intuitivamente, isso ocorre porque qualquer tentativa de deformar γ até uma constante (sem "arrebentar" o laço) deve necessariamente passar pelo "buraco" que é a origem. Apesar do apelo intuitivo desse argumento, a demonstração formal do fato que γ não é homotópica como laço a uma constante não é tão simples; tal resultado será obtido como um corolário do teorema central da próxima seção.

`Observação`. A noção de homotopia (sem restrições adicionais como "extremos fixos" ou "homotopia de laços") no contexto de curvas $\gamma : [a, b] \to M$ é muito pouco interessante: de fato, $[a, b] \times [0, 1] \ni (t, s) \mapsto \gamma\big((1 - s)t + sa\big) \in M$ define uma homotopia de γ até uma aplicação constante, seja lá qual for a curva γ ou o espaço métrico M. Em geral, a noção de homotopia é definida para aplicações entre espaços topológicos quaisquer, e daí a noção de homotopia (sem restrições) passa a ser interessante.

(2) **Integração de formas fechadas e homotopia de curvas**.

O objetivo desta seção é provar o seguinte:

Teorema. *Sejam $U \subset \mathbb{R}^n$ um aberto e $\omega : U \to \mathbb{R}^{n*}$ uma 1-forma fechada (de classe C^1). Dadas curvas $\gamma, \mu : [a, b] \to \mathbb{R}^n$ de classe C^1 por partes, ambas com imagem contida em U, suponha que vale ao menos uma das duas seguintes condições:*

(i) *existe uma homotopia com extremos fixos $H : [a, b] \times [0, 1] \to U$ de $\gamma : [a, b] \to U$ para $\mu : [a, b] \to U$ (neste caso estamos supondo em particular que γ e μ têm os mesmos extremos, obviamente);*

(ii) *existe uma homotopia de laços $H : [a, b] \times [0, 1] \to U$ de $\gamma : [a, b] \to U$ para $\mu : [a, b] \to U$ (neste caso estamos supondo em particular que γ e μ são laços, obviamente).*

Então $\int_\gamma \omega = \int_\mu \omega$.

O teorema acima implica facilmente o seguinte:

Corolário. *Se $U \subset \mathbb{R}^n$ é um aberto simplesmente conexo então toda 1-forma fechada (de classe C^1) em U é exata.*

Demonstração. Se $\gamma : [a, b] \to I\!R^n$ é uma curva fechada (i.e., um laço) de classe C^1 por partes com imagem contida em U então γ é homotópica como laço em U a uma curva constante μ. Como ω é fechada, temos $\int_\gamma \omega = \int_\mu \omega = 0$ e portanto a integral de ω ao longo de qualquer curva fechada em U de classe C^1 por partes é nula. Segue que ω é exata. ■

Corolário. *Se $\omega : U \to I\!R^{n*}$ é uma 1-forma fechada (de classe C^1) num aberto $U \subset I\!R^n$ então, se um laço $\gamma : [a, b] \to U$ de classe C^1 por partes é homotópico como laço em U a uma aplicação constante então $\int_\gamma \omega = 0$.* ■

Corolário. *O laço $\gamma : [0, 2\pi] \to I\!R^2 \backslash \{0\}$ definido por $\gamma(t) = (\cos t, \operatorname{sen} t)$ não é homotópico como laço a uma aplicação constante. Em particular, o "plano furado" $I\!R^2 \setminus \{0\}$ não é simplesmente conexo.*

Demonstração. A 1-forma $\omega(x, y) = \frac{1}{x^2+y^2}(-y\mathrm{d}x + x\mathrm{d}y)$ em U é fechada mas $\int_\gamma \omega \neq 0$.
■

Para provar o teorema principal será útil o seguinte:

Lema. *Sejam K um espaço métrico compacto e M um espaço métrico qualquer. Dada uma função contínua $f : K \to M$ e uma cobertura aberta $M = \bigcup_{i \in I} V_i$ de M então existe $\delta > 0$ com a seguinte propriedade: dado um subconjunto $S \subset K$ de diâmetro menor que δ então $f(S) \subset V_i$ para algum $i \in I$.*

Demonstração. Se $U_i = f^{-1}(V_i)$ então $K = \bigcup_{i \in I} U_i$ é uma cobertura aberta do espaço métrico compacto K e portanto a mesma possui um número de Lebesgue $\delta > 0$. Daí, se $S \subset K$ tem diâmetro menor que δ então existe $i \in I$ com $S \subset U_i = f^{-1}(V_i)$; daí $f(S) \subset V_i$.
■

Antes de demonstrar o teorema, vamos a uma:

`Explicação informal da prova do teorema`. Seja $H : [a, b] \times [0, 1] \to U$ uma homotopia de γ para μ (com extremos fixos ou uma homotopia de laços). Particione o retângulo $[a, b] \times [0, 1]$ em retângulos menores que são mapeados por H em bolas abertas contidas em U; obtemos então uma "grade" em $[a, b] \times [0, 1]$ formada por retângulos pequenos. A cada segmento vertical $\mathfrak{v}$ da grade fazemos corresponder uma curva em $I\!R^n$ que parametriza o segmento cujas extremidades são as imagens por H das extremidades de $\mathfrak{v}$. Para cada segmento horizontal $\mathfrak{h}$ da grade contido no lado $[a, b] \times \{0\}$ ou no lado $[a, b] \times \{1\}$ do retângulo $[a, b] \times [0, 1]$, fazemos corresponder a curva dada pela restrição de H a $\mathfrak{h}$. Se $\mathfrak{h}$ é um segmento horizontal da grade que não está contido nem no lado $[a, b] \times \{0\}$ nem no lado $[a, b] \times \{1\}$ do retângulo $[a, b] \times [0, 1]$ então fazemos corresponder a $\mathfrak{h}$ uma curva em $I\!R^n$ que parametriza o segmento cujas extremidades são as imagens por H das extremidades de $\mathfrak{h}$. Fizemos então corresponder à grade de linhas retas em $[a, b] \times [0, 1]$ uma "grade" de linhas poligonais em U — exceto pelos lados $[a, b] \times \{0\}$ e $[a, b] \times \{1\}$ de $[a, b] \times [0, 1]$, que correspondem às curvas originais γ e μ. Vamos denotar por ρ e $\tilde{\rho}$ as linhas poligonais que correspondem respectivamente aos lados $\{a\} \times [0, 1]$ e $\{b\} \times [0, 1]$ do retângulo $[a, b] \times [0, 1]$. A integral de ω ao longo do laço correspondente à fronteira do retângulo $[a, b] \times [0, 1]$ percorrida no sentido anti-horário é igual à soma das integrais de ω sobre os laços correspondentes aos retângulos menores da grade percorridos no sentido anti-horário, pois "lados comuns a dois retângulos da grade são percorridos em sentidos opostos" e portanto cancelam. Por outro lado, como os laços correspondentes a tais retângulos pequenos estão contidos em bolas contidas em U (nas quais ω é exata), então a integral de ω nos laços correspondentes a tais retângulos pequenos é nula. Obtemos então:

$$\int_\gamma \omega + \int_{\tilde\rho} \omega - \int_\mu \omega - \int_\rho \omega = 0;$$

se H é uma homotopia com extremos fixos então ρ e $\tilde\rho$ são constantes. Se H é uma homotopia de laços então $\rho = \tilde\rho$. Em qualquer caso, a conclusão desejada é obtida. ■

Para escrever a demonstração do teorema de maneira matematicamente mais precisa, vamos introduzir um pouco de notação.

Definição. *Sejam M um espaço métrico e $\gamma : [a,b] \to M$ uma curva. A curva $\tilde\gamma : [0,1] \to M$ definida por:*

$$\tilde\gamma(t) = \gamma\big((1-t)a + tb\big), \quad t \in [0,1],$$

é chamada a reparametrização afim crescente *de γ (definida no intervalo $[0,1]$).*

Como integrais de linha são invariantes por reparametrizações crescentes, em vista da definição acima, não há perda de generalidade em estudar apenas curvas definidas no intervalo $[0,1]$ (veja também o Exercício 2).

Definição. *Sejam M um espaço métrico e $\gamma : [0,1] \to M$ uma curva. A* reparametrização reversa *de γ é a curva $\gamma^{-1} : [0,1] \to M$ definida por:*

$$\gamma^{-1}(t) = \gamma(1-t), \quad t \in [0,1].$$

Se $\mu : [0,1] \to M$ é uma outra curva com $\gamma(1) = \mu(0)$ então a concatenação *de γ e μ é a curva $(\gamma \cdot \mu) : [0,1] \to M$ definida por:*

$$(\gamma \cdot \mu)(t) = \begin{cases} \gamma(2t), & t \in \big[0, \frac{1}{2}\big], \\ \mu(2t-1), & t \in \big[\frac{1}{2}, 1\big]. \end{cases}$$

Obviamente a continuidade de γ implica na continuidade de γ^{-1}; é fácil ver também que a continuidade de γ e de μ implicam na continuidade de $\gamma \cdot \mu$ (recorde item (a), Exercício 8, Aula número 7 — 27/03).

A concatenação de curvas *não* é uma operação associativa: a curva $(\gamma \cdot \mu) \cdot \lambda$ corresponde a "γ e μ percorridas com a velocidade quadruplicada seguidas de λ percorrida com a velocidade dobrada"; por outro lado, a curva $\gamma \cdot (\mu \cdot \lambda)$ corresponde a "γ percorrida com a velocidade dobrada seguida de μ e λ percorridas com a velocidade quadruplicada". É verdade, no entanto, que a diferença entre $(\gamma \cdot \mu) \cdot \lambda$ e $\gamma \cdot (\mu \cdot \lambda)$ é apenas uma questão de parametrização (veja o Exercício 5 (a)). Para facilitar a notação (economizando parênteses), vamos convencionar que $\gamma_1 \cdot \gamma_2 \cdots \gamma_n$ significa $(\cdots(\gamma_1 \cdot \gamma_2)\cdots) \cdot \gamma_n)$.

Prova do Teorema. Sejam $U \subset I\!R^n$ um aberto, $\omega : U \to I\!R^{n*}$ uma 1-forma fechada (de classe C^1) e $\gamma, \mu : [a,b] \to I\!R^n$ curvas de classe C^1 por partes com imagens contidas em U. Suponha que exista uma homotopia $H : [a,b] \times [0,1] \to U$ de γ para μ que seja ou uma homotopia com extremos fixos ou uma homotopia de laços.

Como ω é fechada, sabemos que a restrição de ω a qualquer bola aberta contida em U é exata, pois bolas são convexas. Como U pode ser escrito como uma união de bolas abertas, o lema acima implica que existe $\delta > 0$ tal que para todo subconjunto $S \subset [a,b] \times [0,1]$ com diâmetro menor que δ temos que $H(S) \subset U$ está contido em alguma bola aberta contida em U. Escolha agora partições $a = t_0 < t_1 < \cdots < t_k = b$ e $0 = s_0 < s_1 < \cdots < s_l = 1$ dos intervalos $[a,b]$ e $[0,1]$ respectivamente, de modo que cada retângulo $[t_i, t_{i+1}] \times [s_j, s_{j+1}]$, $i = 0, \ldots, k-1$, $j = 0, \ldots, l-1$, tenha diâmetro menor que δ; em particular, $H\big([t_i, t_{i+1}] \times [s_j, s_{j+1}]\big)$ está contido em alguma bola aberta contida em U.

Para cada $i = 0, \ldots, k-1$ e cada $j = 1, \ldots, l-1$, denote por $\mathfrak{h}_{ij} : [0,1] \to I\!R^n$ a curva $\mathfrak{h}_{ij}(u) = (1-u)H(t_i, s_j) + uH(t_{i+1}, s_j)$, $u \in [0,1]$, i.e., $\mathfrak{h}_{ij}$ é uma parametrização afim do segmento ligando $H(t_i, s_j)$ a $H(t_{i+1}, s_j)$. Para $i = 0, \ldots, k-1$, $j = 0$, definimos $\mathfrak{h}_{ij} : [0,1] \to I\!R^n$ como sendo a reparametrização afim crescente de $\gamma|_{[t_i, t_{i+1}]}$; similarmente, para $i = 0, \ldots, k-1$, $j = l$, definimos $\mathfrak{h}_{ij} : [0,1] \to I\!R^n$ como sendo a reparametrização afim crescente de $\mu|_{[t_i, t_{i+1}]}$. Para $i = 0, \ldots, k$, $j = 0, \ldots, l-1$, definimos $\mathfrak{v}_{ij} : [0,1] \to I\!R^n$ como sendo uma parametrização afim do segmento ligando $H(t_i, s_j)$ a $H(t_i, s_{j+1})$, i.e., $\mathfrak{v}_{ij}(u) = (1-u)H(t_i, s_j) + uH(t_i, s_{j+1})$, $u \in [0,1]$. Defina curvas $\rho, \tilde{\rho} : [0,1] \to I\!R^n$ fazendo:

$$\rho = \mathfrak{v}_{00} \cdot \mathfrak{v}_{01} \cdots \mathfrak{v}_{0(l-1)}, \quad \tilde{\rho} = \mathfrak{v}_{k0} \cdot \mathfrak{v}_{k1} \cdots \mathfrak{v}_{k(l-1)}.$$

Para cada $i = 0, \ldots, k-1$, $j = 0, \ldots, l-1$, considere a curva $R_{ij} : [0,1] \to I\!R^n$ definida por:

$$R_{ij} = \mathfrak{h}_{ij} \cdot \mathfrak{v}_{(i+1)j} \cdot (h_{i(j+1)})^{-1} \cdot (\mathfrak{v}_{ij})^{-1}.$$

Temos que R_{ij} é uma curva fechada de classe C^1 por partes cuja imagem está contida numa bola aberta contida em U; como ω é exata nessa bola, temos que $\int_{R_{ij}} \omega = 0$ para todos $i = 0, \ldots, k-1$, $j = 0, \ldots, l-1$. Além do mais:

$$\begin{aligned} 0 = \sum_{i=0}^{k-1} \sum_{j=0}^{l-1} \int_{R_{ij}} \omega &= \sum_{i=0}^{k-1} \int_{\mathfrak{h}_{i0}} \omega + \sum_{j=0}^{l-1} \int_{\mathfrak{v}_{kj}} \omega - \sum_{i=0}^{k-1} \int_{\mathfrak{h}_{il}} \omega - \sum_{j=0}^{l-1} \int_{\mathfrak{v}_{0j}} \omega \\ &= \int_{\gamma} \omega + \int_{\tilde{\rho}} \omega - \int_{\mu} \omega - \int_{\rho} \omega. \end{aligned}$$

Se H é uma homotopia com extremos fixos então ρ e $\tilde{\rho}$ são curvas constantes; se H é uma homotopia de laços então $\rho = \tilde{\rho}$. Em qualquer caso, temos $\int_{\tilde{\rho}} \omega - \int_{\rho} \omega = 0$ e portanto $\int_{\gamma} \omega = \int_{\mu} \omega$. ■

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Curvas e homotopia.`

1. Dado um espaço métrico M, mostre que as relações "γ é homotópica a μ com extremos fixos" e "γ é homotópica a μ como laço" definem relações de equivalência respectivamente no conjunto de todas as curvas $\gamma : [a,b] \to M$ e no conjunto de todos os laços $\gamma : [a,b] \to M$ [*dica*: a parte menos fácil é a transitividade. Para demonstrá-la, observe que se $H, \widetilde{H} : [a,b] \times [0,1] \to M$ são aplicações contínuas com $H(t,1) = \widetilde{H}(t,0)$ para todo $t \in [a,b]$ então a fórmula:

$$K(t,s) = \begin{cases} H(t,2s), & s \in \big[0, \frac{1}{2}\big], \\ \widetilde{H}(t,2s-1), & s \in \big[\frac{1}{2}, 1\big], \end{cases}$$

$t \in [a,b]$, define uma aplicação contínua $K : [a,b] \times [0,1] \to M$. Use o Exercício 38 da aula número 5 (20/03) para mostrar que K é de fato contínua).

2. Sejam M um espaço métrico e $\gamma, \mu : [a,b] \to M$ curvas. Sejam $\tilde{\gamma}, \tilde{\mu} : [0,1] \to M$ as reparametrizações afins crescentes de γ e μ respectivamente. Mostre que:
   (a) γ é homotópica a μ com extremos fixos se e somente se $\tilde{\gamma}$ é homotópica a $\tilde{\mu}$ com extremos fixos;
   (b) γ é homotópica a μ como laço se e somente se $\tilde{\gamma}$ é homotópica a $\tilde{\mu}$ como laço.

3. Sejam M um espaço métrico e $\gamma, \mu : [0,1] \to M$ curvas com $\gamma(1) = \mu(0)$. Mostre que:
   (a) se $\tilde{\gamma} : [0,1] \to M$ é homotópica a γ com extremos fixos então a concatenação $\tilde{\gamma} \cdot \mu$ é homotópica à $\gamma \cdot \mu$ com extremos fixos;
   (b) se $\tilde{\mu} : [0,1] \to M$ é homotópica a μ com extremos fixos então a concatenação $\gamma \cdot \tilde{\mu}$ é homotópica a $\gamma \cdot \mu$ com extremos fixos;
   (c) se $\tilde{\gamma} : [0,1] \to M$ é homotópica a γ com extremos fixos então $\tilde{\gamma}^{-1}$ é homotópica a γ^{-1} com extremos fixos.

(*dica*: é fácil construir as homotopias; a única coisa não trivial é verificar sua continuidade. Para isso, use o Exercício 38 da aula número 5 — 20/03).

4. Sejam M um espaço métrico, $\gamma : [0,1] \to M$ uma curva e $\sigma : [0,1] \to [0,1]$ uma aplicação contínua. Mostre que:
   (a) se $\sigma(0) = 0$ e $\sigma(1) = 1$ então $\gamma \circ \sigma$ é homotópica a γ com extremos fixos;
   (b) se $\sigma(0) = \sigma(1)$ então $\gamma \circ \sigma$ é homotópica com extremos fixos a uma curva constante.

(*dica*: olhe para $\gamma\big((1-s)t + s\sigma(t)\big)$ no item (a) e para $\gamma\big((1-s)\sigma(t) + s\sigma(0)\big)$ no item (b)).

5. Sejam M um espaço métrico e $\gamma, \mu, \lambda : [0,1] \to M$ curvas com $\gamma(1) = \mu(0)$ e $\mu(1) = \lambda(0)$.
   (a) Determine uma aplicação contínua $\sigma : [0,1] \to [0,1]$ tal que $\big[(\gamma \cdot \mu) \cdot \lambda\big] \circ \sigma = \gamma \cdot (\mu \cdot \lambda)$ e $\sigma(0) = 0$, $\sigma(1) = 1$ (*dica*: é possível escolher σ *afim por partes*).
   (b) Conclua do item (a) e do Exercício 4 que $(\gamma \cdot \mu) \cdot \lambda$ é homotópica com extremos fixos a $\gamma \cdot (\mu \cdot \lambda)$.

(c) Para $x \in M$ denote por $\mathfrak{o}_x : [0,1] \to M$ a *curva constante e igual a* x, i.e., $\mathfrak{o}_x(t) = x$ para todo $t \in [0,1]$. Determine funções contínuas $\sigma, \tilde{\sigma} : [0,1] \to [0,1]$ com $\gamma \cdot \mathfrak{o}_{\gamma(1)} = \gamma \circ \tilde{\sigma}$, $\mathfrak{o}_{\gamma(0)} \cdot \gamma = \gamma \circ \sigma$, $\sigma(0) = \tilde{\sigma}(0) = 0$ e $\sigma(1) = \tilde{\sigma}(1) = 1$ (*dica*: novamente é possível escolher σ e $\tilde{\sigma}$ afins por partes).

(d) Conclua do item (c) e do Exercício 4 que $\gamma \cdot \mathfrak{o}_{\gamma(1)}$ e $\mathfrak{o}_{\gamma(0)} \cdot \gamma$ são homotópicas a γ com extremos fixos.

(e) Determine funções contínuas $\sigma, \tilde{\sigma} : [0,1] \to [0,1]$ tais que $\gamma \circ \sigma = \gamma \cdot \gamma^{-1}$, $\gamma \circ \tilde{\sigma} = \gamma^{-1} \cdot \gamma$, $\sigma(0) = \sigma(1) = 0$ e $\tilde{\sigma}(0) = \tilde{\sigma}(1) = 1$ (*dica*: novamente é possível escolher σ e $\tilde{\sigma}$ afins por partes).

(f) Conclua do item (e) e do Exercício 4 que $\gamma \cdot \gamma^{-1}$ e $\gamma^{-1} \cdot \gamma$ são homotópicas com extremos fixos a curvas constantes.

6. Sejam M, N espaços métricos e $f : M \to N$ uma função contínua. Mostre que se $\gamma, \mu : [a,b] \to M$ são homotópicas com extremos fixos então $f \circ \gamma$ e $f \circ \mu$ são homotópicas com extremos fixos. Similarmente, mostre que se γ e μ são homotópicas como laços então $f \circ \gamma$ e $f \circ \mu$ são homotópicas como laços. Conclua que se dois espaços métricos M e N são homeomorfos então M é simplesmente conexo se e somente se N o for.

7. Sejam M um espaço métrico e $\gamma, \mu : [0,1] \to M$ laços. Suponha que $H : [0,1] \times [0,1] \to M$ é uma homotopia de laços de γ para μ e defina $\lambda : [0,1] \to M$ fazendo $\lambda(s) = H(0,s) = H(1,s)$, $s \in [0,1]$. Mostre que os laços μ e $\lambda^{-1} \cdot \gamma \cdot \lambda$ são homotópicos *com extremos fixos* (*dica*: defina curvas $\mathfrak{l}, \mathfrak{r}, \mathfrak{t}, \mathfrak{b} : [0,1] \to [0,1] \times [0,1]$ fazendo $\mathfrak{l}(u) = (0,u)$, $\mathfrak{r}(u) = (1,u)$, $\mathfrak{t}(u) = (u,1)$, $\mathfrak{b}(u) = (u,0)$; como o retângulo $[0,1] \times [0,1]$ é convexo, as curvas $\mathfrak{l}^{-1} \cdot \mathfrak{b} \cdot \mathfrak{r}$ e $\mathfrak{t}$ são homotópicas com extremos fixos. Aplique o Exercício 6 com $f = H$).

8. Sejam M um espaço métrico, $\gamma : [0,1] \to M$ um laço e $\lambda : [0,1] \to M$ uma curva com $\lambda(0) = \gamma(0)$. Mostre que os laços $\lambda^{-1} \cdot \gamma \cdot \lambda$ e γ são homotópicos *como laços* [*dica*: para $s \in [0,1]$, seja $\lambda_s : [0,1] \to M$ a curva definida por $\lambda_s(u) = \lambda(us)$. Defina uma homotopia H fazendo $H_s = \lambda_s^{-1} \cdot \gamma \cdot \lambda_s$ e mostre a continuidade de H usando o Exercício 38 da aula número 5 (20/03). Observe que λ_0 é uma curva constante e complete a demonstração usando o Exercício 5 para concluir que $\lambda_0^{-1} \cdot \gamma \cdot \lambda_0$ é homotópica a γ com extremos fixos].

9. Dado um espaço métrico M e um laço $\gamma : [a,b] \to M$, mostre que γ é homotópico como laço a uma aplicação constante se e somente se γ é homotópico com extremos fixos a uma aplicação constante (*dica*: use os Exercícios 7 e 5).

10. Dado um espaço métrico M, mostre que as seguintes condições são equivalentes:

   (i) M é simplesmente conexo;

   (ii) todo laço em M é homotópico *com extremos fixos* a um laço constante;

   (iii) dadas curvas $\gamma, \mu : [a,b] \to M$ com os mesmos extremos então γ e μ são homotópicas com extremos fixos.

   Se M é conexo por caminhos, mostre que (i), (ii) e (iii) são também equivalentes a qualquer uma das seguinte condições:

   (iv) existe $x_0 \in M$ tal que todo laço $\gamma : [a,b] \to M$ com $\gamma(a) = \gamma(b) = x_0$ é homotópico com extremos fixos a um laço constante;

(v) existe $x_0 \in M$ tal que todo laço $\gamma : [a, b] \to M$ com $\gamma(a) = \gamma(b) = x_0$ é homotópico como laço a uma aplicação constante.

(*dica*: para (i)⇔(ii) e (iv)⇔(v), use o Exercício 9. Para (ii)⇒(iii) use o Exercício 5, levando em conta que, sob (ii), o laço $\gamma \cdot \mu^{-1}$ é homotópico com extremos fixos a uma constante. (iii)⇒(i) e (i)⇒(iv) são óbvios. Se M é conexo por caminhos, para mostrar (iv)⇒(i) considere um laço $\gamma : [0, 1] \to M$ com $\gamma(0) = \gamma(1) = x_1$ e uma curva $\lambda : [0, 1] \to M$ com $\lambda(0) = x_0$ e $\lambda(1) = x_1$. Use o Exercício 5 e o fato que, sob (iv), $\lambda \cdot \gamma \cdot \lambda^{-1}$ é homotópico com extremos fixos a um laço constante).

O significado dos Exercícios acima.

O Exercício 2 diz que no estudo da teoria de homotopia de curvas e laços pode-se assumir sempre sem perda de generalidade que o intervalo onde as curvas estão definidas é $[0, 1]$. O Exercício 1 diz que homotopia com extremos fixos e homotopia de laços definem relações de equivalência. Vamos denotar por $[\gamma]$ a *classe de equivalência* formada por todas as curvas homotópicas a γ com extremos fixos. O Exercício 3 diz que as operações de composição e inversão de curvas são bem-definidas no espaço quociente formado pelas classes de homotopia $[\gamma]$, i.e., faz sentido definir $[\gamma] \cdot [\mu] = [\gamma \cdot \mu]$ e $[\gamma]^{-1} = \left[\gamma^{-1}\right]$ (se $\gamma(1) = \mu(0)$). O item (a) do Exercício 4 diz que classes de homotopia são invariantes por reparametrização. O Exercício 5 diz que concatenação de curvas é uma operação associativa a nível de classes de homotopia (mas não a nível de curvas!); as curvas constantes funcionam como *elementos neutros* para concatenação e a curva γ^{-1} funciona como um *elemento inverso* de γ. O Exercício 6 diz que aplicações contínuas entre espaços induzem funções nos correspondentes espaços de classes de homotopias de curvas. Os Exercícios 7, 8 e 9 relacionam homotopia de laços com homotopia com extremos fixos: basicamente, dois laços são homotópicos como laços se e somente se um é homotópico com extremos fixos a um *conjugado* do outro.

As classes de homotopia com extremos fixos de curvas em M formam uma estrutura algébrica chamada um *grupóide*; este é chamado o *grupóide fundamental* do espaço M. Fixado $x_0 \in M$ e considerando apenas classes de homotopia (com extremos fixos) de laços $\gamma : [0, 1] \to M$ com $\gamma(0) = \gamma(1) = x_0$, obtemos um *grupo*; este é chamado o *grupo fundamental* do espaço M com ponto base x_0 e é normalmente denotado por $\pi_1(M, x_0)$. Uma curva contínua $\lambda : [0, 1] \to M$ com $\lambda(0) = x_0$, $\lambda(1) = x_1$ induz um isomorfismo entre os grupos fundamentais $\pi_1(M, x_0)$ e $\pi_1(M, x_1)$ dado por $[\gamma] \mapsto \left[\lambda^{-1} \cdot \gamma \cdot \lambda\right]$. Para espaços conexos por caminhos então o grupo fundamental é (a menos de isomorfismos) independente do ponto base. Um espaço é simplesmente conexo justamente quando seu grupo fundamental é trivial.

Em geral, é possível definir a noção de homotopia não apenas para curvas, mas também para aplicações quaisquer entre espaços topológicos. O grupo fundamental de M pode então também ser definido em termos de classes de homotopia de aplicações do círculo em M. A partir daí, pode-se também estudar classes de homotopia de aplicações da *esfera n-dimensional* em M; obtêm-se então o chamado *n-ésimo grupo de homotopia de M*, denotado por $\pi_n(M, x_0)$. A teoria de homotopia é um campo importante dentro da topologia algébrica e nós aqui mal esbarramos no assunto.

Aula número 21 (24/05)

A aula número 21 cobriu o material sobre número de Lebesgue de coberturas (aula número 7 — 27/03, seção 2) e o material sobre a relação entre homotopia de curvas e integrais de linha (aula número 20 — 22/05, seção 2).

Aula número 22 (29/05)

(1) O teorema da função inversa.

O teorema da função inversa é um exemplo típico da idéia básica do Cálculo Diferencial: aproximar funções não lineares por funções lineares. O objetivo principal do teorema da função inversa é relacionar a bijetividade da aproximação linear de uma função (i.e., sua diferencial) com a bijetividade da função original numa região pequena. Começamos primeiro com a implicação fácil do teorema.

Teorema. *Sejam $U, V \subset \mathbb{R}^n$ abertos e $f : U \to V$ uma função bijetora. Suponha que $f : U \to \mathbb{R}^n$ seja diferenciável num ponto $x \in U$ e que $f^{-1} : V \to \mathbb{R}^n$ seja diferenciável no ponto $f(x)$. Então $\mathrm{d}f(x) : \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo e $\mathrm{d}f^{-1}\big(f(x)\big) = \mathrm{d}f(x)^{-1}$ ("a diferencial da inversa é a inversa da diferencial").*

Demonstração. Temos $f^{-1} \circ f = \mathrm{Id}|_U$ e $f \circ f^{-1} = \mathrm{Id}|_V$; aplicando a regra da cadeia obtemos:

$$\mathrm{d}f^{-1}\big(f(x)\big) \circ \mathrm{d}f(x) = \mathrm{Id}, \quad \mathrm{d}f(x) \circ \mathrm{d}f^{-1}\big(f(x)\big) = \mathrm{Id},$$

e a conclusão segue. ■

`Observação`. Na verdade, poderíamos ter dado um enunciado um pouco melhor para o teorema acima: poderíamos ter assumido que U é um aberto de $\mathbb{R}^m$ e que V é um aberto de $\mathbb{R}^n$ (i.e., não assumimos a princípio que seja $m = n$). A condição $m = n$ apareceria então na *tese* do teorema. Para demonstrar esse enunciado mais geral observamos que, sob as hipóteses do teorema, $\mathrm{d}f(x)$ é um isomorfismo de $\mathbb{R}^m$ para $\mathbb{R}^n$, o que implica (da Álgebra Linear!) que $m = n$.

O teorema acima nos diz duas coisas: em primeiro lugar, não há chance alguma de uma função diferenciável admitir uma inversa diferenciável se a sua diferencial em cada ponto não for um isomorfismo. Em segundo lugar, o teorema nos diz como calcular a diferencial da função inversa, *mas é necessário saber a priori que a função inversa é diferenciável*. Nosso objetivo agora é estabelecer condições suficientes para que a função inversa seja diferenciável (obviamente, devemos supor que $\mathrm{d}f(x)$ seja um isomorfismo, antes de mais nada — mas será que isso basta?). Depois, nosso problema será achar condições para que a inversibilidade de $\mathrm{d}f(x)$ implique na inversibilidade de f numa vizinhança de x.

Antes de provar o teorema seguinte, vamos motivar um pouco seu enunciado fazendo um estudo informal do caso de funções de uma variável. Seja então $f : I \to \mathbb{R}$ uma função diferenciável e injetora definida num intervalo $I \subset I\!R$. Suponha que $f'(x) \neq 0$ para algum $x \in I$ (isso corresponde à hipótese que $\mathrm{d}f(x)$ seja um isomorfismo; veja o Exercício 1). Escrevendo $\Delta y = f(x+\Delta x) - f(x)$ então a derivada $f'(x)$ é dada pelo limite $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$; a derivada de f^{-1} no ponto $y = f(x)$ é dada pelo limite $\lim_{\Delta y \to 0} \frac{\Delta x}{\Delta y}$. Como $\frac{\Delta y}{\Delta x}$ tende a $f'(x) \neq 0$ quando $\Delta x \to 0$, temos que $\frac{\Delta x}{\Delta y}$ tende a $\frac{1}{f'(x)}$, *mas quando* $\Delta x \to 0$ (e não, a princípio, quando $\Delta y \to 0$). Para concluir que f^{-1} é diferenciável no ponto $y = f(x)$, precisamos saber então que $\Delta y \to 0$ implica $\Delta x \to 0$; essa hipótese corresponde à *continuidade de* f^{-1} *no ponto* y. Temos então o seguinte:

Teorema. *("teoreminha" da função inversa) Sejam $U, V \subset I\!R^n$ abertos e $f : U \to V$ uma função bijetora; suponha que $f : U \to I\!R^n$ é diferenciável num ponto $x \in U$, que $\mathrm{d}f(x) : I\!R^n \to I\!R^n$ é um isomorfismo e que $f^{-1} : V \to I\!R^n$ é contínua no ponto $f(x) \in V$. Então f^{-1} é diferenciável no ponto $f(x)$ e $\mathrm{d}f^{-1}\big(f(x)\big) = \mathrm{d}f(x)^{-1}$.*

Demonstração. Como f é diferenciável no ponto x, podemos escrever:

$$f(x+h) = f(x) + \mathrm{d}f(x) \cdot h + \rho(h)\|h\|, \tag{1}$$

com $\lim_{h \to 0} \rho(h) = 0$. Tome agora $y = f(x)$ e escreva:

$$f^{-1}(y+k) = f^{-1}(y) + \mathrm{d}f(x)^{-1} \cdot k + \sigma(k)\|k\|; \tag{2}$$

a demonstração ficará completa se mostrarmos que $\lim_{k \to 0} \sigma(k) = 0$.

Para cada $k \in I\!R^n$ com $y + k \in V$ associamos o vetor $h = f^{-1}(y+k) - f^{-1}(y) \in I\!R^n$ (mais precisamente, definimos uma função $k \mapsto \phi(k) = f^{-1}(y+k) - f^{-1}(y)$ e escrevemos h em vez de $\phi(k)$). Resolvendo a igualdade $h = f^{-1}(y+k) - f^{-1}(y)$ para k, obtemos que ($x + h$ pertence a U e que) $k = f(x+h) - f(x)$. De (2) vem então:

$$h = \mathrm{d}f(x)^{-1} \cdot \big[f(x+h) - f(x)\big] + \sigma(k)\|k\|.$$

Usando (1) obtemos $f(x+h) - f(x) = \mathrm{d}f(x) \cdot h + \rho(h)\|h\|$ e daí:

$$h = h + \|h\| \, \mathrm{d}f(x)^{-1} \cdot \rho(h) + \sigma(k)\|k\| \Longrightarrow \sigma(k) = -\frac{\|h\|}{\|k\|} \, \mathrm{d}f(x)^{-1} \cdot \rho(h).$$

A continuidade de f^{-1} no ponto $y = f(x)$ implica que $h \to 0$ quando $k \to 0$, i.e., que $\lim_{k \to 0} f^{-1}(y+k) - f^{-1}(y) = 0$. Como $\lim_{h \to 0} \rho(h) = 0$ e como $\mathrm{d}f(x)^{-1}$ é contínua (pois é linear), segue que também $\lim_{k \to 0} \mathrm{d}f(x)^{-1} \cdot \rho(h) = 0$. Para completar a demonstração, é suficiente então mostrar que o quociente:

$$\frac{\|h\|}{\|k\|} = \frac{\|h\|}{\big\|f(x+h) - f(x)\big\|}$$

é limitado para k numa vizinhança da origem, ou seja, que $\frac{\|f(x+h) - f(x)\|}{\|h\|}$ fica longe de zero para k numa vizinhança da origem. Como a aplicação linear $\mathrm{d}f(x)$ é contínua e não se anula na esfera unitária de $I\!R^n$ (que é compacta), temos:

$$c = \inf_{\|u\|=1} \big\|\mathrm{d}f(x) \cdot u\big\| = \min_{\|u\|=1} \big\|\mathrm{d}f(x) \cdot u\big\| > 0;$$

como $\lim_{h\to 0}\rho(h)=0$, temos que $\big\|\rho(h)\big\| < \frac{c}{2}$ para h em alguma vizinhança Z da origem em $\mathbb{R}^n$. Daí, usando (1), temos que para $h\in Z$:

$$\big\|f(x+h)-f(x)\big\| = \big\|\mathrm{d}f(x)\cdot h+\rho(h)\|h\|\big\| \geq \big\|\mathrm{d}f(x)\cdot h\big\|-\big\|\rho(h)\big\|\|h\| \geq c\|h\|-\frac{c}{2}\|h\| = \frac{c}{2}\|h\|.$$

Como $\lim_{k\to 0} h = \lim_{k\to 0} f^{-1}(y+k)-f^{-1}(y)=0$, temos que $h\in Z$ para k em alguma vizinhança da origem em $\mathbb{R}^n$; nessa vizinhança, $\frac{\|h\|}{\|k\|}\leq\frac{2}{c}$, i.e., a quantidade $\frac{\|h\|}{\|k\|}$ é limitada. Isso completa a demonstração. ■

`Observação`. Se $I\subset\mathbb{R}$ é um intervalo e $f:I\to\mathbb{R}$ é uma função contínua e injetora então sabe-se que $f(I)$ é um intervalo (pelo Teorema do Valor Intermediário) e que a função inversa $f^{-1}:f(I)\to\mathbb{R}$ também é contínua (isso segue dos seguintes resultados da Análise Real: (i) *toda função contínua e injetora num intervalo é monótona*; (ii) *toda função monótona cuja imagem é um intervalo é contínua*; veja *Curso de Análise*, vol. I, Elon Lages Lima, pg. 186). Daí, a continuidade da inversa no "teoreminha da função inversa" é uma hipótese redundante se $n=1$, U é um intervalo e f é contínua em U. Na verdade, se $U\subset\mathbb{R}^n$ é um aberto e $f:U\to\mathbb{R}^n$ é uma função contínua e injetora então $f(U)$ é aberto em $\mathbb{R}^n$ e $f:U\to f(U)$ é um homeomorfismo, i.e., $f^{-1}:f(U)\to\mathbb{R}^n$ é contínua. Isso segue do famoso *Teorema da Invariância do Domínio* — tal teorema é bem mais sofisticado que o resultado de Análise Real mencionado acima. Sua demonstração é normalmente feita utilizando recursos da Topologia Algébrica (mais precisamente, utiliza-se recursos da teoria de *Homologia Singular* de um espaço topológico). Na demonstração do teorema da função inversa adiante, mostraremos a continuidade da função inversa através de recursos muito mais elementares: usaremos no entanto a hipótese mais forte que a função f seja de classe C^1 e que sua diferencial em cada ponto seja um isomorfismo.

`Não-Exemplo`. A função $f:\mathbb{R}\to\mathbb{R}$ dada por $f(x)=x^3$ é um homeomorfismo de classe C^∞ mas $f'(0)=0$, i.e., $\mathrm{d}f(0)$ não é um isomorfismo. Obviamente então f^{-1} não é diferenciável na origem.

A seguinte definição facilita o enunciado do teorema da função inversa:

Definição. *Sejam $U,V\subset\mathbb{R}^n$ abertos e $f:U\to V$ uma função. Dizemos que f é um* difeomorfismo *quando f é bijetora e tanto $f:U\to\mathbb{R}^n$ como $f^{-1}:V\to\mathbb{R}^n$ são diferenciáveis.*

Na prática estaremos mais interessados em bijeções de classe C^k com inversas de classe C^k. Temos o seguinte:

Teorema. *Sejam $U,V\subset\mathbb{R}^n$ abertos e $f:U\to V$ um difeomorfismo. Se $f:U\to\mathbb{R}^n$ é de classe C^k $(0\leq k\leq\infty)$ então também $f^{-1}:V\to\mathbb{R}^n$ é de classe C^k.*

Demonstração. Seja $\mathrm{GL}(n,\mathbb{R})$ o conjunto das aplicações lineares inversíveis de $\mathbb{R}^n$ em $\mathbb{R}^n$ (esse é chamado o *grupo linear geral* de $\mathbb{R}^n$; recorde o complemento à Aula número 4 — 15/03). Temos que $\mathrm{GL}(n,\mathbb{R})$ é um aberto no espaço $\mathrm{Lin}(\mathbb{R}^n,\mathbb{R}^n)$ de todos as aplicações lineares de $\mathbb{R}^n$ em $\mathbb{R}^n$, pois $T\in\mathrm{GL}(n,\mathbb{R})$ se e somente se $\det(T)\neq 0$ e o determinante é uma função contínua em $\mathrm{Lin}(\mathbb{R}^n,\mathbb{R}^n)$ (pois é um polinômio nas entradas da matriz que representa T). Além do mais, a *função inversão* $\mathcal{I}:\mathrm{GL}(n,\mathbb{R})\to\mathrm{Lin}(\mathbb{R}^n,\mathbb{R}^n)$ dada por $\mathcal{I}(T) = T^{-1}$ é de classe C^∞, pois as entradas da matriz que representa T^{-1} são dadas por quocientes de polinômios nas entradas da matriz que representa T (recorde que a inversa de uma matriz quadrada A é dada pela transposta da matriz dos cofatores de A, dividida por $\det(A)$). Se $f : U \to V$ é um difeomorfismo então sabemos que $\mathrm{d}f^{-1}\big(f(x)\big) = \mathrm{d}f(x)^{-1} = \mathcal{I}\big(\mathrm{d}f(x)\big)$ para todo $x \in U$, ou seja, $\mathrm{d}f^{-1}(y) = \mathcal{I}\big[\mathrm{d}f\big(f^{-1}(y)\big)\big]$ para todo $y \in V$. Temos então:

$$\mathrm{d}f^{-1} = \mathcal{I} \circ \mathrm{d}f \circ f^{-1}.$$

Usando indução em k, a igualdade acima e o fato que $\mathcal{I}$ é de classe C^∞ segue agora facilmente que se f é um difeomorfismo de classe C^k então também f^{-1} é de classe C^k. ■

O teorema da função inversa afirma, entre outras coisas, que uma função satisfazendo certas hipóteses possue como imagem um aberto em $\mathbb{R}^n$. Na verdade, essa é a parte mais difícil de demonstrar. De fato, mostrar que a imagem de uma função f é um conjunto aberto envolve um problema de existência de soluções de equações: deve-se mostrar que para y perto de $f(a)$ a equação $f(x) = y$ admite solução. Essa parte do argumento fará parte da demonstração do *Lema da Perturbação da Identidade* que será mostrado logo adiante. Antes de mais nada, no entanto, precisamos de algum princípio de existência de soluções de equações. Esse é o nosso próximo passo.

Definição. *Sejam M um conjunto e $\phi : M \to M$ uma função. Dizemos que $x \in M$ é um* ponto fixo *de ϕ se $\phi(x) = x$.*

Tipicamente equações da forma $f(x) = y$ (com $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$) podem ser transformadas em problemas de ponto fixo: basta escrever $f(x) = y$ sob a forma equivalente $f(x) + x - y = x$. Daí resolver $f(x) = y$ corresponde a determinar um ponto fixo para $\phi(x) = f(x) + x - y$. Essa observação indica a importância do seguinte:

Teorema. *(do ponto fixo de Banach) Sejam (M, d) um espaço métrico completo (não vazio) e $\phi : M \to M$ uma contração. Então ϕ tem um único ponto fixo. Além do mais, para todo $x_0 \in M$, se definirmos uma seqüência $(x_n)_{n \geq 0}$ em M recursivamente fazendo $x_n = \phi(x_{n-1})$, $n \geq 1$, então $(x_n)_{n \geq 0}$ converge para o único ponto fixo de ϕ.*

Demonstração. Em primeiro lugar, é óbvio que ϕ pode ter no máximo um ponto fixo, já que ϕ contrai distâncias: se $x, y \in M$ são pontos fixos de ϕ e $\lambda < 1$ é uma constante de Lipschitz para ϕ então:

$$d(x,y) = d\big(\phi(x), \phi(y)\big) \leq \lambda\, d(x,y) \Longrightarrow 0 \leq (1-\lambda) d(x,y) \leq 0 \Longrightarrow d(x,y) = 0,$$

e portanto $x = y$.

Para completar a demonstração, seja $x_0 \in M$ um ponto qualquer e defina $(x_n)_{n \geq 0}$ recursivamente como no enunciado do teorema. Só precisamos mostrar que $(x_n)_{n \geq 1}$ é uma seqüência de Cauchy; de fato, nesse caso deve existir $x = \lim_{n \to +\infty} x_n$, pois M é completo e além do mais, como ϕ é contínua:

$$\lim_{n \to +\infty} x_n = x \Longrightarrow \lim_{n \to +\infty} \underbrace{\phi(x_n)}_{x_{n+1}} = \phi(x) = x = \lim_{n \to +\infty} x_{n+1},$$

i.e., x é um ponto fixo de ϕ.

Para mostrar que $(x_n)_{n\geq 0}$ é uma seqüência de Cauchy, devemos estimar a distância $d(x_n, x_m)$; começamos estimando $d(x_n, x_{n+1})$. Temos:

$$d(x_n, x_{n+1}) \leq \lambda\, d(x_{n-1}, x_n) \leq \lambda^2 d(x_{n-2}, x_{n-1}) \leq \cdots \leq \lambda^n d(x_0, x_1);$$

daí, para $n < m$:

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &\leq \big(\lambda^n + \lambda^{n+1} + \cdots + \lambda^{m-1}\big) d(x_0, x_1) \\ &\leq \left(\sum_{k=n}^{+\infty} \lambda^k\right) d(x_0, x_1) = \frac{\lambda^n}{1-\lambda}\, d(x_0, x_1); \end{aligned}$$

como $\lim_{n\to+\infty} d(x_0, x_1)\frac{\lambda^n}{1-\lambda} = 0$, a conclusão segue. ■

`Observação`. O teorema acima é um dos grandes "princípios de existência" da matemática. Quando aplicado para operadores integrais em espaços de funções ele produz o famoso *teorema de existência e unicidade para soluções de equações diferenciais*. O teorema acima também é a base matemática para diversos métodos de Cálculo Numérico, como o *Método de Newton para obtenção de raízes de funções* $f : I\!R \to I\!R$.

A idéia principal da demonstração do teorema da função inversa é a seguinte: se f é diferenciável e $\mathrm{d}f(x)$ é um isomorfismo então, numa vizinhança do ponto x, a função f é uma "perturbação pequena" do isomorfismo $\mathrm{d}f(x)$. Queremos então estudar o comportamento de uma perturbação pequena de um isomorfismo $T : I\!R^n \to I\!R^n$; começamos com o caso em que T é igual a identidade no seguinte:

Lema. *(da perturbação da identidade) Se $U \subset I\!R^n$ é um aberto e $g : U \to I\!R^n$ é uma contração então a aplicação $f : U \to I\!R^n$ definida por $f(x) = x + g(x)$, $x \in U$, é um homeomorfismo sobre sua imagem $f(U)$; além do mais, $f(U)$ é um aberto em $I\!R^n$.*

Demonstração. Se $\lambda < 1$ é uma constante de Lipschitz para g então, para todos $x, y \in U$:

$$\begin{aligned} \big\|f(x) - f(y)\big\| &= \big\|x - y + g(x) - g(y)\big\| \geq \|x - y\| - \big\|g(x) - g(y)\big\| \\ &\geq \|x - y\| - \lambda\|x - y\| = (1-\lambda)\|x - y\|. \end{aligned}$$

Como $1 - \lambda > 0$, as desigualdades acima implicam que f é injetora e que sua inversa $f^{-1} : f(U) \to U$ é Lipschitziana com constante $\frac{1}{1-\lambda}$; em particular, f^{-1} é contínua.

Na verdade, a parte não trivial da demonstração é mostrar que $f(U)$ é aberto em $I\!R^n$. Para isso, considere um ponto arbitrário de $f(U)$, digamos $f(a)$ com $a \in U$. Devemos mostrar que se $\big\|y - f(a)\big\|$ é suficientemente pequeno então a equação $f(x) = y$ admite uma solução $x \in U$. Fixe então $y \in I\!R^n$ e vamos tentar determinar condições suficientes sobre y para que $f(x) = y$ tenha solução. Temos:

$$f(x) = y \Longleftrightarrow -f(x) = -y \Longleftrightarrow x - f(x) = x - y \Longleftrightarrow x - f(x) + y = x,$$

i.e., transformamos a equação $f(x) = y$ num problema de ponto fixo! Explicitamente, definindo $h : U \to I\!R^n$ por $h(x) = x - f(x) + y = y - g(x)$ então $f(x) = y$ equivale a $h(x) = x$. Do fato que g é uma contração, segue trivialmente que h também o é — mas cuidado, isso ainda não resolveu o problema! O teorema do ponto fixo de Banach só se aplica para contrações definidas num espaço métrico *completo* e *tomando valores nele mesmo*. Devemos determinar então um subconjunto fechado $M \subset I\!R^n$ contido em U e tal que $h(M) \subset M$ (recorde que, como $I\!R^n$ é completo, um subespaço $M \subset I\!R^n$ é completo se e somente se for um subconjunto fechado). Como no final das contas deveremos escolher y perto de $f(a)$, espera-se que a solução x de $f(x) = y$ esteja perto de a e portanto um candidato natural para M é uma bola fechada centrada em a (contida em U). Vamos então seguir esta estratégia.

Escolha $r_0 > 0$ tal que $\mathrm{B}[a; r_0] \subset U$. Queremos determinar condições suficientes sobre y que tornem possível a escolha de $r > 0$, $r \le r_0$, com a seguinte propriedade: se $\|x-a\| \le r$ então $\big\|h(x) - a\big\| \le r$.

Calculamos então:

$$\begin{aligned}\big\|h(x) - a\big\| &= \big\|y - g(x) - a\big\| = \big\|y - f(a) + g(a) - g(x)\big\| \\ &\le \big\|y - f(a)\big\| + \big\|g(a) - g(x)\big\| \\ &\le \big\|y - f(a)\big\| + \lambda\|x - a\|.\end{aligned}$$

Para concluir que $\big\|h(x) - a\big\| \le r$ (sob $\|x - a\| \le r$), precisamos então da desigualdade:

$$\big\|y - f(a)\big\| \le (1 - \lambda)r;$$

a desigualdade acima será satisfeita se escolhermos $r \ge \frac{1}{1-\lambda}\big\|y - f(a)\big\|$. A restrição $r \le r_0$ nos obriga então a supor:

$$\big\|y - f(a)\big\| \le (1 - \lambda)r_0.$$

Agora tudo se encaixa: se começamos com $y \in I\!R^n$ satisfazendo a desigualdade acima então, definindo $r = r_0$, obteremos que $\|x - a\| \le r$ implica $\big\|h(x) - a\big\| \le r$, i.e., que $h\big(\mathrm{B}[a; r]\big) \subset \mathrm{B}[a; r]$. Aplicando então o teorema do ponto fixo de Banach para a contração h no espaço métrico completo $\mathrm{B}[a; r]$, obtemos $x \in \mathrm{B}[a; r] \subset U$ com $h(x) = x$, i.e., com $f(x) = y$. Juntando tudo vem:

$$\mathrm{B}\big[f(a); (1 - \lambda)r_0\big] \subset f(U),$$

e portanto $f(a)$ é um ponto interior de $f(U)$. ■

Corolário. *(perturbação de um isomorfismo) Sejam $T : I\!R^n \to I\!R^n$ um isomorfismo linear, $U \subset I\!R^n$ um aberto e $g : U \to I\!R^n$ uma função Lipschitziana com constante de Lipschitz menor que $\big\|T^{-1}\big\|^{-1}$ (onde $\|\cdot\|$ denota a norma de operadores). Então, definindo uma aplicação $f : U \to I\!R^n$ por $f(x) = T(x) + g(x)$, $x \in U$, temos que $f : U \to f(U)$ é um homeomorfismo e que $f(U)$ é aberto em $I\!R^n$.*

Demonstração. Temos $f(x) = T\big(x + (T^{-1} \circ g)(x)\big)$, para todo $x \in U$, ou seja:

$$f = T \circ \big(\mathrm{Id}|_U + T^{-1} \circ g\big).$$

Como T é um homeomorfismo de $\mathbb{R}^n$ sobre $\mathbb{R}^n$, a conclusão seguirá do lema de perturbação da identidade se mostrarmos que $T^{-1} \circ g$ é uma contração. Mas isso é muito simples; temos:

$$\big\|T^{-1}\big(g(x)\big) - T^{-1}\big(g(y)\big)\big\| \le k\big\|T^{-1}\big\|\|x - y\|,$$

para todos $x, y \in U$, onde $k < \big\|T^{-1}\big\|^{-1}$ é uma constante de Lipschitz para g. ■

Lema. *Seja $f : U \to \mathbb{R}^n$ uma função de classe C^1 definida num aberto $U \subset \mathbb{R}^n$. Então o conjunto dos pontos $x \in U$ para os quais $\mathrm{d}f(x) : \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo é aberto em $\mathbb{R}^n$.*

Demonstração. O conjunto mencionado no enunciado é nada mais que a imagem inversa do aberto $\mathrm{GL}(n, \mathbb{R})$ em $\mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$ pela função contínua $\mathrm{d}f : U \to \mathrm{Lin}(\mathbb{R}^n, \mathbb{R}^n)$. Segue que esse conjunto é aberto em U e portanto aberto em $\mathbb{R}^n$. ■

Finalmente, estamos em condição de demonstrar o:

Teorema. *(da função inversa) Seja $f : U \to \mathbb{R}^n$ uma função de classe C^k $(1 \le k \le \infty)$ definida num aberto $U \subset \mathbb{R}^n$. Se para um certo $x_0 \in U$ temos que $\mathrm{d}f(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo então existe uma vizinhança aberta V de x_0 em $\mathbb{R}^n$ contida em U tal que $f|_V$ é injetora, $f(V)$ é aberto em $\mathbb{R}^n$ e $f|_V : V \to f(V)$ é um difeomorfismo; em particular, a função inversa $(f|_V)^{-1} : f(V) \to \mathbb{R}^n$ também é de classe C^k.*

Demonstração. É suficiente achar uma vizinhança aberta W de x_0 em U tal que $f(W)$ é aberto em $\mathbb{R}^n$ e $f|_W : W \to f(W)$ é um homeomorfismo. De fato, uma vez encontrada W, definimos $V = \big\{x \in W : \mathrm{d}f(x) \text{ é um isomorfismo}\big\}$ e daí $V \subset \mathbb{R}^n$ é aberto, $f(V) \subset \mathbb{R}^n$ é aberto e $f|_V : V \to f(V)$ é um homeomorfismo; segue do "teoreminha" da função inversa que $(f|_V)^{-1} : f(V) \to \mathbb{R}^n$ é diferenciável e portanto de classe C^k (já que $f|_V$ é de classe C^k). Para encontrar W, vamos usar o princípio da perturbação de um isomorfismo por funções com constante de Lipschitz pequenas. Defina $g : U \to \mathbb{R}^n$ satisfazendo a igualdade:

$$f(x) = \mathrm{d}f(x_0) \cdot x + g(x),$$

para todo $x \in U$; temos então:

$$\mathrm{d}g(x) = \mathrm{d}f(x) - \mathrm{d}f(x_0).$$

Como $\mathrm{d}f$ é contínua, em alguma vizinhança aberta W de x_0 em U teremos:

$$\big\|\mathrm{d}g(x)\big\| \le \frac{1}{2}\big\|\mathrm{d}f(x_0)^{-1}\big\|^{-1}$$

e daí (escolhendo W convexa), podemos aplicar a desigualdade do valor médio para concluir que $g|_W$ é Lipschitziana com constante de Lipschitz menor que $\big\|\mathrm{d}f(x_0)^{-1}\big\|^{-1}$. A conclusão segue. ■

Definição. *Uma aplicação $f : M \to N$ entre espaços métricos M, N é dita* aberta *quando f leva abertos de M em abertos de N, i.e., se para todo $U \subset M$ aberto em M temos que $f(U)$ é aberto em N. Similarmente, dizemos que f é* fechada *quando f leva fechados de M em fechados de N.*

Definição. *Seja $U \subset \mathbb{R}^n$ um aberto. Uma aplicação $f : U \to \mathbb{R}^n$ é dita um* difeomorfismo local *quando todo $x \in U$ admite uma vizinhança aberta $V \subset \mathbb{R}^n$ contida em U tal que $f(V)$ é um aberto de $\mathbb{R}^n$ e $f|_V : V \to f(V)$ é um difeomorfismo.*

É fácil ver que todo difeomorfismo local $f : U \to \mathbb{R}^n$ é uma aplicação aberta. Além do mais, se $f : U \to \mathbb{R}^n$ é um difeomorfismo local injetor então $f(U)$ é aberto em $\mathbb{R}^n$ e $f : U \to f(U)$ é um difeomorfismo (veja também o Exercício 2).

Corolário. *Seja $f : U \to \mathbb{R}^n$ uma aplicação de classe C^k $(1 \le k \le \infty)$ definida num aberto $U \subset \mathbb{R}^n$. Se $\mathrm{d}f(x)$ é um isomorfismo para todo $x \in U$ então f é um difeomorfismo local. Em particular, f é uma aplicação aberta e $f(U)$ é um aberto em $\mathbb{R}^n$. Se f for injetora então $f : U \to f(U)$ é um difeomorfismo (com inverso $f^{-1} : f(U) \to \mathbb{R}^n$ de classe C^k).* ■

`Exemplo`. O exemplo padrão de um difeomorfismo local não injetor é a "mudança das coordenadas polares para as cartesianas", i.e., a aplicação $f : (0, +\infty) \times \mathbb{R} \to \mathbb{R}^2$ definida por $f(r, \vartheta) = (r \cos \vartheta, r \operatorname{sen} \vartheta)$. Temos que f é de classe C^∞ e que a matriz Jacobiana de f no ponto (r, ϑ) tem determinante igual a $r \ne 0$ para todo $(r, \vartheta) \in (0, +\infty) \times \mathbb{R}$. Segue do teorema da função inversa que f é de fato um difeomorfismo local; obviamente f não pode ser injetora pois $f(r, \vartheta) = f(r, \vartheta + 2\pi)$ para todos r, ϑ. Por outro lado, dado $\theta_0 \in \mathbb{R}$ a restrição de f à faixa aberta $(0, +\infty) \times (\theta_0, \theta_0 + 2\pi)$ é um difeomorfismo de classe C^∞ sobre o aberto A_{θ_0} obtido de $\mathbb{R}^2$ removendo a semi-reta fechada $\big\{(t \cos \theta_0, t \operatorname{sen} \theta_0) : t \ge 0\big\}$.

(2) **Derivadas parciais "gordas".**

Definição. *Sejam $U \subset \mathbb{R}^m$ um aberto, $f : U \to \mathbb{R}^n$ uma função, $x \in U$ um ponto e $S \subset \mathbb{R}^m$ um subespaço vetorial. Dizemos que f é* diferenciável no ponto x na direção de S *quando existe uma aplicação linear $T : S \to \mathbb{R}^n$ tal que, definindo uma aplicação r pela fórmula:*

$$f(x + h) = f(x) + T(h) + r(h), \quad h \in S,\ x + h \in U,$$

então $\lim_{h \to 0} \frac{r(h)}{\|h\|} = 0$ (observe que r é definida numa vizinhança aberta da origem em *S!).*

É fácil ver que a aplicação linear $T : S \to \mathbb{R}^n$ considerada na definição acima é única quando existe (pois, para todo $v \in S$, $T(v)$ coincide com a derivada direcional $\frac{\partial f}{\partial v}(x)$). A aplicação linear $T : S \to \mathbb{R}^n$ é chamada então a *derivada direcional de f no ponto x na direção do subespaço S* e é denotada por $\frac{\partial f}{\partial S}(x)$ ou simplesmente por $\partial_S f(x)$.

`Observação`. Sejam $f : U \to \mathbb{R}^n$ uma aplicação definida num aberto $U \subset \mathbb{R}^m$, $x \in U$ um ponto e $S \subset \mathbb{R}^m$ um subespaço. Defina um subconjunto $\widehat{U} \subset S$ e uma aplicação $\hat{f} : \widehat{U} \to \mathbb{R}^n$ fazendo:

$$\widehat{U} = \big\{w \in S : x + w \in U\big\}, \qquad \hat{f}(w) = f(x + w), \quad w \in \widehat{U};$$

temos que $\widehat{U}$ é aberto em S (pois é a pré-imagem do aberto $U \subset \mathbb{R}^m$ pela aplicação contínua $S \ni w \mapsto x + w \in \mathbb{R}^m$) e $0 \in \widehat{U}$. É fácil ver que *f é diferenciável no ponto x na direção de S se e somente se $\widehat{f}$ é diferenciável no ponto* $0 \in \widehat{U}$; nesse caso, a derivada direcional de f no ponto x na direção de S coincide com a diferencial de $\hat{f}$ no ponto 0, i.e., $\frac{\partial f}{\partial S}(x) = \mathrm{d}\hat{f}(0)$.

`Observação`. Quando $S \subset I\!R^m$ é um subespaço unidimensional (digamos, gerado por um vetor $v \in I\!R^m$) então a existência de $\frac{\partial f}{\partial S}(x)$ é equivalente à existência da derivada direcional usual $\frac{\partial f}{\partial v}(x)$; além do mais, a aplicação linear $\frac{\partial f}{\partial S}(x)$ definida em S satisfaz $\frac{\partial f}{\partial S}(x) \cdot v = \frac{\partial f}{\partial v}(x)$. Nesse sentido a definição de derivada na direção de um subespaço dada nesta seção generaliza a noção usual de derivada direcional.

`Exemplo`. Quando $f : U \to I\!R^n$ ($U \subset I\!R^m$ aberto) é diferenciável num ponto $x \in U$ então f é diferenciável no ponto x na direção de *qualquer* subespaço $S \subset I\!R^m$ e a derivada direcional de f no ponto x na direção de S é simplesmente a restrição da diferencial total $\mathrm{d}f(x)$ a S, ou seja:

$$\frac{\partial f}{\partial S}(x) = \mathrm{d}f(x)|_S.$$

Esse é o caso que de fato nos interessa.

Suponha agora que tenhamos escrito $I\!R^m$ como uma soma direta de subespaços $V_1, \ldots, V_k \subset I\!R^m$, ou seja, $I\!R^m = \bigoplus_{i=1}^k V_i$ (recorde que isso significa que todo vetor $v \in I\!R^m$ se escreve de modo único na forma $v = \sum_{i=1}^k v_i$ com cada $v_i \in V_i$; veja Exercícios 2–6, aula número 9 (03/04), para detalhes). Se $U \subset I\!R^m$ é um aberto e $f : U \to I\!R^n$ é uma função então podemos pensar em f como uma função de k variáveis, identificando cada k-upla $(x_1, \ldots, x_k) \in \prod_{i=1}^k V_i$ com o vetor $x = \sum_{i=1}^k x_i \in I\!R^m$. Para $x \in U$, as derivadas direcionais $\frac{\partial f}{\partial V_i}(x) \in \mathrm{Lin}(V_i, I\!R^n)$, $i = 1, \ldots, k$ (se existirem) são chamadas as *derivadas parciais "gordas"* de f com respeito à decomposição em soma direta $I\!R^m = \bigoplus_{i=1}^k V_i$; uma vez subentendida a decomposição $I\!R^m = \bigoplus_{i=1}^k V_i$ escrevemos:

$$\partial_i f(x) = \frac{\partial f}{\partial V_i}(x) \in \mathrm{Lin}(V_i, I\!R^n), \quad i = 1, \ldots, k.$$

É fácil ver que a derivada parcial "gorda" $\partial_i f(x)$ coincide com a diferencial no ponto $x_i \in V_i$ da aplicação $v_i \mapsto f(x_1 + \cdots + x_{i-1} + v_i + x_{i+1} + \cdots + x_k)$ definida num aberto de V_i, i.e., $\partial_i f(x)$ coincide com a diferencial no ponto $x_i \in V_i$ da aplicação obtida de f quando "deixamos livre" apenas a i-ésima variável.

Se f é diferenciável no ponto $x \in U$ então, como já observamos:

$$\partial_i f(x) = \frac{\partial f}{\partial V_i}(x) = \mathrm{d}f(x)|_{V_i},$$

para todo $i = 1, \ldots, k$. Segue então da linearidade de $\mathrm{d}f(x)$ que:

$$\mathrm{d}f(x) \cdot h = \sum_{i=1}^k \partial_i f(x) \cdot h_i,$$

onde $h = \sum_{i=1}^k h_i$ e $h_i \in V_i$, $i = 1, \ldots, k$.

`Observação`. O nome "derivada parcial gorda" *não* é oficial, apesar do fato que o conceito em si é padrão (é comum usar apenas o termo "derivada parcial", sem distinção em relação ao nome usado para as derivadas parciais usuais).

`Exemplo`. Se $\mathbb{R}^m = \bigoplus_{i=1}^k V_i$ e cada V_i é gerado por alguns dos vetores da base canônica de $\mathbb{R}^m$ então a derivada parcial "gorda" $\partial_i f(x)$ é a aplicação linear de V_i em $\mathbb{R}^n$ cuja representação matricial com respeito às bases canônicas é a matriz cujas colunas contém as derivadas parciais de f correspondentes aos vetores da base canônica de $\mathbb{R}^m$ que pertencem a V_i. Observe então que a derivada parcial "gorda" $\partial_i f(x)$ é simplesmente um aglomerado de derivadas parciais comuns; se $k = m$ e cada V_i é gerado somente pelo i-ésimo vetor da base canônica então as derivadas parciais "gordas" podem ser identificadas com as derivadas parciais usuais $\frac{\partial f}{\partial x_i}(x)$.

Em certas situações é também interessante considerar decomposições em soma direta do *contra-domínio* de uma função $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$. Suponha então que $\mathbb{R}^n = \bigoplus_{i=1}^r W_i$. Denote por $\pi_i : \mathbb{R}^n \to W_i$ a i-ésima aplicação de projeção e por $\iota_i : W_i \to \mathbb{R}^n$ a inclusão de W_i em $\mathbb{R}^n$. A aplicação $f_i : U \to W_i$ definida por $f_i = \pi_i \circ f$ é chamada a *i-ésima coordenada* de f com respeito à decomposição $\mathbb{R}^n = \bigoplus_{i=1}^r W_i$. Das igualdades:

$$f = \sum_{i=1}^r \iota_i \circ f_i, \qquad f_i = \pi_i \circ f, \quad i = 1, \ldots, r,$$

segue que f é diferenciável num ponto $x \in U$ se e somente se cada f_i é diferenciável em x e que f é de classe C^k se e somente se cada f_i é de classe C^k; quando f é diferenciável em x segue da regra da cadeia que $\mathrm{d}f_i(x) = \pi_i \circ \mathrm{d}f(x)$ ("a diferencial da i-ésima coordenada é a i-ésima coordenada da diferencial").

`Observação`. Em tudo que estudamos até agora com respeito a Cálculo no $\mathbb{R}^n$, pode-se substituir os espaços Euclideanos $\mathbb{R}^n$ por espaços vetorais reais de dimensão finita arbitrários munidos de normas arbitrárias — em geral nos restringimos a $\mathbb{R}^n$ na exposição apenas por "razões psicológicas". Observamos que na discussão acima, as decomposições em soma direta $\mathbb{R}^m = \bigoplus_{i=1}^k V_i$ e $\mathbb{R}^n = \bigoplus_{i=1}^r W_i$ não estão em princípio *de forma alguma relacionadas com as bases canônicas de $\mathbb{R}^m$ e $\mathbb{R}^n$*; seria então até mais simples visualizar a teoria utlizando espaços vetoriais reais de dimensão finita arbitrários no lugar de $\mathbb{R}^m$ e $\mathbb{R}^n$.

`Exemplo`. (*matrizes de blocos*) Sejam V, W espaços vetoriais reais de dimensão finita e suponha que sejam dadas decomposições em soma direta $V = \bigoplus_{j=1}^k V_j$, $W = \bigoplus_{i=1}^r W_i$. Se $T : V \to W$ é uma transformação linear então para cada $i = 1, \ldots, r$, $j = 1, \ldots, k$ obtemos uma transformação linear $T_{ij} = \pi_i \circ T|_{V_j} \in \mathrm{Lin}(V_j, W_i)$, onde $\pi_i : W \to W_i$ denota a aplicação de projeção na i-ésima coordenada. Reciprocamente, se para cada $i = 1, \ldots, r$, $j = 1, \ldots, k$ for dada uma transformação linear $T_{ij} : V_j \to W_i$ então *existe uma única transformação linear $T : V \to W$ tal que $\pi_i \circ T|_{V_j} = T_{ij}$ para todos i, j*; explicitamente, T é dada por:

$$T(v) = \sum_{i=1}^r \sum_{j=1}^k T_{ij}(v_j),$$

onde $v = \sum_{j=1}^k v_j$ e $v_j \in V_j$, $j = 1, \ldots, k$. As transformações lineares $T_{ij} \in \mathrm{Lin}(V_j, W_i)$ são chamadas as *componentes* de T com respeito às decomposições $V = \bigoplus_{j=1}^k V_j$ e $W = \bigoplus_{i=1}^r W_i$.

Escolha agora uma base em cada espaço V_j e uma base em cada espaço W_i. Concatenando as bases dos espaços V_j obtemos uma base de V e concatenando as bases dos espaços W_i obtemos uma base de W. Se $T \in \mathrm{Lin}(V, W)$ é tal que $\pi_i \circ T|_{V_j} = T_{ij} \in \mathrm{Lin}(V_j, W_i)$ e se A_{ij} é a matriz que representa T_{ij} com respeito às bases fixadas em V_j e W_i então a matriz A que representa T com respeito às bases fixadas em V e W é obtida por *concatenação dos blocos* A_{ij}; escrevemos:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rk} \end{pmatrix}.$$

A descrição dada acima para a matriz A é normalmente conhecida como uma descrição por *matrizes de blocos* (*block-matrices*); a idéia é que descrevemos a matriz A através de submatrizes menores, sem individualizar suas entradas. A grande vantagem da notação de matriz de blocos é a sua compatibilidade com a regra usual de multiplicação de matrizes; de fato, dadas matrizes de blocos:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rk} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1s} \\ B_{21} & B_{22} & \cdots & B_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ B_{k1} & B_{k2} & \cdots & B_{ks} \end{pmatrix},$$

então o produto AB é dado por:

$$AB = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1s} \\ C_{21} & C_{22} & \cdots & C_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ C_{r1} & C_{r2} & \cdots & C_{rs} \end{pmatrix},$$

onde $C_{ij} = \sum_{\lambda=1}^k A_{i\lambda} B_{\lambda j}$ para todos $i = 1, \ldots, r$, $j = 1, \ldots, s$ (*note bem*: não se pode confundir $A_{i\lambda} B_{\lambda j}$ com $B_{\lambda j} A_{i\lambda}$ aqui!).

Se f é uma função definida num aberto de $V = \bigoplus_{j=1}^k V_j$, tomando valores em $W = \bigoplus_{i=1}^r W_i$, diferenciável num ponto x de seu domínio, então $\mathrm{d}f(x) : V \to W$ é uma aplicação linear e sua componente $\pi_i \circ \mathrm{d}f(x)|_{V_j}$ é exatamente a derivada parcial "gorda" $\partial_j f_i(x)$ da i-ésima coordenada $f_i = \pi_i \circ f$ de f. Fixando bases nos espaços V_j, W_i e concatenando-as para obter bases em V e W respectivamente, vemos então que a matriz Jacobiana de f no ponto x (com respeito às bases escolhidas) pode ser descrita como uma matriz de blocos formada pelas matrizes que representam as derivadas parciais "gordas" $\partial_j f_i(x)$.

(3) **O teorema da função implícita**.

O teorema da função implícita fornece condições suficientes para que seja possível resolver para y uma igualdade do tipo "$f(x,y) =$ constante" obtendo (localmente) uma função diferenciável $y = \phi(x)$.

Teorema. *(da função implícita) Seja $f : U \to \mathbb{R}^n$ uma função de classe C^k ($1 \le k \le \infty$) definida num aberto U do produto $\mathbb{R}^m \times \mathbb{R}^n$ e seja $(x_0, y_0) \in U$; escreva $c = f(x_0, y_0)$. Suponha que a derivada parcial "gorda" $\partial_2 f(x_0, y_0) : \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo. Então existem uma vizinhança aberta V de x_0 em $\mathbb{R}^m$, uma vizinhança aberta W de y_0 em $\mathbb{R}^n$ e uma função $\phi : V \to \mathbb{R}^n$ de classe C^k com $\phi(V) \subset W$, $V \times W \subset U$ e tal que para todos $(x, y) \in V \times W$ temos $f(x, y) = c$ se e somente se $y = \phi(x)$. Além do mais, a diferencial da função ϕ no ponto x_0 é dada por:*

$$\mathrm{d}\phi(x_0) = -\partial_2 f(x_0, y_0)^{-1} \circ \partial_1 f(x_0, y_0).$$

Demonstração. Uma vez estabelecida a existência de ϕ, a dedução da fórmula para $\mathrm{d}\phi(x_0)$ é simples: diferenciamos a igualdade $f\big(x, \phi(x)\big) = c$ no ponto x_0 usando a regra da cadeia obtendo:

$$\mathrm{d}f(x_0, y_0) \circ \big(\mathrm{Id}, \mathrm{d}\phi(x_0)\big) = \partial_1 f(x_0, y_0) + \partial_2 f(x_0, y_0) \circ \mathrm{d}\phi(x_0) = 0,$$

já que $\phi(x_0) = y_0$. A conclusão segue.

Procedemos agora com a prova da existência de ϕ. Defina $G : U \to \mathbb{R}^m \times \mathbb{R}^n$ fazendo $G(x, y) = \big(x, f(x, y)\big)$, para todos $(x, y) \in U$. Temos que G é de classe C^k e sua diferencial no ponto $(x_0, y_0) \in U$ é dada por:

$$\mathrm{d}G(x_0, y_0) \cdot (h, k) = \big(h, \partial_1 f(x_0, y_0) \cdot h + \partial_2 f(x_0, y_0) \cdot k\big),$$

para todos $h \in \mathbb{R}^m$, $k \in \mathbb{R}^n$. Usando a fórmula acima e o fato que $\partial_2 f(x_0, y_0)$ é um isomorfismo vê-se facilmente que $\mathrm{d}G(x_0, y_0) : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ é um isomorfismo. Pelo teorema da função inversa, existe uma vizinhança aberta de (x_0, y_0) em U (que podemos tomar como sendo um produto $V_1 \times W$ de um aberto V_1 de $\mathbb{R}^m$ por um aberto W de $\mathbb{R}^n$) tal que $G(V_1 \times W)$ é aberto em $\mathbb{R}^m \times \mathbb{R}^n$ e $G|_{V_1 \times W} : V_1 \times W \to G(V_1 \times W)$ é um difeomorfismo (cujo inverso é de classe C^k). Temos que $G(x_0, y_0) = (x_0, c)$ pertence ao aberto $G(V_1 \times W) \subset \mathbb{R}^m \times \mathbb{R}^n$ e portanto, como a aplicação $x \mapsto (x, c)$ é contínua, existe uma vizinhança aberta V de x_0 em V_1 tal que $(x, c) \in G(V_1 \times W)$ para todo $x \in V$. Defina agora $\phi : V \to \mathbb{R}^n$ fazendo:

$$\phi(x) = \pi_2\big[(G|_{V_1 \times W})^{-1}(x, c)\big],$$

para todo $x \in V$, onde $\pi_2 : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ denota a projeção na segunda coordenada. Obviamente ϕ é de classe C^k, $\phi(V) \subset W$ e $V \times W \subset V_1 \times W \subset U$. Finalmente, para $x \in V$, $y \in W$ temos:

$$f(x, y) = c \Longleftrightarrow G(x, y) = (x, c) \Longleftrightarrow (G|_{V_1 \times W})^{-1}(x, c) = (x, y) \Longleftrightarrow \phi(x) = y. \blacksquare$$

`Observação`. A hipótese que $\partial_2 f(x_0, y_0) : \mathbb{R}^n \to \mathbb{R}^n$ seja um isomorfismo significa que as últimas n colunas da matriz Jacobiana de f no ponto (x_0, y_0) formam uma matriz inversível.

`Observação`. Obviamente o teorema da função implícita continua válido (com a mesma demonstração) se tomarmos no lugar de $\mathbb{R}^{m+n}$ um espaço vetorial real de dimensão finita arbitrário Z, uma decomposição em soma direta qualquer $Z = Z_1 \oplus Z_2$ (em vez de $Z_1 = \mathbb{R}^m$, $Z_2 = \mathbb{R}^n$). Em particular, podemos tomar $Z = \mathbb{R}^{m+n}$ e os subespaços $Z_1, Z_2 \subset Z$ gerados por vetores da base canônica de $\mathbb{R}^{m+n}$, mas *não sendo Z_2 necessariamente gerado pelos n últimos vetores*! Do ponto de vista prático, temos o seguinte: considere um sistema:

$$\begin{cases} f_1(z_1, \ldots, z_{m+n}) = c_1, \\ f_2(z_1, \ldots, z_{m+n}) = c_2, \\ \vdots \\ f_n(z_1, \ldots, z_{m+n}) = c_n, \end{cases}$$

com n equações e $m + n$ incógnitas $z_1, \ldots, z_{m+n}$, sendo cada f_i uma função de classe C^k $(1 \le k \le \infty)$ definida num aberto de $\mathbb{R}^{m+n}$. Suponha que tenha sido encontrada alguma solução $z_0 \in \mathbb{R}^{m+n}$ do sistema e suponha que nesse ponto z_0 podemos encontrar n colunas $j_1, \ldots, j_n$ da matriz Jacobiana de $f = (f_1, \ldots, f_n)$ que formam uma matriz $(n \times n)$ inversível. Aplicamos então o teorema da função implícita com $Z = \mathbb{R}^{m+n}$, $Z_2 \subset Z$ o subespaço gerado pelos vetores $e_{j_1}, \ldots, e_{j_n}$ da base canônica de $\mathbb{R}^{m+n}$ e $Z_1 \subset Z$ o subespaço gerado pelos demais vetores da base canônica de $\mathbb{R}^{m+n}$ (de modo que $Z = Z_1 \oplus Z_2$). Concluímos então que, numa vizinhança de z_0, pode-se "resolver" o sistema escrevendo as n incógnitas $z_{j_1}, \ldots, z_{j_n}$ em função das restantes m incógnitas, utilizando funções de classe C^k.

É interessante fazer agora uma comparação com o caso de sistemas lineares: quando as colunas $j_1, \ldots, j_n$ da matriz de coeficientes de um sistema linear (n equações, $n + m$ incógnitas) formam uma matriz inversível então podemos resolver o sistema (globalmente) escrevendo a solução em *forma paramétrica* em termos das *variáveis livres* $z_{j_1}, \ldots, z_{j_n}$. O teorema da função implícita nos diz então que o mesmo vale para sistemas não lineares, *mas só localmente*!

Dados conjuntos arbitrários A, B e uma função $\phi : A \to B$, recorde que o *gráfico* de ϕ é o conjunto:

$$\mathrm{Gr}(\phi) = \big\{ \big(x, \phi(x)\big) : x \in A \big\} \subset A \times B.$$

`Observação`. Em geral, em cursos de teoria dos conjuntos, aquilo que definimos acima como sendo o gráfico de ϕ não é nada mais do que a própria função ϕ. Embora dentro da teoria moderna dos conjuntos *todos* os objetos matemáticos sejam conjuntos (até os números naturais!), no dia a dia do trabalho matemático, raramente pensa-se em funções como sendo conjuntos. Daí o nome "gráfico de ϕ" para a própria função ϕ.

Podemos agora enunciar o teorema da função implícita da seguinte maneira: *se $f : U \to \mathbb{R}^n$ é uma função de classe C^k $(1 \le k \le \infty)$ definida num aberto $U \subset \mathbb{R}^m \times \mathbb{R}^n$, $(x_0, y_0) \in U$, $f(x_0, y_0) = c \in \mathbb{R}^n$ e $\partial_2 f(x_0, y_0) : \mathbb{R}^n \to \mathbb{R}^n$ é um isomorfismo então existem* *uma vizinhança aberta* V *de* x_0 *em* $\mathbb{R}^m$, *uma vizinhança aberta* W *de* y_0 *em* $\mathbb{R}^n$ *de modo que* $V \times W \subset U$ *e que* $f^{-1}(c) \cap (V \times W)$ *é o gráfico de uma função* $\phi : V \to W$ *de classe* C^k.

A (única) função ϕ cujo gráfico é $f^{-1}(c) \cap (V \times W)$ é dita *definida implicitamente* pela igualdade $f(x, y) = c$ em torno do ponto (x_0, y_0).

Exercícios.

(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Álgebra Linear`.

1. Mostre que uma transformação linear $T : \mathbb{R} \to \mathbb{R}$ é um isomorfismo se e somente se $T(1) \neq 0$. Nesse caso, mostre que $T^{-1}(1) = \frac{1}{T(1)}$.

`Aplicações abertas e difeomorfismos locais`.

2. Sejam M, N espaços métricos e $f : M \to N$ uma função. Dizemos que f é um *homeomorfismo local* quando todo ponto $x \in M$ possui uma vizinhança aberta V em M tal que $f(V)$ é aberto em N e $f|_V : V \to f(V)$ é um homeomorfismo. Mostre que:
   (a) todo homeomorfismo local é uma aplicação aberta e contínua;
   (b) se f é um homeomorfismo local injetor então $f : M \to f(M)$ é um homeomorfismo (e $f(M)$ é aberto em N);
   (c) se f é *localmente injetora* (i.e., injetora numa vizinhança de cada ponto), aberta e contínua então f é um homeomorfismo local;
   (d) se $U \subset \mathbb{R}^m$ é um aberto e $f : U \to \mathbb{R}^m$ é um difeomorfismo local então f também é um homeomorfismo local.
3. Sejam M, N espaços métricos. Mostre que:
   (a) se $f : M \to N$ é uma aplicação aberta e $N' \subset N$ contém a imagem de f então também $f : M \to N'$ é uma aplicação aberta;
   (b) se N' é um aberto de N e $f : M \to N'$ é uma aplicação aberta então também $f : M \to N$ é uma aplicação aberta;
   (c) se M' é um aberto de M e $f : M \to N$ é uma aplicação aberta então também $f|_{M'} : M' \to N$ é uma aplicação aberta;
   (d) repita os itens (a)–(c) trocando "f é uma aplicação aberta" por "f é uma aplicação fechada", "M' é aberto em M" por "M' é fechado em M" e "N' é aberto em N" por "N' é fechado em N".
4. Dados espaços métricos M, N, P, considere o diagrama comutativo:

$$\begin{tikzcd} M \arrow[d, "q"'] \arrow[dr, "f"] & \\ N \arrow[r, "\bar{f}"'] & P \end{tikzcd}$$

   onde q é uma aplicação contínua e sobrejetora que é aberta ou fechada. Mostre que f é contínua se e somente se $\bar{f}$ é contínua (a aplicação $\bar{f}$ diz-se obtida de f por *passagem ao quociente*).

[*dica*: para $A \subset P$ observe que $\bar{f}^{-1}(A) = q\big(f^{-1}(A)\big)$].

`Derivadas parciais ‘‘gordas’’ e matrizes de blocos.`

5. Sejam A uma matriz real $n \times n$, B uma matriz real $m \times m$ e C uma matriz real $m \times n$; considere a matriz real $(n+m) \times (n+m)$ definida em notação de blocos por:

$$X = \begin{pmatrix} A & 0 \\ C & B \end{pmatrix}.$$

Mostre que X é inversível se e somente se A e B o forem; em caso afirmativo, calcule a inversa de X (*dica*: procure uma inversa da forma:

$$X' = \begin{pmatrix} A' & D' \\ C' & B' \end{pmatrix}$$

para X).

Aula número 23 (31/05)

A aula número 23 cobriu parte da demonstração do teorema da função inversa e o material das seções (2) e (3) da aula número 22 — 29/05 (derivadas parciais "gordas" e o teorema da função implícita).

Aula número 24 (05/06)

(1) **Sistemas de coordenadas**.

Esta seção funciona como uma preparação psicológica para a noção de variedade diferenciável e para os enunciados das formas locais das imersões, submersões e para o teorema do posto.

Começamos com uma pergunta: qual deve ser a definição correta para o conceito de "sistema de coordenadas"? Bom, a idéia básica é a seguinte: começamos com um "mundo abstrato" X onde temos uma certa quantidade de "habitantes". Um certo habitante de X deseja usar algum sistema de coordenadas para descrever (ao menos uma parte de) X. Tal habitante deve então associar a cada ponto de X uma n-upla $(x_1, \ldots, x_n)$ de números reais, que seriam as "coordenadas" desse ponto x. Um bom sistema de coordenadas deve ter a propriedade que pontos diferentes possuem coordenadas diferentes (senão seria uma tremenda confusão!). Essa visão caricata do conceito de sistema de coordenadas motiva a seguinte:

Definição. *Seja X um conjunto qualquer. Um* sistema de coordenadas *em X é uma função bijetora $\varphi : U \to \widetilde{U}$, onde U é um subconjunto de X e $\widetilde{U}$ é um subconjunto aberto de $\mathbb{R}^n$ para algum n.*

A exigência de que $\widetilde{U}$ seja *aberto* em $\mathbb{R}^n$ é feita por razões técnicas e é importante na teoria de variedades diferenciáveis. Num primeiro momento, seria razoável exigir apenas que $\widetilde{U}$ fosse um subconjunto arbitrário de $\mathbb{R}^n$.

`Exemplo`. Seja $X = \mathbb{R}^n$. Fazendo $U = \widetilde{U} = \mathbb{R}^n$ e φ igual à aplicação identidade então o sistema de coordenadas $\varphi : U \to \widetilde{U}$ em $X = \mathbb{R}^n$ é chamado o *sistema de coordenadas cartesianas*.

`Exemplo`. Escolha $\theta_0 \in \mathbb{R}$ e seja $U = A_{\theta_0} \subset \mathbb{R}^2$ o aberto de $\mathbb{R}^2$ cujo complementar é a semi-reta fechada $\big\{(t\cos\theta_0, t\,\mathrm{sen}\,\theta_0) : t \ge 0\big\}$. Seja $\widetilde{U} = (0, +\infty) \times (\theta_0, \theta_0 + 2\pi)$ e defina $\varphi : U \to \widetilde{U}$ fazendo $\varphi(x, y) = (\rho, \theta)$, onde $\rho = \sqrt{x^2 + y^2}$ e θ é o único ângulo para (x, y) em $(\theta_0, \theta_0 + 2\pi)$, i.e., θ é o único elemento de $(\theta_0, \theta_0 + 2\pi)$ tal que $x = \rho\cos\theta$ e $y = \rho\,\mathrm{sen}\,\theta$. É fácil ver que a aplicação $\varphi : U \to \widetilde{U}$ é de fato bijetora e é portanto um sistema de coordenadas em $\mathbb{R}^2$; esse é chamado o *sistema de coordenadas polares* (relativo à escolha de θ_0) no plano $\mathbb{R}^2$.

`Exemplo`. Escolha $\theta_0 \in I\!R$ e defina A_{θ_0} como no exemplo anterior. Considere os abertos $U = A_{\theta_0} \times I\!R \subset I\!R^3$, $\widetilde{U} = (0, +\infty) \times (\theta_0, \theta_0 + 2\pi) \times I\!R \subset I\!R^3$ e defina $\varphi : U \to \widetilde{U}$ fazendo $\varphi(x, y, z) = (\rho, \theta, z)$, onde $\rho = \sqrt{x^2 + y^2}$ e θ é o único ângulo para o ponto (x, y) pertencente a $(\theta_0, \theta_0 + 2\pi)$. Temos que $\varphi : U \to \widetilde{U}$ é um sistema de coordenadas no espaço $I\!R^3$ chamado o *sistema de coordenadas cilíndricas* (relativo à escolha de θ_0) no espaço $I\!R^3$.

`Exemplo`. Sejam $U = I\!R^3 \setminus \big((-\infty, 0] \times \{0\} \times I\!R \big)$, $\widetilde{U} = (0, +\infty) \times (-\pi, \pi) \times (0, \pi)$ e $\varphi : U \to \widetilde{U}$ a aplicação definida por $\varphi(x, y, z) = (r, \theta, \phi)$, onde $r \in (0, +\infty)$, $\theta \in (-\pi, \pi)$ e $\phi \in (0, \pi)$ são os únicos escalares para os quais as relações:

$$x = r \cos\theta \operatorname{sen}\phi, \quad y = r \operatorname{sen}\theta \operatorname{sen}\phi, \quad z = r\cos\phi,$$

são satisfeitas (note que $r = \sqrt{x^2 + y^2 + z^2}$). Temos que φ é um sistema de coordenadas em $I\!R^3$ chamado o *sistema de coordenadas esféricas* do espaço $I\!R^3$.

`Exemplo`. Seja $\mathfrak{B} = (b_i)_{i=1}^n$ uma base arbitrária de $I\!R^n$. Seja $\varphi : I\!R^n \to I\!R^n$ a única transformação linear tal que $\varphi(b_i)$ é o i-ésimo vetor da base canônica de $I\!R^n$ para todo $i = 1, \ldots, n$. Temos que φ é um isomorfismo e portanto um sistema de coordenadas (com $U = \widetilde{U} = I\!R^n$) em $I\!R^n$. Note que para todo $x \in I\!R^n$ temos que $\varphi(x)$ coincide precisamente com a n-upla formada pelas coordenadas de x na base $\mathfrak{B}$. Dizemos que φ é um *sistema de coordenadas linear* em $I\!R^n$. Quando $\mathfrak{B}$ é a base canônica, temos que $\varphi = \mathrm{Id}$, i.e., reobtemos as coordenadas cartesianas. Em geral, o sistema de coordenadas φ corresponde à idéia de usar "eixos de coordenadas oblíquos" e "escalas de medida arbitrárias" em cada um dos eixos. Mais geralmente, fixada uma base $\mathfrak{B}$ em $I\!R^n$ e um ponto $\mathcal{O} \in I\!R^n$ então podemos definir um sistema de coordenadas $\varphi : I\!R^n \to I\!R^n$ fazendo $\varphi(x)$ igual às coordenadas de $x - \mathcal{O}$ na base $\mathfrak{B}$. Dizemos então que φ é um *sistema de coordenadas afim com origem* $\mathcal{O}$. Quando $\mathcal{O} = 0$, estamos de volta ao caso de um sistema de coordenadas linear.

A definição de sistema de coordenadas que demos no início da seção é um tanto geral demais para nossos propósitos imediatos. De fato, observe que todos os sistemas de coordenadas mencionados nos exemplos acima se enquadram na seguinte definição mais restrita.

Definição. *Um* sistema de coordenadas de classe C^k $(1 \le k \le \infty)$ em $I\!R^n$ *é um difeomorfismo* $\varphi : U \to \widetilde{U}$ *de classe* C^k*, onde tanto* U *como* $\widetilde{U}$ *são abertos em* $I\!R^n$*. Por um* sistema de coordenadas de classe C^0 *em* $I\!R^n$ *entendemos simplesmente um homeomorfismo* $\varphi : U \to \widetilde{U}$*, onde* U *e* $\widetilde{U}$ *são abertos em* $I\!R^n$*.*

Observe que todos os exemplos mencionados acima são sistemas de coordenadas de classe C^∞ em $I\!R^n$.

Para finalizar, apresentamos alguns exemplos um pouco diferentes (que não correspondem a sistemas de coordenadas em $I\!R^n$).

`Exemplo`. Denote por S^2 a *esfera unitária bidimensional*, ou seja:

$$S^2 = \big\{ (x, y, z) \in I\!R^3 : x^2 + y^2 + z^2 = 1 \big\}.$$

Seja $U \subset S^2$ o aberto (relativo a S^2) definido por:

$$U = S^2 \setminus \big(\{0\} \times [0, +\infty) \times I\!R \big),$$

i.e., U é o complementar em S^2 de um meridiano fechado. Definimos uma aplicação $\varphi : U \to I\!R^2$ fazendo $\varphi(x, y, z) = (\theta, \phi)$ onde θ é a "longitude" de (x, y, z) e ϕ é a "latitude" de (x, y, z); mais explicitamente, $\theta \in (-\pi, \pi)$, $\phi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ são unicamente determinados pelas relações:

$$x = \cos\phi \operatorname{sen}\theta, \quad y = -\cos\phi\cos\theta, \quad z = \operatorname{sen}\phi.$$

Temos que φ é uma bijeção sobre o aberto $\widetilde{U} = (-\pi, \pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ em $I\!R^2$. Portanto φ é um sistema de coordenadas na esfera unitária S^2.

`Exemplo`. Seja V um espaço vetorial real de dimensão $n < +\infty$ e seja $\mathfrak{B} = (b_i)_{i=1}^n$ uma base para V. Existe uma única aplicação linear $\varphi : V \to I\!R^n$ que leva o vetor b_i sobre o i-ésimo vetor e_i da base canônica de $I\!R^n$. A aplicação φ é um isomorfismo que leva cada vetor $v \in V$ sobre a n-upla que contém as coordenadas de v na base $\mathfrak{B}$. Temos que $\varphi : V \to I\!R^n$ é um sistema de coordenadas em V; diz-se que φ é um *sistema de coordenadas linear* no espaço vetorial V. Na verdade, o presente exemplo é apenas uma pequena generalização do exemplo onde mencionamos sistemas de coordenadas lineares em $I\!R^n$ (veja também o Exercício 2 para uma generalização dos sistemas de coordenadas afins). Observe no entanto que se V é um espaço vetorial real arbitrário de dimensão n então não há um *sistema de coordenadas canônico* em V (por isso um espaço vetorial real genérico de dimensão 3 é um modelo mais adequado para o "espaço físico" do que $I\!R^3$, já que o "espaço físico" não possui uma base canônica — na verdade, *espaços afins* de dimensão 3 são um modelo ainda melhor, já que o "espaço físico" não possui sequer uma *origem* canônica).

`Observação`. Quem já estudou um pouco de teoria de cardinais em cursos de teoria dos conjuntos sabe que existe uma bijeção $\varphi : S^2 \to I\!R$ da esfera unitária S^2 sobre a reta real $I\!R$ (isso segue por exemplo do teorema de Schröder–Bernstein e do fato que $I\!R^3$ tem a mesma cardinalidade que $I\!R$). Tal bijeção φ é a rigor um sistema de coordenadas em S^2 pela nossa definição geral, apesar do fato que esse sistema de coordenadas φ deve parecer "um tanto estranho". Quando estudarmos a noção de variedade diferenciável faremos algumas restrições adicionais sobre a noção de sistema de coordenadas que eliminam patologias desagradáveis como essa.

Definição. *Sejam X, Y conjuntos e $\varphi : U \subset X \to \widetilde{U} \subset I\!R^m$, $\psi : V \subset Y \to \widetilde{V} \subset I\!R^n$ sistemas de coordenadas para X e Y respectivamente. Seja $f : X \to Y$ uma função tal que $f(U) \subset V$ e considere a função $\tilde{f} : \widetilde{U} \to \widetilde{V}$ dada pela composição $\tilde{f} = \psi \circ f \circ \varphi^{-1}$. Dizemos que $\tilde{f}$ é a função que* representa *f com respeito aos sistemas de coordenadas φ e ψ.*

A relação entre f e $\tilde{f}$ é representada pelo seguinte diagrama comutativo:

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\varphi"', "\cong"] & V \arrow[d, "\psi", "\cong"'] \\ \widetilde{U} \arrow[r, "\tilde{f}"'] & \widetilde{V} \end{tikzcd}$$

o símbolo $\cong$ foi usado para indicar que φ e ψ são bijeções. No Exercício 1 pedimos para vocês relacionarem a noção acima com a noção usual da Álgebra Linear de "matrizes que representam aplicações lineares em bases".

(2) **A versão linear do teorema do posto**.

Em álgebra linear é muito estudado o problema de diagonabilidade de operadores lineares $T : V \to V$, onde V é um espaço vetorial de dimensão finita. O problema consiste em achar uma base de V de modo que a matriz que representa T nessa base seja diagonal. Tal tarefa não é sempre realizável, i.e., existem operadores que não são diagonalizáveis. A vantagem básica de diagonalizar um operador linear é basicamente óbvia: quer-se um sistema de coordenadas no qual T seja descrito de maneira simples.

Vamos tratar aqui um problema muito mais simples do que o da diagonalização: dada uma transformação linear $T : V \to W$ (com V, W espaços vetoriais possivelmente distintos, de dimensão finita), queremos encontrar bases de V e W que tornem a representação matricial de T o mais "simples" possível. Note que mesmo quando $V = W$ tal problema é mais simples do que o problema usual de diagonalização; de fato, permitimos aqui que sejam usadas *bases diferentes* no domínio e no contra-domínio de T.

Temos o seguinte:

Teorema. *Sejam V, W espaços vetoriais com $\dim(V) = m < +\infty$ e $\dim(W) = n < +\infty$. Dada uma transformação linear $T : V \to W$ então existem bases $\mathfrak{B}$ e $\mathfrak{B}'$ para V e W respectivamente de modo que a matriz $[T]_{\mathfrak{B}\mathfrak{B}'}$ que representa T com respeito a tais bases é dada (em notação de blocos) por:*

$$[T]_{\mathfrak{B}\mathfrak{B}'} = \begin{pmatrix} \mathrm{I}_k & 0_{k\times(m-k)} \\ 0_{(n-k)\times k} & 0_{(n-k)\times(m-k)} \end{pmatrix},$$

onde I_k denota a matriz identidade $k \times k$ e $0_{\alpha\times\beta}$ denota a matriz nula $\alpha \times \beta$. Além do mais, o número k é precisamente o posto de T (i.e., a dimensão de $\mathrm{Im}(T)$).

Demonstração. Em primeiro lugar, se tais bases existirem então k deve coincidir com o posto de T pois o posto de T deve ser igual ao posto da matriz $[T]_{\mathfrak{B}\mathfrak{B}'}$ (que é k). Vamos agora mostrar a existência das bases $\mathfrak{B}$ e $\mathfrak{B}'$. Escolha uma base qualquer de $\mathrm{Ker}(T)$ e complete-a até uma base de V; obtemos então uma base $\mathfrak{B} = (b_i)_{i=1}^m$ de V tal que $(b_i)_{i=k+1}^m$ é uma base de $\mathrm{Ker}(T)$. Temos que $(b_i)_{i=1}^k$ é uma base para um subespaço $S \subset V$ tal que $V = S \oplus \mathrm{Ker}(T)$. Daí T leva S isomorficamente sobre $T(V) = \mathrm{Im}(T) \subset W$ (veja Exercício 3); concluímos então que $b'_i = T(b_i)$, $i = 1, \ldots, k$ nos dá uma base para a imagem de T. Escolha agora $\mathfrak{B}' = (b'_i)_{i=1}^n$ como sendo um completamento qualquer de $(b'_i)_{i=1}^k$ até uma base de W. Segue agora facilmente que $[T]_{\mathfrak{B}\mathfrak{B}'}$ assume a forma desejada. ■

`Observação`. Usando o resultado do Exercício 1, vemos que se $\mathfrak{B}$ e $\mathfrak{B}'$ são bases como no enunciado do teorema acima e se $\varphi_{\mathfrak{B}} : V \to \mathbb{R}^m$, $\varphi_{\mathfrak{B}'} : W \to \mathbb{R}^n$ são os correspondentes sistemas de coordenadas então temos um diagrama comutativo:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\varphi_{\mathfrak{B}}"', "\cong"] & W \arrow[d, "\cong"', "\varphi_{\mathfrak{B}'}"] \\ \mathbb{R}^m \arrow[r, "F"'] & \mathbb{R}^n \end{tikzcd}$$

onde $F : I\!R^m \to I\!R^n$ é dada por:

$$F(x_1, \ldots, x_k, \ldots, x_m) = (x_1, \ldots, x_k, \underbrace{0, \ldots, 0}_{n-k \text{ zeros}}),$$

para todo $x = (x_1, \ldots, x_m) \in I\!R^m$. Encontramos então sistemas de coordenadas (lineares) em V e W que tornam a representação de T (ou seja, F) bem simples!

`Observação`. Se $T : V \to W$ é injetora então $k = m \le n$ e o teorema nos dá sistemas de coordenadas nos quais a representação de T é dada por:

$$F(x_1, \ldots, x_m) = (x_1, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m \text{ zeros}}).$$

Nesse caso, podemos fazer ainda uma pequena melhora no enunciado do teorema: é possível *para toda* base $\mathfrak{B}$ de V encontrar uma base $\mathfrak{B}'$ de W na qual $[T]_{\mathfrak{B}\mathfrak{B}'}$ tem a forma desejada. De fato, se $\mathfrak{B} = (b_i)_{i=1}^m$ é uma base qualquer de V então $b'_i = T(b_i)$, $i = 1, \ldots, m$ é um conjunto linearmente independente e portanto pode ser completado a uma base $\mathfrak{B}'$ para W. Segue que $[T]_{\mathfrak{B}\mathfrak{B}'}$ é dada por:

$$[T]_{\mathfrak{B}\mathfrak{B}'} = \begin{pmatrix} \mathrm{I}_m \\ 0_{(n-m)\times m} \end{pmatrix}.$$

`Observação`. Se $T : V \to W$ é sobrejetora então $k = n \le m$ e o teorema nos dá sistemas de coordenadas nos quais a representação de T é dada por:

$$F(x_1, \ldots, x_m) = (x_1, \ldots, x_n).$$

Nesse caso, podemos também fazer uma pequena melhora no enunciado do teorema: é possível *para toda* base $\mathfrak{B}'$ de W encontrar uma base $\mathfrak{B}$ de V na qual $[T]_{\mathfrak{B}\mathfrak{B}'}$ tem a forma desejada. De fato, se $\mathfrak{B}' = (b'_i)_{i=1}^n$ é uma base qualquer para W, escolha um subespaço $S \subset V$ com $V = S \oplus \mathrm{Ker}(T)$ (veja o Exercício 0); daí $T|_S : S \to W$ é um isomorfismo e portanto $b_i = (T|_S)^{-1}(b'_i)$, $i = 1, \ldots, n$ nos dá uma base de S. Seja $(b_i)_{i=n+1}^m$ uma base qualquer de $\mathrm{Ker}(T)$, de modo que $\mathfrak{B} = (b_i)_{i=1}^n$ é uma base de V. Segue que $[T]_{\mathfrak{B}\mathfrak{B}'}$ é dada por:

$$[T]_{\mathfrak{B}\mathfrak{B}'} = \begin{pmatrix} \mathrm{I}_n & 0_{n\times(m-n)} \end{pmatrix}.$$

`Observação`. Uma aplicação linear $T : V \to W$ que é injetora ou sobrejetora é muitas vezes chamada uma *aplicação linear de posto máximo*. Isso se deve ao seguinte fato simples: se $\dim(V) \le \dim(W)$ então o maior posto possível para uma transformação linear $T : V \to W$ é $\dim(V)$ e T tem posto $\dim(V)$ se e somente se for injetora; além do mais, se $\dim(V) \ge \dim(W)$ então o maior posto possível para $T : V \to W$ é $\dim(W)$ e T tem posto $\dim(W)$ se e somente se for sobrejetora.

(3) **A forma local das imersões**.

Nosso objetivo agora é generalizar os resultados da seção anterior para o caso de transformações não lineares (mas diferenciáveis). Começamos com a generalização do teorema da seção anterior no caso de transformações lineares injetoras.

Definição. *Seja $f : U \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$. Se f é diferenciável num ponto $x \in U$ e se a transformação linear $\mathrm{d}f(x) : \mathbb{R}^m \to \mathbb{R}^n$ é injetora então dizemos que f é uma* imersão *no ponto x. Se f é diferenciável em U e se $\mathrm{d}f(x)$ é injetora para todo $x \in U$ então dizemos simplesmente que f é uma* imersão.

Obviamente só é possível que $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ seja uma imersão num ponto $x \in U$ se $m \le n$.

Teorema. *(forma local das imersões) Suponha que $f : U \to \mathbb{R}^n$ é uma função de classe C^k $(1 \le k \le \infty)$ definida num aberto $U \subset \mathbb{R}^m$ e suponha que f é uma imersão num ponto $x_0 \in U$. Então existem abertos $V \subset \mathbb{R}^m$, W, $\widetilde{W} \subset \mathbb{R}^n$ e um difeomorfismo $\varphi : W \to \widetilde{W}$ de classe C^k com $x_0 \in V \subset U$, $f(V) \subset W$ e de modo que:*

$$(\varphi \circ f)(x_1, \ldots, x_m) = (x_1, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m \text{ zeros}}),$$

para todo $x = (x_1, \ldots, x_m) \in V$.

Demonstração. Seja $S \subset \mathbb{R}^n$ um subespaço (de dimensão $n - m$) tal que:

$$\mathbb{R}^n = \mathrm{Im}\big(\mathrm{d}f(x_0)\big) \oplus S.$$

Defina uma aplicação $G : U \times S \to \mathbb{R}^n$ fazendo:

$$G(x, y) = f(x) + y,$$

para todos $x \in U$, $y \in S$. Obviamente G é de classe C^k e a diferencial:

$$\mathrm{d}G(x_0, 0) : \mathbb{R}^m \oplus S \longrightarrow \mathbb{R}^n$$

é dada por:

$$\mathrm{d}G(x_0, 0) \cdot (h, k) = \mathrm{d}f(x_0) \cdot h + k,$$

para todos $h \in \mathbb{R}^m$, $k \in S$. Segue facilmente do fato que $\mathrm{d}f(x_0)$ é injetora e de $\mathbb{R}^n = \mathrm{Im}\big(\mathrm{d}f(x_0)\big) \oplus S$ que $\mathrm{d}G(x_0, 0)$ é um isomorfismo. Pelo teorema da função inversa, G leva uma vizinhança aberta de $(x_0, 0)$ em $U \times S$ (que podemos escolher da forma $V \times V'$, com $V \subset U$ e $V' \subset S$ abertos) difeomorficamente sobre uma vizinhança aberta $W = G(V \times V')$ de $G(x_0, 0) = f(x_0)$ em $\mathbb{R}^n$. Escolha agora um isomorfismo qualquer $T_0 : S \to \mathbb{R}^{n-m}$ e considere o isomorfismo $T : \mathbb{R}^m \oplus S \to \mathbb{R}^n$ definido por $T(x, y) = \big(x, T_0(y)\big)$, $x \in \mathbb{R}^m$, $y \in S$. Temos agora que a aplicação $\varphi : W \to \mathbb{R}^n$ definida por $\varphi = T \circ (G|_{V \times V'})^{-1}$ é um difeomorfismo de classe C^k sobre o aberto $\widetilde{W} = T(V \times V') = V \times T_0(V')$ em $\mathbb{R}^n$. Além do mais, se $x \in V$ então $(x, 0) \in V \times V'$ e:

$$G(x, 0) = f(x) \in W = G(V \times V');$$

finalmente, temos:

$$\varphi\big(f(x)\big) = \big(T \circ (G|_{V \times V'})^{-1}\big)\big(f(x)\big) = T(x, 0) = (x, 0) = (x_1, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m \text{ zeros}}),$$

para todo $x = (x_1, \ldots, x_m) \in V$. Isso completa a demonstração. ∎

O teorema acima nos dá sistemas de coordenadas nos quais uma imersão pode (num aberto pequeno) ser representada de maneira mais simples (na verdade, só é necessário escolher um sistema de coordenadas no *contra-domínio* da função — no domínio podemos continuar usando as coordenadas cartesianas). Segue então que para demonstrar propriedades locais de imersões é possível supor sem perda de generalidade que a imersão em questão é da forma $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0, \ldots, 0)$. Temos os seguintes corolários imediatos desse princípio.

Corolário. *Se uma função $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ de classe C^1 é uma imersão num ponto $x_0 \in U$ então f é ainda uma imersão numa vizinhança aberta de x_0 em U.*

Demonstração. Observe que $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0, \ldots, 0)$ é uma imersão. ∎

Na verdade o corolário acima é também uma conseqüência simples da continuidade de $\mathrm{d}f$ e do Exercício 7.

Corolário. *Uma imersão $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ de classe C^1 é localmente injetora, i.e., todo ponto de U possui uma vizinhança onde f é injetora.*

Demonstração. Observe que $(x_1, \ldots, x_m) \mapsto (x_1, \ldots, x_m, 0, \ldots, 0)$ é injetora. ∎

`Observação`. Intuitivamente, se f é uma imersão então a imagem de f "possui a mesma dimensão" (num sentido que será feito preciso no futuro) que o domínio de f. A forma local das imersões confirma essa idéia intuitiva.

(4) **A forma local das submersões**.

Definição. *Seja $f : U \to \mathbb{R}^n$ uma função definida num aberto $U \subset \mathbb{R}^m$. Se f é diferenciável num ponto $x \in U$ e se a transformação linear $\mathrm{d}f(x) : \mathbb{R}^m \to \mathbb{R}^n$ é sobrejetora então dizemos que f é uma* submersão no ponto x. *Se f é diferenciável em U e se $\mathrm{d}f(x)$ é sobrejetora para todo $x \in U$ então dizemos simplesmente que f é uma* submersão.

Obviamente só é possível que $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ seja uma submersão num ponto $x \in U$ se $m \geq n$.

Teorema. *(forma local das submersões) Seja $f : U \to \mathbb{R}^n$ uma função de classe C^k $(1 \leq k \leq \infty)$ definida num aberto $U \subset \mathbb{R}^m$ e suponha que f é uma submersão num ponto $z_0 \in U$. Então existem abertos $V, \widetilde{V} \subset \mathbb{R}^m$ e um difeomorfismo $\varphi : V \to \widetilde{V}$ de classe C^k de modo que $z_0 \in V \subset U$ e:*

$$\big(f \circ \varphi^{-1}\big)(v_1, \ldots, v_m) = (v_1, \ldots, v_n),$$

para todo $v = (v_1, \ldots, v_m) \in \widetilde{V}$.

Demonstração. Seja $S \subset \mathbb{R}^m$ um subespaço (de dimensão n) tal que:

$$\mathbb{R}^m = \mathrm{Ker}\big(\mathrm{d}f(z_0)\big) \oplus S.$$

Defina uma aplicação $G : U \to \mathbb{R}^n \oplus \mathrm{Ker}\big(\mathrm{d}f(z_0)\big)$ fazendo:

$$G(x, y) = \big(f(x, y), x\big),$$

para todos $x \in \operatorname{Ker}\big(\mathrm{d}f(z_0)\big)$, $y \in S$ tais que $(x,y) \in U \subset \operatorname{Ker}\big(\mathrm{d}f(z_0)\big) \oplus S = \mathbb{R}^m$. Obviamente G é de classe C^k e sua diferencial no ponto $z_0 = (x_0, y_0) \in \operatorname{Ker}\big(\mathrm{d}f(z_0)\big) \oplus S$ é dada por:

$$\mathrm{d}G(x_0,y_0)\cdot(h,k) = \big(\partial_1 f(x_0,y_0)\cdot h + \partial_2 f(x_0,y_0)\cdot k, h\big),$$

para todos $h \in \operatorname{Ker}\big(\mathrm{d}f(z_0)\big)$, $k \in S$. Como $\partial_2 f(x_0,y_0) = \mathrm{d}f(z_0)|_S : S \to \mathbb{R}^n$ é um isomorfismo (veja o Exercício 3), segue facilmente que:

$$\mathrm{d}G(x_0,y_0) : \mathbb{R}^m = \operatorname{Ker}\big(\mathrm{d}f(z_0)\big) \oplus S \longrightarrow \mathbb{R}^n \oplus \operatorname{Ker}\big(\mathrm{d}f(z_0)\big),$$

é um isomorfismo. Pelo teorema da função inversa, G leva uma vizinhança aberta V de $z_0 = (x_0, y_0)$ em U difeomorficamente sobre uma vizinhança aberta V' de $G(z_0) = \big(f(z_0), x_0\big)$ em $\mathbb{R}^n \oplus \operatorname{Ker}\big(\mathrm{d}f(z_0)\big)$. Escolha um isomorfismo qualquer $T_0 : \operatorname{Ker}\big(\mathrm{d}f(z_0)\big) \to \mathbb{R}^{m-n}$ e considere o isomorfismo $T : \mathbb{R}^n \oplus \operatorname{Ker}\big(\mathrm{d}f(z_0)\big) \to \mathbb{R}^m$ definido por $T(u,x) = \big(u, T_0(x)\big)$, $u \in \mathbb{R}^n$, $x \in \operatorname{Ker}\big(\mathrm{d}f(z_0)\big)$. Temos agora que $\widetilde{V} = T(V')$ é um aberto de $\mathbb{R}^m$ e que a aplicação $\varphi : V \to \widetilde{V}$ definida por:

$$\varphi = T \circ G|_V,$$

é um difeomorfismo de classe C^k. Para finalizar, seja $v = (v_1, \ldots, v_m) \in \widetilde{V}$. Temos que $T^{-1}(v) = (u,x) \in V'$, onde $u = (v_1, \ldots, v_n) \in \mathbb{R}^n$ e $x \in \operatorname{Ker}\big(\mathrm{d}f(z_0)\big)$ satisfaz $T_0(x) = (v_{n+1}, \ldots, v_m)$. Daí:

$$\varphi^{-1}(v) = (G|_V)^{-1}(u,x) = (x,y),$$

onde $y \in S$ é caracterizado pelo fato que $(x,y) \in V$ e $f(x,y) = u$. A conclusão agora é obtida observando que:

$$\big(f \circ \varphi^{-1}\big)(v_1, \ldots, v_m) = \big(f \circ \varphi^{-1}\big)(v) = f(x,y) = u = (v_1, \ldots, v_n). \blacksquare$$

O teorema acima nos dá sistemas de coordenadas nos quais uma submersão pode (num aberto pequeno) ser representada de maneira mais simples (na verdade, só é necessário escolher um sistema de coordenadas no *domínio* da função — no contra-domínio podemos continuar usando as coordenadas cartesianas). Segue então que para demonstrar propriedades locais de submersões é possível supor sem perda de generalidade que a submersão em questão é da forma $(v_1, \ldots, v_m) \mapsto (v_1, \ldots, v_n)$. Temos os seguintes corolários imediatos desse princípio.

Corolário. *Se uma função $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ de classe C^1 é uma submersão num ponto $z_0 \in U$ então f é ainda uma submersão numa vizinhança aberta de z_0 em U.*

Demonstração. Observe que $(v_1, \ldots, v_m) \mapsto (v_1, \ldots, v_n)$ é uma submersão. $\blacksquare$

Na verdade o corolário acima é também uma conseqüência simples da continuidade de $\mathrm{d}f$ e do Exercício 7.

Corolário. *Uma submersão $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ de classe C^1 é uma aplicação aberta.*

Demonstração. Observe que a aplicação $(v_1, \ldots, v_m) \mapsto (v_1, \ldots, v_n)$ (e também sua restrição a um aberto qualquer de $\mathbb{R}^m$) é uma aplicação aberta (veja os Exercícios 11 e 12). ■

(5) **O teorema do posto**.

O próximo teorema generaliza tanto a forma local das imersões quanto a forma local das submersões.

Teorema. *(do posto) Seja $f : U \to \mathbb{R}^n$ uma função de classe C^k ($1 \le k \le \infty$) definida num aberto $U \subset \mathbb{R}^m$. Suponha que o posto de $\mathrm{d}f(x)$ é (constante e) igual a r para todo $x \in U$, para algum $r = 0, \ldots, \min\{m, n\}$. Então para todo $z_0 \in U$ existem abertos $V, \widetilde{V} \subset \mathbb{R}^m$, $W, \widetilde{W} \subset \mathbb{R}^n$ e difeomorfismos $\varphi : V \to \widetilde{V}$, $\psi : W \to \widetilde{W}$ de classe C^k com $z_0 \in V \subset U$, $f(V) \subset W$ e:*

$$\psi\big[f\big(\varphi^{-1}(v_1, \ldots, v_m)\big)\big] = (v_1, \ldots, v_r, \underbrace{0, \ldots, 0}_{n-r \text{ zeros}}),$$

para todo $v = (v_1, \ldots, v_m) \in \widetilde{V}$.

Demonstração. Seja $S \subset \mathbb{R}^n$ um subespaço (de dimensão $n - r$) tal que:

$$\mathbb{R}^n = \mathrm{Im}\big(\mathrm{d}f(z_0)\big) \oplus S. \tag{1}$$

Segue então da continuidade de $\mathrm{d}f$ e do lema a seguir que $\mathrm{Im}\big(\mathrm{d}f(z)\big) + S = \mathbb{R}^n$ para todo z em alguma vizinhança aberta V_0 de z_0 em U; como $\mathrm{Im}\big(\mathrm{d}f(z)\big)$ tem dimensão r, obtemos:

$$\mathbb{R}^n = \mathrm{Im}\big(\mathrm{d}f(z)\big) \oplus S, \tag{2}$$

para todo z pertencente a V_0. Seja $\pi : \mathbb{R}^n \to \mathrm{Im}\big(\mathrm{d}f(z_0)\big)$ a aplicação de projeção correspondente à soma direta (1). Segue de (2) e do Exercício 4 que π leva $\mathrm{Im}\big(\mathrm{d}f(z)\big)$ isomorficamente sobre $\mathrm{Im}\big(\mathrm{d}f(z_0)\big)$. Concluímos então que a aplicação:

$$\pi \circ f|_{V_0} : V_0 \longrightarrow \mathrm{Im}\big(\mathrm{d}f(z_0)\big)$$

é uma submersão (de classe C^k), já que $\mathrm{d}(\pi \circ f|_{V_0})(z) = \pi \circ \mathrm{d}f(z)$, para todo $z \in V_0$. Observe também que a injetividade da restrição de π a $\mathrm{Im}\big(\mathrm{d}f(z)\big)$ implica que:

$$\mathrm{Ker}\big(\pi \circ \mathrm{d}f(z)\big) = \mathrm{Ker}\big(\mathrm{d}f(z)\big), \tag{3}$$

para todo $z \in V_0$.

Escolha um isomorfismo qualquer $T : \mathrm{Im}\big(\mathrm{d}f(z_0)\big) \to \mathbb{R}^r$; obviamente $T \circ \pi \circ f|_{V_0}$ é ainda uma submersão. Pela forma local das submersões, existem abertos $V, \widetilde{V} \subset \mathbb{R}^m$ e um difeomorfismo $\varphi : V \to \widetilde{V}$ de classe C^k com $z_0 \in V \subset V_0$ e:

$$\big(T \circ \pi \circ f \circ \varphi^{-1}\big)(v_1, \ldots, v_m) = (v_1, \ldots, v_r), \tag{4}$$

para todo $v = (v_1, \ldots, v_m) \in \widetilde{V}$; podemos também supor que $\widetilde{V}$ é da forma $\widetilde{V} = \widetilde{V}_1 \times \widetilde{V}_2$, onde $\widetilde{V}_1$ é um aberto de $\mathbb{R}^r$ e $\widetilde{V}_2$ é um aberto conexo de $\mathbb{R}^{m-r}$. Diferenciando (4) num ponto $v = \varphi(z) \in \widetilde{V}$ e aplicando ao i-ésimo vetor e_i da base canônica de $\mathbb{R}^m$ obtemos:

$$\big[T \circ \pi \circ \mathrm{d}f(z) \circ \mathrm{d}\varphi(z)^{-1}\big] \cdot e_i = 0, \quad i = r+1, \ldots, m,$$

para todo $z \in V$. Como T é um isomorfismo, usando (3) e a fórmula acima concluímos que $\mathrm{d}\varphi(z)^{-1} \cdot e_i \in \mathrm{Ker}\big(\mathrm{d}f(z)\big)$, $i = r+1, \ldots, m$ e portanto:

$$\mathrm{d}\big(f \circ \varphi^{-1}\big)\big(\varphi(z)\big) \cdot e_i = \big[\mathrm{d}f(z) \circ \mathrm{d}\varphi(z)^{-1}\big] \cdot e_i = 0, \quad i = r+1, \ldots, m,$$

para todo $z \in V$. Segue que para todo $u \in \widetilde{V}_1 \subset \mathbb{R}^r$, a função $f \circ \varphi^{-1}(u, \cdot)$ definida no aberto conexo $\widetilde{V}_2 \subset \mathbb{R}^{m-r}$ possui diferencial identicamente nula e portanto é constante; isso significa que $f \circ \varphi^{-1}$ *não depende* das últimas $m - r$ variáveis, i.e., existe uma função $\alpha : \widetilde{V}_1 \to \mathbb{R}^n$ de classe C^k com:

$$f \circ \varphi^{-1}(u, u') = \alpha(u), \tag{5}$$

para todos $u \in \widetilde{V}_1$, $u' \in \widetilde{V}_2$ (para definir α, escolha qualquer $u'_0 \in \widetilde{V}_2$ e ponha $\alpha(u) = f \circ \varphi^{-1}(u, u'_0)$, $u \in \widetilde{V}_1$). Considere as coordenadas $\alpha_1 : \widetilde{V}_1 \to \mathrm{Im}\big(\mathrm{d}f(z_0)\big)$ e $\alpha_2 : \widetilde{V}_1 \to S$ de α com respeito à decomposição $\mathbb{R}^n = \mathrm{Im}\big(\mathrm{d}f(z_0)\big) \oplus S$. A igualdade (4) nos diz que:

$$T\big(\alpha_1(u)\big) = u, \tag{6}$$

para todo $u \in \widetilde{V}_1$. Definimos agora o difeomorfismo $\psi : W \to \widetilde{W}$ de classe C^k fazendo $W = T^{-1}(\widetilde{V}_1) \times S \subset \mathrm{Im}\big(\mathrm{d}f(z_0)\big) \oplus S = \mathbb{R}^n$, $\widetilde{W} = \widetilde{V}_1 \times \mathbb{R}^{n-r} \subset \mathbb{R}^n$ e:

$$\psi(w, w') = \Big(T(w), T'\big[w' - \alpha_2\big(T(w)\big)\big]\Big), \quad w \in T^{-1}(\widetilde{V}_1),\ w' \in S,$$

onde $T' : S \to \mathbb{R}^{n-r}$ é um isomorfismo qualquer. Segue agora de (5) e de (6) que $f(V) \subset W$ e que:

$$\big(\psi \circ f \circ \varphi^{-1}\big)(u, u') = (u, 0),$$

para todos $u \in \widetilde{V}_1 \subset \mathbb{R}^r$, $u' \in \widetilde{V}_2 \subset \mathbb{R}^{m-r}$. ■

Lema. *Sejam V, W espaços vetoriais reais de dimensão finita e $S \subset W$ um subespaço. Então o conjunto das aplicações lineares $T : V \to W$ tais que $\mathrm{Im}(T) + S = W$ é aberto em $\mathrm{Lin}(V, W)$.*

Demonstração. Considere a aplicação $\lambda : \mathrm{Lin}(V, W) \to \mathrm{Lin}(V \oplus S, W)$ definida por:

$$\lambda(T)|_V = T, \quad \lambda(T)|_S = \text{inclusão de } S \text{ em } W,$$

para todo $T \in \mathrm{Lin}(V, W)$. Temos que λ é contínua (pois é afim) e que $\mathrm{Im}(T) + S = W$ se e somente se $\lambda(T)$ é sobrejetora. A conclusão segue do fato que o conjunto das aplicações lineares sobrejetoras de $V \oplus S$ em W é aberto em $\mathrm{Lin}(V \oplus S, W)$ (veja Exercício 7). ■

O teorema do posto nos diz que toda aplicação f cuja diferencial tem posto constante num aberto pode ser escrita de maneira mais simples em sistemas de coordenadas apropriados (ao contrário das formas locais das imersões e das submersões, é possível que tenhamos que mudar as coordenadas *tanto no domínio como no contra-domínio* de f para obter a representação desejada para a função f).

Exercícios.
(não é para entregar, mas é bom dar uma olhada e quem tiver problemas me procura).

`Conjuntos e funções.`

-1. Sejam X, Y conjuntos e $f : X \to Y$ uma função.
   (a) Se X é não vazio, mostre que f é injetora se e somente se admite uma *inversa à esquerda*, i.e., se e somente se existe uma função $g : Y \to X$ tal que $g \circ f$ é igual à identidade de X.
   (b) Mostre que f é sobrejetora se e somente se admite uma *inversa à direita*, i.e., se e somente se existe uma função $g : Y \to X$ tal que $f \circ g$ é igual à identidade de Y.

`Álgebra Linear.`

0. Seja V um espaço vetorial de dimensão finita. Mostre que todo subespaço $S \subset V$ admite um *subespaço complementar*, i.e., um subespaço $S' \subset V$ com $V = S \oplus S'$ (na verdade, esse resultado também vale para $\dim(V) = +\infty$, mas a demonstração depende do *Lema de Zorn*).
1. Sejam V, W espaços vetoriais reais de dimensão finita e $\mathfrak{B} = (b_i)_{i=1}^m$, $\mathfrak{B}' = (b_i')_{i=1}^n$ bases para V e para W respectivamente. Denote por:

$$\varphi_{\mathfrak{B}} : V \longrightarrow I\!R^m, \quad \varphi_{\mathfrak{B}'} : W \longrightarrow I\!R^n,$$

respectivamente os sistemas de coordenadas em V e W associados a $\mathfrak{B}$ e a $\mathfrak{B}'$, i.e., $\varphi_{\mathfrak{B}}$ é o isomorfismo que leva $\mathfrak{B}$ sobre a base canônica de $I\!R^m$ e $\varphi_{\mathfrak{B}'}$ é o isomorfismo que leva $\mathfrak{B}'$ sobre a base canônica de $I\!R^n$. Dada uma transformação linear $T : V \to W$, denote por A a matriz real $n \times m$ que representa T com respeito às bases $\mathfrak{B}$ e $\mathfrak{B}'$; denote também por $L_A : I\!R^m \to I\!R^n$ a *aplicação de multiplicação por* A, i.e., $L_A(x) = Ax$ para todo $x \in I\!R^m$, onde interpretamos x como uma matriz coluna $m \times 1$. Mostre que o seguinte diagrama:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "\varphi_{\mathfrak{B}}"', "\cong"] & W \arrow[d, "\cong"', "\varphi_{\mathfrak{B}'}"] \\ I\!R^m \arrow[r, "L_A"'] & I\!R^n \end{tikzcd}$$

comuta, i.e., mostre que $\varphi_{\mathfrak{B}'} \circ T = L_A \circ \varphi_{\mathfrak{B}}$. Isso significa que L_A é a função que representa T com respeito aos sistemas de coordenadas $\varphi_{\mathfrak{B}}$ e $\varphi_{\mathfrak{B}'}$!
2. Sejam V um espaço vetorial real e P um conjunto; seja dada também uma aplicação $\rho : V \times P \to P$ satisfazendo as seguinte propriedades:
   (i) $\rho\big(v, \rho(w,p)\big) = \rho(v+w, p)$ para todos $v, w \in V$, $p \in P$;
   (ii) $\rho(0,p) = p$ para todo $p \in P$;
   (iii) se para algum $v \in V$, $p \in P$ temos $\rho(v,p) = p$ então $v = 0$;
   (iv) para todos $p, q \in P$ existe $v \in V$ com $\rho(v,p) = q$.

A trinca (P, V, ρ) é dita um *espaço afim* e V é dito o *espaço vetorial paralelo* a tal espaço afim. Tipicamente pensa-se em P como um "conjunto de pontos" e, para $p \in P$,

$v \in V$, escreve-se $p+v$ em vez de $\rho(v,p)$, i.e., diz-se que $\rho(v,p)$ é a *soma do vetor v com o ponto p*. Mostre que dados $\mathcal{O} \in P$ e $\mathfrak{B} = (b_i)_{i=1}^n$ uma base para V então para cada $p \in P$ existe um único $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ tal que $p = \mathcal{O} + \sum_{i=1}^n x_i b_i$; definindo $\varphi(p) = x$, mostre que obtêm-se uma bijeção $\varphi : P \to \mathbb{R}^n$. Diz-se que φ é o *sistema de coordenadas afim* em P com *origem* $\mathcal{O}$ e "eixos" $(b_i)_{i=1}^n$.

`Observação`. Para quem já estudou um pouco de teoria de ação de grupos: as condições impostas acima sobre $\rho : V \times P \to P$ dizem que ρ é uma *ação livre e transitiva* do grupo abeliano aditivo $(V,+)$ no conjunto P (livre = "sem pontos fixos").

3. Sejam V, W espaços vetoriais e $T : V \to W$ uma transformação linear. Mostre que as seguintes condições são equivalentes sobre um subespaço $S \subset V$:
   (i) $V = \mathrm{Ker}(T) \oplus S$;
   (ii) $T|_S : S \to \mathrm{Im}(T)$ é um isomorfismo.

[*dica*: supondo (ii), para mostrar que $V = \mathrm{Ker}(T) + S$, tome $v \in V$ e olhe para o vetor $(T|_S)^{-1}\big(T(v)\big) \in S$].

4. Seja V um espaço vetorial e sejam V_1, V_2, V_2' subespaços de V tais que:
$$V = V_1 \oplus V_2, \quad V = V_1 \oplus V_2'.$$
   Denote por $\pi : V \to V_2$ a projeção em V_2 relativa à decomposição $V = V_1 \oplus V_2$ e por $\pi' : V \to V_2'$ a projeção em V_2' relativa à decomposição $V = V_1 \oplus V_2'$. Mostre que:
$$\pi'|_{V_2} : V_2 \longrightarrow V_2' \quad \text{e} \quad \pi|_{V_2'} : V_2' \longrightarrow V_2$$
   são isomorfismos mutuamente inversos.

5. Se $A \in \mathrm{M}_{m\times n}(\mathbb{R})$ é uma matriz real $m \times n$ e se $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, m\}$, $J = \{j_1, \ldots, j_l\} \subset \{1, \ldots, n\}$ são subconjuntos não vazios então denotamos por A_{IJ} a submatriz $k \times l$ de A que possui apenas as linhas $i_1, \ldots, i_k$ de A e as colunas $j_1, \ldots, j_l$ de A. Se $k = l$ então dizemos que $\det(A_{IJ})$ é um *determinante menor de ordem k de A*. Mostre que A possui posto k se e somente se A possui um determinante menor não nulo de ordem k e todo determinante menor de ordem $k+1$ é nulo [*dica*: Use o fato que o "posto coluna" e o "posto linha" de uma matriz coincidem (veja Exercício 5, item (f), Aula número 18 — 15/05); o posto é por definição igual ao "posto linha" e ao "posto coluna". Se A tem posto k escolha k colunas linearmente independentes em A e use o fato que a submatriz de A constituída por essas k colunas possui "posto linha" igual a k. Mostre também que se A possui um determinante menor não nulo de ordem $k+1$ então o posto de A é maior ou igual a $k+1$].

6. Mostre que se $A \in \mathrm{M}_{m\times n}(\mathbb{R})$ possui posto k então A possui uma vizinhança em $\mathrm{M}_{m\times n}(\mathbb{R})$ formada só por matrizes de posto maior ou igual a k (diz-se então que o posto é uma função *semi-contínua inferiormente*).

[*dica*: pelo Exercício 5, A possui um determinante menor não nulo de ordem k. Use a continuidade da função determinante para concluir que numa vizinhança pequena de A em $\mathrm{M}_{m\times n}(\mathbb{R})$ toda matriz possuirá também um determinante menor não nulo de ordem k].

7. Dados espaços vetoriais reais de dimensão finita V, W, mostre que os conjuntos:

$$\big\{T \in \mathrm{Lin}(V, W) : T \text{ é injetora}\big\} \quad \text{e} \quad \big\{T \in \mathrm{Lin}(V, W) : T \text{ é sobrejetora}\big\},$$

   são abertos em $\mathrm{Lin}(V, W)$ (*dica*: o posto de uma aplicação linear coincide com o posto de uma matriz que a representa numa base qualquer. Use o Exercício 6).

8. Dados espaços vetoriais V, W, mostre que uma aplicação linear $T : V \to W$ é injetora se e somente se T admite uma *inversa linear à esquerda*, i.e., se e somente se existe uma aplicação linear $S : W \to V$ com $S \circ T$ igual à identidade de V (*dica*: se T é injetora então $T : V \to \mathrm{Im}(T)$ é um isomorfismo. Defina $S : W \to V$ como sendo uma extensão linear arbitrária de $T^{-1} : \mathrm{Im}(T) \to V$).

9. Dados espaços vetoriais V, W, mostre que uma aplicação linear $T : V \to W$ é sobrejetor se e somente se T admite um *inverso linear à direita*, i.e., se e somente se existe uma aplicação linear $S : W \to V$ com $T \circ S$ igual à identidade de W (*dica*: se Z é um subespaço de V com $V = \mathrm{Ker}(T) \oplus Z$ então, pelo Exercício 3, $T|_Z : Z \to W$ é um isomorfismo. Defina $S = (T|_Z)^{-1}$).

`Observação`. O resultado dos Exercícios 8 e 9 produzem uma solução interessante para o Exercício 7, usando também a continuidade das aplicações $T \mapsto S \circ T$, $T \mapsto T \circ S$ e o fato que o conjunto dos isomorfismos de um espaço V em si mesmo é aberto em $\mathrm{Lin}(V, V)$ (para achar essa solução, generalize os Exercícios 8 e 9, mostrando que:

$$T \in \mathrm{Lin}(V, W) \text{ injetora} \Longleftrightarrow \text{existe } S \in \mathrm{Lin}(W, V) \text{ com } S \circ T \text{ um isomorfismo},$$
$$T \in \mathrm{Lin}(V, W) \text{ sobrejetora} \Longleftrightarrow \text{existe } S \in \mathrm{Lin}(W, V) \text{ com } T \circ S \text{ um isomorfismo},$$

para todo $T \in \mathrm{Lin}(V, W)$).

10. Sejam V, W espaços vetoriais reais de dimensão finita e $S \subset V$ um subespaço. Mostre que o conjunto das aplicações lineares $T : V \to W$ tais que $S \cap \mathrm{Ker}(T) = \{0\}$ é aberto em $\mathrm{Lin}(V, W)$ (*dica*: $S \cap \mathrm{Ker}(T) = \{0\}$ se e somente se $T|_S \in \mathrm{Lin}(S, W)$ é injetora; use o Exercício 7).

`Aplicações abertas`.

11. Sejam M, N espaço métricos e considere $M \times N$ munido da métrica produto. Mostre que as projeções $M \times N \to M$ e $M \times N \to N$ são aplicações abertas (*dica*: use o Exercício 20, item (c), aula número 5 — 20/03).

12. Sejam M, N espaço métricos e $f : M \to N$ uma função. Suponha que todo ponto $x \in M$ possui uma vizinhança aberta V em M tal que $f|_V : V \to N$ é uma aplicação aberta. Mostre que f é uma aplicação aberta.

Aula número 25 (07/06)

(1) Variedades diferenciáveis.

Uma variedade diferenciável é um "mundo abstrato" onde pode-se estudar Cálculo Diferencial. O seguinte esquema é basicamente auto-explicativo:

Espaços vetoriais	$\longleftrightarrow$	"mundo abstrato" onde se estuda transformações lineares
Espaços topológicos	$\longleftrightarrow$	"mundo abstrato" onde se estuda limite e continuidade
Espaços de medida	$\longleftrightarrow$	"mundo abstrato" onde se estuda integração
Variedades diferenciáveis	$\longleftrightarrow$	"mundo abstrato" onde se estuda Cálculo Diferencial

A definição geral de variedade diferenciável é talvez um pouco abstrata demais para um primeiro contato com o assunto — apresentaremos tal definição mais tarde. Nesse curso nos restringiremos à noção de *subvariedades do espaço Euclideano* $\mathbb{R}^n$, que é mais intuitiva. Subvariedades k-dimensionais de $\mathbb{R}^n$ generalizam o conceito de superfície regular em $\mathbb{R}^3$ (o caso de superfícies regulares de $\mathbb{R}^3$ é reobtido com $k = 2$, $n = 3$).

Existem várias definições equivalentes para o conceito de subvariedade k-dimensional de $\mathbb{R}^n$ e tais equivalências serão discutidas adiante. A seguir apresentamos a nossa definição. No que segue, identificaremos $\mathbb{R}^k$ com o subespaço de $\mathbb{R}^n$ gerado pelos k primeiros vetores da base canônica, para $k = 0, 1, \ldots, n$; mais explicitamente, escrevemos:

$$\mathbb{R}^k = \big\{(x_1, \ldots, x_k, 0, \ldots, 0) \in \mathbb{R}^n : x_1, \ldots, x_k \in \mathbb{R}\big\}.$$

Definição. *Seja $M \subset \mathbb{R}^n$ um subconjunto não vazio. Dizemos que M é uma* subvariedade k-dimensional de $\mathbb{R}^n$ *de classe C^r $(1 \le r \le \infty)$ se para todo $x \in M$ existe um difeomorfismo $\varphi : A \to \widetilde{A}$ de classe C^r com A, $\widetilde{A}$ abertos em $\mathbb{R}^n$, $x \in A$ e:*

$$\varphi(A \cap M) = \widetilde{A} \cap \mathbb{R}^k.$$

A definição acima deve ser encarada da seguinte maneira: o subespaço $\mathbb{R}^k \subset \mathbb{R}^n$ é o "modelo padrão" de subvariedade k-dimensional de $\mathbb{R}^n$. Um outro subconjunto $M \subset \mathbb{R}^n$ é chamado uma subvariedade k-dimensional se em torno de cada ponto de M podemos encontrar um sistema de coordenadas de classe C^r de modo que os "usuários" desse sistema de coordenadas "vejam" M como $\mathbb{R}^k \subset \mathbb{R}^n$.

`Exemplo`. Seja M um subespaço vetorial de $\mathbb{R}^n$ de dimensão k. É fácil ver que existe um isomorfismo linear $T : \mathbb{R}^n \to \mathbb{R}^n$ tal que $T(M) = \mathbb{R}^k$ (escolha uma base de M, estende-a a uma base de $\mathbb{R}^n$ e defina T mandando tal base sobre a base canônica de $\mathbb{R}^n$). Daí $A = \widetilde{A} = \mathbb{R}^n$ são abertos em $\mathbb{R}^n$, $\varphi = T$ é um difeomorfismo de classe C^∞.

To be continued...