CÁLCULO II - MAT 147 - FEAUSP - 2º Semestre de 2012

Professor Oswaldo Rio Branco de Oliveira

MÁXIMOS E MÍNIMOS CONDICIONADOS E MULTIPLICADORES DE LAGRANGE

Definições: Seja $f : \mathrm{Dom}(f) \to \mathbb{R}$, com $\mathrm{Dom}(f)$ um aberto em $\mathbb{R}^n$, n arbitrário em $\mathbb{N}$. Dizemos que o ponto P_0 em $\mathrm{Dom}(f)$ é:

- Ponto crítico, ou estacionário, de f se $\vec{\nabla} f(P_0) = \vec{0}$.
- Ponto de máximo [mínimo] local de f se existe uma bola aberta $B(P_0; r)$, de centro P_0 e raio $r > 0$, tal que $f(P_0) \geq f(P)$ $[f(P_0) \leq f(P)]$, para todo ponto P na intersecção $B(P_0; r) \cap \mathrm{Dom}(f)$.
- Extremante local de f se é ponto de máximo [mínimo] local de f.
- Extremante local de f em um subconjunto A de $\mathrm{Dom}(f)$, se P_0 está em A e P_0 é ponto de máximo[mínimo] local da função f restrita a A.

Teorema 1. *Seja Ω um aberto em $\mathbb{R}^2$. Sejam f e g duas funções em $C^1(\Omega; \mathbb{R})$ e $L = \{(x, y) : g(x, y) = 0\}$, a curva de nível zero de g, com $\vec{\nabla} g(x, y) \neq \vec{0}$, para todo (x, y) em L. Seja $P_0 = (x_0, y_0)$ em L um ponto de máximo, ou mínimo, local de f em L. Então, existe λ em $\mathbb{R}$ tal que*

$$\vec{\nabla} f(P_0) = \lambda \vec{\nabla} g(P_0). \tag{1.1}$$

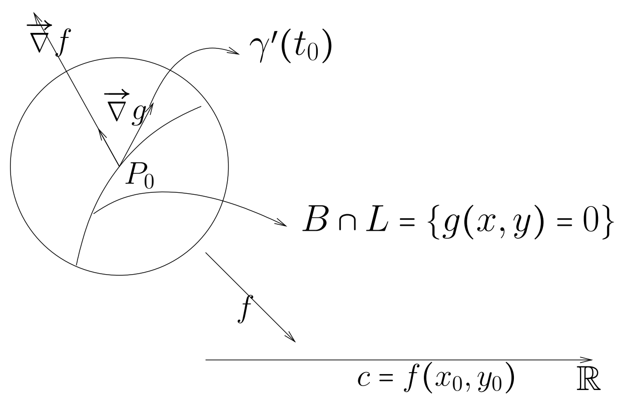

Figura 1: Teorema de Lagrange no Plano

Prova. Vide Figura 1. Localmente, no ponto P_0, pelo Teorema da Função Implícita segue que L é parametrizável como uma curva $\gamma(t)$, com t em uma vinhança de zero, com $\gamma(0) = P_0$ e $\gamma'(0) \neq \overrightarrow{0}$. Assim, a função $(f \circ \gamma)(t)$ têm extremante em $t = 0$. Donde segue $(f\circ\gamma)'(0) = 0$ e pela regra da cadeia obtemos $\overrightarrow{\nabla} f(P_0) \perp \gamma'(0)$. Porém, também são válidas as relações:

$$g \circ \gamma \equiv 0\,, \quad \overrightarrow{\nabla} g(P_0) \perp \gamma'(0)\,, \quad \gamma'(0) \neq \overrightarrow{0} \quad \text{e} \quad \overrightarrow{\nabla} g(P_0) \neq \overrightarrow{0}\,.$$

Portanto, o vetor $\overrightarrow{\nabla} f(P_0)$ é paralelo ao vetor não nulo $\overrightarrow{\nabla} g(P_0)$. Donde concluímos a tese ■

Interpretações Geométricas para o Teorema de Lagrange (acima).

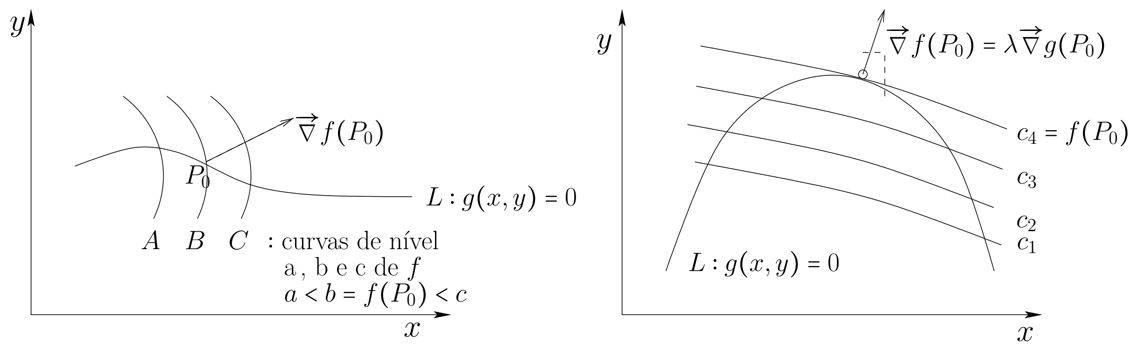

Figura 2: Ilustrações 2A e 2B (complementares) para o Teorema de Lagrange

Interpretação para Figura 2-A (à esquerda).

Suponha que $\overrightarrow{\nabla} f(P_0) \neq \overrightarrow{0}$ é oblíquo a L em P_0. Então, a curva C de nível $f(P_0)$, de f, cruza a curva L e as curvas de nível c, de f, com $c \approx f(P_0)$ [i.e., c próximo a $f(P_0)$], cujos gráficos próximos a C, também cruzam L e, orientando-as na direção de crescimento de c (a direção do gradiente) vemos que P_0 não é ponto de máximo, nem ponto de mínimo, local de f em L, contra a hipótese. Portanto, o gradiente $\overrightarrow{\nabla} f(P_0)$ é ortogonal a L, assim como o gradiente $\overrightarrow{\nabla} g(P_0)$. Logo, o vetor $\overrightarrow{\nabla} f(P_0)$ é paralelo a $\overrightarrow{\nabla} g(P_0)$.

Interpretação para Figura 2-B (à direita): Representemos as curvas de nível $c_1 < c_2 < c_3 < c_4 = f(x_0, y_0) = f(P_0)$ no sentido de crescimento do gradiente de f (pois f cresce na direção do gradiente). É claro que o valor máximo de f sobre a curva $L = \{(x,y) : g(x,y) = 0\}$, corresponde ao maior valor c tal que a curva de nível $f(x,y) = c$ intercepta a curva L. Em tal ponto P_0 de intersecção tais curvas tem mesma reta tangente e assim mesma reta normal, as quais a priori tem direção $\vec{\nabla} f(P_0)$ e $\vec{\nabla} g(P_0)$ (Figura 2-B), respectivamente. Logo, tais vetores são paralelos e o vetor $\vec{\nabla} f(P_0)$ é um múltiplo de $\vec{\nabla} g(P_0)$ já que este é não nulo.

Observações. Mantenhamos a notação do Teorema 1 e sua demonstração.

(1) Suponhamos que P_0 é ponto crítico de f e $\lambda = 0$. Então, temos $\vec{\nabla} f(P_0) = \vec{0} = \lambda \vec{\nabla} g(P_0)$ apesar que não existe um número λ tal que $\vec{0} \neq \vec{\nabla} g(P_0) = \lambda \vec{\nabla} f(P_0)$. Neste particular sentido, concluímos que a equação (1.1) é a melhor possível.

(2) Pontos críticos de f pertencentes a L, são candidatos a extremantes locais de f restrita a L.

(3) Se $\vec{\nabla} f \neq \vec{0}$ para todo (x,y), localmente parametrizamos a curva de nível $c = f(P_0)$, de f, por uma curva δ, com $\delta' \neq \vec{0}$. Esta curva tem vetor tangente ortogonal a $\vec{\nabla} f(P_0)$. Logo, $\delta'(P_0) \perp \vec{\nabla} g(P_0)$. Assim, no ponto P_0, visto que os vetores tangentes δ' e γ' são não nulos e ortogonais a $\vec{\nabla} g(P_0)$, vemos que δ' e γ' são múltiplos um do outro. Portanto, as curvas $L = \{(x,y) :\ g(x,y) = 0\}$ e a de nível $c = f(P_0)$ da função f, tangenciam-se no ponto P_0.

(4) Os extremantes surgem do sistema abaixo, dado por uma equação vetorial e uma equação escalar,

$$\begin{cases} \vec{\nabla} f(x,y) = \lambda \vec{\nabla} g(x,y) \\ g(x,y) = 0, \end{cases}$$

equivalente ao sistema abaixo, que apresenta três equações escalares e três incógnitas x, y e λ:

$$\begin{cases} \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x} \\ \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y} \\ g(x,y) = 0. \end{cases}$$

(5) Na equação (1.1), é possível eliminar o parâmetro λ trivialmente (mas a conveniência depende do problema). De fato, a resolubilidade da equação $\vec{\nabla} f(P_0) = \lambda \vec{\nabla} g(P_0)$ equivale à resolubilidade da equação

$$\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix} (P_0) \;=\; 0.$$

A questão da determinação dos extremantes locais de f sobre a curva $\{(x,y) : g(x,y) = 0\}$ é usualmente referida como o problema da determinação dos máximos e mínimos de f sujeita à restrição $g(x,y) = 0$ ou, dos máximos e mínimos condicionados de f sujeita à condição $g(x,y) = 0$.

Para aplicarmos o Método dos Multiplicadores de Lagrange são necessários os conceitos abaixo.

Definições Topológicas. Sejam A e K dois subconjuntos de $\mathbb{R}^n$.

- A é fechado se seu complementar $\mathbb{R}^n \setminus \{A\} = A^c$ é aberto.
- K é compacto se é fechado e limitado.

Segue um teorema fundamental para estudo de máximos e mínimos.

Teorema (Weierstrass). *Seja $f : K \to \mathbb{R}$ uma função contínua, K um subconjunto compacto em $\mathbb{R}^n$. Então, f assume valor máximo absoluto e valor mínimo absoluto em K.*

Prova. Inicialmente, mostremos que f é limitada.

- f é limitada. Suponhamos, por contradição, que f é ilimitada. Visto que K é limitado, temos que K está contido em um cubo fechado $C = C_0$ em $\mathbb{R}^n$, de faces paralelas aos planos coordenados, e de aresta de comprimento L. Particionemos este cubo em 2^n cubos fechados de faces paralelas aos planos coordenados, de forma natural e trivial. É óbvio que f é ilimitada sobre a intersecção de K com ao menos destes 2^n cubos. Seja C_1 tal cubo. Então, iterando tal argumentação construímos uma sequência C_n, n em $\mathbb{N}$, de cubos fechados com faces paralelas aos planos coordenados e de arestas de comprimento $L/2^n$, tal que o cubo C_{j+1} está contido em C_j, para todo $j \geq 1$, e tal que f é ilimitada sobre a intersecção $K \cap C_n$. Então, pelo Princípio dos Intervalos Encaixantes [vide *Um Curso de Cálculo*, H. L. Guidorizzi, Vol 1, 5ª ed, p. 19] temos que a intersecção de todos estes cubos é dada por um único ponto p em $\mathbb{R}^n$. Ainda, escolhendo para cada n em $\mathbb{N}$ um ponto x_n na intersecção $C_n \cap K$, temos que $|x_n - p| \leq L\sqrt{2}/2^n$. Logo, $x_n \to p$ se $n \to +\infty$. Portanto, como K é fechado, o ponto p pertence a K. Então, como f é contínua em p, temos que existe uma bola aberta $B(p;r)$, centrada em p, com raio $r > 0$, tal que f é limitada sobre $B(p;r) \cap K$. Obviamente existe um cubo C_N contido em tal bola $B(p;r)$. Entretanto, f é ilimitada sobre a intersecção $C_N \cap K$ $\lightning$

- Existência dos pontos de máximo e mínimo. Como $f(K)$ é limitado em $\mathbb{R}$, pela propriedade do supremo temos que existem $m = \inf f(K)$, o ínfimo de $f(K)$, e $m = \sup f(K)$, o supremo de $f(K)$. Suponhamos, por contradição, que $f(x) < M$, para todo x em K. Então, a função

$$\frac{1}{M - f(x)}\,,\; x \text{ em } K\,,$$

é contínua e, como consequência da definição de supremo, ilimitada $\lightning$ Assim, existe b em K tal que $f(b) = M$. Analogamente, existe a em K tal que $f(a) = m$ ■

Exemplo 1. Determine a curva de nível da função $f(x,y) = x^2 + 16y^2$ que seja tangente à curva (um ramo de hipérbole) $xy = 1$, $x > 0$ e $y > 0$. Determine também o ponto de tangência.

Solução.

No ponto (x_0, y_0) em que as curvas se tangenciam seus vetores tangentes são paralelos e seus vetores normais também. O vetor normal à curva de nível é o gradiente $\vec{\nabla} f$ e o vetor normal à hipérbole $xy = 1$ é o vetor gradiente da função $g(x,y) = xy$, $\vec{\nabla} g(x_0,y_0) = \langle y_0, x_0 \rangle$, que é não nulo pois $x_0 y_0 = 1$. Temos então, $\vec{\nabla} f(x_0,y_0) = \lambda \vec{\nabla} g(x_0,y_0)$ para algum $\lambda \in \mathbb{R}$ e,

$$\langle 2x_0, 32y_0 \rangle = \lambda \langle y_0, x_0 \rangle\ .$$

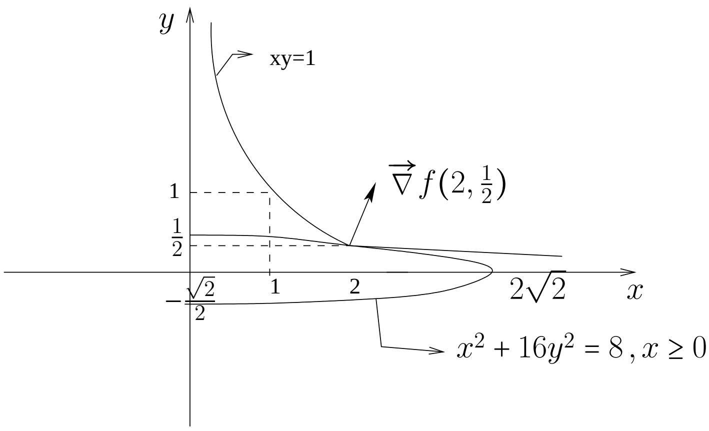

Figura 3: Ilustração ao Exemplo 1

Logo, $2x_0.32y_0 = \lambda y_0.\lambda x_0$ e, como $x_0 y_0 = 1$, obtemos $64 = \lambda^2$ e $\lambda = \pm 8$. Porém, como temos $x_0 > 0$ e $y_0 > 0$, a possibilidade $\lambda = -8$ é inaceitável. Desta forma obtemos $x_0 = 4y_0$ e portanto, $1 = x_0 y_0 = 4y_0^2$ e assim encontramos $y_0 = \frac{1}{2}$ e $x_0 = 2$ e o ponto de tangência

$$(x_0, y_0) = \left(2, \frac{1}{2}\right).$$

A curva de nível é a elipse dada por $x^2 + 16y^2 = f(2, \frac{1}{2}) = 4 + 4 = 8$■

Exemplo 2. Analise os máximos e mínimos de

$$f(x,y) = x^2 - 2xy + 3y^2\,, \quad \text{com a restrição} \quad x^2 + 2y^2 = 1.$$

Solução.

Consideremos a elipse $E = \{(x,y) : x^2 + 2y^2 = 1\}$ e também a função $g(x,y) = x^2 + 2y^2 - 1$. Então, $E = g^{-1}(0)$ e $\overrightarrow{\nabla} g = \langle 2x, 4y \rangle \neq \overrightarrow{0}$ sobre E. Como f é contínua e E é compacto, pelo Teorema de Weierstrass f assume máximo e mínimo sobre E. Pelo Teorema dos Multiplicadores de Lagrange segue que para tais extremantes existe λ em $\mathbb{R}$ tal que

$$\begin{cases} \overrightarrow{\nabla} f(x,y) = \lambda \overrightarrow{\nabla} g(x,y) \\ x^2 + 2y^2 = 1 \end{cases} \qquad \text{ou,} \qquad \begin{cases} \langle 2x - 2y, -2x + 6y \rangle = \lambda \langle 2x, 4y \rangle \\ x^2 + 2y^2 = 1\ . \end{cases}$$

Eliminando o parâmetro λ na primeira equação do sistema à direita temos,

$$\begin{vmatrix} 2x-2y & 6y-2x \\ 2x & 4y \end{vmatrix} = 0 \qquad \Rightarrow \qquad 2y(x-y) - x(3y-x) = 0 \ .$$

Porém, temos $0 = 2(xy-y^2) - (3xy-x^2) = (x-2y)(x+y)$. Assim, obtemos as soluções $P = (2y, y)$ ou $Q = (-y, y)$. Como P e Q pertencem à elipse concluímos que

$$\begin{cases} (2y)^2 + 2y^2 = 1 \\ \text{ou} \\ y^2 + 2y^2 = 1 \end{cases} \Rightarrow \begin{cases} y = \pm\frac{1}{\sqrt{6}} \\ \text{ou} \\ y = \pm\frac{1}{\sqrt{3}} \ . \end{cases} \Rightarrow \begin{cases} P = \pm\left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) \\ \text{ou} \\ Q = \pm\left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right) \ . \end{cases}$$

Computemos f em P e em Q. Como $f(-x,-y) = f(x,y)$ temos,

$$\begin{aligned} f\left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) &= f\left(-\frac{2}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right) &= \frac{1}{2} \\ f\left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right) &= f\left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) &= 2. \end{aligned}$$

Os pontos $\pm\left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}\right)$ são de máximo e os pontos $\pm\left(\frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)$ são de mínimo$\blacksquare$

Teorema 2. *Sejam f e g funções em $C^1(\Omega;\mathbb{R})$, Ω um aberto em $\mathbb{R}^3$. Consideremos a superfície de nível $S = \{P = (x,y,z) \text{ em } \Omega : g(P) = 0\}$, com $\vec{\nabla} g(P) \neq \vec{0}$ para todo P em S. Suponhamos que um ponto P_0 em S é extremante local de f restrita a S. Então, existe um real λ tal que*

$$\vec{\nabla} f(P_0) = \lambda \vec{\nabla} g(P_0) \ .$$

Interpretação. Análoga a dada no caso planar. Se $\vec{\nabla} f(P_0) \neq \vec{0}$ é oblíquo a S em P_0 (i.e., ao plano π tangente a S em P_0), então a superfície S' de nível $f(P_0)$, de f, cruza S em P_0 e as superfícies de nível c, de f, para valores de c próximos de $f(P_0)$, também cruzam S e orientando-as na direção $\vec{\nabla} f(P_0)$, de crescimento de c, vemos que P_0 não é extremante de f, contra a hipótese. Logo, concluímos que $\vec{\nabla} f(P_0)$ é ortogonal a S, assim como $\vec{\nabla} g(P_0) \neq \vec{0}$. Donde deduzimos que os vetores $\vec{\nabla} f(P_0)$ e $\vec{\nabla} g(P_0)$ são paralelos.

Prova.

Pelo Teorema 2 das Funções Implícitas, S é localmente em $P_0 = (x_0, y_0, z_0)$ (isto é, em uma bola aberta contendo P_0), dada pelo gráfico de uma função $z = \varphi(x, y)$, com o par (x, y) numa vizinhança de (x_0, y_0) e $\varphi(x_0, y_0) = z_0$. Pelas hipóteses segue que dada uma curva arbitrária γ no gráfico de φ, satisfazendo $\gamma(0) = P_0$, a função $(f \circ \gamma)$ tem máximo ou mínimo em $t = 0$. Portanto, o vetor gradiente $\overrightarrow{\nabla} f(P_0)$ é ortogonal ao vetor tangente $\gamma'(0)$. Consequentemente, o vetor $\overrightarrow{\nabla} f(P_0)$ é ortogonal ao plano π, tangente ao gráfico de φ no ponto P_0. É claro que o vetor $\overrightarrow{\nabla} g(P_0)$ também é ortogonal ao gráfico de φ. Logo, ao plano π.

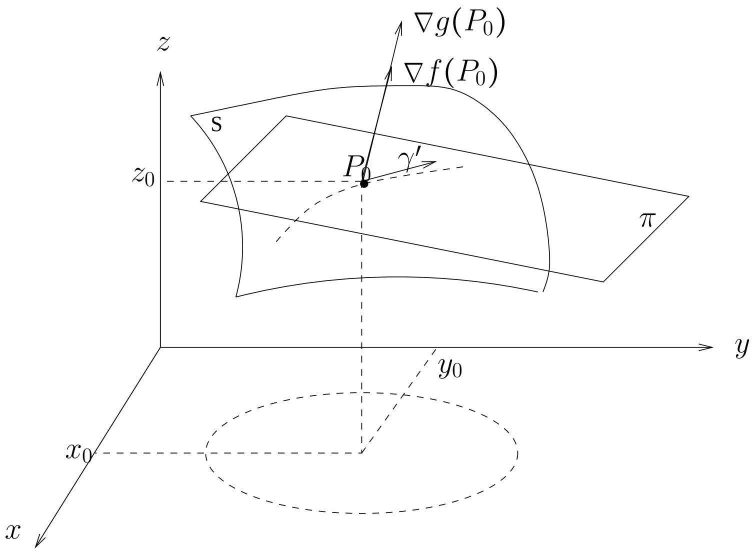

Figura 4: Ilustração ao Teorema 2.

Desta forma, os vetores $\overrightarrow{\nabla} f(P_0)$ e $\overrightarrow{\nabla} g(P_0)$ são paralelos e, como o vetor $\overrightarrow{\nabla} g$ não se anula, concluímos a tese ■

Observação 6. No Teorema 2 também é simples eliminar o parâmetro λ e, novamente, a conveniência depende do problema. A resolubilidade da equação $\overrightarrow{\nabla} f(P_0) = \lambda \overrightarrow{\nabla} g(P_0)$ equivale a resolubilidade de

$$\begin{vmatrix} \overrightarrow{i} & \overrightarrow{j} & \overrightarrow{k} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial x} & \frac{\partial g}{\partial z} \end{vmatrix} (P_0) = \overrightarrow{0} .$$

Exemplo 3. Verifique que existe um único ponto P no plano $3x+2y+z=12$ cuja soma dos quadrados das distâncias de P a $(0,0,0)$ e a $(1,1,1)$ é mínima. Determine-o.

Solução.

Geometricamente, se K é um disco fechado e limitado que contém os pontos dados e intersecta o plano π dado, o ponto procurado pertence a $K\cap\pi$ e é então o mínimo absoluto da função (quadrado da distância)

$$D(x,y,z)=(x-0)^2+(y-0)^2+(z-0)^2+(x-1)^2+(y-1)^2+(z-1)^2$$

restrita ao compacto $K\cap\pi$. O Teorema de Weierstrass assegura a existência de tal mínimo. Pelo Teorema 2 os extremantes de $D(x,y,z)$ com a restrição $3x+2y+z=12$ satisfazem:

$$\vec{\nabla} D(x,y,z)=\lambda\langle 3,2,1\rangle\ ,\ \lambda \text{ em } \mathbb{R}\,,$$

ou, eliminando o parâmetro λ,

$$\vec{0}=\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2x+2(x-1) & 2y+2(y-1) & 2z+2(z-1) \\ 3 & 2 & 1 \end{vmatrix}=2\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 2x-1 & 2y-1 & 2z-1 \\ 3 & 2 & 1 \end{vmatrix}.$$

Desta forma é claro que desenvolvendo o segundo determinante pela primeira linha temos, $(2y-1-4z+2)\,\vec{i}-(2x-1-6z+3)\,\vec{j}+(4x-2-6y+3)\,\vec{k}=\vec{0}$ e portanto, $2y=4z-1$ e $x=3z-1$. Donde, substituindo na equação do plano, obtemos $3(3z-1)+(4z-1)+z=12$. Finalmente,

$$z=\frac{16}{14}\,,\ y=\frac{25}{14}\,,\ x=\frac{34}{14}\ \text{ e } P_0=\left(\frac{34}{14},\frac{25}{14},\frac{16}{14}\right).$$

Atenção. Neste caso podemos eliminar λ mais simplesmente com as equações

$$\frac{2x+2(x-1)}{3}=\frac{2y+2(y-1)}{2}=\frac{2z+2(z-1)}{1}\quad(=\lambda)\qquad\blacksquare$$

Exemplo 4. Consideremos o plano $\pi : ax + by + cz + d = 0$, $a^2 + b^2 + c^2 \neq 0$.

(a) Determine o ponto $P_1 = (x_1, y_1, z_1)$ no plano π mais próximo à origem.

(b) Com o método utilizado em (a) mostre que a distância de todo ponto $P_0 = (x_0, y_0, z_0)$ ao plano é dada pela fórmula:

$$\frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}} \ .$$

1ª Solução.

(a) Geometricamente sabemos que existe tal ponto $P_1 = (x_1, y_1, z_1)$ e que $P_1 = (0, 0, 0)$ se $d = 0$. Passemos à determinação do mínimo da função $\Phi(x, y, z) = x^2 + y^2 + z^2$, o quadrado da distância, sujeita à restrição $ax + by + cz + d = 0$. Utilizando multiplicadores de Lagrange temos,

$$\overrightarrow{\nabla}\Phi(P_1) = \langle 2x_1, 2y_1, 2z_1 \rangle = \lambda \langle a, b, c \rangle \ .$$

Logo, $P_1 = (x_1, y_1, z_1) = \frac{\lambda}{2}(a, b, c)$ e, substituindo tais coordenadas para P_0 na equação para π temos $0 = ax_1 + by_1 + cz_1 + d = \frac{\lambda}{2}(a^2 + b^2 + c^2) + d$. Donde, $\frac{\lambda}{2} = -\frac{d}{a^2+b^2+c^2}$. Assim, o ponto de π mais próximo à origem é

$$P_1 = -\frac{d}{a^2 + b^2 + c^2}(a, b, c) \ .$$

(b) Geometricamente é claro que existe em π o ponto $P_2 = (x_2, y_2, z_2)$ mais próximo de $P_0 = (x_0, y_0, z_0)$. Claramente, P_2 é ponto de mínimo absoluto da função $\Psi(x, y, z) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2$ restrita ao plano $ax + by + cz + d = 0$, que avalia o quadrado da distância dos pontos do plano π ao ponto P_0. Aplicando multiplicadores de Lagrange temos,

$$\overrightarrow{\nabla}\Psi(P_2) = \langle 2(x_2 - x_0), 2(y_2 - y_0), 2(z_2 - z_0) \rangle = \mu \langle a, b, c \rangle \ .$$

Portanto, $\overrightarrow{P_0P_2} = \frac{\mu}{2}\langle a, b, c \rangle$ e a distância de P_2 ao plano π, dada pelo módulo do vetor $\overrightarrow{P_0P_2}$, é então $\frac{|\mu|}{2}|\langle a, b, c \rangle|$. Visto que $P_2 = P_0 + \frac{\mu}{2}(a, b, c)$ pertence a π temos $a(x_0 + \frac{\mu}{2}a) + b(y_0 + \frac{\mu}{2}b) + c(z_0 + \frac{\mu}{2}c) + d = 0$. Donde,

$$\frac{\mu}{2}(a^2 + b^2 + c^2) = -(ax_0 + by_0 + cz_0 + d).$$

Finalmente, a distância procurada é

$$\frac{|\mu|}{2}|\langle a, b, c \rangle| = \frac{\frac{|\mu|}{2}|(a^2 + b^2 + c^2)}{\sqrt{a^2 + b^2 + c^2}} = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

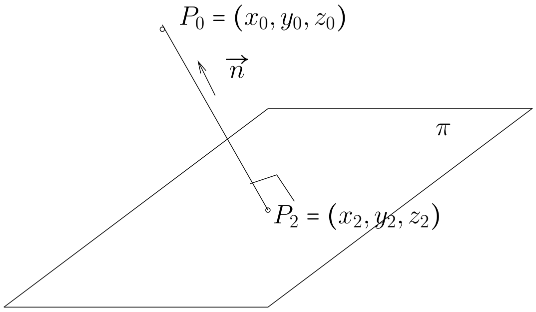

Figura 5: A distância de um ponto a um plano.

2ª Solução (Geométrica, simples e sem multiplicadores de Lagrange):

Se P_2, pertencente a π, é o pé da perpendicular por P_0 ao plano π temos que $\overrightarrow{P_0P_2}$ é paralelo ao vetor normal a π, $\overrightarrow{n_\pi} = \langle a, b, c\rangle \neq \overrightarrow{0}$. Logo, existe λ em $\mathbb{R}$ tal que $\overrightarrow{P_0P_2} = \lambda \langle a, b, c\rangle$ e portanto a distância de P_0 a π é $|\overrightarrow{P_0P_2}|$. Porém, da identidade

$$P_2 = P_0 + \overrightarrow{P_0P_2} = P_0 + \lambda \langle a, b, c\rangle \quad ,$$

segue que

$$0 = a(x_0 + \lambda a) + b(y_0 + \lambda b) + c(z_0 + \lambda c) + d = 0\,,$$

e assim, $\lambda(a^2 + b^2 + c^2) = -(ax_0 + by_0 + cz_0 + d)$ e $\lambda = -\frac{ax_0 + by_0 + cz_0 + d}{a^2+b^2+c^2}$. Donde,

$$|\overrightarrow{P_0P_2}| = |\lambda|\,|\langle a, b, c\rangle| = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}} \qquad \blacksquare$$

Recordemos que: se $\vec{u}$, $\vec{v}$ e $\vec{w}$ são vetores em $\mathbb{R}^3$ tais que $\vec{u} \times \vec{v} \neq \vec{0}$ e $\vec{w}$ é ortogonal a $\vec{u} \times \vec{v}$ então $\vec{w}$ é combinação linear de $\vec{u}$ e $\vec{v}$.

Teorema 3. *Consideremos três funções f, g e h em $C^1(\Omega; \mathbb{R})$, com Ω um aberto em $\mathbb{R}^3$, e a intersecção de duas superfícies de nível zero*

$$L = \big\{P = (x, y, z) \text{ em } \Omega : g(P) = 0 \text{ e } h(P) = 0\big\}.$$

Admitamos que $\vec{\nabla} g(x, y, z)$ e $\vec{\nabla} h(x, y, z)$ são L.I. em todo (x, y, z) em L. Suponhamos que $P_0 = (x_0, y_0, z_0)$ é um extremante local de f sobre L. Então, existem dois números reais λ_1 e λ_2 tais que

$$\vec{\nabla} f(P_0) = \lambda_1 \vec{\nabla} g(P_0) + \lambda_2 \vec{\nabla} h(P_0).$$

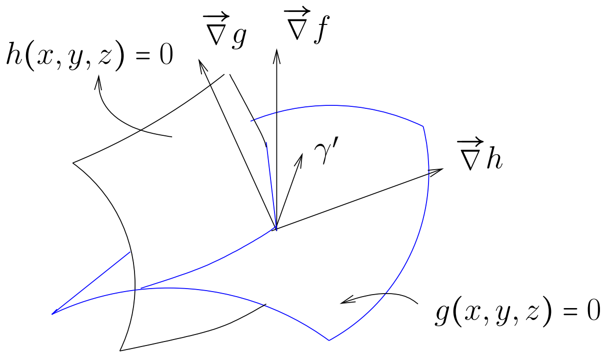

Figura 6: Ilustração ao Teorema 3.

Prova. Podemos supor $g_z(P_0) \neq 0$ pois, por hipótese, são L.I as linhas de

$$M = \begin{bmatrix} g_x(P_0) & g_y(P_0) & g_z(P_0) \\ h_x(P_0) & h_y(P_0) & h_z(P_0) \end{bmatrix}.$$

Se $\begin{vmatrix} g_x(P_0) & g_z(P_0) \\ h_x(P_0) & h_z(P_0) \end{vmatrix} = 0 = \begin{vmatrix} g_y(P_0) & g_z(P_0) \\ h_y(P_0) & h_z(P_0) \end{vmatrix}$, segue que as primeiras colunas de M são múltiplas da terceira. Logo, existem reais α e β tais que

$$M = \begin{bmatrix} \alpha g_z(P_0) & \beta g_z(P_0) & g_z(P_0) \\ \alpha h_z(P_0) & \beta h_z(P_0) & h_z(P_0) \end{bmatrix}.$$

Portanto, as linhas de M são múltiplas de $\begin{bmatrix} \alpha & \beta & 1 \end{bmatrix}$ e L.D. Absurdo!

Assim, supomos sem perda de generalidade que o menor de ordem 2 determinado pelas duas últimas colunas de M é não nulo. Então, pelo Teorema 3 das Funções Implícitas segue que localmente, em P_0, determinamos as coordenadas y e z em função de x e assim concluímos que L é (localmente) uma curva $\gamma(x) = \big(x, y(x), z(x)\big)$, satisfazendo $\gamma(x_0) = P_0$ e $\gamma'(x_0) \neq \vec{0}$. Como a imagem de γ está contida na superfície de nível zero da função g, segue que o gradiente $\vec{\nabla} g(P_0)$ é ortogonal ao vetor tangente $\gamma'(x_0)$ e, analogamente, o gradiente $\vec{\nabla} h(P_0)$ é ortogonal a $\gamma'(x_0)$. Vide Figura 6. Portanto, $\gamma'(x_0)$ é paralelo ao produto vetorial $\vec{\nabla} g(P_0) \times \vec{\nabla} h(P_0)$.

Por hipótese, a função $(f \circ \gamma)(x)$ tem um extremante local em x_0. Logo, pela regra da cadeia temos $\vec{\nabla} f(P_0) \perp \gamma'(x_0)$, com $\gamma'(x_0) \neq \vec{0}$. Assim, $\vec{\nabla} f(P_0)$ é ortogonal ao produto vetorial $\vec{\nabla} g(P_0) \times \vec{\nabla} h(P_0)$. Donde segue que o vetor $\vec{\nabla} f(P_0)$ é uma combinação linear dos vetores $\vec{\nabla} g(P_0)$ e $\vec{\nabla} h(P_0)$ ■

Observação 7. Uma vez mais é elementar eliminar os parâmetros λ_1 e λ_2 (sendo a conveniência circunstancial). De fato, no Teorema 3 temos a identidade $\vec{\nabla} f(P_0) = \lambda_1 \vec{\nabla} g(P_0) + \lambda_2 \vec{\nabla} h(P_0)$ se, e somente se,

$$\begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial z} \end{vmatrix} (P_0) \;=\; 0.$$

Observação 8. Os sistemas encontrados nos teoremas acima são redutíveis a uma única equação vetorial, eliminando equações de restrição e acrescentando variáveis. Para o Teorema 3, e analogamente para os demais, definimos a função $\mathcal{L}$ em cinco variáveis livres (vide Exemplo 6, a seguir), ou não condicionadas,

$$\mathcal{L}(x, y, z, \lambda, \mu) = f(x, y, z) - \lambda g(x, y, z) - \mu\, h(x, y, z).$$

As soluções do sistema mencionado são os pontos de máximo ou mínimo não-condicionados de $\mathcal{L}$, os quais se encontram entre seus (de $\mathcal{L}$) pontos críticos e satisfazem,

$$\vec{\nabla}\mathcal{L}(x, y, z, \lambda, \mu) = \langle f_x - \lambda g_x - \mu\, h_x, f_y - \lambda g_y - \mu\, h_y, f_z - \lambda g_z - \mu h_z, -g, -h\rangle = \vec{0},$$

que é um problema sem condições de restrição e simétrico no sentido que as variáveis tem igual importância. Esta formatação do problema é elegante e trivialmente generalizável para uma função f com qualquer número de variáveis e com qualquer número de restrições.

Definição. A variável λ (ou μ) é um multiplicador de Lagrange.

Em Economia, Geometria Diferencial, Cálculo das Variações, etc. o multiplicador λ é convenientemente interpretado e tem importância por si só.

Exemplo 5. Determine os pontos mais afastados da origem e com coordenadas sujeitas às restrições $x^2 + 4y^2 + z^2 = 4$ e $x + y + z = 1$.

Solução.

Determinemos os pontos que maximizam a função $D(x, y, z) = x^2 + y^2 + z^2$, o quadrado da distância de (x, y, z) a $(0, 0, 0)$ com as restrições $g(x, y, z) = 0$

e $h(x,y,z) = 0$, com $g(x,y,z) = x + y + z - 1$ e $h(x,y,z) = x^2 + 4y^2 + z^2 - 4$. Tais pontos pertencem à intersecção do elipsóide $g = 0$ com o plano $h = 0$, a qual é um compacto K de $\mathbb{R}^3$. Como D é contínua, D assume máximo e mínimo em K. Pela Observação 7 ao Teorema 3, se $P_0 = (x_0, y_0, z_0)$ é extremante local de D sobre $K = \{(x,y,z) : g = 0 \text{ e } h = 0\}$ temos

$$\begin{vmatrix} 2x_0 & 2y_0 & 2z_0 \\ 1 & 1 & 1 \\ 2x_0 & 8y_0 & 2z_0 \end{vmatrix} = 4 \begin{vmatrix} x_0 & y_0 & z_0 \\ 1 & 1 & 1 \\ x_0 & 4y_0 & z_0 \end{vmatrix} = 0.$$

Logo, $0 = x_0(z_0 - 4y_0) - y_0(z_0 - x_0) + z_0(4y_0 - x_0) = 3y_0(z_0 - x_0)$ e temos $y_0 = 0$ ou $z_0 = x_0$. Se $y_0 = 0$ obtemos

$$\begin{cases} x + z = 1 \\ x^2 + z^2 = 4\,, \end{cases}$$

$$x^2 + (1-x)^2 = 4 \iff 2x^2 - 2x - 3 = 0 \iff x = \frac{1 \pm \sqrt{7}}{2}.$$

Assim, $P_1 = \left(\frac{1+\sqrt{7}}{2}, 0, \frac{1-\sqrt{7}}{2}\right)$ e $P_2 = \left(\frac{1-\sqrt{7}}{2}, 0, \frac{1+\sqrt{7}}{2}\right)$ são candidatos a extremantes. Se $z_0 = x_0$ temos,

$$\begin{cases} 2x + y = 1 \\ 2x^2 + 4y^2 = 4 \end{cases} \iff \begin{cases} y = 1 - 2x \\ x^2 + 2y^2 = 2 \end{cases}$$

$$x^2 + 2(1-2x)^2 = 2 \iff 9x^2 - 8x = 0 \iff x = 0 \text{ ou } x = \frac{8}{9}\,.$$

Neste caso os candidatos são $P_3 = (0,1,0)$ e $P_4 = \left(\frac{8}{9}, -\frac{7}{9}, \frac{8}{9}\right)$.

Comparando os valores $D(P_1) = D(P_2) = 4$, $D(P_3) = 1$ e $D(P_4) = \frac{177}{81}$ constatamos que P_1 e P_2 são os pontos desejados ■

Exemplo 6. Determine o ponto sobre a reta de intersecção dos planos $x + 2y + z = 1$ e $-3x - y + 2z = 4$ que está mais próximo à origem.

Solução.

Minimizemos $x^2+y^2+z^2$ com as restrições $x+2y+z-1 = 0$ e $-3x-y+2z-4 = 0$. Definindo

$$\mathcal{L}(x,y,z,\lambda,\mu) = x^2 + y^2 + z^2 - \lambda(x + 2y + z - 1) - \mu(-3x - y + 2z - 4)\,,$$

resolvamos o sistema

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} & = 2x - \lambda + 3\mu = 0\,, \\ \frac{\partial \mathcal{L}}{\partial y} & = 2y - 2\lambda + \mu = 0\,, \\ \frac{\partial \mathcal{L}}{\partial z} & = 2z - \lambda - 2\mu = 0\,, \\ \frac{\partial \mathcal{L}}{\partial \lambda} & = -(x + 2y + z - 1) = 0\,, \\ \frac{\partial \mathcal{L}}{\partial \mu} & = (3x + y - 2z + 4) = 0\ . \end{cases}$$

Das três primeiras equações temos,

$$x = \frac{\lambda - 3\mu}{2}, \qquad y = \frac{2\lambda - \mu}{2}, \qquad z = \frac{\lambda + 2\mu}{2},$$

que substituídas na quarta e na quinta equações (e simplificando) fornecem

$$\begin{cases} 6\lambda - 3\mu & = 2 \\ 3\lambda - 14\mu & = -8, \end{cases}$$

e $\lambda = \frac{52}{75}$ e $\mu = \frac{54}{75}$. Donde concluímos, $x = -\frac{11}{15}$, $y = \frac{5}{15}$ e $z = \frac{16}{15}$■

No espaço $\mathbb{R}^n$, chamamos um subespaço vetorial de dimensão $n-1$ de hiperplano (se $n = 3$, um hiperplano é um plano passando pela origem).

Exemplo 7. Seja $f : \mathbb{R}^n \to \mathbb{R}$ diferenciável e tal que a restrição de f a um hiperplano π tem um máximo em um ponto P no hiperplano. Verifique que dado um vetor arbitrário $\vec{v}$ unitário e paralelo a π então temos

$$\frac{\partial f}{\partial \vec{v}}(P) = \nabla f(P) \cdot \vec{v} = 0.$$

Isto é, $\nabla f(P)$ é ortogonal a π. Supondo $\vec{n}$ o vetor normal ao hiperplano π conclua que existe λ em $\mathbb{R}$ tal que

$$\nabla f(P) = \lambda \vec{n}\,.$$

Seja Ω um aberto em $\mathbb{R}^n$, n fixo em $\mathbb{N}$. Como consequência do Teorema da Função Implícita para uma função g em $C^1(\Omega;\mathbb{R})$, mostremos que vale um resultado análogo aos teoremas 1 e 3, no espaço n-dimensional.

Teorema 4. *Sejam f e g, ambas em $C^1(\Omega;\mathbb{R})$, e $S = \{x \in \Omega : g(x) = 0\}$, a superfície de nível zero de g, com $\overrightarrow{\nabla} g(x) \neq \overrightarrow{0}$ para todo x em S. Seja P, em S, um extremante local de f restrita a S. Então, existe λ em $\mathbb{R}$ tal que*

$$\overrightarrow{\nabla} f(P) = \lambda \overrightarrow{\nabla} g(P)\,.$$

Interpretação. A superfície solução S da equação $g(x) = 0$ possui hiperplano tangente π e para toda curva γ em S, $\gamma(0) = P$, a composta $f \circ \gamma$ tem um extremante em P. Donde, $\nabla f(P)$ é ortogonal ao vetor tangente $\gamma'(0)$ em π. Assim, $\nabla f(P)$ e $\nabla g(P)$ são ortogonais a π. A conclusão é então trivial.

Prova.

Indiquemos por $x = (x', x_n)$, com $x' = (x_1, \ldots, x_{n-1})$ em $\mathbb{R}^{n-1}$ e x_n em $\mathbb{R}$, a variável em $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$. Analogamente, $P = (p_1, ..., p_n) = (p', p_n)$. Suponhamos, sem perda de generalidade, $\frac{\partial g}{\partial x_n}(P) \neq 0$. Pelo Teorema da Função Implícita concluímos que S é, localmente no ponto P, o gráfico de uma função $\varphi(x')$ de classe C^1.

Descrevamos S, localmente em P, como $S = \Big\{\big(x', \varphi(x')\big) : x' \text{ pertence a } A\Big\}$, onde A é um aberto em $\mathbb{R}^{n-1}$, com p' em A e $\big(p', \varphi(p')\big) = P$. É então válida a identidade

$$g\Big(x_1, ..., x_{n-1}, \varphi(x_1, ..., x_{n-1})\Big) = 0\,, \ \text{para todo } (x_1, ..., x_{n-1}) \text{ em } A.$$

Assim, pela regra da cadeia e omitindo os pontos de aplicação x' e $\big(x', \varphi(x')\big)$,

$$\begin{bmatrix} \frac{\partial g}{\partial x_1} & \ldots & \frac{\partial g}{\partial x_n} \end{bmatrix} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ 0 & . & \ldots & 0 \\ 0 & . & \ldots & 0 \\ 0 & . & \ldots & 1 \\ \frac{\partial \varphi}{\partial x_1} & \frac{\partial \varphi}{\partial x_2} & \ldots & \frac{\partial \varphi}{\partial x_{n-1}} \end{bmatrix}_{n \times (n-1)} = \begin{bmatrix} 0 & \ldots & 0 \end{bmatrix},$$

onde a matriz nula indicada é de tamanho $1 \times (n-1)$. É claro que as $n-1$ colunas da matriz $n \times (n-1)$ acima são linearmente independentes em todo ponto x' em A. É fácil constatar que o vetor ∇g é, localmente em S, ortogonal aos vetores correspondentes às $n-1$ colunas citadas.

Por hipótese, a função $f\big(x', \varphi(x')\big)$ tem um extremante em $p' = (p_1, ..., p_{n-1})$. Logo, suas derivadas parciais se anulam em p'. Sabidamente, tais derivadas parciais correspondem às colunas da matriz jacobiana de $f\big(x', \varphi(x')\big)$, cuja expressão é análoga a da matriz jacobiana da função $g\big(x', \varphi(x')\big)$ dada acima, bastando trocar g por f.

Assim, a função f satisfaz a mesma equação matricial que g, no ponto $\big(p', \varphi(p')\big) = P$. Portanto, $\nabla f(P)$ e $\nabla g(P)$ são ortogonais a um mesmo conjunto de $n-1$ vetores linearmente independentes em $\mathbb{R}^n$. Logo, os vetores $\nabla f(P)$ e $\nabla g(P) \neq \overrightarrow{0}$ são paralelos. Donde, concluímos a tese ■

Para demonstrarmos o próximo teorema, necessitamos do lema abaixo cuja prova é agradavelmente simples ainda que não concisa.

Lema 5. *Seja M uma matriz em $M_{m\times n}(\mathbb{R})$, com $m \leq n$, tal que suas m linhas são linearmente independentes. Então, existem m colunas de M tais que o determinante da matriz $m \times m$ formado por tais colunas é não zero.*

Prova. Escrevamos

$$M_1 = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix}.$$

A primeira linha de M_1 é não nula e assim, existe uma j-ésima coluna com primeiro elemento $a_{1j} = \lambda_{1j} \neq 0$. Então, multiplicando a primeira linha por um número conveniente e somando-a à segunda linha obtemos uma nova matriz cuja segunda linha apresenta o número 0 na j-ésima coluna. As m linhas desta nova matriz são também linearmente independentes e todos os seus menores (determinantes) de ordem m são iguais aos correspondentes menores (determinantes) originais. Iterando tal procedimento, multiplicamos a primeira linha por sucessivos números convenientes e a somamos ordenadamente às demais linhas e obtemos uma matriz M_2 satisfazendo as condições: o primeiro elemento na j-ésima coluna é $\lambda_{1j} = a_{1j} \neq 0$, os demais elementos na j-ésima coluna são nulos, suas linhas são linearmente independentes e todos os seus menores de ordem m são iguais aos respectivos menores originais da matriz M_1.

Escrevamos,

$$M_2 = \begin{bmatrix} b_{11} & \dots & \lambda_{1j} & \dots & b_{1n} \\ . & \dots & 0 & \dots & . \\ . & \dots & 0 & \dots & . \\ . & \dots & 0 & \dots & . \\ b_{m1} & \dots & 0 & \dots & b_{mn} \end{bmatrix}.$$

A seguir, notemos que a segunda linha de M_2 é não nula e assim, existe uma k-ésima coluna de M_2, onde $k \neq j$, com um elemento $b_{2k} = \lambda_{2k} \neq 0$. Então, argumentando analogamente ao parágrafo anterior, encontramos a próxima matriz M_3 (v. abaixo) satisfazendo: o segundo elemento em sua k-ésima coluna é $\lambda_{2k} = b_{2k} \neq 0$, os demais elementos na k-ésima coluna são nulos (a operação de multiplicar a segunda linha de M_2 por uma constante e então adicioná-la à primeira linha não altera o elemento λ_{1j} presente na j-ésima coluna da primeira linha de M_2), a j-ésima coluna da matriz M_3 é igual a j-ésima coluna da matriz M_2, suas linhas são linearmente independentes e, ainda, todos os seus menores (determinantes) de ordem m são iguais aos respectivos menores de M_2. Escrevamos,

$$M_3 = \begin{bmatrix} c_{11} & . & . & 0 & . & . & \lambda_{1j} & . & . & . & c_{1n} \\ . & . & . & \lambda_{2k} & . & . & 0 & . & . & . & . \\ . & . & . & 0 & . & . & 0 & . & . & . & . \\ . & . & . & 0 & . & . & 0 & . & . & . & . \\ c_{m1} & . & . & 0 & . & . & 0 & . & . & . & c_{mn} \end{bmatrix}.$$

Então, iterando tal processo encontramos as m colunas desejadas ■

Aproveitando a ocasião, vale a pena notar que o Lema 5 implica trivialmente o **Teorema do Posto:** *Se M é uma matriz em $M_{m\times n}(\mathbb{R})$ então, o número p de linhas linearmente independentes de M é igual ao número de colunas linearmente independentes e existe um menor (determinante) de M, de ordem p e não nulo.* Ainda, o Teorema do Posto implica o **Teorema do Núcleo e da Imagem**. Vide demonstrações no Apêndice 1.

Teorema 6. *Seja f em $C^1(\mathbb{R}^n;\mathbb{R})$ e $g = (g_1,\dots,g_m)$ em $C^1(\mathbb{R}^n;\mathbb{R}^m)$. Suponhamos que f tem um extremante local no ponto P em $S = g^{-1}(0,\dots,0)$.*

Suponhamos também que o conjunto $\{\nabla g_1, \ldots, \nabla g_m\}$, onde $m < n$, é linearmente independente em $\mathbb{R}^n$ em todo ponto. Então, existem m números reais $\lambda_1, \ldots, \lambda_m$ tais que

$$\nabla f(P) = \lambda_1 \nabla g_1(P) + \cdots + \lambda_m \nabla g_m(P).$$

Interpretação. Consideremos, para cada $j = 1, \ldots, m$, o hiperplano em $\mathbb{R}^n$ tangente à superfície $g_j^{-1}(0)$ no ponto P. A intersecção destes hiperplanos é o espaço vetorial V tangente a S e de dimensão $n - m$. O espaço vetorial $V^\perp$ ortogonal a V, tem dimensão m. O conjunto $\{\nabla g_1(P), \ldots, \nabla g_m(P)\}$ é ortogonal a V e portanto uma base de $V^\perp$. Ainda, dada uma curva arbitrária γ em S, com $\gamma(0) = P$, temos que a função $f \circ \gamma$ tem um extremante em $t = 0$ e assim, $\nabla f(P)$ é ortogonal a $\gamma'(0)$. Logo, o vetor $\nabla f(P)$ é ortogonal a V e uma combinação linear de $\nabla g_1(P), \ldots, \nabla g_m(P)$.

Prova.

Escrevamos $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^m$, com $k = n - m$. Indiquemos um ponto em $\mathbb{R}^k \times \mathbb{R}^m$ por (x, y), com $x \in \mathbb{R}^k$ e y em $\mathbb{R}^m$. Analogamente, escrevemos $P = (x_0, y_0)$ com x_0 em $\mathbb{R}^k$ e y_0 em $\mathbb{R}^m$. Pelo Lema 5 acima podemos supor que a matriz formada pelos gradientes das funções $g_{j's}$ tem a forma

$$\left[\begin{array}{ccccc|ccccc}
\frac{\partial g_1}{\partial x_1} & . & . & . & \frac{\partial g_1}{\partial x_k} & \frac{\partial g_1}{\partial y_1} & . & . & . & \frac{\partial g_1}{\partial y_m} \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . \\
\frac{\partial g_m}{\partial x_1} & . & . & . & \frac{\partial g_m}{\partial x_k} & \frac{\partial g_m}{\partial y_1} & . & . & . & \frac{\partial g_m}{\partial y_m}
\end{array}\right]_{m \times n}$$

na qual a matriz $\left[\frac{\partial g_i}{\partial y_j}(x, y)\right]$ é inversível. Então, pelo Teorema da Função Implícita, em uma vizinhança de P a superfície S é o gráfico de uma função $\varphi : A \to \mathbb{R}^m$ de classe C^1, onde A é um aberto contendo x_0. Então temos,

$$g\big(x, \varphi(x)\big) = 0\,, \text{ para todo } x \text{ em } A.$$

Aplicando a regra da cadeia obtemos então a equação matricial

$$\begin{bmatrix} \frac{\partial g_1}{\partial x_1} & . & . & \frac{\partial g_1}{\partial x_k} & \frac{\partial g_1}{\partial y_1} & . & . & \frac{\partial g_1}{\partial y_m} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ \frac{\partial g_m}{\partial x_1} & . & . & \frac{\partial g_m}{\partial x_k} & \frac{\partial g_m}{\partial y_1} & . & . & \frac{\partial g_m}{\partial y_m} \end{bmatrix} \begin{bmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & 1 \\ \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & . & . & \frac{\partial \varphi_1}{\partial x_m} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{\partial \varphi_m}{\partial x_1} & \frac{\partial \varphi_m}{\partial x_2} & . & . & \frac{\partial \varphi_m}{\partial x_m} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}_{m \times k}.$$

Fixemos um ponto $\big(x, \varphi(x)\big)$ no gráfico de φ. Consideremos V o sub-espaço vetorial de $\mathbb{R}^n$ gerado pelas k colunas linearmente independentes da matriz $(k+m) \times k$ na equação matricial acima e $V^\perp$ o sub-espaço vetorial de $\mathbb{R}^n$ que é ortogonal a V. Sabidamente, a dimensão de $V^\perp$ é $n - k = m$. Pela citada equação matricial vemos que o conjunto linearmente independente $\{\nabla g_1, ..., \nabla g_m\}$ está contido em $V^\perp$ e é, portanto, uma base de $V^\perp$.

Por hipótese, a função $f\big(x, \varphi(x)\big)$ tem um extremante em x_0. Logo, suas derivadas parciais se anulam em x_0. Sabidamente, tais derivadas parciais correspondem às colunas da matriz jacobiana de $f\big(x, \varphi(x)\big)$ no ponto x_0,

$$\begin{bmatrix} \frac{\partial f}{\partial x_1} & . & . & \frac{\partial f}{\partial x_k} & \frac{\partial f}{\partial y_1} & . & . & \frac{\partial f}{\partial y_m} \end{bmatrix} \begin{bmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & . & . & . & 1 \\ \frac{\partial \varphi_1}{\partial x_1} & \frac{\partial \varphi_1}{\partial x_2} & . & . & \frac{\partial \varphi_1}{\partial x_m} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{\partial \varphi_m}{\partial x_1} & \frac{\partial \varphi_m}{\partial x_2} & . & . & \frac{\partial \varphi_m}{\partial x_m} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}_{1 \times k}.$$

Segue então que o vetor $\nabla f(P)$ é ortogonal a V. Portanto, $\nabla f(P)$ é uma combinação linear dos vetores da base $\mathcal{C} = \{\nabla g_1(P), \ldots, \nabla g_m(P)\}$ de $V^\perp$ $\blacksquare$.

Máximos e mínimos, locais e absolutos, de uma função f em $C^1(K;\mathbb{R})$, K um compacto.

Definições topológicas. Seja A um subconjunto de $\mathbb{R}^n$.

- O ponto P de A, é um ponto interior de A se existe uma bola aberta, não vazia, centrada em P e contida em A. Isto é, se existe $r > 0$ tal que $B(P;r)$ está contida em A.
- O ponto P de $\mathbb{R}^n$ é um ponto de fronteira de A se toda bola aberta, não vazia, centrada em P intersecta A e também $\mathbb{R}^n \setminus A = A^C$, o complementar de A. Isto é, se para todo $r > 0$, temos

$$B(P;r) \cap A \neq \varnothing \quad \text{e também} \quad B(P;r) \cap A^C \neq \varnothing.$$

- O interior de A, $\mathrm{int}(A)$, é o conjunto dos pontos interiores de A.
- A fronteira de A, ∂A, é o conjunto dos pontos de fronteira de A.

Dado P em A, temos uma só das possibilidades: ou $P \in \mathrm{int}(A)$, ou $P \in \partial A$.

Notemos que:

(A) Pelo Teorema de Weierstrass f assume máximo e mínimo, absolutos.

(B) Os pontos de máximo e mínimo locais e interiores a K são pontos críticos de f. Isto é, em tais pontos o gradiente se anula. Assim, adotamos o procedimento abaixo.

(1) Restringindo f a $\mathrm{int}(K)$ determinamos os pontos críticos, candidatos a extremantes.

(2) Encontramos os possíveis pontos de máximo e mínimo de f sobre a fronteira, ∂K, ou elementarmente ou por multiplicadores de Lagrange.

(3) O máximo e o mínimo absolutos estão entre os valores de f nos pontos acima obtidos.

Exemplo 8. Dada $f(x,y) = 6x^2 + 18xy + 4y^2 - 6x - 10y + 5$, determine os extremantes e os máximos e mínimos locais e absolutos de f no quadrado

$$K = \{(x,y) : |x| \le 1 \text{ e } |y| \le 1\}.$$

Solução. Os pontos críticos de f são dados pela equação

$$\overrightarrow{\nabla} f = \langle 12x + 18y - 6, 18x + 8y - 10\rangle = \overrightarrow{0} \;\Rightarrow\; 2x + 3y - 1 = 0 \text{ e } 9x + 4y - 5 = 0.$$

Logo, o único ponto crítico no interior de K é $P_0 = \left(\frac{11}{19}, -\frac{1}{19}\right)$. A matriz hessiana de f é,

$$\mathcal{H}_f = \begin{bmatrix} 12 & 18 \\ 18 & 8 \end{bmatrix}$$

e então, $f_{xx}(P_0) = 12 > 0$ e o determinante hessiano $H_f(P_0)$ é negativo. Logo, P_0 é ponto de sela e f não tem máximo ou mínimo local em $\text{int}(D)$.

Assim, o máximo e o mínimo absolutos pertencem à fronteira de K,

$$\partial K = \{-1\} \times [-1,1] \;\cup\; \{1\} \times [-1,1] \;\cup\; [-1,1] \times \{-1\} \;\cup\; [-1,1] \times \{1\}\ .$$

Na figura abaixo as setas indicam a direção do vetor ortogonal a ∂K.

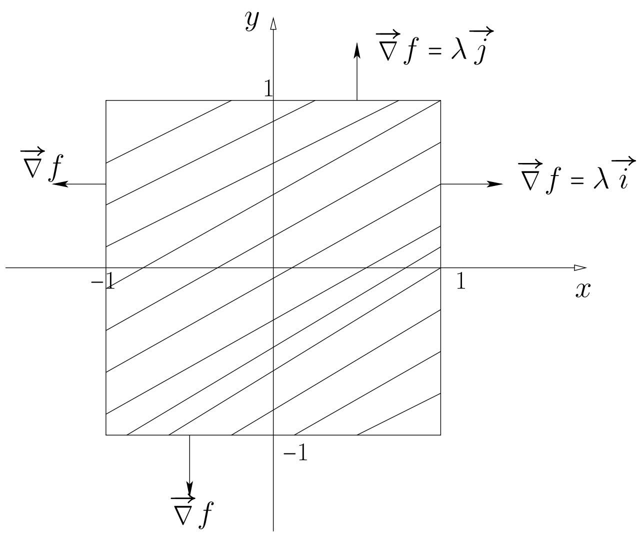

Figura 7: Ilustração ao Exemplo 7

Os extremantes locais na fronteira de K, mas não um vértice, satisfazem (v. Teorema 2):

Em $\{-1\}\times]-1,1[$ temos $x=-1$ e $0=f_y=-18+8y-10$; logo, $y=7/2$, possibilidade que descartamos.

Em $\{1\}\times]-1,1[$ temos $x=1$ e $0=f_y=18+8y-10$; logo, $y=-1$ e $P_1=(1,-1)$ que não pertence ao segmento considerado, é um vértice e será analizado à parte.

Em $]-1,1[\times\{-1\}$ temos $y=-1$ e $0=f_x=12x-18-6$; logo, $x=2$, que também descartamos.

Em $]-1,1[\times\{1\}$ temos $y=1$ e $0=f_x=12x+18-6$; logo, $x=-1$ e $P_2=(-1,1)$ que não pertence ao segmento considerado, é um vértice e será analizado à parte.

Finalmente, os pontos de máximo e o mínimo se encontram entre o vértices. Temos, $f(1,1)=17$, $f(-1,-1)=49$, $f(1,-1)=+1$, e $f(-1,+1)=-7$.

Resposta. Não existem máximo ou mínimo locais e interiores. Ainda, $f(-1,+1)=-7$ é o mínimo absoluto e $f(-1,-1)=49$ é o máximo absoluto■

Exemplo 9. Determine os pontos de máximo e mínimo absoluto da função

$$f(x,y,z)=x^2+y^2+z^2+x+y$$

sobre o compacto $M=\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2\leq 4 \text{ e } z\geq 1\}$.

Solução. Pelo Teorema de Weierstrass a função f admite pontos de máximo e mínimo absolutos em M. Se tais pontos estiverem no interior de M então eles são pontos críticos. Os pontos críticos de f satisfazem

$$\vec{\nabla} f=\langle 2x+1\,,2y+1\,,2z\rangle=\langle 0\,,0\,,0\rangle$$

e portanto o único ponto crítico $(-1/2\,,-1/2\,,0)$ não pertence ao interior de M e assim os extremantes de f pertencem à fronteira de M, ∂M. Então, um extremante P_0 arbitrário pertence a

$$M_1=\Big\{(x,y,z) : x^2+y^2+z^2=4 \text{ e } z>1\Big\} \quad \text{ou} \quad M_2=\Big\{(x,y,1) : x^2+y^2<3\Big\} \text{ ou}$$

$$M_3=\{(x,y,z) : x^2+y^2+z^2=4 \text{ e } z=1\}.$$

Se $P_0 = (x_0, y_0, z_0)$ estiver em M_1, por Multiplicadores de Lagrange temos, notando que $\langle 2x_0\,, 2y_0\,, 2z_0\rangle \neq \overrightarrow{0}$ sobre M_1,

$$\overrightarrow{\nabla} f = \langle 2x_0 + 1\,, 2y_0 + 1\,, 2z_0\rangle = \lambda\,\langle 2x_0\,, 2y_0\,, 2z_0\rangle$$

e portanto $2z_0 = \lambda 2z_0$ e então : ou temos $\lambda = 1$ ou temos $z_0 = 0$. Se $z_0 = 0$, então P_0 não pertence a M_1 e descartamos tal possibilidade. Se $\lambda = 1$ então temos $2x_0 + 1 = 2x_0$, uma contradição!

Se P_0 pertence a M_2 (onde temos $z_0 = 1$) então segue

$$\langle 2x_0 + 1, 2y_0 + 1, 2z_0\rangle = \lambda\,\langle 0, 0, 1\rangle\,.$$

Logo, encontramos o candidato $P_1 = (-1/2, -1/2, 1)$.

Se P_0 pertence a M_3 então P_0 é extremante condicionado de

$$\begin{cases} f(x,y,z) = x^2 + y^2 + x + y + z\,, \\ x^2 + y^2 + z^2 = 4 \\ z = 1. \end{cases}$$

Pelo Método dos Multiplicadores de Lagrange obtemos, notando que os gradientes $\langle 2x_0 + 1, 2y_0 + 1, 2z_0\rangle$ e $\langle 0,0,1\rangle$ são L.I. sobre M_3,

$$\nabla f(P_0) = \lambda_1\,\langle 2x_0 + 1, 2y_0 + 1, 2z_0\rangle + \lambda_2\,\langle 0, 0, 1\rangle\,.$$

Logo,

$$\begin{vmatrix} 2x_0 + 1 & 2y_0 + 1 & 2 \\ 2x_0 & 2y_0 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 0$$

Donde segue $2y_0(2x_0 + 1) - 2x_0(2y_0 + 1) = 0$ e $y_0 = x_0$. Substituindo na equação da esfera segue $2x_0^2 = 3$ e $x_0 = \pm\sqrt{6}/2$. Desta forma, encontramos os candidatos $P_2 = (\frac{\sqrt{6}}{2}, \frac{\sqrt{6}}{2}, 1)$ e $P_3 = (-\frac{\sqrt{6}}{2}, -\frac{\sqrt{6}}{2}, 1)$.

Como $f(P_1) = \frac{1}{2}$, $f(P_2) = 4 + \sqrt{6}$ e $f(P_3) = 4 - \sqrt{6}$, concluímos que P_1 é ponto de mínimo absoluto e P_2 é ponto de máximo absoluto $\blacksquare$

Exemplo 10 (Um Argumento Importante).

Argumento. Seja $S = \{(x,y,z) : \varphi(x,y,z) = 0\}$ a superfície de nível 0 de uma função $\varphi : \Omega \to \mathbb{R}$, Ω um aberto de $\mathbb{R}^3$, com φ de classe C^1 e

$$\vec{\nabla}\varphi(x,y,z) \neq \vec{0}\,, \quad \text{para todo } (x,y,z) \text{ em } S.$$

Seja $P_1 = (x_1, y_1, z_1)$ em $\mathbb{R}^3$ tal que P_1 não pertence a S. Suponhamos que P_0 é o ponto em S que realiza a distância do ponto P_1 à superfície S [isto é, vale a desigualdade $|P_1 - P_0| \leq |P_1 - P|$, para todo P em S]. Então temos,

$$\overrightarrow{P_0P_1} = \lambda \vec{\nabla}\varphi(P_0) \text{ para algum } \lambda \in \mathbb{R}\,.$$

Donde segue, $\overrightarrow{P_0P_1} \perp S$. Consequentemente obtemos: $\overrightarrow{P_0P_1} \perp \gamma'(t_0)$, se γ é uma curva arbitrária de classe C^1, em S, tal que $\gamma(t_0) = P_0$.

Prova. Seja $D(x,y,z) = (x-x_1)^2 + (y-y_1)^2 + (z-z_1)^2$ definida em $\mathbb{R}^3$. Considerando um disco fechado $\overline{D}(P_1; R)$ centrado em P_1 e com raio $R > 0$ suficientemente grande tal que $K = \overline{D}(P_1; R) \cap S \neq \varnothing$, temos que K é compacto (pois fechado e limitado). Pelo Teorema de Weierstrass a função D restrita a K assume um valor mínimo em algum ponto $P \in S$. É claro que P é o ponto P_0 procurado, que realiza a distância de P_1 à superfície S.

Então, como D restrita a S assume valor mínimo em $P_0 \in S$, pelo Método dos Multiplicadores de Lagrange é válida a equação:

$$\vec{\nabla}D(P_0) = \lambda \vec{\nabla}\varphi(P_0)\,, \text{ para algum } \lambda \in \mathbb{R}\,.$$

Computemos o gradiente de $D(x,y,z)$. É fácil ver que

$$\vec{\nabla}D = \langle 2(x-x_1)\,, 2(y-y_1)\,, 2(z-z_1)\rangle\,,$$

e portanto,

$$\vec{\nabla}D(P_0) = 2\,\langle x_0 - x_1\,, y_0 - y_1\,, z_0 - z_1\rangle \;=\; -2\overrightarrow{P_0P_1}\,,$$

e assim temos (notemos, abaixo, que a constante -2 é absorvível por λ)

$$-2\overrightarrow{P_0P_1} = \lambda \vec{\nabla}\varphi(P_0) \quad \blacksquare$$

Uma Aplicação de Multiplicadores de Lagrange

Notações. Identifiquemos vetores em $\mathbb{R}^n$ com matrizes-colunas em $M_{n\times 1}(\mathbb{R})$:

$$X = (x_1, \ldots, x_n) \equiv \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A aplicação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ determinada por A em $M_{n\times n}(\mathbb{R})$ é

$$T(X) = AX\,,\ \ X \text{ em } \mathbb{R}^n.$$

Seja A em $M_n(\mathbb{R})$. Então, um real λ é auto-valor de A se existe $\vec{v}$ tal que

$$A\vec{v} = \lambda\vec{v}\ \ .$$

Seja A em $M_n(\mathbb{R})$ e simétrica. A forma quadrática associada a A é a aplicação

$$Q : \mathbb{R}^n \to \mathbb{R}\,, \quad Q(X) = X^t AX = AX \cdot X \quad .$$

Lema. *Se Q é a forma quadrática associada à matriz simétrica A então*

$$\vec{\nabla} Q(X) = 2AX \quad .$$

Prova.

Se $A = [a_{ij}]$ então $AX = \Big(\sum_{j=1}^{n} a_{1j}x_j\,, \ldots., \sum_{j=1}^{n} a_{nj}x_j\Big)$ e assim, como $a_{ij} = a_{j\,i}$,

$$Q(X) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij}x_j = 2\sum_{1\le i<j\le n} a_{ij}x_ix_j \ + \ \sum_{i=1}^{n} a_{ii}x_i^2\ .$$

Logo,

$$\frac{\partial Q}{\partial x_k} \ = \ 2\sum_{j\neq k} a_{kj}x_j \ + \ 2a_{kk}x_k = 2\sum_{j=1}^{n} a_{kj}x_j \quad \Longrightarrow \quad \vec{\nabla} Q(X) = 2AX \quad \blacksquare$$

A forma quadrática Q é definida positiva se temos $Q(X) > 0$, $\forall X \neq 0$. Analogamente, Q é definida negativa se temos $Q(X) < 0$, para todo $X \neq 0$.

Corolário. *Sejam M, o máximo, e m, o mínimo, da forma quadrática Q associada à matriz simétrica A sobre a esfera unitária $\{X \text{ em } \mathbb{R}^n : |X| = 1\}$.*

(a) Os números M e m são, respectivamente, o maior e o menor auto-valores (reais) de A.

(b) $m|X|^2 \leq Q(X) \leq M|X|^2$, para todo X em $\mathbb{R}^n$.

(c) Q é definida positiva se e só se os auto-valores de A são estritamente positivos.

(d) Q é definida negativa se e só se os auto-valores de A são estritamente negativos.

Prova.[1]

(a) Por ser contínua, a forma quadrática Q assume valor máximo e valor mínimo sobre a esfera compacta $S^{n-1} = \{X \in \mathbb{R}^n : |X| = 1\}$.

Pelo Teorema dos Multiplicadores de Lagrange, concluímos que para cada ponto de máximo e de mínimo X na esfera unitária $S^{n-1} = g^{-1}(0)$, com $g(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2 - 1$, existe algum λ em $\mathbb{R}$ tal que

$$\vec{\nabla} Q(x_1, \ldots, x_n) = \lambda \vec{\nabla} g(x_1, \ldots, x_n),$$

e então, para tais pontos e pelo lema acima temos

$$2AX = \lambda 2X \ \Rightarrow \ AX = \lambda X \ ,$$

e assim, se X_M, onde $|X_M| = 1$, é tal que $Q(X_M) = M$ e λ_M é um número real tal que $AX_M = \lambda_M X_M$, temos

$$M = Q(X_M) = AX_M \cdot X_M = \lambda_M |X_M|^2 = \lambda_M \ ,$$

e analogamente temos as identidades

$$AX_m = \lambda_m X_m, \quad |X_m| = 1 \ \text{ e } \ m = Q(X_m) = AX_m \cdot X_m = \lambda_m.$$

[1] É provado em Álgebra Linear que toda matriz A simétrica e real de ordem n tem n auto-valores reais: as raízes, com suas multiplicidades, do polinômio característico $p(\lambda) = \det(A - \lambda I)$, I a matriz identidade de $M_n(\mathbb{R})$. Mas, não utilizaremos tal fato.

Ainda mais , se λ é um auto-valor real de A então é claro que existe um vetor $\overrightarrow{v}$, com $|\overrightarrow{v}| = 1$, tal que $A(\overrightarrow{v}) = \lambda\overrightarrow{v}$. Logo, $m \leq Q(\overrightarrow{v}) \leq M$ e, ainda, $Q(\overrightarrow{v}) = (A\overrightarrow{v}) \cdot \overrightarrow{v} = \lambda\overrightarrow{v} \cdot \overrightarrow{v} = \lambda|\overrightarrow{v}|^2 = \lambda$. Donde finalmente seguem as desigualdades

$$m \leq \lambda \leq M.$$

(b) Se $\overrightarrow{v} \neq \vec{0}$ então,

$$m \leq Q\Big(\frac{\overrightarrow{v}}{|\vec{v}|}\Big) \leq M \implies m \leq A\Big(\frac{\overrightarrow{v}}{|\vec{v}|}\Big) \cdot \frac{\overrightarrow{v}}{|\vec{v}|} \leq M \Longrightarrow m \leq \frac{Q(\overrightarrow{v})}{|\vec{v}|^2} \leq M \ .$$

(c) e (d) Seguem trivialmente de (b) ■

Apêndice - Teorema do Posto e Teorema do Núcleo e da Imagem.

Primeira Parte: Lema 5 implica os Teoremas do Posto e o do Núcleo e da Imagem

Teorema do Posto. *Seja M em $M_{m\times n}(\mathbb{R})$. Seja p (o posto de M) o número máximo de linhas linearmente independentes de M. Então,*

(a) O número máximo de colunas linearmente independentes de M é p.

(b) O máximo entre as ordens dos menores não nulos de M é p.

Prova. É claro que $p \leq m$. Como as linhas de M são vetores em $\mathbb{R}^n$ e em $\mathbb{R}^n$ há no máximo n vetores linearmente independentes, segue que $p \leq n$. Os casos M em $M_{m\times n}(\mathbb{R})$ e M^t em $M_{n\times m}(\mathbb{R})$ são análogos. Logo, basta mostrarmos que M tem p colunas linearmente independentes e um menor (determinante) de ordem p que não é nulo.

(a), (b) Consideremos $\overline{M}$ em $M_{p\times n}(\mathbb{R})$ dada por um conjunto de p linhas linearmente independentes de M, ordenadas (naturalmente) segundo a ordem herdada de M. Os menores de $\overline{M}$ são menores de M. Pelo Lema 5, $\overline{M}$ apresenta um menor de ordem p não nulo.

Portanto, a matriz M tem um menor de ordem p não nulo. Por uma propriedade de determinantes, as p colunas deste menor de ordem p são linearmente independentes. É fácil ver que as p colunas de M correspondentes às p colunas deste menor são linearmente independentes.

Pela propriedade de determinantes citada acima temos que M não tem um menor de ordem superior a p e não nulo$\blacksquare$

Abaixo, identificamos $\mathbb{R}^n$ com $M_{n\times 1}(\mathbb{R})$ e $\mathbb{R}^m$ com $M_{m\times 1}(\mathbb{R})$.

Teorema do Núcleo e da Imagem. *Seja $T : \mathbb{R}^n \to \mathbb{R}^m$ definida por $TX = MX$, onde M é uma matriz em $M_{m\times n}(\mathbb{R})$ e X está em $\mathbb{R}^n$. Então,*

$$dim\,Ker(T) + dim\,Im(T) = n.$$

Prova. Sabidamente, a imagem de T é gerada pelas colunas de M. É fácil ver que o subespaço $\ker(T)$ é ortogonal ao sub-espaço de $\mathbb{R}^n$ gerado pelas linhas de M. Se p é o posto de M, pelo Teorema do Posto temos

$$\begin{cases} \dim \operatorname{Im}(T) = p \\ \dim \operatorname{Ker}(T) = n - p \end{cases} \blacksquare$$

Segunda Parte: Teorema do Núcleo e da Imagem implica Teorema do Posto.

Mantida a notação da primeira parte consideramos as aplicações lineares

$$\begin{cases} T : \mathbb{R}^n \to \mathbb{R}^m, \text{ por } T(X) = MX, \text{ onde } M \in M_{m\times n}(\mathbb{R}), \\ \\ T^t : \mathbb{R}^m \to \mathbb{R}^n, \text{ por } T^tY = M^tY \text{ onde } M^t \in M_{n\times m}(\mathbb{R}) \text{ é a transposta de } M. \end{cases}$$

Notemos que $Y \in \operatorname{Ker}(T^t)$ se e somente se $M^tY = 0$. Isto é, $Y \in \operatorname{Ker}(T^t)$ se e somente se Y é ortogonal às linhas de M^t (i.e., colunas de M). Logo,

$$\ker(T^t) = [\operatorname{Im}(T)]^{\perp}.$$

Porém, pelo teorema do núcleo e da imagem temos

$$\dim \operatorname{Im}(T^t) = m - \dim \ker(T^t).$$

Concluímos então que

$$\dim \operatorname{Im}(T^t) = m - [m - \dim \operatorname{Im}(T)] = \dim \operatorname{Im}(T).$$

Logo, o número de colunas linearmente independentes de M é igual ao número de linhas linearmente independentes de M $\blacksquare$

REFERÊNCIAS

1. Apostol, T. M., *Análisis Matemático*, Editorial Reverté, 1960.
2. Guidorizzi, H. L., *Um Curso de Cálculo*, Vol 1 e 2, 5ª ed., Ed. LTC, 2002.
3. Lima, E., *Curso de Análise* , Vol 2., IMPA, 2009.
4. Mackiw, G. *A Note on the Equality of the Column and Row Rank of a Matrix*, Mathematics Magazine, Vol. 68, No. 4 (Oct., 1995), pp. 285-286.
5. Simmons, G. F., *Cálculo com Geometria Analítica*, Vol 2, McGraw-Hill, 1988.
6. Wardlaw, W. P., *Row Rank Equals Column Rank*, Mathematics Magazine, Vol. 78, No. 4 (Oct. 2005), pp. 316-318.