

O Hessiano
Polinômios de Taylor de Ordem 1 em Duas Variáveis

Recordamos o mínimo necessário neste capítulo sobre fórmulas de Taylor em uma variável.

Lema 1. Seja $f \in C^2([a, b])$ e $x_0, x \in [a, b]$, $x_0 \neq x$. Existe ξ, ξ entre x_0 e x , $\xi \neq x_0$ e $\xi \neq x$, tal que:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2}(x - x_0)^2.$$

Prova:

Existe, é óbvio, um único λ real e dependendo de x_0 e x tal que

$$(*) \quad f(x) = f(x_0) + f'(x_0)(x - x_0) + \lambda(x - x_0)^2.$$

Definamos então a função,

$$\varphi(t) = f(t) - f(x_0) - f'(x_0)(t - x_0) - \lambda(t - x_0)^2.$$

Temos, $\varphi(x_0) = 0$ e, por (*), $\varphi(x) = 0$. Logo, pelo TVM, $\exists c, c$ entre x_0 e x , $c \neq x_0$, $c \neq x$, tal que

$$0 = \varphi'(c) = f'(c) - f'(x_0) - 2\lambda(c - x_0) \implies 2\lambda = \frac{f'(c) - f'(x_0)}{c - x_0}.$$

Pelo TVM, aplicado a $f' \in C^1([a, b])$, existe ξ, ξ entre x_0 e c , $\xi \neq x_0$ e $\xi \neq c$, tal que,

$$\frac{f'(c) - f'(x_0)}{c - x_0} = f''(\xi) \implies \lambda = \frac{f''(\xi)}{2!} \quad \blacksquare$$

Proposição 1. Seja $f \in C^2(\Omega)$, $P_o = (x_o, y_o) \in \Omega$ e $\vec{v} = \langle h, k \rangle \neq \vec{0}$. Consideremos a restrição $\varphi = \varphi_{\vec{v}}$ (isto é, φ depende de \vec{v}), de f sobre um segmento por P_o , na direção \vec{v} e contido em Ω :

$$\varphi(t) = f(x_o + th, y_o + tk), \quad t \in (-r, r), \quad r > 0, \quad \text{e } B(P_o; r|\vec{v}|) \subset \Omega.$$

Temos,

$$(a) \quad \varphi'(t) = \frac{\partial f}{\partial x}(x_o + th, y_o + tk)h + \frac{\partial f}{\partial y}(x_o + th, y_o + tk)k.$$

$$(b) \quad \varphi''(t) = \frac{\partial^2 f}{\partial x^2}(x_o + th, y_o + tk)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_o + th, y_o + tk)hk + \frac{\partial^2 f}{\partial y^2}(x_o + th, y_o + tk)k^2.$$

$$(c) \quad \varphi'(0) = \frac{\partial f}{\partial x}(P_o)h + \frac{\partial f}{\partial y}(P_o)k = \vec{\nabla} f(P_o) \cdot \vec{v}, \quad \varphi'(0) = \frac{\partial f}{\partial \vec{v}}(P_o), \quad \text{se } |\vec{v}| = 1$$

$$(d) \quad \varphi''(0) = \frac{\partial^2 f}{\partial x^2}(P_o)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(P_o)hk + \frac{\partial^2 f}{\partial y^2}(P_o)k^2, \quad \varphi''(0) = \frac{\partial^2 f}{\partial \vec{v}^2}(P_o), \quad \text{se } |\vec{v}| = 1.$$

$$(e) \quad \varphi''(0) = |\vec{v}|^2 \frac{\partial^2 f}{\partial \vec{w}^2}(P_o), \quad \text{se } \vec{w} = \frac{\vec{v}}{|\vec{v}|} \text{ é o versor de } \vec{v}.$$

Prova:

(a) Pela regra da cadeia temos

$$\varphi' = \frac{d}{dt} \left\{ f(x_o + th, y_o + tk) \right\} = \frac{\partial f}{\partial x}(x_o + th, y_o + tk)h + \frac{\partial f}{\partial y}(x_o + th, y_o + tk)k = \frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k,$$

(b) Diferenciando a fórmula obtida em (a) e utilizando que pelo Teorema de Schwartz as derivadas mistas de $f \in C^2$ comutam, isto é, $f_{xy} = f_{yx}$, temos

$$\begin{aligned} \varphi'' &= \frac{d}{dt} \left\{ \frac{d}{dt} \left[f(x_o + th, y_o + tk) \right] \right\} = \frac{d}{dt} \left\{ \frac{\partial f}{\partial x}(x_o + th, y_o + tk) \right\} h + \frac{d}{dt} \left\{ \frac{\partial f}{\partial y}(x_o + th, y_o + tk) \right\} k \\ &= \left[\frac{\partial^2 f}{\partial x^2} h + \frac{\partial^2 f}{\partial y \partial x} k \right] h + \left[\frac{\partial^2 f}{\partial x \partial y} h + \frac{\partial^2 f}{\partial y^2} k \right] k = \frac{\partial^2 f}{\partial x^2} h^2 + 2 \frac{\partial^2 f}{\partial x \partial y} hk + \frac{\partial^2 f}{\partial y^2} k^2. \end{aligned}$$

(c) Segue trivialmente de (a) e da fórmula : $\frac{\partial f}{\partial \vec{v}} = \vec{\nabla} f \cdot \vec{v}$, se $|\vec{v}| = 1$.

(d) Segue de (b).

(e) Seja $\vec{w} = \frac{\vec{v}}{|\vec{v}|} = \left\langle \frac{h}{|\vec{v}|}, \frac{k}{|\vec{v}|} \right\rangle$. Temos, por (d), a relação entre $\varphi''_{\vec{w}}(0)$ e $\varphi''_{\vec{v}}(0)$:

$$\frac{\partial f}{\partial \vec{w}^2}(P_o) = \varphi''_{\vec{w}}(0) = \frac{\partial^2 f}{\partial x^2}(P_o) \frac{h^2}{|\vec{v}|^2} + 2 \frac{\partial^2 f}{\partial x \partial y}(P_o) \frac{h}{|\vec{v}|} \frac{k}{|\vec{v}|} + \frac{\partial^2 f}{\partial y^2}(P_o) \frac{k^2}{|\vec{v}|^2} = \frac{1}{|\vec{v}|^2} \varphi''(0),$$

Teorema 1 (Polinômio de Taylor de Ordem 1 Com Resto de Lagrange). Dada $f \in C^2(\Omega)$, $P = (x, y)$ e $P_o = (x_o, y_o)$ ambos em Ω tais que $\overline{PP_o} \subset \Omega$, existe $\bar{P} = (\bar{x}, \bar{y}) \in \overline{PP_o}$ tal que,

$$\begin{aligned} f(x, y) &= f(x_o, y_o) + \frac{\partial f}{\partial x}(x_o, y_o)(x - x_o) + \frac{\partial f}{\partial y}(x_o, y_o)(y - y_o) + \\ &+ \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(\bar{x}, \bar{y})(x - x_o)^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\bar{x}, \bar{y})(x - x_o)(y - y_o) + \frac{\partial^2 f}{\partial y^2}(\bar{x}, \bar{y})(y - y_o)^2 \right]. \end{aligned}$$

Prova Restringindo f ao segmento $\overline{PP_o}$, para $\varphi(t) = f(x_o + th, y_o + tk) = f((x_o, y_o) + t\vec{v})$, $t \in [0, 1]$, com $\vec{v} = \langle h, k \rangle = \langle x - x_o, y - y_o \rangle$, pelo Lema 1 existe $\bar{t} \in (0, 1)$ tal que

$$(*) \quad \varphi(1) = \varphi(0) + \varphi'(0) + \frac{\varphi''(\bar{t})}{2}.$$

Temos, $\varphi(1) = f(x, y)$, $\varphi(0) = f(x_o, y_o)$ e, pela Proposição 1, itens (c) e (d),

$$\begin{cases} \varphi'(0) &= \frac{\partial f}{\partial x}(x_o, y_o)h + \frac{\partial f}{\partial y}(x_o, y_o)k, \\ \varphi''(\bar{t}) &= \frac{\partial^2 f}{\partial x^2}(x_o + \bar{t}h, y_o + \bar{t}k)h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(x_o + \bar{t}h, y_o + \bar{t}k)hk + \frac{\partial^2 f}{\partial y^2}(x_o + \bar{t}h, y_o + \bar{t}k)k^2. \end{cases}$$

Substituindo no sistema acima $(\bar{x}, \bar{y}) = (x_o + \bar{t}h, y_o + \bar{t}k)$, $h = x - x_o$, $k = y - y_o$ e em (*) os valores encontrados para $\varphi(1)$, $\varphi(0)$, $\varphi'(0)$ e $\varphi''(\bar{t})$, segue a tese ■

Adendo ao Teorema 1. Com a notação $(x, y) = (x_o + h, y_o + k)$ temos,

$$f(x_o + h, y_o + k) = f(x_o, y_o) + \frac{\partial f}{\partial x}(x_o, y_o)h + \frac{\partial f}{\partial y}(x_o, y_o)k + \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(\bar{x}, \bar{y})h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\bar{x}, \bar{y})hk + \frac{\partial^2 f}{\partial y^2}(\bar{x}, \bar{y})k^2 \right].$$

Matriz Hessiana e Maximos e Mınimos em Duas Variaveis

Definicoes. Seja $f : \Omega \rightarrow \mathbb{R}$, Ω aberto em \mathbb{R}^2 . Classificamos $P_o \in \Omega$, relativamente a f , como

- (a) **ponto de maximo [mınimo] local** se existe uma bola aberta $B(P_o; r) \subset \Omega$, $r > 0$, tal que $f(P_o) \geq f(P)$ [$f(P_o) \leq f(P)$] para todo $P \in B(P_o; r)$; se tal desigualdade e estrita para $P \in B(P_o; r)$, $P \neq P_o$, P_o e ponto de maximo [mınimo] local **estrito**.
- (b) **ponto de maximo [mınimo] global, ou absoluto**, se $f(P_o) \geq f(P)$ [$f(P_o) \leq f(P)$], $\forall P \in \Omega$; se tal desigualdade e estrita para $P \in \Omega \setminus \{P_o\}$, P_o e, em adicao, **estrito**.
- (c) **extremante local [absoluto]** se e um ponto de maximo, ou de mınimo, local [absoluto].
- (d) **ponto crıtico, ou estacionario, de f** , supondo $f \in C^1(\Omega)$, se

$$\frac{\partial f}{\partial x}(P_o) = \frac{\partial f}{\partial y}(P_o) = 0.$$

- (e) **ponto de sela**, se P_o e ponto crıtico de f mas nao de maximo ou mınimo, locais.

Observacao 1: Um ponto de maximo, ou mınimo, local de $f \in C^1(\Omega)$ e sempre um ponto crıtico.

Corolario 1. Com mesmas hipoteses e notacao da Proposicao 1 temos:

- (i) Se P_o e ponto de maximo [mınimo] local de f na direcao \vec{v} , $|\vec{v}| = 1$; isto e, $t = 0$ e maximo [mınimo] local de $\varphi_{\vec{v}}$, entao,

$$\frac{\partial f}{\partial \vec{v}}(P_o) = \vec{\nabla} f(P_o) \cdot \vec{v} = 0 \quad ; \quad \frac{\partial^2 f}{\partial \vec{v}^2}(P_o) \leq 0 \quad \left[\frac{\partial^2 f}{\partial \vec{v}^2}(P_o) \geq 0 \right].$$

- (ii) Se P_o e ponto de maximo [mınimo] local de f entao $\vec{\nabla} f(P_o) = \vec{0}$ e,

$$\frac{\partial^2 f}{\partial \vec{v}^2}(P_o) = \frac{\partial^2 f}{\partial x^2}(P_o)h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(P_o)hk + \frac{\partial^2 f}{\partial y^2}(P_o)k^2 \leq 0 \quad (\geq 0), \quad \forall \vec{v} \in \mathbb{R}^2, |\vec{v}| = 1.$$

Prova:

- (i) Segue da analise das derivadas de ordens 1 e 2 da funcao de uma variavel $\varphi_{\vec{v}}$ e do Lema 1.
- (ii) Consequencia imediata da Proposicao 1, de (i) e da Observacao 1, acima ■

Corolario 2. Seja $f \in C^2(\Omega)$, $P_o = (x_o, y_o)$ um seu ponto crıtico e $B(P_o; r) \subset \Omega$, $r > 0$. Dado $\vec{v} = \langle h, k \rangle$, $|\vec{v}| < r$, existe (\bar{x}, \bar{y}) no segmento unindo os pontos P_o e $P_o + \vec{v}$ tal que

$$f(x_o + h, y_o + k) - f(x_o, y_o) = \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(\bar{x}, \bar{y})h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(\bar{x}, \bar{y})hk + \frac{\partial^2 f}{\partial y^2}(\bar{x}, \bar{y})k^2 \right].$$

Prova: Consequencia imediata do Teorema 1 e da definicao de ponto crıtico ■

Mantendo a notacao, analisemos $Q_P(h, k) = f_{xx}(P)h^2 + 2f_{xy}(P)hk + f_{yy}(P)k^2$, $(h, k) \in \mathbb{R}^2$.

Observação 2. Dados $a, b, c \in \mathbb{R}$, sejam $H = ac - b^2$ e a **forma quadrática** $Q: \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$z = Q(h, k) = ah^2 + 2bhk + ck^2 = \begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}, \quad \vec{v} = \langle h, k \rangle \in \mathbb{R}^2.$$

(i) Se $a \neq 0$, vale a fatoração $z = Q(h, k) = a\left[\left(h + \frac{b}{a}k\right)^2 + \frac{H}{a^2}k^2\right]$.

(ii) Se $a \neq 0$, o gráfico de $Q: \mathbb{R}^2 \rightarrow \mathbb{R}$ é um parabolóide do tipo:

- se $H > 0$, elíptico ou circular, eixo Oz e concavidade para cima [baixo] se $a > 0$ [$a < 0$].
- se $H < 0$, hiperbólico com sela em $(0, 0)$.
- se $H = 0$, cilíndrico.

(iii) Se $a = c = 0$ e $b \neq 0$ (logo, $H < 0$), o gráfico de Q é um parabolóide hiperbólico.

(iv) Ainda, $z = Q(h, k)$ troca de sinal se e só se $H < 0$.

Se o gráfico de Q é um parabolóide: elíptico ou circular, $0 = Q(0, 0)$ é valor mínimo/máximo estrito e absoluto; se cilíndrico $0 = Q(0, 0)$ é valor mínimo/máximo não estrito mas absoluto.

Prova:

(i) Pondo $a \in \mathbb{R}$ em evidência e completando quadrados obtemos,

$$Q(h, k) = a\left(h^2 + \frac{2bhk}{a} + \frac{ck^2}{a}\right) = a\left[\left(h + \frac{bk}{a}\right)^2 - \frac{b^2k^2}{a^2} + \frac{ack^2}{a^2}\right] = a\left[\left(h + \frac{bk}{a}\right)^2 + \frac{ac - b^2}{a^2}k^2\right].$$

(ii), (iii): Conseqüências triviais de (i). O ítem (iv) segue de (i), (ii) e (iii) ■

Definição. Se M é a matriz (simétrica) 2×2 em Obs. 2, Q é a **forma quadrática associada a M** .

Comentário. Identificando \mathbb{R}^2 com o espaço das matrizes-coluna $M_{2 \times 1}(\mathbb{R})$, indicamos a matriz

das coordenadas de \vec{v} em relação à base canônica ordenada de \mathbb{R}^2 , $\{\vec{e}_1, \vec{e}_2\}$, como $[\vec{v}] = \begin{bmatrix} h \\ k \end{bmatrix}$.

Então, se $[\vec{v}]^t = [h \ k] \in M_{1 \times 2}(\mathbb{R})$ é a transposta da matriz $[\vec{v}]$, mantendo as notações acima temos, $Q(h, k) = [\vec{v}]^t M [\vec{v}] = M [\vec{v}] \cdot [\vec{v}]$, onde “ \cdot ” denota o produto escalar em \mathbb{R}^2 .

Proposição 2. Dada $f \in C^2(\Omega)$ e $P \in \Omega$, P um ponto crítico de f , seja para $(h, k) \in \mathbb{R}^2$,

$$Q_P(h, k) = f_{xx}(P)h^2 + 2f_{xy}(P)hk + f_{yy}(P)k^2 = \begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} f_{xx}(P) & f_{xy}(P) \\ f_{yx}(P) & f_{yy}(P) \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix}.$$

Denotando $\vec{v} = \langle h, k \rangle \neq \vec{0}$ e definindo $Hf(P) = f_{xx}(P)f_{yy}(P) - [f_{xy}(P)]^2$ temos

- (i) $Q_P(h, k) = |\vec{v}|^2 \frac{\partial^2 f}{\partial \vec{v}^2}(P)$, $\vec{\omega} = \frac{\vec{v}}{|\vec{v}|}$ o versor de \vec{v} .
- (ii) $\frac{\partial^2 f}{\partial \vec{v}^2}(P) > 0$, $\forall \vec{v}$ unitário, se e somente se $Hf(P) > 0$ e $f_{xx}(P) > 0$.
- (iii) $\frac{\partial^2 f}{\partial \vec{v}^2}(P) < 0$, $\forall \vec{v}$ unitário, se e somente se $Hf(P) > 0$ e $f_{xx}(P) < 0$.
- (iv) $\frac{\partial^2 f}{\partial \vec{v}^2}(P)$ troca de sinal, em relação a \vec{v} unitário, se e somente se, $Hf(P) < 0$.

Prova:

(i) Segue da Proposição 1, ítem (e).

(ii), (iii) e (iv) Seguem imediatamente de (i) e da Observação 2 ■

Definições: Seja $f \in C^2(\Omega)$ e $P \in \Omega$.

- A **matriz hessiana** de f em P , denotada $\mathcal{H}f(P) \in M_{2 \times 2}(\mathbb{R}) = [a_{ij}]$, é dada por $a_{11} = f_{xx}(P)$, $a_{12} = a_{21} = f_{xy}(P)$ e $a_{22} = f_{yy}(P)$.
- O **hessiano** de f em P , $Hf(P)$, é o determinante da matriz hessiana $\mathcal{H}f(P)$.
- A **forma quadrática associada a f , em P** , indicada Q_P , é a associada à matriz $\mathcal{H}f(P)$.

Teorema 2 (Teste do Hessiano). Seja $f \in C^2(\Omega)$, $P_o = (x_o, y_o) \in \Omega$, P_o um ponto crítico de f e,

$$\mathcal{H}f(P_o) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(P_o) & \frac{\partial^2 f}{\partial y \partial x}(P_o) \\ \frac{\partial^2 f}{\partial x \partial y}(P_o) & \frac{\partial^2 f}{\partial y^2}(P_o) \end{bmatrix}, \quad Hf(P_o) = \det \mathcal{H}f(P_o).$$

- Se $Hf(P_o) > 0$ e $f_{xx}(P_o) > 0$, P_o é um ponto de mínimo local estrito.
- Se $Hf(P_o) > 0$ e $f_{xx}(P_o) < 0$, P_o é um ponto de máximo local estrito.
- Se $Hf(P_o) < 0$ então P_o é um ponto de sela.
- Se $Hf(P_o) = 0$, P_o pode ser de qualquer um dos tipos acima.

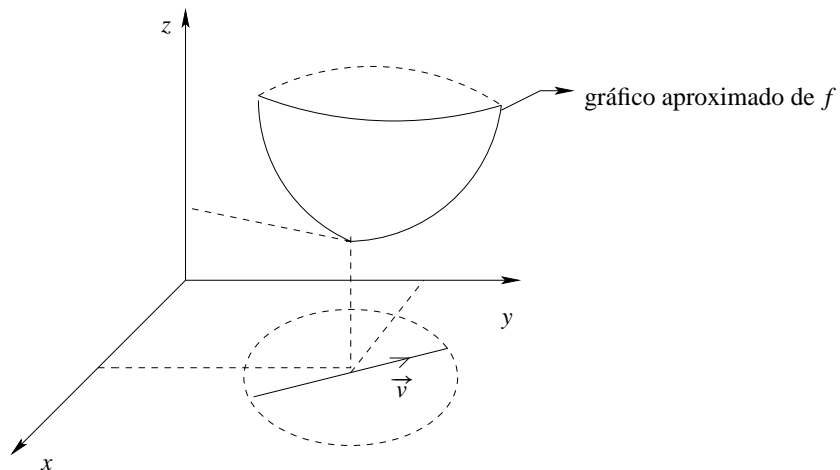


Figura 1: Caso em que $f_{xx}(P_o) > 0$ e $Hf(P_o) > 0$.

Prova.

- Como $f \in C^2$, temos $Hf = f_{xx}f_{yy} - f_{xy}^2 > 0$ e $f_{xx} > 0$ numa bola $B(P_o; r)$, se $r > 0$ e suficientemente pequeno. Pelo Corolário 2, dado $\vec{v} = \langle h, k \rangle \neq \vec{0}$, $|\vec{v}| < r$, existe $\bar{P} = (\bar{x}, \bar{y})$ no segmento unindo P_o e $P_o + \vec{v}$, e portanto $\bar{P} \in B(P_o; r)$, tal que

$$\begin{aligned} f(P_o + \vec{v}) - f(P_o) &= \frac{1}{2} [f_{xx}(\bar{P})h^2 + 2f_{xy}(\bar{P})hk + f_{yy}(\bar{P})k^2] = \\ &= \frac{1}{2} \left[f_{xx}(\bar{P}) \left(h + \frac{f_{xy}(\bar{P})}{f_{xx}(\bar{P})} k \right)^2 + \frac{Hf(\bar{P})}{f_{xx}(\bar{P})} k^2 \right] > 0. \end{aligned}$$

- Basta aplicar o ítem (a) à função $-f$.
- Por contradição. Se P_o é ponto de máximo [mínimo] então, pelo Corolário 1, $\frac{\partial^2 f}{\partial v^2}(P_o) \geq 0$, $\forall |\vec{v}| = 1$ $\left[\frac{\partial^2 f}{\partial v^2}(P_o) \leq 0, \forall |\vec{v}| = 1 \right]$, e pela Prop. 2(iv) $Hf(P_o) \geq 0$ \nexists
- Vide exemplos 1, 2 e 3, abaixo ■

Exemplo 1. A função $f(x, y) = x^4 + y^4$ é tal que $(0, 0)$ é ponto de mínimo absoluto estrito, e o valor mínimo é 0. É, ainda, o único ponto crítico e f e suas derivadas parciais se anulam nele.

Exemplo 2. A função $f(x, y) = x^3 + y^3$ é tal que f e suas derivadas parciais se anulam em $(0, 0)$, que é o único ponto crítico. Porém, é fácil ver, $(0, 0)$ não é extremante local e é um ponto de sela.

Exemplo 3. Seja $f(x, y) = ax^2 + by^2 + cxy + dx + ey + l$, com a, b, c, d e l em \mathbb{R} , e $a^2 + b^2 + c^2 \neq 0$. Se P_o é extremante local então P_o é extremante global (absoluto) (vide Exercício 2, p. 315, “Um Curso de Cálculo”, H. L. Guidorizzi, vol 4, 5ª ed.).

Prova:

Seja $\vec{v} = \langle h, k \rangle \in \mathbb{R}^2$, $\vec{v} \neq \vec{0}$, e $\vec{\omega} = \frac{\vec{v}}{|\vec{v}|}$ seu versor.

Sendo $P_o = (x_o, y_o)$ um extremante local ele é um ponto crítico e assim,

$$(1) \quad \vec{0} = \vec{\nabla} f(x_o, y_o) = \langle 2ax_o + cy_o + d, 2by_o + cx_o + e \rangle.$$

Computando f sobre a reta por P_o e direção \vec{v} , por (1) obtemos,

$$\begin{aligned} f(x_o + th, y_o + tk) &= a(x_o + ht)^2 + b(y_o + kt)^2 + c(x_o + ht)(y_o + kt) + d(x_o + ht) + e(y_o + kt) + l \\ &= (ah^2 + chk + bk^2)t^2 + t[(2ax_o + cy_o + d)h + (2by_o + cx_o + e)k] + \\ &\quad + (ax_o^2 + by_o^2 + cx_o y_o + dx_o + ey_o + l) = \\ &= (ah^2 + chk + bk^2)t^2 + f(x_o, y_o). \end{aligned}$$

Donde, encontramos

$$(2) \quad f(x_o + th, y_o + tk) - f(x_o, y_o) = (ah^2 + chk + bk^2)t^2.$$

1ª solução (simples):

Analisando (2), é claro que seu primeiro membro é positivo (negativo) para $|t| \leq r$, para algum $r > 0$, e qualquer \vec{v} tal que $|\vec{v}| = 1$ se, e só se, o é também para qualquer $t \in \mathbb{R}$ e qualquer \vec{v} tal que $|\vec{v}| = 1$. Logo, P_o é mínimo (máximo) local se e só se P_o é mínimo (máximo) global.

2ª solução (reconhecendo conceitos):

Derivando (2) em relação a t obtemos [$f(x_o, y_o)$ é constante], pela Proposição 1(e),

$$(3) \quad |\vec{v}|^2 \frac{\partial^2 f}{\partial \vec{\omega}^2}(P_o) = \frac{d}{dt} \left\{ \frac{d}{dt} \{ f(x_o + th, y_o + tk) \} \right\} (0) = 2(ah^2 + chk + bk^2),$$

e, substituindo (3) em (2),

$$(4) \quad f(x_o + th, y_o + tk) - f(x_o, y_o) = \frac{|\vec{v}|^2}{2} \frac{\partial^2 f}{\partial \vec{\omega}^2}(P_o) t^2.$$

Ainda, sendo P_o localmente um ponto de mínimo ou máximo temos, respectivamente,

$$(5) \quad \frac{\partial^2 f}{\partial \vec{\omega}^2}(P_o) \geq 0, \quad \forall \vec{\omega} = \frac{\vec{v}}{|\vec{v}|} \neq \vec{0}, \quad \text{ou} \quad \frac{\partial^2 f}{\partial \vec{\omega}^2}(P_o) \leq 0, \quad \forall \vec{\omega} = \frac{\vec{v}}{|\vec{v}|} \neq \vec{0}.$$

Pelas desigualdades em (5), a diferença $f(x_o + th, y_o + tk) - f(x_o, y_o)$ em (4) não muda de sinal segundo $t \in \mathbb{R}$ e $\vec{v} \in \mathbb{R}^2$. Logo, P_o é ponto de mínimo, ou de máximo, global.

3ª solução (sugerida): Aplique o teorema, o Teste do Hessiano ■

Exemplo 4. Uma função $f \in C^2(\mathbb{R}^2)$ com um só ponto crítico, mínimo local mas não global:

$$f(x, y) = x^2 + (1 - x)^3 y^2 .$$

Resolução:

Temos

$$\vec{\nabla} f(x, y) = \langle 2x - 3(1 - x)^2 y^2, 2(1 - x)^3 y \rangle .$$

Logo, $\vec{\nabla} f(x, y) = (0, 0)$ implica $(1 - x)^3 y = 0$ e assim, ou $y = 0$, e portanto $x = 0$, ou $x = 1$ e neste caso não é possível resolver $2x + 3(1 - x)^3 y^2 = 0$. Logo, $(0, 0)$ é o único ponto crítico.

Ainda, $f_{xx} = 2 + 6(1 - x)^2 y^2$, $f_{xy} = -6(1 - x)^2 y$, $f_{yy} = 2(1 - x)^3$ e então

$$\mathcal{H}f(0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} ,$$

e $(0, 0)$ é mínimo local. Ainda, $f(2, 3) = -5 < f(0, 0) = 0$ e $(0, 0)$ não é mínimo global ■

Exemplo 5. Analisemos os pontos críticos de $f(x, y) = x^5 + y^4 - 5x - 32y - 3$. Impondo

$$\vec{\nabla} f = \langle 5x^4 - 5, 4y^3 - 32 \rangle = \langle 0, 0 \rangle ,$$

obtemos os pontos críticos $(1, 2)$ e $(-1, 2)$. Ainda,

$$\mathcal{H}f(x, y) = \begin{bmatrix} 20x^3 & 0 \\ 0 & 12y^2 \end{bmatrix} , \quad Hf(x, y) = 240x^3 y^2 .$$

Logo, $(1, 2)$ é ponto de mínimo local e $(-1, 2)$ é ponto de sela ■

Exemplo 6. Determine a distância entre as retas

$$r : x = 1 + \lambda, y = 1 + 6\lambda, z = 2\lambda, \quad \lambda \in \mathbb{R}$$

$$s : x = 1 + 2\mu, y = 5 + 15\mu, z = -2 + 6\mu, \quad \mu \in \mathbb{R} .$$

Resolução (existem também soluções via geometria vetorial ou multiplicadores de Lagrange):

Consideremos os pontos arbitrários P e Q sobre r e s , respectivamente:

$$\begin{cases} P = (1 + \lambda, 1 + 6\lambda, 2\lambda), & \lambda \in \mathbb{R} \\ Q = (1 + 2\mu, 5 + 15\mu, -2 + 6\mu), & \mu \in \mathbb{R} . \end{cases}$$

O quadrado da distância entre P e Q , $|\overrightarrow{QP}|^2$, é dado pela expressão,

$$D(\lambda, \mu) = (\lambda - 2\mu)^2 + (-4 + 6\lambda - 15\mu)^2 + (2\lambda + 2 - 6\mu)^2 = 41\lambda^2 + 265\mu^2 - 208\lambda\mu - 40\lambda + 96\mu + 20 .$$

Com $\vec{\nabla} D(\lambda, \mu) = \langle 82\lambda - 208\mu - 40, 530\mu - 208\lambda + 96 \rangle$, e os pontos críticos de D dados por:

$$\begin{cases} 41\lambda - 104\mu - 20 = 0 \\ -104\lambda + 265\mu + 48 = 0 \end{cases} \implies (\lambda, \mu) = \left(\frac{44}{7}, \frac{16}{7} \right).$$

A matriz hessiana de D no ponto $P_0 = \left(\frac{44}{7}, \frac{16}{7} \right)$ é,

$$\mathcal{H}(D)(P_0) = \begin{bmatrix} 82 & -208 \\ -208 & 530 \end{bmatrix},$$

com $\frac{\partial^2 D}{\partial \lambda^2}(P_0) = 82 > 0$ e hessiano $H(D)(P_0) = 82 \times 530 - (208)^2 = 196 > 0$.

Logo, P_0 é ponto de mínimo local de D e, como é o único ponto crítico de D , que mede a distância entre dois pontos arbitrários das retas r e s , segue que estas não são paralelas e portanto ou são concorrentes ou são reversas. Ainda mais, geometricamente segue que P_0 é de mínimo global.

Substituindo os valores encontrados para λ e μ obtemos,

$$P = (1 + \lambda, 1 + 6\lambda, 2\lambda) = \left(\frac{51}{7}, \frac{271}{7}, \frac{88}{7} \right) \text{ e}$$

$$Q = (1 + 2\mu, 5 + 15\mu, -2 + 6\mu) = \left(\frac{39}{7}, \frac{275}{7}, \frac{82}{7} \right).$$

Então, a distância ente r e s é:

$$|\vec{PQ}| = \sqrt{\frac{144}{49} + \frac{16}{49} + \frac{36}{49}} = \sqrt{\frac{196}{49}} = \frac{14}{7} = 2.$$

2ª Resolução (via geometria vetorial):

As retas r e s não são paralelas e portanto ou são concorrentes ou são reversas.

Um ponto $P = (x, y, z) \in \mathbb{R}^3$ pertence ao plano π que contém a reta r e é paralelo à reta s se e somente se [note que $(1, 1, 0) \in r$]:

$$0 = \begin{vmatrix} x-1 & y-1 & z-0 \\ 1 & 6 & 2 \\ 2 & 15 & 6 \end{vmatrix} = 6(x-1) - 2(y-1) + 3z = 6x - 2y + 3z - 4.$$

A distância procurada é então a distância de qualquer ponto de s ao plano π . Escolhendo $(1, 5, -2) \in s$ obtemos (utilizando a fórmula para a distância):

$$d = \frac{|6 \cdot 1 - 2 \cdot 5 + 3 \cdot (-2) - 4|}{\sqrt{6^2 + 2^2 + 3^2}} = \frac{14}{7} = 2 \quad \blacksquare$$

O exemplo a seguir é de G. Peano (1884), corrigindo um erro no livro de J. A. Serret.

Exemplo 7. Seja $z = f(x, y) = (y - x^2)(y - 2x^2)$, $(x, y) \in \mathbb{R}^2$. Verifique:

- (a) $(0, 0)$ é o único ponto crítico.
- (b) O teste da derivada segunda é inconclusivo para classificar tal ponto crítico.
- (c) f restrita a qualquer reta $y = mx$ pela origem tem neste ponto um mínimo local.
- (d) f não conserva sinal em nenhuma vizinhança de $(0, 0)$, o qual é então ponto de sela.

Resolução:

- (a) É fácil ver que $\vec{\nabla} f = \langle 2x(4x^2 - 3y), -3x^2 + 2y \rangle = \langle 0, 0 \rangle$ se e só se $(x, y) = (0, 0)$.
- (b) Temos $f_{xx} = 24x^2 - 6y$, $f_{xy} = f_{yx} = -6x$ e $f_{yy} = 2$. Logo, o hessiano em $(0, 0)$ é $Hf(0, 0) = 0$.
- (c) Seja $\varphi(x) = f(x, mx) = (mx - x^2)(mx - 2x^2) = 2x^4 - 3mx^3 + m^2x^2$, m uma constante real. Temos, $\varphi'(x) = 8x^3 - 9mx^2 + 2m^2x$, $\varphi''(x) = 24x^2 - 18mx + 2m^2$, $\varphi'(0) = 0$ e $\varphi''(0) = m^2 > 0$, se $m \neq 0$. Logo, $x = 0$ é ponto de mínimo local de φ , se $m \neq 0$. Se $m = 0$ temos $f(x, 0) = 2x^4$ e é claro que $x = 0$ é ponto de mínimo de φ .
- (d) Nas regiões, do plano, $\{y > 2x^2\}$, $\{x^2 < y < 2x^2\}$ e $\{y < x^2\}$ temos $f > 0$, $f < 0$ e $f > 0$, respectivamente. Vide Figura abaixo. Note também, vide Figura abaixo, que para $x \approx 0$, com $x \neq 0$, os pontos (x, mx) da reta pertencem à região em que $f > 0$.

Exemplo 8. Seja $f : \Omega \rightarrow \mathbb{R}$, Ω um aberto de \mathbb{R}^2 , $Hf = f_{xx}f_{yy} - f_{xy}^2$, $P_0 = (x_0, y_0)$ um ponto crítico de f tal que $f_{xx}(P_0) \neq 0$ e $\mathcal{H}f = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$. Supondo no ítem (a) as matrizes em P_0 verifique:

- (a) Se $M = M(x, y) = \begin{bmatrix} 1 & 0 \\ -\frac{f_{xy}}{f_{xx}} & 1 \end{bmatrix}$ e $N = M^{-1} = \begin{bmatrix} 1 & 0 \\ \frac{f_{xy}}{f_{xx}} & 1 \end{bmatrix}$ então $\det M = \det N = 1$ e

$$M \mathcal{H}f M^t = \begin{bmatrix} f_{xx} & 0 \\ 0 & \frac{Hf}{f_{xx}} \end{bmatrix} \quad \text{e} \quad \mathcal{H}f = N \begin{bmatrix} f_{xx} & 0 \\ 0 & \frac{Hf}{f_{xx}} \end{bmatrix} N^t.$$

- (b) Aplicando (a) e o Corolário 2 dê uma outra demonstração para
 - Se $f_{xx}(P_0) > 0$ e $\frac{Hf}{f_{xx}}(P_0) > 0$ então P_0 é ponto de mínimo local estrito.
 - Se $f_{xx}(P_0) < 0$ e $\frac{Hf}{f_{xx}}(P_0) < 0$ então P_0 é ponto de máximo local estrito.

Matriz Hessiana e Maximos e Minimos em Tres Variaveis

Esta secao e facilmente generalizavel para o caso de matrizes simetricas de ordem $n \in \mathbb{N}^*$.

Proposicao 3. Seja $A \in M_3(\mathbb{R})$ uma matriz simetrica real de ordem 3,

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix},$$

tal que os **menores principais** de ordens 1, 2 e 3,

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} = \det[a_{ij}]_{1 \leq i, j \leq 2} \quad \text{e} \quad \Delta_3 = \det[a_{ij}]_{1 \leq i, j \leq 3} = \det A,$$

respectivamente, sao nao nulos. Entao, existe $P \in M_3(\mathbb{R})$ satisfazendo $\det P = 1$ e

$$P^t A P = \begin{bmatrix} \Delta_1 & 0 & 0 \\ 0 & \frac{\Delta_2}{\Delta_1} & 0 \\ 0 & 0 & \frac{\Delta_3}{\Delta_2} \end{bmatrix} = D.$$

Prova:

Construiremos D a partir de A pelo metodo da eliminacao de Gauss e de operacoes elementares realizadas pela multipliccao por matrizes (elementares) com determinante 1. Lembremos que multiplicar uma linha ou coluna por um numero e entao adiciona-la a uma outra linha ou coluna, respectivamente, e uma operacao elementar que nao altera o determinante de uma matriz.

Na **etapa 1** inicialmente multiplicamos a 1a linha de A por $-\frac{a_{12}}{a_{11}}$ e somamos a 2a linha e, ainda, multiplicamos a 1a linha por $-\frac{a_{13}}{a_{11}}$ e somamos a 3a linha. Efetuamos tais operacoes, que comutam, multiplicando a matriz A a esquerda, respectivamente, pelas matrizes

$$P_1 = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{a_{12}}{a_{11}} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{e} \quad P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{a_{13}}{a_{11}} & 0 & 1 \end{bmatrix}.$$

A seguir, na matriz obtida efetuamos operacoes nas colunas correspondendo as feitas nas linhas: multiplicamos a 1a coluna por $-\frac{a_{12}}{a_{11}}$ e somamos na 2a coluna e multiplicamos a 1a coluna por $-\frac{a_{13}}{a_{11}}$ e somamos na 3a coluna. Efetuamos tais operacoes multiplicando a matriz a direita por P_1^t e P_2^t . Resumindo as quatro operacoes obtemos,

$$P_2 P_1 \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} P_1^t P_2^t = D_1 = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} \end{bmatrix},$$

onde $a_{22}^{(1)}$, $a_{23}^{(1)}$ e $a_{33}^{(1)}$ sao os coeficientes que surgem ao final da etapa 1.

Devido as operacoes realizadas os respectivos menores principais das matrizes simetricas A e D_1 sao iguais. Consequentemente temos,

$$a_{11} a_{22}^{(1)} = \Delta_2 \implies a_{22}^{(1)} = \frac{\Delta_2}{\Delta_1}.$$

Na **etapa 2** multiplicamos a segunda linha de D_1 por $-\frac{a_{23}^{(1)}}{a_{22}^{(1)}}$ e somamos à terceira linha, representando tal operação pela matriz P_3 , e completamos efetuando a operação correspondente nas colunas de D_1 , representada (a operação) pela matriz P_3^t . Obtemos então,

$$P_3 D_1 P_3^t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{a_{23}^{(1)}}{a_{22}^{(1)}} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -\frac{a_{23}^{(1)}}{a_{22}^{(1)}} \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(2)} & 0 \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix} = D_2 .$$

Analogamente à etapa 1, devido às operações efetuadas os menores principais de D_2 : a_{11} , $a_{11}a_{22}^{(2)}$ e $a_{11}a_{22}^{(2)}a_{33}^{(2)}$ são iguais aos respectivos menores principais de A . Assim sendo, obtemos

$$a_{11} = \Delta_1, \quad a_{11}a_{22}^{(2)} = \Delta_2, \quad a_{11}a_{22}^{(2)}a_{33}^{(2)} = \Delta_3,$$

$$a_{22}^{(2)} = \frac{\Delta_2}{\Delta_1} \quad \text{e} \quad a_{33}^{(2)} = \frac{\Delta_3}{\Delta_2} .$$

Por fim, notemos que $P^t A P = D_2$ com $P^t = P_3 P_2 P_1$ e $\det P = 1$ ■

Lembrete: $A, B \in M_n(\mathbb{R})$ são **congruentes** se existe P inversível em $M_n(\mathbb{R})$ tal que $P^t A P = B$.

Notação: No que segue, identificamos vetores em \mathbb{R}^3 com matrizes-colunas em $M_{3 \times 1}(\mathbb{R})$:

$$X = (x_1 \ x_2 \ x_3) \equiv \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} .$$

Corolário 3. Com as hipóteses e notação da Proposição 3, seja $Q(X) = X^t A X$, $X \in \mathbb{R}^3$.

- (a) Se $\Delta_1 > 0$, $\Delta_2 > 0$ e $\Delta_3 > 0$ então Q é definida positiva.
- (b) Se, alternando sinais, $\Delta_1 < 0$, $\Delta_2 > 0$ e $\Delta_3 < 0$ então Q é definida negativa.
- (c) Não ocorrendo (a) ou (b) [mas, $\Delta_{i's} \neq 0$], existem X_1 e X_2 tais que $Q(X_1) > 0$ e $Q(X_2) < 0$.

Prova:

Pela Proposição 3 existe $P \in M_3(\mathbb{R})$ inversível, $P^t A P = D$, $D = [d_{ij}]$ uma matriz diagonal e $d_1 = d_{11} = \Delta_1$ e $d_i = d_{ii} = \frac{\Delta_i}{\Delta_{i-1}}$, se $2 \leq i \leq n$. Ainda mais, se $X \in \mathbb{R}^3$ existe $Y \in \mathbb{R}^3$ tal que $X = P Y$ e,

$$Q(X) = X^t A X = (P Y)^t A P Y = Y^t P^t A P Y = Y^t D Y .$$

Logo, é fácil ver,

- (a) Q é positiva definida $\iff d_i > 0 \ (\forall 1 \leq i \leq n) \iff \Delta_i > 0 \ (\forall 1 \leq i \leq n)$.
- (b) Q é definida negativa $\iff d_i < 0 \ \forall 1 \leq i \leq n) \iff \Delta_1 < 0, \Delta_2 > 0$ e $\Delta_3 < 0$.
- (c) Existem elementos na diagonal de D com sinais contrários. Logo, existe $Y_1 \in \mathbb{R}^3$ tal que $Y_1^t D Y_1 > 0$ e existe $Y_2 \in \mathbb{R}^3$ tal que $Y_2^t D Y_2 < 0$. Assim, basta tomarmos $X_i = P Y_i$, $i = 1, 2$ ■

Definição O sinal de $x \in \mathbb{R}$ é $+1$ se $x > 0$, -1 se $x < 0$ e 0 se $x = 0$. Se $x, y \in \mathbb{R}$ são tais que $x > 0$ e $y < 0$, x e y tem **sinais opostos**. Se $x \geq 0$, x é **positivo** e, se $x \leq 0$, x é **negativo**. Se $x > 0$, x é **estritamente positivo** e, se $x < 0$ então x é **estritamente negativo**.

Teorema 3. Dada $f \in C^2(\Omega)$ e P_0 um seu ponto crítico, consideremos os números:

$$f_{xx}(P_0) \quad , \quad H_1 f(P_0) = H_1 = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix} (P_0) \quad \text{e} \quad Hf(P_0) = H = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{vmatrix} (P_0).$$

(a) Supondo-os não nulos temos,

- (i) Se os três são (estritamente) positivos, P_0 é ponto de mínimo local estrito.
- (ii) Se (alternando sinais) $f_{xx} < 0$, $H_1 > 0$ e $H < 0$, P_0 é ponto de máximo local estrito.
- (iii) Se (i) e (ii) não ocorrem, P_0 é ponto de sela.

(b) Se existirem números na diagonal principal de Hf com sinais opostos, P_0 é ponto de sela.

Esta afirmação é válida independentemente de f_{xx} ou H_1 ou H se anularem ou não em P_0 .

Prova:

Analogamente à Proposição 2 definimos para $\vec{v} = \langle v_1, v_2, v_3 \rangle \equiv [v_1 \ v_2 \ v_3] \in \mathbb{R}^3$ e $P \in \Omega$,

$$Q_P(\vec{v}) = [v_1 \ v_2 \ v_3] \begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

com as entradas da matriz 3×3 , a **matriz hessiana** $\mathcal{H}f$ de f , avaliadas em P .

- (a) (i) Por continuidade, existe uma bola aberta $B(P_0; r)$, $r > 0$, em que f_{xx} , $H_1 f$ e Hf são estritamente positivas. Então, pelo Corolário 3 temos $Q_P(\vec{v}) > 0$, $\forall \vec{v} \neq \vec{0}$, e por um resultado trivialmente análogo ao já provado em duas variáveis, na Proposição 2(i),

$$\frac{\partial^2 f}{\partial \vec{v}^2}(P) = Q_P(\vec{v}) > 0, \quad \forall P \in B(P_0; r) \quad \text{e} \quad \forall \vec{v}, \quad \text{com} \quad |\vec{v}| = 1.$$

Então, fixado um versor \vec{v} é elementar que a restrição de f sobre o segmento $P_0 + t\vec{v}$, $|t| < r$, tem um ponto de mínimo estrito em P_0 . Logo, $f(P_0 + t\vec{v}) > f(P_0)$ se $0 < |t| < r$. Assim, como \vec{v} é versor arbitrário, obtemos

$$f(P) > f(P_0), \quad \forall P \in B(P_0; r), \quad P \neq P_0.$$

- (ii) Segue do item (a)(i) aplicado à função $-f$.
- (iii) Pelo Corolário 3(c) existem \vec{v}_1 e \vec{v}_2 tais que $Q_{P_0}(\vec{v}_1) > 0$ e $Q_{P_0}(\vec{v}_2) < 0$ e escolhendo seus versores se preciso, supomos $|\vec{v}_1| = 1$ e $|\vec{v}_2| = 1$. Então, $\frac{\partial^2 f}{\partial \vec{v}_1^2}(P_0) = Q_{P_0}(\vec{v}_1) > 0$, $\frac{\partial^2 f}{\partial \vec{v}_2^2}(P_0) = Q_{P_0}(\vec{v}_2) < 0$, f restrita ao segmento $P_0 + t\vec{v}_1$, $|t|$ pequeno, tem em P_0 um ponto de mínimo estrito e, f restrita ao segmento $P_0 + t\vec{v}_2$, $|t|$ pequeno, tem em P_0 um ponto de máximo estrito. Logo, P_0 não é extremante local e é um ponto de sela.

(b) Óbvio ■

Comentário Tal teorema tem análogos nas variáveis y e z , se ao invés de $f_{xx}(P_0) \neq 0$, tivermos $f_{yy}(P_0)$ ou $f_{zz}(P_0)$ não zero. Se $f_{xx} = f_{yy} = f_{zz} = 0$, em P_0 , e $\mathcal{H}f(P_0) \neq 0$ é claro que $Q(r, t, s) = 2\delta rs + 2\epsilon rt + 2\psi st$ muda de sinal e P_0 é ponto de sela.

Exemplo 9. Estude com relação a máximos e mínimos locais a função

$$f(x, y, z) = x^3 + y^3 + z^3 - 3x - 3y - 3z + 2.$$

Resolução:

Temos,

$$\vec{\nabla} f = \langle 3(x^2 - 1), 3(y^2 - 1), 3(z^2 - 1) \rangle,$$

e 8 pontos críticos:

$$P_o = (x_o, y_o, z_o) = (\pm 1, \pm 1, \pm 1).$$

Como as derivadas parciais em uma variável são funções independentes das demais variáveis, as derivadas mistas f_{xy} , f_{xz} e f_{yz} são nulas e a matriz hessiana de f é uma matriz diagonal:

$$Hf(P_o) = \begin{vmatrix} 6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z \end{vmatrix}, \quad H_1(P_o) = \begin{vmatrix} 6x & 0 \\ 0 & 6y \end{vmatrix}.$$

Pelo Teste do Hessiano para uma função de três variáveis, os pontos a seguir, em que a diagonal troca de sinal são de sela :

$$P_1 = (1, -1, -1), P_2 = (1, 1, -1), P_3 = (1, -1, 1), P_4 = (-1, 1, 1), P_5 = (-1, 1, -1), P_6 = (-1, -1, 1).$$

Em $P_7 = (1, 1, 1)$ temos $H_1 f > 0$, $Hf > 0$ e $f_{xx} > 0$; logo, $(1, 1, 1)$ é ponto de mínimo local.

Em $P_8 = (-1, -1, -1)$, como $H_1 f > 0$, $Hf < 0$ e $f_{xx} < 0$, $(-1, -1, -1)$ é ponto de máximo local ■

Exemplo 10. Estude com relação a máximos e mínimos locais, e pontos de sela, a função

$$f(x, y, z) = \frac{x^5}{5} + y^4 + z^4 - \frac{x^3}{3} - 2y^2.$$

Resolução:

Os pontos críticos satisfazem $\vec{\nabla} f(x, y, z) = \langle x^4 - x^2, 4y^3 - 4y, 4z^3 \rangle = \langle 0, 0, 0 \rangle$. Logo,

$$x^2(x^2 - 1) = 0, 4y(y^2 - 1) = 0, z = 0.$$

Isto é,

$$P_1 = (0, 0, 0), P_2 = (0, -1, 0), P_3 = (0, 1, 0),$$

$$P_4 = (1, 0, 0), P_5 = (-1, 0, 0), P_6 = (1, 1, 0),$$

$$P_7 = (-1, 0, 0), P_8 = (-1, -1, 0), P_9 = (-1, 1, 0).$$

Temos, $f_{xx} = 4x^3 - 2x = 2x(2x^2 - 1)$, $f_{yy} = 12y^2 - 4 = 4(3y^2 - 1)$, $f_{zz} = 12z^2$ e derivadas mistas nulas. Analisemos as matrizes em P_i (note que $z = 0$), $1 \leq i \leq 9$,

$$\mathcal{H}(f)(P_i) = \begin{bmatrix} 2x(2x^2 - 1) & 0 & 0 \\ 0 & 4(3y^2 - 1) & 0 \\ 0 & 0 & f_{zz} = 0 \end{bmatrix},$$

$$\mathcal{H}_1(f)(P_i) = \begin{bmatrix} 2x(2x^2 - 1) & 0 \\ 0 & 4(3y^2 - 1) \end{bmatrix}.$$

Como o (determinante) hessiano é zero, analisemos os sinais na diagonal da matriz hessiana. Pelo Teorema 2(d), os pontos críticos em que a diagonal de $\mathcal{H}f$ troca de sinal são de sela. Assim, no ponto P_4 temos $f_{xx} = 2$ e $f_{yy} = -4$; em P_8 e P_9 temos, $f_{xx} = -2$ e $f_{yy} = 8$.

Os pontos P_i , $i = 1, 2, 3$, tem a forma $P_i = (0, y_i, 0)$ com $y_i = 0, -1$ ou 1 , respectivamente, e são de sela: $x = 0$ não é máximo ou mínimo local da restrição $\varphi_i(x) = f(x, y_i, 0) = x^3(\frac{x^2}{5} - \frac{1}{3}) + (y_i^4 - 2y_i^2)$ pois $(\frac{x^2}{5} - \frac{1}{3}) < 0$ se $x \approx 0$ e x^3 é positivo à direita de zero e negativo à esquerda. Isto é, a diferença

$$f(x, y_i, 0) - f(0, y_i, 0) = x^3(\frac{x^2}{5} - \frac{1}{3})$$

é positiva ou negativa conforme x se aproxima de 0 pela direita ou pela esquerda.

O ponto $P_7 = (-1, 0, 0)$ é de sela pois $\varphi(z) = f(-1, 0, z) = z^4 + \frac{2}{15}$ têm mínimo local estrito em $z = 0$ e $\psi(y) = f(-1, y, 0) = \frac{2}{15} + y^4 - 2y^2$, $\psi'' = 12y^2 - 4$ têm máximo local estrito em $y = 0$.

Os pontos $P_5 = (1, -1, 0)$ e $P_6 = (1, 1, 0)$ são de mínimo local pois as funções de uma variável, $\frac{x^5}{5} - \frac{x^3}{3}$, $y^4 - 2y^2$ e z^4 têm mínimo local em $x = 1$, $y = \pm 1$ e $z = 0$, respectivamente, e considerando-as funções de três variáveis, x, y, z , elas têm mínimo local em $(1, \pm 1, 0)$ e, a soma das três, que é f , têm mínimo local em $(1, \pm 1, 0)$.

Resposta: Pontos de sela: $P_1, P_2, P_3, P_4, P_7, P_8$ e P_9 . Pontos de mínimo local: P_5 e P_6 ■

Matriz Hessiana e Máximos e Mínimos em Várias Variáveis

Supondo $f \in C^2(\Omega)$, Ω aberto em \mathbb{R}^n , e considerando a matriz (simétrica) hessiana de f em P_0

$$\mathcal{H}f = \mathcal{H}f(P_0) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j} (P_0) \right]_{1 \leq i, j \leq n} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix},$$

seja Δ_k , $1 \leq k \leq n$, o **menor principal** de ordem k dado pelo determinante da matriz $k \times k$ formada pelas primeiras k linhas e k colunas de $\mathcal{H}f$.

Teorema 4. Dada $f \in C^2(\Omega)$ e P_0 um ponto crítico, consideremos os números $\Delta_1, \Delta_2, \dots, \Delta_n$.

- (a) Supondo-os não nulos temos,
- (i) Se todos são (estritamente) positivos, P_0 é ponto de mínimo local estrito.
 - (ii) Se (alternando sinais) $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0$, etc, P_0 é ponto de máximo local estrito.
 - (iii) Se (i) e (ii) não ocorrem P_0 , é ponto de sela.
- (b) Se existirem números na diagonal principal de $\mathcal{H}f$ com sinais opostos, P_0 é ponto de sela.

Prova: Com a argumentação no Teorema 3, é claro que existe $P^t \in M_n(\mathbb{R})$, com P^t um produto finito de matrizes que realizam a operação de multiplicar uma linha por um número e então soma-la em uma outra linha, tal que $P^t \mathcal{H}f(P_0) P = D$, $D = [d_{ij}]_{1 \leq i, j \leq n}$ uma matriz diagonal. Como os menores principais não mudam com cada operação temos $d_{11} = f_{xx} = \Delta_1$, $d_{11}d_{22} = \Delta_2$ e $d_{22} = \frac{\Delta_2}{\Delta_1}, \dots, d_{jj} = \frac{\Delta_j}{\Delta_{j-1}}$. O restante da prova é análogo ao Teorema 3. ■

No que segue aplicamos a fórmula de Taylor de ordem 2 e Multiplicadores de Lagrange para expressar o Teste da Derivada Segunda segundo o conceito de auto-valor de uma matriz.

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

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

Definição: A aplicação linear $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ determinada pela matriz $A \in M_{n \times n}(\mathbb{R})$ é dada por

$$T(X) = AX.$$

Definição: Se $A \in M_n(\mathbb{R})$ então $\lambda \in \mathbb{R}$ é um **auto-valor de A** se existe $\vec{v} \in \mathbb{R}^n$ tal que

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

Definição: Se $A \in M_{n \times n}(\mathbb{R})$ é simétrica a **forma quadrática associada a A** é a aplicação

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

Lema 4. Se Q é a forma quadrática associada à matriz simétrica $A \in M_{n \times n}(\mathbb{R})$ então

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

Prova:

Se $A = [a_{ij}]$ então $AX = \left(\sum_{j=1}^n a_{1j}x_j, \dots, \sum_{j=1}^n a_{nj}x_j \right)$ e então, como $a_{ij} = a_{ji}$,

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

Logo,

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

Corolário 4. Sejam M , o máximo, e m , o mínimo, de Q sobre a esfera unitária $\{X \in \mathbb{R}^n : |X| = 1\}$.

- (a) Os números M e m são, respectivamente, o maior e o menor auto-valores (reais) de A .
- (b) $m|X|^2 \leq Q(X) \leq M|X|^2$, $\forall X \in \mathbb{R}^n$.
- (c) Q é definida positiva se e só se todos os auto-valores de A são estritamente positivos.
- (d) Q é definida negativa se e só se todos os auto-valores de A são estritamente negativos.

Demonstração¹:

- (a) Por ser contínua Q assume máximo e mínimo sobre o compacto $S^{n-1} = \{X \in \mathbb{R}^n : |X| = 1\}$. Pelo Teorema dos Multiplicadores de Lagrange, para cada ponto de máximo e de mínimo X na esfera unitária $S^{n-1} = g^{-1}(0)$, com $g(x_1, \dots, x_n) = x_1^2 + \dots + x_n^2 - 1$, existe $\lambda \in \mathbb{R}$ tal que

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

e então, para tais pontos e pelo Lema 4 temos

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

e assim, se $X_M, |X_M| = 1$, é tal que $Q(X_M) = M$ e $\lambda_M \in \mathbb{R}$ é 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 $AX_m = \lambda_m X_m, |X_m| = 1$ e $m = Q(X_m) = AX_m \cdot X_m = \lambda_m$.

Ainda, se $\lambda \in \mathbb{R}$ é auto-valor de A , é claro que existe $\vec{v}, |\vec{v}| = 1$, tal que $A(\vec{v}) = \lambda \vec{v}$. Logo, $m \leq Q(\vec{v}) \leq M$ e, ainda, $Q(\vec{v}) = (A\vec{v}) \cdot \vec{v} = \lambda \vec{v} \cdot \vec{v} = \lambda |\vec{v}|^2 = \lambda$; donde, $m \leq \lambda \leq M$.

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

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

- (c) e (d) Seguem trivialmente de (b) \blacksquare

¹É 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.

Lema 5. Seja $f \in C^2(B(a; r))$, com $B(a; r) \subset \mathbb{R}^n$, $r > 0$, a um ponto estacionário (isto é, ponto crítico) de f e $\vec{v} = \langle v_1, \dots, v_n \rangle \in \mathbb{R}^n$ tal que $|\vec{v}| < r$. Então,

$$f(a + \vec{v}) = f(a) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j + |\vec{v}|^2 E(a; \vec{v}), \quad \text{com } \lim_{\vec{v} \rightarrow 0} E(a; \vec{v}) = 0.$$

Prova:

Fixado $\vec{v} = \langle v_1, \dots, v_n \rangle \neq \vec{0} \in \mathbb{R}^n$, $|\vec{v}| < r$, consideremos o seu versor $\vec{\omega} = \langle \omega_1, \dots, \omega_n \rangle = \frac{\vec{v}}{|\vec{v}|}$. Aplicando a $\varphi(t) = f(a + t\vec{\omega})$, $t \in [0, r]$, a Fórmula de Taylor com Resto Infinitesimal temos

$$(*) \quad \varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2}t^2 + t^2 E(0; t), \quad \lim_{t \rightarrow 0} E(0; t) = 0,$$

com, analogamente ao mostrado no Teorema 2 e no Corolário 1 para funções em duas variáveis,

$$\varphi(0) = f(a), \quad \varphi'(0) = \left\langle \vec{\nabla} f(a), \vec{\omega} \right\rangle = \frac{\partial f}{\partial \vec{\omega}}(a) = 0 \quad \text{e} \quad \varphi''(0) = \frac{\partial^2 f}{\partial^2 \vec{\omega}}(a) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \omega_i \omega_j.$$

Assim, substituindo tais expressões em (*) obtemos a equação

$$f(a + t\vec{\omega}) = f(a) + \frac{t^2}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \omega_i \omega_j + t^2 E(0; t); \quad \lim_{t \rightarrow 0} E(0; t) = 0,$$

da qual com a substituições $t = |\vec{v}|$ e $\omega = \frac{\vec{v}}{|\vec{v}|}$ e notando que $|\vec{v}|^2 \omega_i \omega_j = v_i v_j$ concluímos a tese ■

Teorema 5 Sejam $f \in C^2(B(a; r))$, com $B(a; r) \subset \mathbb{R}^n$, $r > 0$, a um ponto estacionário de f e

$$Q(\vec{v}) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a) v_i v_j, \quad \vec{v} = \langle v_1, \dots, v_n \rangle \in \mathbb{R}^n.$$

Então,

- (a) Se os auto-valores de $\mathcal{H}f(a)$ são estritamente positivos, f tem um mínimo local em a .
- (b) Se os auto-valores de $\mathcal{H}f(a)$ são estritamente negativos, f tem um máximo local em a .
- (c) Se $\mathcal{H}f(a)$ tem auto-valores positivos e negativos, ambos estritamente, a é um ponto de sela.

Prova:

- (a) Se $m > 0$ é o menor auto-valor de A , pelo Corolário 4 (b) temos $Q(\vec{v}) \geq m|\vec{v}|^2$, $\forall \vec{v} \in \mathbb{R}^n$. Ainda, pelo Lema 5 e sua notação existe $\delta > 0$ tal que $|E(a; \vec{v})| < \frac{m}{4}$ se $0 < |\vec{v}| < \delta$ e então,

$$f(a + \vec{v}) - f(a) = \frac{1}{2} Q(\vec{v}) + |\vec{v}|^2 E(a; \vec{v}) \geq \frac{m}{2} |\vec{v}|^2 - \frac{m}{4} |\vec{v}|^2 = \frac{m}{4} |\vec{v}|^2 > 0.$$

- (b) Segue de (a), trocando f por $-f$.
- (c) Basta notar que se λ é auto-valor de A e \vec{v} é um vetor unitário tal que $A\vec{v} = \lambda\vec{v}$ então,

$$\frac{\partial^2 f}{\partial \vec{v}^2}(a) = Q(\vec{v}) = \lambda \vec{v} \cdot \vec{v} = \lambda |\vec{v}|^2 = \lambda.$$

Logo, havendo auto-valores com sinais opostos, sobre um segmento por a a restrição de f assume em a um valor mínimo estrito e sobre um outro segmento por a , f assume em a um valor máximo estrito. Logo, a não é um extremante local de f e é um ponto de sela ■

O resultado abaixo permite mostrar a “volta” do Teorema 4. Isto é, se Q é uma forma definida positiva então os menores principais $\Delta_1, \dots, \Delta_n$ são estritamente positivos.

Corolário 4: Se $A \in M_{n \times n}(\mathbb{R})$ é simétrica e real e $Q(X) = AX \cdot X > 0, \forall X \neq 0$, então $\det A > 0$.

Prova:

Pelo Corolário 3 o polinômio característico de A , $p(\lambda) = \det(A - \lambda I)$, I a matriz identidade de ordem n , tem ao menos uma raiz real e todas as suas raízes reais pertencem a $(0, +\infty)$. Ainda mais, seu monômio de maior grau é $(-1)^n \lambda^n$ e assim,

$$\lim_{\lambda \rightarrow -\infty} p(\lambda) = \lim_{\lambda \rightarrow -\infty} (-1)^n \lambda^n = +\infty, \text{ seja } n \text{ par ou ímpar.}$$

Então, como não há raízes reais em $(-\infty, 0]$ segue que $p(0) = \det A$ é estritamente positivo ■

Estratégia para a determinação de máximos e mínimos, locais e absolutos.

Definições topológicas: Abaixo, \mathbb{R}^n é \mathbb{R} ou \mathbb{R}^2 ou \mathbb{R}^3 . Seja $a \subset \mathbb{R}^n$.

- A é **fechado** se seu complementar é aberto.
- $A = K \subset \mathbb{R}^n$ é **compacto** se é fechado e limitado.
- O **interior** de A , $\text{int}(A)$, é o conjunto dos pontos X em A tais que existe uma bola aberta centrada em X e contida em A : $B(X; r) \subset A, r > 0$.
- A **fronteira** de A , indicada por ∂A , é o conjunto dos pontos $X \in \mathbb{R}^n$ tais que toda bola aberta centrada em X intersecta A e A^C , o **complementar de A** . Isto é, $\forall r > 0, B(X; r) \cap A \neq \emptyset$ e $B(X; r) \cap A^C \neq \emptyset$.

Obs Dado $X \in A \subset \mathbb{R}^n$, X é um **ponto interior**, $X \in \text{int}(A)$, ou **de fronteira**, $X \in \partial A$.

Def. Dada $f : \text{Dom}(f) \rightarrow \mathbb{R}$ e $A \subset \text{Dom}(f) \subset \mathbb{R}^n$, um **extremante local de f em A** é um ponto $P_o \in A$ que é **máximo, ou mínimo, local de f restrita a A** . Isto é, existe $r > 0$ tal que P_o é máximo, ou mínimo, de f restrita ao conjunto $B(P_o; r) \cap A$.

Segue, sem a demonstração, um resultado fundamental para estudo de máximos e mínimos.

Teorema (Weierstrass) Dada $f \in C(K)$, K compacto em \mathbb{R}^n , f assume valor máximo absoluto e mínimo absoluto em K . Isto é, existem P_1 e $P_2 \in K$ tais que

$$f(P_1) \leq f(X) \leq f(P_2), \quad \forall X \in K.$$

Determinação de máximos e mínimos, locais e absolutos, para $f \in C^1(K)$:

- (1) Restringindo f a $\text{int}(K)$ determinamos os pontos críticos, candidatos a extremantes.
- (2) Investigamos os possíveis pontos de máximo e mínimo de f sobre a fronteira, ∂K , ou elementarmente ou (veremos) por multiplicadores de Lagrange.
- (3) Os extremos absolutos estão entre os valores de f nos pontos encontrados.