CÁLCULO II

FEA-USP - 2º SEMESTRE de 2012

Professor Oswaldo Rio Branco de Oliveira

O HESSIANO
POLINÔMIOS DE TAYLOR DE ORDEM A 1 EM DUAS VARIÁVEIS

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

Lema 1. *Seja* $f \in C^2([a,b])$ *e dois pontos* x_0 *e* x *distintos e ambos em* $[a,b]$. *Então, existe ao menos um ponto* ξ, *com* 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 número real λ dependendo de x_0 e x tal que

$$(1.1) \qquad f(x) = f(x_0) + f'(x_0)(x-x_0) + \lambda(x-x_0)^2.$$

Definamos então a função,

$$\varphi(t) = f(t) - f(x_0) - f'(x_0)(t-x_0) - \lambda(t-t_0)^2.$$

Temos, $\varphi(x_0) = 0$ e, pela equação (1.1), $\varphi(x) = 0$. Logo, pelo TVM existe c, com c entre x_0 e x, $c \neq x_0$ e $c \neq x$, tal que

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

Pelo TVM aplicado a f', existe ξ, com ξ entre x_0 e c, $\xi \neq x_0$ e $\xi \neq c$, tal que

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

Abaixo, Ω indica um conjunto aberto em $\mathbb{R}^2$.

Proposição 2. *Sejam* $f \in C^2(\Omega)$, $P_0 = (x_0,y_0)$ *um ponto em* Ω *e* $\vec{v} = \langle h,k\rangle \neq \vec{0}$. *Consideremos a restrição* $\varphi = \varphi_{\vec{v}}$ *(isto é,* φ *depende do vetor* $\vec{v}$*), de* f *sobre um segmento por* P_0, *na direção* $\vec{v}$ *e contido em* Ω:

$$\varphi(t) = f(x_0+th, y_0+tk)\,, \quad \textit{onde } t \in (-r,r)\,,\ r>0 \ \textit{ e } \ B(P_0\,;\, r|\vec{v}|) \subset \Omega\ .$$

Temos,

$$(a) \quad \varphi'(t) = \frac{\partial f}{\partial x}(x_0+th, y_0+tk)h + \frac{\partial f}{\partial y}(x_0+th, y_0+tk)k.$$

$$(b) \quad \varphi''(t) = \frac{\partial^2 f}{\partial x^2}(x_0+th, y_0+tk)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0+th, y_0+tk)hk + \frac{\partial^2 f}{\partial y^2}(x_0+th, y_0+tk)k^2.$$

$$(c) \quad \varphi'(0) = \frac{\partial f}{\partial x}(P_0)h + \frac{\partial f}{\partial y}(P_0)k = \vec{\nabla} f(P_0)\cdot\vec{v}\ , \qquad \varphi'(0) = \frac{\partial f}{\partial \vec{v}}(P_0)\,,\ se\,|\vec{v}| = 1.$$

$$(d) \quad \varphi''(0) = \frac{\partial^2 f}{\partial x^2}(P_0)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(P_0)hk + \frac{\partial^2 f}{\partial y^2}(P_0)k^2\,, \quad \varphi''(0) = \frac{\partial^2 f}{\partial \vec{v}^2}(P_0)\,,\ se\ |\vec{v}| = 1.$$

Prova.

(a) Pela regra da cadeia temos

$$\varphi' = \frac{d}{dt}\Big\{f(x_0+th\,,y_0+tk)\Big\} = \frac{\partial f}{\partial x}(x_0+th\,,y_0+tk)h + \frac{\partial f}{\partial y}(x_0+th\,,y_0+tk)k = \frac{\partial f}{\partial x}h+\frac{\partial f}{\partial y}k\,,$$

(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

$$\begin{aligned}
\varphi'' &= \frac{d}{dt}\Big\{\frac{d}{dt}\Big[f(x_0+th\,,y_0+tk)\Big]\Big\} \\
&= \frac{d}{dt}\Big[\frac{\partial f}{\partial x}(x_0+th\,,y_0+tk)\Big]h + \frac{d}{dt}\Big[\frac{\partial f}{\partial y}(x_0+th\,,y_0+tk)\Big]k \\
&= \Big[\frac{\partial^2 f}{\partial x^2}h+\frac{\partial^2 f}{\partial y \partial x}k\Big]h + \Big[\frac{\partial^2 f}{\partial x \partial y}h+\frac{\partial^2 f}{\partial y^2}k\Big]k = \frac{\partial^2 f}{\partial x^2}h^2 + 2\frac{\partial^2 f}{\partial x \partial y}hk + \frac{\partial^2 f}{\partial y^2}k^2.
\end{aligned}$$

(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) ■

Teorema 3 (Polinômio de Taylor de Ordem 1 com Resto de Lagrange). *Sejam f em $C^2(\Omega)$ e dois pontos $P=(x,y)$ e $P_0=(x_0,y_0)$, ambos em Ω, com o segmento $\overline{PP_0}\subset\Omega$. Então, existe um ponto $\overline{P}=(\overline{x},\overline{y})$ em $\overline{PP_0}$ e satisfazendo*

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) +$$

$$+ \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(\overline{x},\overline{y})(x-x_0)^2 + 2\frac{\partial^2 f}{\partial x\partial y}(\overline{x},\overline{y})(x-x_0)(y-y_0) + \frac{\partial^2 f}{\partial y^2}(\overline{x},\overline{y})(y-y_0)^2\right].$$

Prova.

Restringindo f ao segmento $\overline{PP_0}$, dada $\varphi(t)=f(x_0+th,y_0+tk)=f[(x_0,y_0)+t\vec{v}]$, $t\in[0,1]$, com $\vec{v}=\langle h,k\rangle=\langle x-x_0,y-y_0\rangle$, pelo Lema 1 existe $\overline{t}\in(0,1)$ tal que

$$\varphi(1)=\varphi(0)+\varphi'(0)+\frac{\varphi''(\overline{t})}{2}. \tag{3.1}$$

Temos, $\varphi(1)=f(x,y)$, $\varphi(0)=f(x_0,y_0)$ e, pela Proposição 2, itens (c) e (d),

$$\begin{cases} \varphi'(0) &= \frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k, \\ \\ \varphi''(\overline{t}) &= \frac{\partial^2 f}{\partial x^2}(x_0+\overline{t}h,y_0+\overline{t}k)h^2 + 2\frac{\partial^2 f}{\partial x\partial y}(x_0+\overline{t}h,y_0+\overline{t}k)hk + \frac{\partial^2 f}{\partial y^2}(x_0+\overline{t}h,y_0+\overline{t}k)k^2. \end{cases}$$

Substituindo no sistema acima $(\overline{x},\overline{y})=(x_0+\overline{t}h,y_0+\overline{t}k)$, $h=x-x_0$, $k=y-y_0$ e em (3.1) os valores achados para $\varphi(1)$, $\varphi(0)$, $\varphi'(0)$ e $\varphi''(\overline{t})$, obtemos a tese ■

Adendo. Com a notação $(x,y)=(x_0+h,y_0+k)$ temos,

$$\begin{aligned} f(x_0+h,y_0+k) &= f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)h+\frac{\partial f}{\partial y}(x_0,y_0)k+ \\ &+\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(\overline{x},\overline{y})h^2+2\frac{\partial^2 f}{\partial x\partial y}(\overline{x},\overline{y})hk+\frac{\partial^2 f}{\partial y^2}(\overline{x},\overline{y})k^2\right]. \end{aligned}$$

Matriz Hessiana e Máximos e Mínimos em Duas Variáveis

Definições. Seja $f : \Omega \to \mathbb{R}$, com Ω aberto em $\mathbb{R}^2$. Classificamos um ponto P_0 em Ω, relativamente à função f, como

(a) ponto de máximo [mínimo] local se existe uma bola aberta $B(P_0; r) \subset \Omega$, onde $r > 0$, tal que $f(P_0) \geq f(P)$ $[f(P_0) \leq f(P)]$ para todo $P \in B(P_0; r)$; se tal desigualdade é estrita para $P \in B(P_0; r)$, com $P \neq P_0$, dizemos que P_0 é ponto de máximo [mínimo] local estrito.

(b) ponto de máximo [mínimo] global, ou absoluto, se $f(P_0) \geq f(P)$ $[f(P_0) \leq f(P)]$, para todo $P \in \Omega$; se tal desigualdade é estrita para $P \in \Omega \setminus \{P_0\}$, dizemos que P_0 é, em adição, estrito.

(c) extremante local [absoluto] se é um ponto de máximo, ou de mínimo, local [absoluto].

(d) ponto crítico, ou estacionário, de f, supondo f em $C^1(\Omega)$, se

$$\frac{\partial f}{\partial x}(P_0) = \frac{\partial f}{\partial y}(P_0) = 0 .$$

(e) ponto de sela, se P_0 é ponto crítico de f mas não de máximo ou mínimo, locais.

Observação 4. Um ponto de máximo, ou mínimo, local de uma função f, com f de classe C^1 em um aberto, é sempre um ponto crítico.

Corolário 5. *Com mesmas hipóteses e notação da Proposição 2 temos:*

(i) Se P_0 é ponto de máximo [mínimo] local de f na direção $\overrightarrow{v}$, $|\vec{v}| = 1$; isto é, $t = 0$ é máximo [mínimo] local de $\varphi_{\vec{v}}$, então,

$$\frac{\partial f}{\partial \vec{v}}(P_0) = \overrightarrow{\nabla} f(P_0) \cdot \overrightarrow{v} = 0 \quad ; \quad \frac{\partial^2 f}{\partial \vec{v}^2}(P_0) \leq 0 \quad \left[\frac{\partial^2 f}{\partial \vec{v}^2}(P_0) \geq 0\right].$$

(ii) Se P_0 é ponto de máximo [mínimo] local de f então $\overrightarrow{\nabla} f(P_0) = \overrightarrow{0}$ e,

$$\frac{\partial^2 f}{\partial \vec{v}^2}(P_0) = \frac{\partial^2 f}{\partial x^2}(P_0)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(P_0)hk + \frac{\partial^2 f}{\partial y^2}(P_0)k^2 \leq 0 \; [\geq 0], \forall |\overrightarrow{v}| = 1.$$

Prova.

(i) Segue da análise das derivadas de ordens 1 e 2 da função de uma variável $\varphi_{\vec{v}}$ e do Lema 1.

(ii) Consequência imediata da Proposição 2, de (i) e da Observação 4, acima ■

Corolário 6. *Seja $f \in C^2(\Omega)$, $P_0 = (x_0, y_0)$ um seu ponto crítico e $B(P_0; r) \subset \Omega$, com $r > 0$. Dado $\overrightarrow{v} = \langle h, k \rangle$, com $|\overrightarrow{v}| < r$, existe um ponto $(\overline{x}, \overline{y})$ no segmento unindo os pontos P_0 e $P_0 + \overrightarrow{v}$ tal que*

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(\overline{x}, \overline{y})h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\overline{x}, \overline{y})hk + \frac{\partial^2 f}{\partial y^2}(\overline{x}, \overline{y})k^2\right].$$

Prova. Consequência imediata do Teorema 3 e da definição de ponto crítico ■

Estudemos a forma quadrática $Q_P = Q_P(h, k) = f_{xx}(P)h^2 + 2f_{xy}(P)hk + f_{yy}(P)k^2$.

Observação 7. Dados $a, b, c \in \mathbb{R}$, sejam $H = ac - b^2$ e a forma quadrática $Q : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$z = Q(h, k) = ah^2 + 2bhk + ck^2 = \begin{bmatrix} h & k \end{bmatrix}\begin{bmatrix} a & b \\ b & c \end{bmatrix}\begin{bmatrix} h \\ k \end{bmatrix}.$$

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

(ii) Se $a \neq 0$, o gráfico de $Q : \mathbb{R}^2 \to \mathbb{R}$ 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 $H < 0$, o gráfico de Q é um parabolóide hiperbólico.

(iv) A função $z = Q(h, k)$ troca de sinal se e somente se $H < 0$.

Se o gráfico de Q é um parabolóide elíptico ou circular então $0 = Q(0, 0)$ é valor mínimo/máximo estrito e absoluto. Se o gráfico de Q é um parabolóide cilíndrico então $0 = Q(0, 0)$ 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) Consequência trivial de (i).

(iii) Segue de (ii), se $a^2 + c^2 \neq 0$. Se $a = c = 0$, segue da expressão para $Q(h,k)$.

(iv) Segue de (i), (ii) e (iii) ■

Definição. Se M é a matriz (simétrica) 2×2 dada na Observação 7 então Q é a forma quadrática associada a M.

Comentário. Identificando $\mathbb{R}^2$ com o espaço das matrizes-coluna $M_{2\times1}(\mathbb{R})$, 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 $[\vec{v}]^t = [h \;\; k] \in M_{1\times2}(\mathbb{R})$ 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 "·" denota o produto escalar em $\mathbb{R}^2$.

Proposição 8. *Seja $f \in C^2(\Omega)$, P um ponto crítico de f, e a forma quadrática*

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

Seja $\overrightarrow{v} = \langle h, k \rangle$ unitário e $Hf(P) = f_{xx}(P)f_{yy}(P) - [f_{xy}(P)]^2$. Temos,

(i) $Q_P(h,k) = \frac{\partial^2 f}{\partial \vec{v}^2}(P)$.

(ii) $\frac{\partial^2 f}{\partial \vec{v}^2}(P) > 0$*, para todo* $\overrightarrow{v}$ *unitário, se somente se* $Hf(P) > 0$ *e* $f_{xx}(P) > 0$.

(iii) $\frac{\partial^2 f}{\partial \vec{v}^2}(P) < 0$*, para todo* $\overrightarrow{v}$ *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* $\overrightarrow{v}$ *unitário, se e somente se,* $Hf(P) < 0$.

Prova.

(i) Segue da Proposição 2 (d).

(ii) , (iii) e (iv) Seguem imediatamente de (i) e da Observação 7 ■

Definições. Seja $f \in C^2(\Omega)$ e P um ponto em Ω.

- A matriz hessiana de f em P, denotada $\mathcal{H}f(P) \in M_{2x2}(\mathbb{R}) = [a_{ij}]$, 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, no ponto P, indicada Q_P, é a forma quadrática associada à matriz hessiana $\mathcal{H}f(P)$.

Teorema 9 (Teste do Hessiano). *Seja $f \in C^2(\Omega)$, um ponto $P_0 = (x_0, y_0)$ crítico de f em Ω, a matriz hessiana*

$$\mathcal{H}f(P_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(P_0) & \frac{\partial^2 f}{\partial y \partial x}(P_0) \\ \frac{\partial^2 f}{\partial x \partial y}(P_0) & \frac{\partial^2 f}{\partial y^2}(P_0) \end{bmatrix} \quad e \quad Hf(P_0) = \det \mathcal{H}f(P_0).$$

(a) *Se $Hf(P_0) > 0$ e $f_{xx}(P_0) > 0$ então P_0 é um ponto de mínimo local estrito.*

(b) *Se $Hf(P_0) > 0$ e $f_{xx}(P_0) < 0$ então P_0 é um ponto de máximo local estrito.*

(c) *Se $Hf(P_0) < 0$ então P_0 é um ponto de sela.*

(d) *Se $Hf(P_0) = 0$ então P_0 pode ser de qualquer um dos tipos acima.*

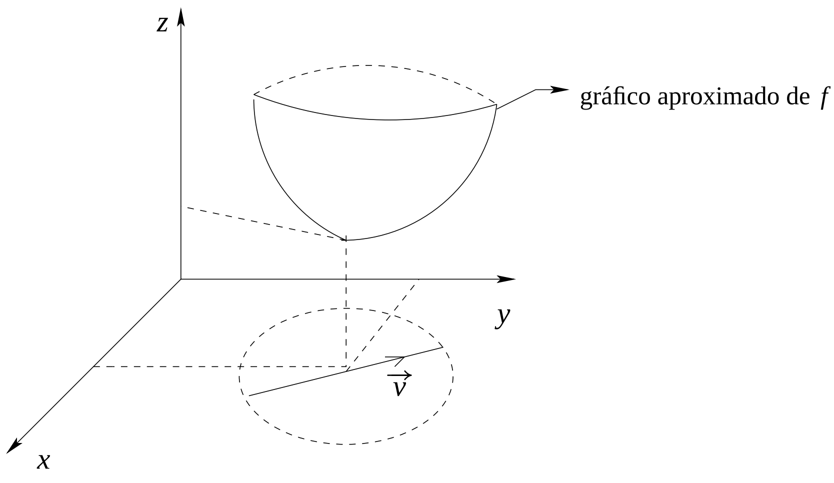

Figura 1: Caso em que $f_{xx}(P_0) > 0$ e $Hf(P_0) > 0$.

Prova.

(a) Como $f \in C^2$, temos $Hf = f_{xx}f_{yy} - f_{xy}^2 > 0$ e $f_{xx} > 0$ numa bola $B(P_0; r)$, se $r > 0$ e suficientemente pequeno. Pelo Corolário 6, dado $\vec{v} = \langle h, k \rangle \neq \vec{0}$, com $|\vec{v}| < r$, existe $\overline{P} = (\overline{x}, \overline{y})$ no segmento unindo P_0 e $P_0 + \vec{v}$, e portanto $\overline{P} \in B(P_0; r)$, tal que

$$\begin{aligned} f(P_0 + \vec{v}) - f(P_0) &= \tfrac{1}{2}\left[f_{xx}(\overline{P})h^2 + 2f_{xy}(\overline{P})hk + f_{yy}(\overline{P})k^2 \right] = \\ &= \tfrac{1}{2}\left[f_{xx}(\overline{P})\left(h + \tfrac{f_{xy}(\overline{P})}{f_{xx}(\overline{P})}k\right)^2 + \tfrac{Hf(\overline{P})}{f_{xx}(\overline{P})}k^2 \right] > 0\,. \end{aligned}$$

(b) Basta aplicar o ítem (a) à função $-f$.

(c) Por contradição. Se P_0 é ponto de máximo/mínimo então, pelo Corolário 5 obtemos $\frac{\partial^2 f}{\partial \vec{v}^2}(P_0) \geq 0$, para todo $|\vec{v}| = 1$ $\left[\frac{\partial^2 f}{\partial \vec{v}^2}(P_0) \leq 0\right.$, para todo $\left.|\vec{v}| = 1\right]$, e pela Proposição 8 (iv) concluímos $Hf(P_0) \geq 0$ $\lightning$

(d) 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_0 é extremante local então P_0 é extremante global (absoluto) (vide Exercício 2, p. 315, "Um Curso de Cálculo", H. L. Guidorrizzi, vol 4, 5ª edição).

Prova.

Seja $\vec{v} = \langle h, k \rangle \in \mathbb{R}^2$. Sendo $P_0 = (x_0, y_0)$ um extremante local ele é um ponto crítico e assim,

$$(1) \qquad \vec{0} = \vec{\nabla} f(x_0, y_0) = \langle 2ax_0 + cy_0 + d, 2by_0 + cx_0 + e \rangle\,.$$

Computando f em $(x_0 + h, y_0 + k)$ obtemos

$$\begin{aligned} f(x_0 + h, y_0 + k) &= a(x_0 + h)^2 + b(y_0 + k)^2 + c(x_0 + h)(y_0 + k) + d(x_0 + h) + e(y_0 + k) + l \\ &= (ah^2 + chk + bk^2) + (2ax_0 + cy_0 + d)h + (2by_0 + cx_0 + e)k + \\ &+ (ax_0^2 + by_0^2 + cx_0y_0 + dx_0 + ey_0 + l) = \\ &= (ah^2 + chk + bk^2) + f(x_0, y_0)\,. \end{aligned}$$

Donde, encontramos

$$(2)\quad f(x_0+h,y_0+k)-f(x_0,y_0)=ah^2+chk+bk^2\ .$$

Solução (sugerida): Aplique o Teste do Hessiano ■

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

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

Resolução.

Temos

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

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

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

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

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

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

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

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

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

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 (há soluções via geometria vetorial ou multiplicadores de Lagrange).

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

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

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 $P_0 = (\frac{44}{7}\ ,\ \frac{16}{7})$ 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, mede a distância (ao quadrado) 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 deduzimos que P_0 é ponto 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:

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

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 $(1, 1, 0)$ pertence a 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 $(1, 5, -2)$ em s obtemos (utilizando a fórmula para a distância):

$$d = \frac{|6.1 - 2.5 + 3.(-2) - 4|}{\sqrt{6^2 + 2^2 + 3^2}} = \frac{14}{7} = 2 \blacksquare$$

O exemplo a seguir é de G. Peano (1884), corrigindo um engano 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.

Solução.

(a) É fácil ver que $\vec{\nabla} f = \langle 2x(4x^2 - 3y), -3x^2 + 2y \rangle = \langle 0, 0 \rangle$ se e somente se $(x, y) = (0, 0)$.

(b) Temos $f_{xx} = 24x^2 - 6y$, $f_{xy} = f_{yx} = -6x$ e $f_{yy} = 2$. Logo, o determinante hessiano em $(0, 0)$ é $Hf(0, 0) = 0$.

(c) Seja $\varphi(x) = f(x, mx) = (mx - x^2)(mx - 2x^2) = 2x^4 - 3mx^3 + m^2x^2$, com m uma constante real. Vales as fórmulas, $\varphi'(x) = 8x^3 - 9mx^2 + 2m^2x$, $\varphi''(x) = 24x^2 - 18mx + 2m^2$, $\varphi'(0) = 0$ e $\varphi''(0) = m^2 > 0$, se $m \neq 0$. Logo, o ponto $x = 0$ 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 então ponto de mínimo de φ.

(d) Nas regiões $\{(x,y) : y > 2x^2\}$, $\{(x,y) : x^2 < y < 2x^2\}$ e $\{(x,y) : y < x^2\}$ 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 $f > 0$.

Indiquemos a transposta de uma matriz $A = [a_{ij}] \in M_{n\times m}(\mathbb{R})$, $1 \leq i \leq n$ e $1 \leq j \leq m$, por $A^t \in M_{m\times n}(\mathbb{R})$, onde $A^t = [b_{rs}]$ e $b_{rs} = a_{sr}$ se $1 \leq r \leq m$ e $1 \leq s \leq n$. Dizemos que uma matriz quadrada $A \in M_{n\times n}(\mathbb{R})$ é simétrica se $A^t = A$.

Exemplo 8. Seja $f : \Omega \to \mathbb{R}$, 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), a seguir, 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 as matrizes M e N tem determinante 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 Máximos e Mínimos em Três Variáveis

Estas seção é facilmente generalizável para o caso de matrizes simétricas de ordem $n \in \mathbb{N}^*$.

Proposição 10. *Seja* $A \in M_3(\mathbb{R})$ *uma matriz simétrica 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,*

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

respectivamente, são não nulos. Então, existe $P \in M_3(\mathbb{R})$ *tal que* $\det P = 1$ *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 método da eliminação de Gauss e de operações elementares realizadas pela multiplicação por matrizes (elementares) com determinante 1. Lembremos que multiplicar uma linha ou coluna por um número e então adicioná-la a uma outra linha ou coluna, respectivamente, é uma operação elementar que não altera o determinante de uma matriz.

Na etapa 1 inicialmente multiplicamos a 1ª linha de A por $-\frac{a_{12}}{a_{11}}$ e somamos à 2ª linha e, ainda, multiplicamos a 1ª linha por $-\frac{a_{13}}{a_{11}}$ e somamos à 3ª linha. Efetuamos tais operações, que comutam, multiplicando a matriz A à 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 operações nas colunas correspondendo as feitas nas linhas: multiplicamos a 1ª coluna por $-\frac{a_{12}}{a_{11}}$ e somamos na 2ª coluna e multiplicamos a 1ª coluna por $-\frac{a_{13}}{a_{11}}$ e somamos na 3ª coluna. Efetuamos tais operações multiplicando a matriz à direita por P_1^t e P_2^t. Resumindo as quatro operações obtemos,

$$P_2P_1\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} P_1^tP_2^t = D_1 = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} \end{bmatrix},$$

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

Devido às operações realizadas os respectivos menores principais das matrizes simétricas A e D_1 são 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,

$$P_3D_1P_3^t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{a_{23}^{(1)}}{a_{22}^{(1)}} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -\frac{a_{23}^{(1)}}{a_{22}^{(1)}} \\ 0 & 1 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(2)} & 0 \\ 0 & 0 & a_{33}^{(2)} \end{bmatrix} = D_2.$$

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

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

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

Por fim, notemos que $P^tAP = D_2$ com $P^t = P_3P_2P_1$ e $\det P = 1$ ■

Lembrete. Duas matrizes quadradas A e B, ambas em $M_n(\mathbb{R})$, são congruentes se existe P inversível em $M_n(\mathbb{R})$ tal que $P^tAP = B$.

Notação. A seguir, 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 11. *Com as hipóteses e a notação na Proposição 10, consideremos a forma quadrática $Q(X) = X^tAX$, com X em $\mathbb{R}^3$. Valem as propriedades abaixo.*

(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) [porém, com $\Delta_{i's}$ todos não nulos], então existem dois vetores X_1 e X_2 tais que $Q(X_1) > 0$ e $Q(X_2) < 0$.

Prova.

Pela Proposição 10 existe uma matriz inversível P em $M_3(\mathbb{R})$ tal que $P^tAP = D$, onde $D = [d_{ij}]$ é a matriz diagonal definida por $d_1 = d_{11} = \Delta_1$ e $d_i = d_{ii} = \frac{\Delta_i}{\Delta_{i-1}}$, se $2 \leq i \leq n$. Ainda mais, dado X em $\mathbb{R}^3$, existe Y em $\mathbb{R}^3$ tal que $X = PY$ e

$$Q(X) = X^tAX \;=\; (PY)^tAPY \;= Y^tP^tAPY \;= Y^tDY.$$

Logo, é fácil ver,

(a) Q é positiva definida $\Leftrightarrow d_i > 0 \;\; (\forall 1 \leq i \leq n) \;\Leftrightarrow\; \Delta_i > 0 \;\; (\forall 1 \leq i \leq n)$.

(b) Q é definida negativa $\Leftrightarrow d_i < 0 \; (\forall 1 \leq i \leq n) \;\Leftrightarrow\; \Delta_1 < 0, \Delta_2 > 0$ e $\Delta_3 < 0$.

(c) Existem elementos na diagonal de D com sinais contrários. Portanto, existe Y_1 em $\mathbb{R}^3$ tal que $Y_1^tDY_1 > 0$ e existe Y_2 em $\mathbb{R}^3$ tal que $Y_2^tDY_2 < 0$. Assim, basta tomarmos $X_i = PY_i$, com $i = 1, 2$ ■

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 12. *Dada $f \in C^2(\Omega)$ e P_0 um seu ponto crítico, consideremos*

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

(a) *Supondo-os não nulos temos,*

  (i) *Se os três são (estritamente) positivos, então P_0 é ponto de mínimo local estrito.*

  (ii) *Se (alternando sinais) temos $f_{xx} < 0$, $H_1 > 0$ e $H < 0$, então P_0 é ponto de máximo local estrito.*

  (iii) *Se (i) e (ii) não ocorrem, então P_0 é ponto de sela.*

(b) *Se existirem números na diagonal principal de Hf com sinais opostos, então 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 8 definimos para $\overrightarrow{v} = \langle v_1, v_2, v_3 \rangle$ em $\mathbb{R}^3$ e P em Ω,

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

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 11 temos $Q_P(\overrightarrow{v}) > 0$, $\forall \overrightarrow{v} \neq \vec{0}$, e por um resultado trivialmente análogo ao já provado em duas variáveis, na Proposição 8(i),

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

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

$$f(P) > f(P_0)\,, \text{para tod } P \in B(P_0; r), \text{ com } P \neq P_0.$$

(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, $\frac{\partial^2 f}{\partial^2 \vec{v_1}}(P_0) = Q_{P_0}(\vec{v_1}) > 0$, $\frac{\partial^2 f}{\partial^2 \vec{v_2}}(P_0) = Q_{P_0}(\vec{v_2}) < 0$, 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. Por outro lado, se tivermos $f_{xx} = f_{yy} = f_{zz} = 0$, 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

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

Solução.

Temos,

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

e 8 pontos críticos:

$$P_0 = (x_0, y_0, z_0) = (\pm 1\,, \pm 1\,, \pm 1)\ .$$

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:

$$Hf(P_0) = \begin{vmatrix} 6x & 0 & 0 \\ 0 & 6y & 0 \\ 0 & 0 & 6z \end{vmatrix} \quad , \quad H_1(P_0) = \begin{vmatrix} 6x & 0 \\ 0 & 6y \end{vmatrix}.$$

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

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

Solução.

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

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

Isto é,

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

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

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

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 $z = 0$), $1 \le i \le 9$,

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

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

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 minímo 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 $(\frac{x^2}{5} - \frac{1}{3}) \approx -\frac{1}{3} < 0$ 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 três 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 $(x,y,z) \in \mathbb{R}^3$, as três 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.

$$\begin{cases} \text{Pontos de sela:} & P_1, P_2, P_3, P_4, P_7, P_8 \text{ e } P_9. \\ \text{Pontos de mínimo local:} & P_5 \text{ e } P_6 \blacksquare \end{cases}$$

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

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

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

(a) Supondo-os não nulos temos,

(i) Se todos são (estritamente) positivos, então P_0 é ponto de mínimo local estrito.

(ii) Se (alternando sinais) temos $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, etc, então P_0 é ponto de máximo local estrito.

(iii) Se (i) e (ii) não ocorrem, então P_0 é ponto de sela.

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

Prova.

Com a argumentação no Teorema 12, 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 $d_{22} = \frac{\Delta_2}{\Delta_1}$,...,$d_{jj} = \frac{\Delta_j}{\Delta_{j-1}}$. O restante da prova é análogo ao Teorema 12 ■

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 $M_{n\times 1}(\mathbb{R})$:

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

Definições. A aplicação linear $T : \mathbb{R}^n \to \mathbb{R}^n$ associada à matriz $A \in M_{n\times n}(\mathbb{R})$ é dada por

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

Um número real λ é um auto-valor de A se existe X em $\mathbb{R}^n$ tal que

$$AX = \lambda X.$$

Se $A \in M_{n\times n}(\mathbb{R})$ é simétrica, então a forma quadrática associada a A é dada por

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

Lema 14. *Se Q é a forma quadrática associada à matriz simétrica A então*

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

Prova.

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

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

Logo,

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

Corolário 15. *Sejam M e m, o máximo e o mínimo, respectivamente, da forma quadrática Q sobre a esfera unitária $\{X \in \mathbb{R}^n : |X| = 1\}$. Então,*

(a) M e m são, respectivamente, o maior e o menor auto-valor (real) de A.

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

(c) Q é definida positiva se e só se os auto-valores de A são maiores que zero.

(d) Q é definida negativa se e só se os auto-valores de A são menores que zero.

Prova[1].

(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 $g(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2 - 1$, existe $\lambda \in \mathbb{R}$ tal que

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

e então, para tais pontos e pelo Lema 14 temos

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

e assim, se X_M, com $|X_M| = 1$, é tal que $Q(X_M) = M$ e λ_M em $\mathbb{R}$ 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 o número real λ é auto-valor de A, é claro que existe $\overrightarrow{v}$ unitário tal que $A(\overrightarrow{v}) = \lambda \overrightarrow{v}$. Neste caso temos $m \leq Q(\overrightarrow{v}) \leq M$ e, ainda, $Q(\overrightarrow{v}) = (A\overrightarrow{v}) \cdot \overrightarrow{v} = \lambda \overrightarrow{v} \cdot \overrightarrow{v} = \lambda |\overrightarrow{v}|^2 = \lambda$. Donde concluímos, $m \leq \lambda \leq M$.

(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) ∎

[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.

Lema 16. *Seja $f \in C^2(B(a;r))$, com $B(a;r) \subset \mathbb{R}^n$ e $r > 0$, a um ponto estacionário (isto é, ponto crítico) de f e $\overrightarrow{v} = \langle v_1, \dots, v_n\rangle \in \mathbb{R}^n$ tal que $|\overrightarrow{v}| < r$. Então,*

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

Prova.

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

$$(16.1) \qquad \varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2}t^2 + t^2 E(0;t), \quad \lim_{t\to 0} E(0;t) = 0,$$

com, analogamente ao mostrado no Teorema 9 e no Corolário 5 para funções em duas variáveis,

$$\varphi(0) = f(a), \quad \varphi'(0) = \left\langle \overrightarrow{\nabla} f(a), \overrightarrow{\omega} \right\rangle = \frac{\partial f}{\partial \overrightarrow{\omega}}(a) = 0 \text{ e}$$

$$\varphi''(0) = \frac{\partial^2 f}{\partial^2 \overrightarrow{\omega}}(a) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \omega_i \omega_j.$$

Assim, substituindo tais expressões em (16.1) obtemos a equação

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

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

Teorema 17. *Sejam* $f \in C^2(B(a;r))$, *com* $B(a;r) \subset \mathbb{R}^n$ *e* $r > 0$, *a um ponto estacionário de* f *e*

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

São válidas as propriedades abaixo.

(a) Se os auto-valores de $\mathcal{H}f(a)$ *são estritamente positivos, então* f *tem um mínimo local em* a.

(b) Se os auto-valores de $\mathcal{H}f(a)$ *são estritamente negativos, então* f *tem um máximo local em* a.

(c) Se $\mathcal{H}f(a)$ *tem auto-valores positivos e negativos, ambos estritamente, então* a *é um ponto de sela.*

Prova.

(a) Se $m > 0$ é o menor auto-valor de A, pelo Corolário 15 (b) obtemos $Q(\vec{v}) \geq m|\vec{v}|^2$, para todo $\vec{v}$ em $\mathbb{R}^n$. Ainda, pelo Lema 16 e sua notação, vemos que 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 13. Isto é, se Q é uma forma definida positiva então os menores principais $\Delta_1,....,\Delta_n$ são estritamente positivos.

Corolário 18. *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 15 o polinômio característico de A, $p(\lambda) = \det(A - \lambda I)$, I a matriz identidade de ordem n, tem ao menos uma raíz 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,

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

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 é compacto se A é fechado e limitado. Escrevemos então $A = K$.
- O interior de A, $int(A)$, 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 $X \in R^n$ tais que toda bola aberta centrada em X intersecta A e A^C, o complementar de A. Isto é, $\forall r > 0$, $B(X;r) \cap A \neq \varnothing$ e $B(X;r) \cap A^C \neq \varnothing$.

Obs. Dado $X \in A$, X é um ponto interior [$X \in int(A)$] ou de fronteira [$X \in \partial A$].

Definição. Dada $f : Dom(f) \to \mathbb{R}$ e $A \subset Dom(f) \subset \mathbb{R}^n$, um extremante local de f em A é um ponto $P_0 \in A$ que é máximo, ou mínimo, local dea função f restrita a A. Isto é, existe $r > 0$ tal que P_0 é máximo, ou mínimo, de f restrita ao conjunto $B(P_0; r) \cap A$.

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 e mínimo absolutos em K. Isto é, existem P_1 e P_2, em K, tais que

$$f(P_1) \leq f(X) \leq f(P_2)\ , \quad \text{para todo } X \text{ em } K\ .$$

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

(1) Restringindo f a $int(K)$ 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.

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. Simmons, G. F., *Cálculo com Geometria Analítica*, Vol 2, McGraw-Hill, 1988.
5. Hairer, E., and Wanner, G., *Analysis by Its History*, Springer-Verlag, 1996.