CÁLCULO III - MAT 216 - IFUSP - primeiro semestre de 2014
Professor Oswaldo Rio Branco de Oliveira
`http://www.ime.usp.br/~oliveira` oliveira@ime.usp.br

O HESSIANO EM DUAS E VÁRIAS VARIÁVEIS

1 - Introdução

Lema 1. *Seja* $\varphi \in C^2([a,b])$ *e dois pontos* t_0 *e* t *distintos e ambos em* $[a,b]$. *Então, existe ao menos um ponto* ξ, *com* ξ *entre* t_0 *e* t, $\xi \neq t_0$ *e* $\xi \neq t$, *tal que*

$$\varphi(t) = \varphi(t_0) + \varphi'(t_0)(t-t_0) + \frac{\varphi''(\xi)}{2}(t-t_0)^2.$$

Prova. Existe, é óbvio, um único número λ, dependendo de t_0 e t. tal que

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

Definamos a função $F(s) = \varphi(s) - \varphi(t_0) - \varphi'(t_0)(s-t_0) - \lambda(s-t_0)^2$.

Temos $F(t_0) = 0 = F(t)$. Pelo TVM existe c entre (estritamente) t_0 e t tal que

$$0 = F'(c) = \varphi'(c) - \varphi'(t_0) - 2\lambda(c-t_0) \quad \Longrightarrow \quad 2\lambda = \frac{\varphi'(c)-\varphi'(t_0)}{c-t_0}.$$

Pelo TVM aplicado a φ', existe ξ, com ξ entre t_0 e c, $\xi \neq t_0$ e $\xi \neq c$, tal que

$$\frac{\varphi'(c)-\varphi'(t_0)}{c-t_0} = \varphi''(\xi) \quad \Longrightarrow \quad \lambda = \frac{\varphi''(\xi)}{2!} \blacksquare$$

A seguir, para simplificar, obtemos fórmulas em torno da origem $O = (0,0)$ no plano cartesiano. Sejam $\{\overrightarrow{e_1}, \overrightarrow{e_2}\}$ a base canônica de $\mathbb{R}^2$ e $r > 0$.

Teorema 2 (Polinômio de Taylor de Ordem 1, Resto de Lagrange). *Sejam* $f \in C^2(B(O;r))$, $\overrightarrow{v} = (h,k)$ *com* $|\overrightarrow{v}| < r$ *e o ponto* $P = O + \overrightarrow{v}$. *Definamos*

$$\varphi(t) = f(th,tk)\,, \quad \textit{para } t \textit{ em um intervalo aberto contendo } [-1,1].$$

Temos,

$$(a)\quad \varphi'(t) = f_x(th,tk)h + f_y(th,tk)k = \frac{\partial f}{\partial \vec{v}}(th,tk).$$

$$(b)\quad \varphi''(t) = f_{xx}(th,tk)h^2 + 2f_{xy}(th,tk)hk + f_{yy}(th,tk)k^2 = \frac{\partial^2 f}{\partial \vec{v}^2}(th,tk).$$

$$(c)\quad f(O+\overrightarrow{v}) = f(O) + \frac{\partial f}{\partial \vec{v}}(O) + \frac{1}{2}\frac{\partial^2 f}{\partial \vec{v}^2}(\overline{P}), \quad \textit{com } \overline{P} \textit{ no segmento } \overline{OP}.$$

Prova.

(a) Imediata, pela regra da cadeia e pela definição de $\partial f/\partial\vec{v}$.

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

$$\varphi''(t) = \Big[\frac{\partial^2 f}{\partial x^2}h + \frac{\partial^2 f}{\partial y\partial x}k\Big]h + \Big[\frac{\partial^2 f}{\partial x\partial y}h + \frac{\partial^2 f}{\partial y^2}k\Big]k \; = \; \frac{\partial^2 f}{\partial x^2}h^2 + 2\frac{\partial^2 f}{\partial x\partial y}hk + \frac{\partial^2 f}{\partial y^2}k^2.$$

(c) Basta avaliar $\varphi(1)$, $\varphi(0)$, $\varphi'(0)$ e usar o Lema 1 e os itens (a) e (b)$\blacksquare$

Exercício 1 Sejam $f \in C^2(\mathbb{R}^2;\mathbb{R})$, um ponto P e $\overrightarrow{v} = (h,k)$. Então,

$$f(P + \overrightarrow{v}) = f(P) + \langle\nabla f(P),(h,k)\rangle + \frac{1}{2}\left[f_{xx}(\bar{P})h^2 + 2f_{xy}(\bar{P})hk + f_{yy}(\bar{P})k^2\right],$$

para algum $\overline{P}$ no segmento unindo os pontos P e $P + \overrightarrow{v}$.

2 - O Hessiano 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) de máximo [mínimo] local se $f(P_0) \geq f(P)$ $[f(P_0) \leq f(P)]$ para todo $P \in B(P_0;r) \subset \Omega$, para algum $r > 0$; se a desigualdade é estrita para todo $P \in B(P_0;r) \setminus \{P_0\}$, então P_0 é de máximo [mínimo] local estrito.

(b) 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 todo ponto $P \in \Omega \setminus \{P_0\}$, dizemos que P_0 é, em adição, estrito.

(c) extremante local [absoluto] se é de máximo, ou 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) de sela, se é ponto crítico de f mas não de máximo ou mínimo, locais.

Atenção. 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.

Se $|\overrightarrow{v}| = 1$, então $\frac{\partial f}{\partial \vec{v}}(P)$ é chamada a derivada direcional (normalizada) de f, no ponto P e na direção $\overrightarrow{v}$.

Corolário 3. *Seja $f \in C^2(B(P_0;r))$, com $P_0 = (x_0, y_0)$ um ponto crítico de f. Dado $\overrightarrow{v} = (h,k)$, com $|\overrightarrow{v}| < r$, existe um ponto $(\bar{x}, \bar{y})$ no segmento unindo os pontos P_0 e $P_0 + \overrightarrow{v}$ tal que*

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

Prova. Segue, trivialmente, do Exercício 1 e da definição de ponto crítico■

Estudemos a forma quadrática $Q_P = Q_P(h,k) = f_{xx}(P)h^2 + 2f_{xy}(P)hk + f_{yy}(P)k^2$. Antes, vejamos algumas definições e notações apropriadas.

A transposta de uma matriz $A = (a_{ij}) \in M_{n\times m}(\mathbb{R})$, $1 \leq i \leq n$ e $1 \leq j \leq m$, é a matriz $A^T \in M_{m\times n}(\mathbb{R})$, onde $A^T = (b_{rs})$ e $b_{rs} = a_{sr}$ se $1 \leq r \leq m$ e $1 \leq s \leq n$.

Uma matriz (quadrada) é simétrica se $A^T = A$.

Fixemos $\{\overrightarrow{e_1}, \overrightarrow{e_2}\}$ a base canônica de $\mathbb{R}^2$. Seja O a origem de $\mathbb{R}^2$.

Identificando $\mathbb{R}^2 \equiv M_{2\times 1}(\mathbb{R})$, indicamos um vetor $v = h\overrightarrow{e_1} + k\overrightarrow{e_2}$ por $v = \begin{pmatrix} h \\ k \end{pmatrix}$.

Assim, temos o (vetor) transposto $v^T = (h \quad k)$.

Teorema 4. Dados $a, b, c \in \mathbb{R}$, sejam $H = ac - b^2$ e a forma quadrática

$$z = Q(v) = ah^2 + 2bhk + ck^2 = \begin{pmatrix} h & k \end{pmatrix}\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix}, \text{ com } v = \begin{pmatrix} h \\ k \end{pmatrix} \text{ em } \mathbb{R}^2.$$

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

(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 na origem O.
- se $H = 0$, cilíndrico.

(iii) Se $a = c = 0$ e $b \neq 0$ (logo, $H < 0$), parabolóide hiperbólico.

(iv) Ainda, a função $z = Q(v)$ troca de sinal se e somente se $H < 0$.

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

Prova.

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

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

(ii) , (iii): Consequências triviais de (i).

(iv) O caso $a \neq 0$ segue por (i). Se $a = 0$, a função $Q(v) = 2bhk + ck^2$ troca sinal se e só se $b \neq 0$. Isto é, se e somente se $H = -b^2 < 0$■

Se M é a matriz simétrica 2×2 no Teorema 4, então Q é chamada de forma quadrática associada a M. Então, obtemos as fórmulas

$$Q(v) = v^T M v = \langle Mv, v\rangle\,,\ \text{onde } v = \begin{pmatrix} h \\ k \end{pmatrix},$$

com $\langle \cdot, \cdot \rangle$ indicando o produto escalar em $\mathbb{R}^2$.

Definições. Sejam $f \in C^2(\Omega)$ e P um ponto crítico de f. A matriz hessiana e o hessiano, ambos de f e em P, são

$$\mathcal{H}f(P) = \begin{pmatrix} f_{xx}(P) & f_{xy}(P) \\ f_{yx}(P) & f_{yy}(P) \end{pmatrix} \quad \text{e} \quad Hf(P) = \det \mathcal{H}f(P).$$

A forma quadrática associada a f, no ponto P, indicada Q_P, é a forma quadrática associada à matriz hessiana $\mathcal{H}f(P)$.

Proposição 5. *Sejam $f \in C^2(\Omega)$, um ponto P em Ω e $\vec{v} = \begin{pmatrix} h \\ k \end{pmatrix}$ em $\mathbb{R}^2$. Então,*

$$Q_P(v) = v^T \mathcal{H}f(P) v = \langle \mathcal{H}f(P)v, v\rangle = \frac{\partial^2 f}{\partial \vec{v}^2}(P).$$

Prova. Segue da Proposição 2■

Teorema 6 (Teste do Hessiano). *Seja $f \in C^2$, um ponto crítico P de f,*

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

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

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

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

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

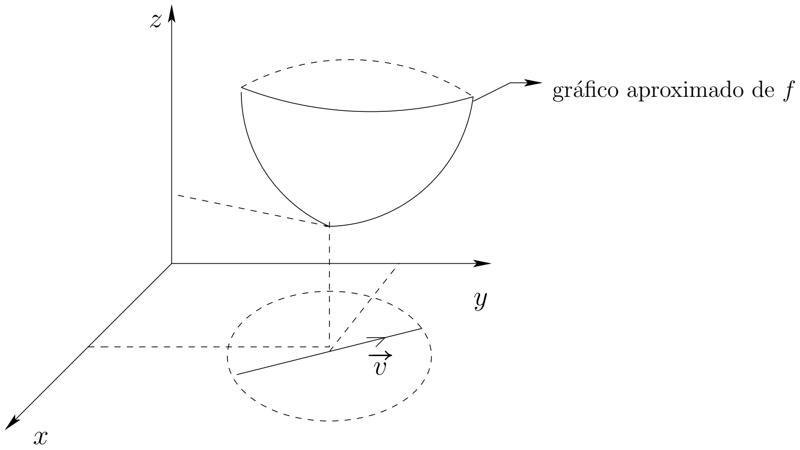

Figura 1: Caso em que $f_{xx}(P) > 0$ e $Hf(P) > 0$.

Prova.

(a) Como $f \in C^2$, temos $Hf = f_{xx}f_{yy} - f_{xy}^2 > 0$ e $f_{xx} > 0$ numa bola $B(P; r)$, se $r > 0$ e suficientemente pequeno. Pelo Corolário 3, dado $\overrightarrow{v} = h\overrightarrow{e_1} + k\overrightarrow{e_2}$, com $0 < |\overrightarrow{v}| < r$, existe $\overline{P}$ no segmento unindo P e $P + \overrightarrow{v}$ tal que

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

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

(c) Pelo Teorema 4(iv) segue que [vide Proposição 5] $Q_P(\overrightarrow{v}) = \frac{\partial^2 f}{\partial \vec{v}^2}(P)$ troca de sinal segundo $\overrightarrow{v}$. Logo, P não é extremante local. É ponto de sela.

(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).

Prova. Seja $\vec{v} = (h\,,k) \in \mathbb{R}^2$. O ponto $P_0 = (x_0, y_0)$ é crítico e

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

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

Para completar a solução: aplique o Teste do Hessiano■

Exemplo 4. Um único ponto crítico, minimo local mas não global. Seja

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

Solução. Temos $\vec{\nabla} f(x,y) = (2x - 3(1-x)^2y^2\,, 2(1-x)^3y)$ e $f(0,0) = 0$. Logo, $\vec{\nabla} f = (0,0)$ implica $(1-x)^3y = 0$ e assim, ou $y = 0$ [e $x = 0$], ou $x = 1$ [mas, $2x + 3(1-x)^3y^2 = 0$ não tem solução]. Logo, só $(0,0)$ é 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

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

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

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

$$\vec{\nabla} f = \big(5x^4 - 5\,,\,4y^3 - 32\big) = (0\,,0)\,,$$

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

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

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

Exemplo 6. Determine a distância entre as retas

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

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

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

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

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

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

$$D(\lambda\,,\mu) = (\lambda - 2\mu)^2 + (-4 + 6\lambda - 15\mu)^2 + (2\lambda + 2 - 6\mu)^2 =$$

$$= 41\lambda^2 + 265\mu^2 - 208\lambda\mu - 40\lambda + 96\mu + 20\quad .$$

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

$$\begin{cases} 41\lambda - 104\mu - 20 &= 0 \\ -104\lambda + 265\mu + 48 &= 0 \end{cases} \quad \Longrightarrow \quad (\lambda\,,\mu) = \Big(\frac{44}{7}\,,\frac{16}{7}\Big)\,.$$

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

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

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,

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

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

Então, a distância ente r e s é:

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

2 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):

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

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

(a) $O = (0,0)$ é o único ponto crítico.

(b) O teste da derivada segunda é inconclusivo para classificar O.

(c) f restrita a qualquer reta $y = mx$ por O tem em O um mínimo local.

(d) f não conserva sinal nas vizinhanças de O, que é ponto de sela.

Solução.

(a) É claro que $\vec{\nabla} f = (2x(4x^2 - 3y), -3x^2 + 2y) = O$ se e só se $(x, y) = O$.

(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. Valem 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. Note também que para $x \approx 0$, com $x \neq 0$, os pontos (x, mx) da reta pertencem à região em que $f > 0\blacksquare$

No Exemplo 8, identificamos vetores (e pontos) em $\mathbb{R}^2$ com matrizes-colunas em $M_{2\times 1}(\mathbb{R})$ e fixamos a base canônica de $\mathbb{R}^2$, $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ e $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Exemplo 8. Seja $f \in C^2(\Omega;\mathbb{R})$, no aberto Ω de $\mathbb{R}^2$, com ponto crítico O. Sejam $Hf = f_{xx}f_{yy} - f_{xy}^2$ e $\mathcal{H}f = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}$, para P arbitrário em Ω.

(a) Suponhamos $f_{xx}(P) \neq 0$ e as matrizes abaixo avaliadas em P. Verifique: definindo $M = \begin{pmatrix} 1 & 0 \\ -\frac{f_{xy}}{f_{xx}} & 1 \end{pmatrix}$ e $N^T = M^{-1} = \begin{pmatrix} 1 & 0 \\ \frac{f_{xy}}{f_{xx}} & 1 \end{pmatrix}$ obtemos

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

(b) Com a representação diagonal para $\mathcal{H}f$ e o Corolário 4 mostre que

- Se $f_{xx}(O) > 0$ e $\frac{Hf}{f_{xx}}(O) > 0$ então O é ponto de mínimo local estrito.
- Se $f_{xx}(O) < 0$ e $\frac{Hf}{f_{xx}}(O) < 0$ então O é ponto de máximo local estrito.

(c) Se $Hf(O) < 0$, então O é ponto de sela.

Dicas para (c). Sejam $a = f_{xx}(O)$, $b = f_{xy}(O)$, $c = f_{yy}(O)$ e $N = N(O)$.

- No caso $a \neq 0$, considere os vetores u e v, ambos em $\mathbb{R}^2$, definidos por $Nu = e_1$ e $Nv = e_2$, e mostre que $\frac{\partial^2 f}{\partial u^2}(O) = a$ e $\frac{\partial^2 f}{\partial v^2}(O) = \frac{Hf(O)}{a}$.
- O caso $b \neq 0$ segue do caso acima, considerando $g(x,y) = f(y,x)$.
- Se $a = c = 0$, para $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, avalie $\frac{\partial^2 f}{\partial u^2}(O)$ e $\frac{\partial^2 f}{\partial v^2}(O)$.

Outra sugestão para (c). Compute $\frac{\partial^2 f}{\partial v^2}(O) = v^T \mathcal{H}f(O) v$ nos casos:

- Se $a \neq 0$, para $v = \begin{pmatrix} -b \\ a \end{pmatrix}$.
- Se $c \neq 0$, para $v = \begin{pmatrix} c \\ -b \end{pmatrix}$.
- Se $a = c = 0$, para $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ e para $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

3 - O Hessiano em Três Variáveis

É simples generalizar esta seção para matrizes simétricas de ordem $n \in \mathbb{N}^*$.

Proposição 7. *Seja* $A \in M_3(\mathbb{R})$ *uma matriz simétrica,*

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

com menores principais *de ordens 1, 2 e 3,*

$$\Delta_1 = a_{11}\,, \quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} \quad e \quad \Delta_3 = \det A,$$

não nulos. Então, existe $M \in M_3(\mathbb{R})$ *tal que* $\det M = 1$ *e*

$$M^T A\ M \ = \begin{pmatrix} \Delta_1 & 0 & 0 \\ 0 & \frac{\Delta_2}{\Delta_1} & 0 \\ 0 & 0 & \frac{\Delta_3}{\Delta_2} \end{pmatrix} = \ D\ .$$

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 1 linha de A por $-\frac{a_{12}}{a_{11}}$ e somamos à 2 linha e, ainda, multiplicamos a 1 linha por $-\frac{a_{13}}{a_{11}}$ e somamos à 3 linha. Efetuamos tais operações, que comutam, multiplicando a matriz A à esquerda, respectivamente, pelas matrizes

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{12}}{a_{11}} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{e} \qquad M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{a_{13}}{a_{11}} & 0 & 1 \end{pmatrix}.$$

A seguir, na matriz obtida efetuamos operações nas colunas correspondendo as feitas nas linhas: multiplicamos a 1 coluna por $-\frac{a_{12}}{a_{11}}$ e somamos na 2 coluna e multiplicamos a 1 coluna por $-\frac{a_{13}}{a_{11}}$ e somamos na 3 coluna. Efetuamos tais operações multiplicando a matriz à direita por M_1^T e M_2^T.

Resumindo as quatro operações obtemos,

$$M_2 M_1 \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} M_1^T M_2^T = D_1 = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} \end{pmatrix},$$

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,

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

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 M_3, e completamos efetuando a operação correspondente nas colunas de D_1, representada (a operação) pela matriz M_3^T. Obtemos então,

$$M_3 D_1 M_3^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -\frac{a_{23}^{(1)}}{a_{22}^{(1)}} & 1 \end{pmatrix} \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22}^{(1)} & a_{23}^{(1)} \\ 0 & a_{23}^{(1)} & a_{33}^{(1)} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -\frac{a_{23}^{(1)}}{a_{22}^{(1)}} \\ 0 & 1 & 1 \end{pmatrix} =$$

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

Analogamente à etapa 1, devido às operações efetuadas, os três menores principais de D_2 são iguais a seus respectivos em A e D_1. Isto é,

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

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

Por fim, notemos que $M^T AM = D_2$ com $M^T = M_3 M_2 M_1$ e $\det M = 1$■

Corolário 8. Definindo $N^T = M^{-1}$, temos

$$A = N^T DN, \text{ com } \det N = 1.$$

Prova. Trivial■

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

Notação. Fixemos $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}\}$, a base canônica do espaço vetorial tridimensional $\mathbb{R}^3$. Seja O a origem no espaço cartesiano tri-dimensional $\mathbb{R}^3$.

Identifiquemos vetores (e pontos) em $\mathbb{R}^3$ com matrizes-colunas em $M_{3\times1}(\mathbb{R})$:

$$\overrightarrow{v} = v_1\overrightarrow{e_1} + v_2\overrightarrow{e_2} + v_3\overrightarrow{e_3} \equiv \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = v.$$

Observemos que $v^T = (v_1, v_2, v_3)$.

Com tal identificação, dada $A \in M_{3\times3}(\mathbb{R})$ definimos a aplicação linear

$$T : \mathbb{R}^3 \to \mathbb{R}^3,$$

por

$$T(\overrightarrow{v}) = Av, \text{ para } \overrightarrow{v} \text{ em } \mathbb{R}^3.$$

Identificando $T \equiv A$, escrevemos

$$A(\overrightarrow{v}) = Av.$$

Dada um função $f = f(x, y, z)$ de classe C^2, sua matriz hessiana é

$$\mathcal{H}f = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{pmatrix}.$$

Sejam x e y números reais.

- O sinal de x é $+1$ se $x > 0$, -1 se $x < 0$ e 0 se $x = 0$.
- Se $x > 0$ e $y < 0$, então x e y tem sinais opostos.
- Se $x \geq 0$, então x é positivo.
- Se $x \leq 0$, então x é negativo.
- Se $x > 0$, então x é estritamente positivo.
- Se $x < 0$, então x é estritamente negativo.

Teorema 9. *Seja $f \in C^2(\Omega)$, com Ω aberto, ponto crítico O. Sejam os números*

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

(a) *Supondo estes três números não nulos, valem as propriedades a seguir.*

  (i) *Se eles são (estritamente) positivos, O é ponto de mínimo local estrito.*

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

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

(b) *Se ocorre qualquer das condições abaixo em O, tal ponto é de sela.*

  (i) *Existem números na diagonal principal de $\mathcal{H}f$ com sinais opostos.*

  (ii) *Ou $H_1 < 0$ ou (por analogias) $f_{xx}f_{zz} - f_{xz}^2 < 0$ ou $f_{yy}f_{zz} - f_{yz}^2 < 0$.*

Prova.

(a) Por continuidade, f_{xx}, $H_1 f$ e Hf não se anulam numa bola $B(O;r)$, $r > 0$. Seja $\overrightarrow{v}$ com $0 < |\overrightarrow{v}| < r$. Analogamente ao $\mathbb{R}^2$, existe $\overline{P} \in B(O;r)$ tal que
$$f(O + \overrightarrow{v}) = f(O) + \frac{v^T \mathcal{H}f(\overline{P})v}{2}.$$
Pela Proposição 7 temos $\mathcal{H}f(\overline{P}) = N^T D N$, onde $N = N(\overline{P})$ e $D = D(\overline{P})$, sendo que N é uma matriz inversível e $D = (d_{ij})_{1\le i,j\le n}$ é uma matriz diagonal com $d_{11} = f_{xx}$, $d_{22} = \frac{H_1}{f_{xx}}$ e $d_{33} = \frac{H}{H_1}$. Logo,
$$f(O + \overrightarrow{v}) = f(O) + \frac{(Nv)^T D(Nv)}{2}.$$

  (i) Segue da identidade imediatamente acima.

  (ii) Segue do item (a)(i) aplicado à função $-f$.

  (iii) Sejam $D = D(O)$ e $N = N(O)$. A diagonal de D tem elementos com sinais opostos e $\mathcal{H}f(O) = N^T D N$. Como N é inversível, para cada $i = 1,2,3$ existe um vetor ϵ_i tal que $N\epsilon_i = e_i$. Assim,
$$\frac{\partial^2 f}{\partial \vec{\epsilon}_i}(O) = \epsilon_i^T \mathcal{H}f(O)\epsilon_i = (N\epsilon_i)^T D(N\epsilon_i) = e_i^T D e_i = d_{ii}.$$

Logo, O não é ponto de máximo nem mínimo local. É de sela.

(b) (i) Trivial, pois f_{xx}, f_{yy} e f_{zz} tem a forma $\partial^2 f/\partial v^2$ para $v = e_1, e_2, e_3$.

(ii) Definindo $F(x,y) = f(x,y,0)$, temos $\nabla F(0,0) = 0$ com hessiano $HF(0,0) = H_1 f(0,0,0) < 0$. Pelo caso bi-dimensional, $(0,0)$ é ponto de sela de F. Donde, $O = (0,0,0)$ é ponto de sela de f■

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, $\nabla f = \left(3(x^2-1)\,,3(y^2-1)\,,3(z^2-1)\right)$ e e 8 pontos críticos:

$$P = (\pm 1\,,\pm 1\,,\pm 1).$$

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. Donde,

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

Pelo Teste do Hessiano os pontos abaixo (observe a diagonal) são de sela:

$$(1,-1,-1)\,,(1,1,-1)\,,\ (1,-1,1)\,,\ (-1,1,1)\,,\ (-1,,1,-1)\,,(-1,-1,1).$$

O ponto $(1,1,1)$, com $H_1 f > 0$, $Hf > 0$ e $f_{xx} > 0$, é ponto de mínimo local. O ponto $(-1,-1,-1)$, com $H_1 f > 0$, $Hf < 0$ e $f_{xx} < 0$, é de máximo local■

Exemplo 10. Estude com relação a extremantes locais e pontos de sela,

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

Solução. Temos $\vec{\nabla} f = (x^4 - x^2, 4y^3 - 4y, 4z^3)$. Pontos criticos:

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

$$\text{Isto é, } \begin{cases} P_1 = (0,0,0)\ , & P_2 = (0,-1,0)\ , & P_3 = (0,1,0), \\ P_4 = (1,0,0)\ , & P_5 = (1,-1,0)\ , & P_6 = (1,1,0), \\ P_7 = (-1,0,0)\ , & P_8 = (-1,-1,0)\ , & P_9 = (-1,1,0). \end{cases}$$

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. Vejamos as matrizes em P_i, $1 \leq i \leq 9$ (com $z = 0$):

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

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

Como o (determinante) hessiano é zero, vejamos os sinais na diagonal de $\mathcal{H}f$. Os pontos críticos em que a diagonal de $\mathcal{H}f$ troca de sinal são de sela. Em 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, pela ordem, 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/negativa conforme x se aproxima de 0 pela direita/esquerda.

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

Resp: Pontos de mínimo local: P_5 e P_6. Pontos de sela: os outros $P_{i's}$■

4 - O Hessiano em Várias Variáveis

Seja Ω um aberto em $\mathbb{R}^n$ tal que Ω contém a origem O.

Teorema 10. *Sejam $f \in C^2(\Omega; \mathbb{R})$, com ponto crítico O, e a matriz hessiana*

$$\mathcal{H}f = \mathcal{H}f(O) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & \cdots & \frac{\partial^2 f}{\partial x_n x_n} \end{pmatrix} (O).$$

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 $\mathcal{H}f(O)$.*

(a) Supondo tais números não nulos temos,

(i) Se eles são (estritamente) positivos, O é ponto de mínimo local estrito.

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

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

(b) Se ocorre qualquer das condições abaixo em O, este é ponto de sela.

(i) Existem números na diagonal principal de $\mathcal{H}f$ com sinais opostos.

(ii) $f_{x_i x_i} f_{x_j x_j} - f_{x_i x_j}^2 < 0$ para algum par de índices distintos i e j. [Em particular, se (na diagonal) $f_{x_i x_i} = 0$ e $f_{x_j x_j} = 0$ porém $f_{x_i x_j} \neq 0$.]

Prova.

Argumentando analogamente ao Teorema 9, encontramos uma matriz $N \in M_n(\mathbb{R})$, com N um produto finito de matrizes que realizam a operação de multiplicar uma linha por um número e então somá-la em uma outra linha, tal que $\mathcal{H}f(P_0) = N^T D N$, $D = (d_{ij})_{1 \leq i,j \leq n}$ 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 $d_{22} = \frac{\Delta_2}{\Delta_1}, \ldots, d_{jj} = \frac{\Delta_j}{\Delta_{j-1}}$. O restante da prova é análogo ao Teorema 9∎

5 - Uma Aplicação em Álgebra Linear[1]

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

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

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

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

Um número real λ é um auto-valor de A se existe X em $\mathbb{R}^n$ tal que $AX = \lambda X$. Se a matriz $A \in M_{n\times n}(\mathbb{R})$ é simétrica, a forma quadrática associada a A é

$$Q : \mathbb{R}^n \to \mathbb{R}, \text{ com } Q(X) = X^T AX = \langle AX, X\rangle .$$

Lema 11. *Se Q é a forma quadrática associada à matriz simétrica A então*

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

Prova.

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

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

Logo,

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

[1] É bastante fácil provar via Álgebra Linear (mais o Teorema Fundamental da Álgebra e o conceito de produto interno complexo) que, dada A simétrica e real de ordem n, existe em $\mathbb{R}^n$ uma base ortonormal de auto-vetores de A. Os respectivos auto-valores são as raízes, com suas multiplicidades, do polinômio característico $p(\lambda) = \det(A - \lambda I)$, com I a matriz identidade de $M_n(\mathbb{R})$. Mas, não utilizaremos tal fato. Vide Apostol [1], pp. 136-137.

Corolário 12. *Sejam M, o máximo, e m, o mínimo, da forma quadrática Q_A sobre a esfera $\{X \in \mathbb{R}^n : |X| = 1\}$. Consideremos os aulto-valores de A.*

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

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

(c) Q é definida positiva se e só se os auto-valores são maiores que 0.

(d) Q é definida negativa se e só se os auto-valores são menores que 0.

Prova .

(a) Por continuidade Q assume máximo e mínimo no 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

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

e então, para tais pontos e pelo Lema 11 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

$$M = Q(X_M) = \langle AX_M, X_M \rangle = \lambda_M |X_M|^2 = \lambda_M.$$

Por analogia, $AX_m = \lambda_m X_m$, $|X_m| = 1$ e $m = Q(X_m) = \lambda_m$.

Ainda, se o número real λ é auto-valor de A, é claro que existe $\vec{v}$ unitário tal que $A(\vec{v}) = \lambda \vec{v}$. Neste caso temos $m \leq Q(\vec{v}) \leq M$ e, ainda, $Q(\vec{v}) = \langle A\vec{v}, \vec{v} \rangle = \langle \lambda \vec{v}, \vec{v} \rangle = \lambda |\vec{v}|^2 = \lambda$. Donde concluímos, $m \leq \lambda \leq M$.

(b) Se $\vec{v} \neq \vec{0}$ então,

$$m \leq Q\Big(\frac{\vec{v}}{|\vec{v}|}\Big) \leq M \implies m \leq \Big\langle A\Big(\frac{\vec{v}}{|\vec{v}|}\Big), \frac{\vec{v}}{|\vec{v}|} \Big\rangle \leq M \implies m \leq \frac{Q(\vec{v})}{|\vec{v}|^2} \leq M \ .$$

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

Lema 13. *Seja $f \in C^2(B(a;r))$, com $B(a;r) \subset \mathbb{R}^n$ e $r > 0$, onde a é um ponto crítico de f e $\vec{v} = (v_1, \ldots, v_n) \in \mathbb{R}^n$ tal que $|\vec{v}| < r$. Então,*

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

Prova.

Seja $\vec{\omega} = (\omega_1, \ldots, \omega_n) = \frac{\vec{v}}{|\vec{v}|}$. Aplicando à função $\varphi(t) = f(a+t\vec{\omega})$, t variando em $(-r,r)$, a Fórmula de Taylor com Resto Infinitesimal temos

$$(13.1) \qquad \varphi(t) = \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2}t^2 + t^2 E(0;t), \quad \lim_{t\to 0} E(0;t) = 0,$$

com, analogamente ao mostrado no plano,

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

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

Substituindo tais expressões em (13.1) obtemos

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

Por fim, basta substituir $t = |\vec{v}|$, $\omega = \frac{\vec{v}}{|\vec{v}|}$ e notar a identidade $|\vec{v}|^2\omega_i\omega_j = v_i v_j$ ∎

Teorema 14. *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*

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

São válidas as seguintes afirmações sobre os auto-valores de $\mathcal{H}f(a)$.

(a) Se todos são estritamente positivos, f tem um mínimo local em a.

(b) Se todos são estritamente negativos, f tem um máximo local em a.

(c) Se houver auto-valores com sinais opostos, então a é um ponto de sela.

Prova.

(a) Se $m > 0$ é o menor auto-valor de $\mathcal{H}f(a)$, pelo Corolário 12(b) temos $Q(v) \geq m|v|^2$, para todo v em $\mathbb{R}^n$. Ainda, pelo Lema 13 e sua notação, vemos que existe $\delta > 0$ tal que $|E(a;v)| < \frac{m}{4}$ se $0 < |v| < \delta$ e então,

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

(b) Segue de (a), trocando f por $-f$.

(c) Se λ é auto-valor de $\mathcal{H}f(a)$ e v é vetor unitário e $\mathcal{H}f(a)v = \lambda v$ então,

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

Logo, $\partial^2 f/\partial\vec{v}^2(a)$ troca de sinal e então a é um ponto de sela∎

Corolário 15. *Se A é simétrica e $Q(X) = \langle AX, X\rangle > 0$, $\forall X \neq 0$, então os menores principais $\Delta_1, \ldots, \Delta_n$, de A, são estritamente positivos.*

Prova. Por indução em n. O caso $n = 1$ é trivial.

Suponhamos a afirmação válida para n. Consideremos o caso $n + 1$. Então, $Q(e_1) = \Delta_1 > 0$. Analogamente à Proposição 7, existe N inversível tal que

$$A = N^TBN, \quad \text{sendo } B = \left(\begin{array}{c|c} \Delta_1 & 0 \\ \hline 0 & A_1 \end{array}\right)$$

simétrica [logo, A_1 é matriz simétrica $n \times n$] e com mesmos menores principais que A. É fácil ver que Q_{A_1} é definida positiva (pois N é inversível). Por hipótese de indução, os menores principais de A_1 são estritamente positivos. Como $\Delta_1 > 0$, os menores principais de B (e de A) também∎

REFERÊNCIAS

1. Apostol, T. M., *Análisis Matemático*, Editorial Reverté, 1960.
2. Apostol, T. M., *Cálculo*, Vol 2., Editora Reverté, 1999.
3. Guidorizzi, H. L., *Um Curso de Cálculo*, Vol 1 e 2, 5 ed., Ed. LTC.
4. Lima, E., *Curso de Análise* , Vol 2., IMPA, 2009.
5. Simmons, G. F., *Cálculo com Geometria Analítica*, Vol 2.
6. Hairer, E., and Wanner, G., *Analysis by Its History*, Springer, 1996.

Departamento de Matemática, Universidade de São Paulo,
e-mail: oliveira@ime.usp.br