Ano 2015

O HESSIANO EM VÁRIAS VARIÁVEIS

Professor Oswaldo Rio Branco de Oliveira

`http://www.ime.usp.br/~oliveira` e-mail: oliveira@ime.usp.br

1 - Preliminares

Lema 1. *Seja* $\varphi : [0,1] \to \mathbb{R}$ *uma função de classe* C^2*. Então, existe ao menos um* $\bar{t}$ *em* $(0,1)$ *tal que*

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

Prova.

Segue da fórmula de Taylor com resto de Lagrange♣

Neste texto, Ω é um aberto em $\mathbb{R}^n$ e O é a origem em $\mathbb{R}^n$.

Fixemos $\{e_1, \ldots, e_n\}$ a base canônica de $\mathbb{R}^n$.

A transposta de uma matriz $A = (a_{ij})$ em $M_{n\times m}(\mathbb{R})$, onde $1 \leq i \leq n$ e $1 \leq j \leq m$, é a matriz A^T em $M_{m\times n}(\mathbb{R})$, onde

$$A^T = (b_{kl}) \ \text{ e } \ b_{kl} = a_{lk} \text{ se } 1 \leq k \leq m \text{ e } 1 \leq l \leq n.$$

Dizemos que A em $M_n(\mathbb{R}) = M_{n\times n}(\mathbb{R})$ é simétrica se $A^T = A$.

Identificando $\mathbb{R}^n \equiv M_{n\times 1}(\mathbb{R})$, indicamos $v = v_1e_1 + \cdots + v_ne_n$ em $\mathbb{R}^n$, com os $v_{i's}$ em $\mathbb{R}$, por uma matriz-coluna v cuja transposta é $v^T = (v_1 \cdots v_n)$.

Dado x em $\mathbb{R}^n$, definimos a norma de x por $|x| = \sqrt{x_1^2 + \cdots + x_n^2}$.

Dados $x \in \mathbb{R}^n$ e $r > 0$, definimos a bola aberta de centro x e raio r por

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

Lema 2. *Seja* $f : B(O;r) \to \mathbb{R}$ *de classe* C^2, *onde* $r > 0$. *Seja* v *em* $B(O;r)$. *Definimos*

$$\varphi(t) = f(tv)\,, \quad \text{para } t \text{ em um intervalo aberto contendo } [0,1].$$

Temos,

(a) $\varphi'(t) = \sum_{i=1}^{n} f_{x_i}(tv)v_i = \frac{\partial f}{\partial v}(tv)$ *e* $\varphi''(t) = \sum_{i,j=1}^{n} f_{x_i x_j}(tv)v_i v_j = \frac{\partial^2 f}{\partial v^2}(tv)$.

(b) $f(v) = f(O) + \frac{\partial f}{\partial v}(O) + \frac{1}{2}\frac{\partial^2 f}{\partial v^2}(\overline{p})$,

para algum $\overline{p}$ *no segmento unindo* O *e* v.

Prova.

(a) Imediata, pela regra da cadeia.

(b) Basta computar $\varphi(1)$, $\varphi(0)$, $\varphi'(0)$ e aplicar o Lema 1 e o item (a)♣

Definições. Sejam $f : \Omega \to \mathbb{R}$ e p em Ω. Temos as seguintes classificações para o ponto p, com relação a f.

(a) Ponto de máximo [mínimo] local se $f(p) \geq f(x)$ [$f(p) \leq f(x)$] para todo x em $B(p;r) \subset \Omega$, para algum $r > 0$. Se tal desigualdade é estrita para todo x em $B(p;r) \setminus \{p\}$, então p é ponto de máximo [mínimo] local estrito.

(b) Extremante local se p é um ponto de máximo, ou mínimo, local.

(c) Ponto crítico, se $\nabla f(p) = 0$ e f é de classe C^1.

(d) Ponto de sela, se p é ponto crítico de f mas não um extremante local.

Todo extremante local de uma função de classe C^1 é um ponto crítico.

Sejam $f \in C^2(\Omega;\mathbb{R})$ e p um ponto crítico de f. A matriz hessiana e o (determinante) hessiano, ambos de f e em p, são

$$Hf(p) = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ \vdots & & & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{pmatrix}(p) \quad \text{e} \quad \det Hf(p).$$

Pelo Teorema de Schwarz, a matriz hessiana é simétrica em todos os pontos do aberto Ω.

Lema 3. *Seja* $f \in C^2(B(O;r))$, *onde* $r > 0$, *com ponto crítico* O.

(a) *Para quaisquer* p *em* $B(O;r)$ *e* v *em* $\mathbb{R}^n$, *temos*

$$v^T Hf(p)v = \frac{\partial^2 f}{\partial v^2}(p).$$

(b) *Dado* v *em* $B(O;r)$, *existe* $\overline{p}$ *no segmento unindo* O *e* $O+v$ *tal que*

$$f(v) = f(O) + \frac{v^T Hf(\overline{p})v}{2}.$$

Prova.

(a) Segue do Lema 2(a) e da definição de $Hf(p)$.

(b) Segue do Lema 2(b), da definição de ponto crítico e do item (a)♣

2 - O Hessiano em Várias Variáveis

Sejam A em $M_n(\mathbb{R})$ e $\lambda \neq 0$. A operação elementar que consiste em multiplicar por λ a i-ésima linha de A e então somá-la à j-ésima linha, $i \neq j$, mantendo todas as linhas de A que não a j-ésima, é representada pelo produto matricial

$$EA$$

com $E = (\epsilon_{kl})_{1\leq k,l\leq n}$ em $M_n(\mathbb{R})$ e definida por $\epsilon_{ji} = \lambda$, $\epsilon_{11} = \cdots = \epsilon_{nn} = 1$, e $\epsilon_{kl} = 0$ nos demais casos.

Destaquemos que os respectivos menores principais de EA e A são iguais e $\det E = 1$.

Analogamente, a operação elementar que consiste em multiplicar por λ a i-ésima coluna de A e então somá-la à j-ésima coluna, $i \neq j$, mantendo todas as colunas de A que não a j-ésima, é representada pelo produto matricial

$$AE^T.$$

Teorema 4. *Seja A em $M_n(\mathbb{R})$ e simétrica, com todos os menores principais*

$$\Delta_1 = a_{11}\ ,\ \ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix}, \ldots, \Delta_k = \det(a_{ij})_{1\leq i,j\leq k}, \ldots, \Delta_n = \det A$$

não nulos. Então, existe M em $M_n(\mathbb{R})$ satisfazendo

$$M\,A\,M^T \ = \begin{pmatrix} \Delta_1 & 0 & \cdots & 0 \\ 0 & \frac{\Delta_2}{\Delta_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\Delta_n}{\Delta_{n-1}} \end{pmatrix} = \ D \ \textit{e} \ \det M = 1.$$

Ainda, definindo $N^T = M^{-1}$ temos $A = N^T D N$ com $\det N = 1$.

Prova.

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

Suponhamos a afirmação válida para $n-1$ e consideremos o caso n.

Seja $i = 2, \ldots, n$ e E_i a matriz representando a operação elementar que multiplica a primeira linha de por

$$-\frac{a_{1i}}{a_{11}}$$

e a soma à i-ésima linha de A. A matriz E_i^T representa a operação elementar que multiplica a primeira coluna de A por

$$-\frac{a_{1i}}{a_{11}}$$

e a soma à i-ésima coluna de A. Temos $\det E_i = 1$. Seja $M_1 = E_n \cdots E_2$. Os menores principais de A e da matriz simétrica

$$B = M_1 A M_1^T = E_n \cdots E_2 A E_2^T \cdots E_n^T$$

são iguais e

$$M_1 A M_1^T = B = \left(\begin{array}{c|c} \Delta_1 & 0 \\ \hline 0 & B_1 \end{array}\right), \ \text{com} \ \det M_1 = \det(E_n \cdots E_2) = 1.$$

A matriz B_1, em $M_{n-1}(\mathbb{R})$, é simétrica e tem menores principais não nulos. Por hipótese de indução, existe N_1, em $M_{n-1}(\mathbb{R})$ e com $\det N_1 = 1$, satisfazendo

$$N_1 B_1 N_1^T = D_1,$$

onde D_1 é uma matriz diagonal com mesmos menores principais que B_1.

Seja

$$M_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & N_1 \end{array}\right)$$

em $M_n(\mathbb{R})$ e tal que $\det M_2 = 1$.

Segue então, utilizando $N_1 B_1 N_1^T = D_1$,

$$\begin{aligned} M_2 B M_2^T &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & N_1 \end{array}\right)\left(\begin{array}{c|c} \Delta_1 & 0 \\ \hline 0 & B_1 \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & N_1^T \end{array}\right) \\ &= \left(\begin{array}{c|c} \Delta_1 & 0 \\ \hline 0 & D_1 \end{array}\right) \\ &= D. \end{aligned}$$

Então, D é uma matriz diagonal. Como D_1 e B_1 tem mesmos menores principais, segue que D e B também. Portanto, D e A também.

Definindo $M = M_2 M_1$ temos

$$\begin{aligned} MAM^T &= M_2 M_1 A M_1^T M_2^T \\ &= M_2 B M_2^T \\ &= D, \end{aligned}$$

com $\det M = 1$.

Definindo $N^T = M^{-1}$, obtemos

$$N^T D N = A \clubsuit$$

Teorema 5. (Teste do Hessiano). *Sejam f em $C^2(\Omega;\mathbb{R})$, com ponto crítico O, e os menores principais $\Delta_1,\ldots,\Delta_n$ da matriz hessiana Hf nos pontos de Ω.*

(a) *Se tais menores, computados em O, são não nulos então O é um ponto de*

(i) *mínimo local estrito, se temos $\Delta_1 > 0$, $\Delta_2 > 0$, $\Delta_3 > 0$, etc.,*

(ii) *máximo local estrito, se temos $\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0$, etc.,*

(iii) *sela, se não ocorrem as possibilidades em (i) e (ii) para tais menores.*

(b) *O é um ponto de sela, se ocorre qualquer das condições abaixo no ponto O.*

(i) *A diagonal principal de Hf contém números com sinais opostos.*

(ii) *Temos $f_{x_ix_i}f_{x_jx_j} - f^2_{x_ix_j} < 0$ para algum par i,j, com $i \neq j$ [em particular, se na diagonal principal temos $f_{x_ix_i} = f_{x_jx_j} = 0$, porém $f_{x_ix_j} \neq 0$].*

Prova.

(a) Por continuidade, tais menores não se anulam em alguma bola $B(O;r)$, com $r > 0$. Seja $v \in B(O;r)$. Pelo Lema 3 e o Teorema 4 temos

$$f(v) = f(O) + \frac{v^T Hf(\overline{p})v}{2}, \text{ para algum } \overline{p} \in B(O;r),$$

com $Hf(\overline{p}) = N^T DN$, 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

$$d_{11} = \Delta_1 \text{ e } d_{ii} = \frac{\Delta_i}{\Delta_{i-1}} \text{ para } i = 2,\ldots,n.$$

Logo,

$$f(v) = f(O) + \frac{(Nv)^T D(Nv)}{2}.$$

(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 $Hf(O) = N^T DN$. Como N é inversível, para cada $i = 1, \dots, n$ existe um vetor ϵ_i tal que $N\epsilon_i = e_i$. Assim,

$$\begin{aligned}\frac{\partial^2 f}{\partial \epsilon_i^2}(O) &= \epsilon_i^T Hf(O)\epsilon_i \\ &= (N\epsilon_i)^T D(N\epsilon_i) \\ &= e_i^T De_i \\ &= d_{ii}.\end{aligned}$$

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

(b) (i) Trivial, pois

$$f_{x_i x_i} = \frac{\partial^2 f}{\partial e_i^2}, \text{ para } i = 1, \dots, n.$$

(ii) Reordenando variáveis, se preciso, podemos supor $i = 1$ e $j = 2$. Sejam $a = f_{x_1x_1}(O)$, $b = f_{x_1x_2}(O) = f_{x_2x_1}(O)$ e $c = f_{x_2x_2}(O)$. Por hipótese temos $h = ac - b^2 < 0$.

Mostremos que

$$\frac{\partial^2 f}{\partial v^2}(O) \text{ muda de sinal segundo } v.$$

- Caso $a \neq 0$. Dado $v^T = (-b, a, 0, \dots, 0)$, os sinais de

$$\frac{\partial^2 f}{\partial v^2}(O) = v^T Hf(O)v = a(ac - b^2) = ah \quad \text{e} \quad \frac{\partial^2 f}{\partial e_1^2}(O) = a$$

  são opostos.

- Caso $c \neq 0$. Segue do caso $a \neq 0$, reordenando variáveis.

- Caso $a = c = 0$. Logo, $b \neq 0$. Sejam u e v em $\mathbb{R}^n$, dados por $u^T = (1, 1, 0, \dots, 0)$ e $v^T = (1, -1, 0, \dots, 0)$. Os sinais

$$\frac{\partial^2 f}{\partial u^2}(O) = v^T Hf(O)v = 2b \quad \text{e} \quad \frac{\partial^2 f}{\partial v^2}(O) = -2b$$

  são opostos♣

REFERÊNCIAS

1. Apostol, T. M., *Análisis Matemático*, Editorial Reverté, 1960.
2. Guidorizzi, H. L., *Um Curso de Cálculo*, Vol 1 e 2, 5 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.

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