MULTIPLICADORES DE LAGRANGE EM VÁRIAS VARIÁVEIS

Professor Oswaldo Rio Branco de Oliveira

http://www.ime.usp.br/~oliveira (2015-2019) oliveira@ime.usp.br

Definição. Seja M uma matriz em $M_{m\times n}(\mathbb{R})$ e $k \in \mathbb{N}$ tal que $0 < k \leq m$ e $0 < k \leq n$. Um menor de ordem k de M é o determinante de uma matriz quadrada de ordem k obtida pela remoção de $m-k$ linhas e $n-k$ colunas da matriz M.

Definição. Seja M em $M_{m\times n}(\mathbb{R})$. O posto coluna/linha de M é o número máximo de colunas/linhas linearmente independentes de M.

Lema. *Seja M uma matriz em $M_{m\times n}(\mathbb{R})$, com $m \leq n$ e com suas m linhas LI. Então, a matriz M tem um menor de ordem m não nulo (e M tem posto m).*

Prova.

- Escrevamos
$$M_1 = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}.$$
A primeira linha de M_1 contém um elemento $a_{1j} = \lambda_{1j} \neq 0$. Multiplicando a primeira linha por um número conveniente e somando-a à segunda linha, temos uma nova matriz cuja segunda linha apresenta o número 0 na j-ésima coluna. As m linhas desta nova matriz são também L. I. e todos os seus menores (determinantes) de ordem m são iguais aos correspondentes menores (determinantes) originais. Iterando, multiplicamos a primeira linha por sucessivos números convenientes e a somamos ordenadamente às demais linhas e obtemos uma matriz M_2 satisfazendo as condições: *o primeiro elemento na j-ésima coluna é $\lambda_{1j} = a_{1j} \neq 0$, os demais elementos na j-ésima coluna são nulos, suas linhas são L. I. e todos os seus menores de ordem m são iguais aos respectivos menores da matriz M_1.*

  Escrevamos,
$$M_2 = \begin{bmatrix} b_{11} & \cdots & \lambda_{1j} & \cdots & b_{1n} \\ b_{21} & \cdots & 0 & \cdots & b_{2n} \\ \vdots & & \vdots & & \vdots \\ b_{m1} & \cdots & 0 & \cdots & b_{mn} \end{bmatrix}.$$
A segunda linha de M_2 contém um elemento $b_{2k} = \lambda_{2k} \neq 0$ com $k \neq j$.

Então, analogamente ao feito acima, obtemos a matriz M_3 (vide abaixo) satisfazendo: *o segundo elemento em sua k-ésima coluna é $\lambda_{2k} = b_{2k} \neq 0$, os demais elementos na k-ésima coluna são nulos (a operação de multiplicar a segunda linha de M_2 por uma constante e então somá-la à primeira linha não muda o elemento λ_{1j} na j-ésima coluna da primeira linha de M_2), a j-ésima coluna de M_3 é igual a j-ésima coluna de M_2, suas linhas são LI e, ainda, todos os seus menores de ordem m são iguais aos respectivos menores de M_2.* Escrevamos,

$$M_3 = \begin{bmatrix} c_{11} & \cdots & 0 & \cdots & \lambda_{1j} & \cdots & c_{1n} \\ c_{21} & & \lambda_{2k} & & 0 & & c_{2n} \\ c_{31} & & 0 & & 0 & & c_{3n} \\ \vdots & & \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & 0 & \cdots & 0 & \cdots & c_{mn} \end{bmatrix}.$$

Por fim, iterando tal processo encontramos as m colunas desejadas♣

Seja $f : \Omega \to \mathbb{R}$, onde Ω é aberto em $\mathbb{R}^n$. Um ponto P_0 em X é dito:

- Crítico, ou estacionário, da função f se

$$\nabla f(P_0) = 0.$$

- Um máximo [mínimo] local da função f se existe um raio $r > 0$, tal que temos

$$f(P_0) \geq f(P) \ [f(P_0) \leq f(P)], \text{ para todo } P \text{ na bola aberta } B(P_0; r) \subset \Omega.$$

- Extremante local de f, se P_0 é ponto de máximo [mínimo] local de f.

- Extremante local de f em um subconjunto S de Ω, se P_0 está em S e P_0 é ponto de máximo [mínimo] de f restrita a $B(P; r) \cap S$, para algum $r > 0$.

O resultado abaixo é local, por praticidade o enunciamos em todo o espaço. Escrevamos

$$\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m = \{(x,y) : x \in \mathbb{R}^n \text{ e } y \in \mathbb{R}^m\}.$$

Teorema. *Sejam $F : \mathbb{R}^{n+m} \to \mathbb{R}$ diferenciável e $g = (g_1, \ldots, g_m)$ em $C^1(\mathbb{R}^{n+m}; \mathbb{R}^m)$. Suponhamos que o campo escalar F tem um extremante local no ponto P pertencente à superfície de nível $S = g^{-1}(0, \ldots, 0)$ e que $\{\nabla g_1, \ldots, \nabla g_m\}$ é L. I. em todo ponto de alguma vizinhança de P. Então, temos*

$$\nabla F(P) = \lambda_1 \nabla g_1(P) + \cdots + \lambda_m \nabla g_m(P),$$

para certos $\lambda_1, \ldots, \lambda_m$ em $\mathbb{R}$.

Prova.

- Idéia/Interpretação. O vetor $\nabla F(P)$ é ortogonal a S pois P é um extremante de F. O vetor $\nabla g_1(P)$ é ortogonal a S, pois $S \subset g^{-1}(0)$. Portanto, $\{\nabla F(p), \nabla g_1(P), \ldots, \nabla g_m(P)\}$ é ortogonal a S. A superfície S tem dimensão n (pelo teorema da função implícita) e o espaço tangente a S, no ponto P, tem dimensão n. Logo, o espaço ortogonal a S tem dimensão m. Portanto, $\{\nabla F(p), \nabla g_1(P), \ldots, \nabla g_m(P)\}$ é LD. Donde $\nabla F(P)$ é uma combinação linear dos vetores linearmente independentes $\nabla g_1(P), \ldots, \nabla g_m(P)$.

- A prova propriamente dita. Suponhamos $P = (0,0)$.

  A matriz $Jg(0,0)$ tem posto m e reordenando as variáveis $x_1, \ldots, x_n, y_1, \ldots, y_m$ podemos supor sem perda de generalidade que a matriz quadrada

  $$\left[\frac{\partial g}{\partial y}(0,0)\right] = \left[\frac{\partial g_i}{\partial y_j}(x,y)(0,0)\right]_{\substack{1 \le i \le m \\ 1 \le j \le m}}$$

  é inversível.

  O Teorema da Função Implícita garante existir uma função $\varphi : \Omega \to \mathbb{R}^m$ de classe C^1 e satisfazendo

  $$\begin{cases} g\big(x, \varphi(x)\big) = 0, \text{ para todo } x \text{ no aberto } \Omega \text{ em } \mathbb{R}^n, \\ \text{e} \\ \varphi(0) = 0 \in \mathbb{R}^m. \end{cases}$$

Seja $\{e_1, \ldots, e_n\}$ a base usual de $\mathbb{R}^n$. As curvas

$$\gamma_i(t) = \big(te_i; \varphi(te_i)\big), \text{ com } i = 1, \ldots, n,$$

contidas em S e definidas para $|t| \neq 0$ pequeno o suficiente, satisfazem:

$$\gamma_i(0) = P, \ \gamma_i'(0) \neq 0 \text{ e } \{\gamma_1'(0), \ldots, \gamma_n'(0)\} \text{ é linearmente independente.}$$

Os vetores $\gamma_i'(0)$ e $\nabla g_j(0)$ são ortogonais, para $i = 1, \ldots, n$ e $j = 1, \ldots, m$. É então trivial ver que

$$\{\gamma_1'(0), \ldots, \gamma_n'(0), \nabla g_1(0), \ldots, \nabla g_m(0)\}$$

é uma base de $\mathbb{R}^{n+m}$. Ainda, o vetor $\nabla F(0)$ é ortogonal a cada $\gamma_i'(0)$. Logo,

$$\nabla F(0) \text{ é gerado por } \nabla g_1(0), \ldots, \nabla g_m(0) \clubsuit$$

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

Com o Método dos Multiplicadores de Lagrange utilizamos os conceitos topológicos e o teorema abaixo.

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

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

Teorema (Weierstrass). *Seja $f : K \to \mathbb{R}$ contínua no compacto K de $\mathbb{R}^n$. Então, f assume valor máximo absoluto e valor mínimo absoluto em K.*

Oswaldo Rio Branco de Oliveira

UMA APLICAÇÃO DE MULTIPLICADORES DE LAGRANGE: EM ÁLGEBRA LINEAR

Fixemos a base canônica de $\mathbb{R}^n$ e indiquemos o produto interno de dois vetores em $\mathbb{R}^n$ por $\langle \cdot, \cdot \rangle$.

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

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

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

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

Um real λ é autovalor de A se existe v tal que $Av = \lambda v$. Suponhamos que A é simétrica. A forma quadrática associada a A é a aplicação

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

Lema. *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 desta forma, como $a_{ij} = a_{j\,i}$, encontramos

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

Logo,

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

Corolário. *Sejam M, o máximo, e m, o mínimo, da forma quadrática Q associada à matriz simétrica A sobre a esfera unitária*

$$S^{n-1} = \{X \in \mathbb{R}^n : |X| = 1\}.$$

São válidas as seguintes propriedades.

(a) *Os números M e m são, respectivamente, o maior e o menor autovalores (reais) de A.*

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

(c) *Q é definida positiva se e só se os autovalores de A são maiores que* 0.

(d) *Q é definida negativa se e só se os autovalores de A são menores que* 0.

Prova.[1]

(a) Por continuidade, Q assume máximo e mínimo na esfera compacta S^{n-1}.

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

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

e então, para tais pontos e pelo lema acima temos

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

Sejam X_M e λ_M, com $|X_M| = 1$, $Q(X_M) = M$ e $AX_M = \lambda_M X_M$. Então,

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

Analogamente,

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

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

Ainda, se λ é um autovalor real de A, existe v, com $|v| = 1$, tal que $A(v) = \lambda v$. Logo, $m \leq Q(v) \leq M$ e, ainda, $Q(v) = \langle Av, v \rangle = \langle \lambda v, v \rangle = \lambda |v|^2 = \lambda$. Donde seguem as desigualdades

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

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

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

(c) e (d) Seguem trivialmente de (b)♣

REFERÊNCIAS

1. Apostol, T. M., *Análisis Matemático*, Editorial Reverté, 1960.
2. Fitzpatrick, P., *Advanced Calculus*, 2nd ed., AMS, 2006.
3. Lima, E., *Curso de Análise* , Vol 2., IMPA, 2009.
4. Sallum, E. M., Murakami, L. S. I. e Silva, J. P., *Cálculo Diferencial Geométrico no* $\mathbb{R}^n$, Instituto de Matemática e Estatística da USP, 2009.

Departamento de Matemática, Universidade de São Paulo
São Paulo, SP - Brasil
oliveira@ime.usp.br
`http://www.ime.usp.br/~oliveira`