

Análise

— Extremos livres e condicionados ————— 2010'11 ————

Extremos livres

A determinação dos extremos locais de uma função real de n variáveis reais, $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ definida num aberto D , passa pela resolução de dois problemas:

Problema 1

Determinação dos *pontos críticos* da função, ou seja, das soluções do sistema

$$\left\{ \begin{array}{l} \frac{\partial f}{\partial x_1}(x_1, x_2, \dots, x_n) = 0 \\ \frac{\partial f}{\partial x_2}(x_1, x_2, \dots, x_n) = 0 \\ \dots \quad \dots \quad \dots \\ \frac{\partial f}{\partial x_n}(x_1, x_2, \dots, x_n) = 0 \end{array} \right. , \quad (1)$$

que constituem os únicos candidatos a pontos extremantes da função.

Problema 2

Para cada ponto crítico A , encontrado anteriormente como solução do sistema (1), estuda-se o sinal de $f(X) - f(A)$, com X um ponto arbitrário de alguma bola $B(A, \varepsilon)$ centrada em A .

1. Em casos especiais, através de uma análise local da função em torno do ponto A , é “fácil” concluir se o sinal daquela diferença é constante em alguma $B(A, \varepsilon)$. Assim:

- (a) se existir uma bola aberta $B(A, \varepsilon)$ tal que $f(X) - f(A) \geq 0$, $\forall X \in B(A, \varepsilon)$, então A é um minimizante local de f ;

- (b) se existir uma bola aberta $B(A, \varepsilon)$ tal que $f(X) - f(A) \leq 0, \forall X \in B(A, \varepsilon)$, então A é um maximizante local de f ;
- (c) se em qualquer bola $B(A, \varepsilon)$ existirem X, Y tais que $f(X) - f(A) \geq 0$ e $f(Y) - f(A) \leq 0$, então A não é extremante de f (trata-se de um *ponto de sela*).

2. Quando esta análise local não é conclusiva, procede-se da seguinte forma:

- Determinam-se as sucessivas derivadas direccionais de f em A , de ordem cada vez maior, até encontrar a primeira que não se anule indenticamente no ponto A . Suponhamos que a ordem desta derivada é p , isto é, que se tem

$$\frac{\partial^p f}{\partial \vec{v}^p}(A) \neq 0 \text{ para algum } \vec{v} \in \mathbb{R}^n \setminus \{0\},$$

$$\frac{\partial^k f}{\partial \vec{v}^k}(A) = 0 \text{ para todo } \vec{v} \in \mathbb{R}^n \setminus \{0\} \text{ e todo } k \leq p-1.$$

Note que, como se assume que f é de classe C^l com $l > p$, encontrar a ordem da primeira derivada direccional não nula, equivale a encontrar a ordem da primeira derivada parcial não nula (ver adiante).

- De seguida, recorre-se ao seguinte resultado (f é de classe C^p em D):

Teorema [Caso geral]

- (a) Se p é ímpar então A não é ponto extremante de f .
- (b) Se p é par e
 - i. $\frac{\partial^p f}{\partial \vec{v}^p}(A) < 0, \forall \vec{v} \in \mathbb{R}^n \setminus \{0\}$, então A é um maximizante local de f ;
 - ii. $\frac{\partial^p f}{\partial \vec{v}^p}(A) > 0, \forall \vec{v} \in \mathbb{R}^n \setminus \{0\}$, então A é um minimizante local de f ;
 - iii. $\exists \vec{u}, \vec{w} \in \mathbb{R}^n \setminus \{0\}$ tais que $\frac{\partial^p f}{\partial \vec{u}^p}(A) > 0, \frac{\partial^p f}{\partial \vec{w}^p}(A) < 0$, então A não é extremante de f ;

- iv. $\exists \vec{u} \in \mathbb{R}^n \setminus \{0\}$ tal que $\frac{\partial^p f}{\partial \vec{u}^p}(A) = 0$, e para $\vec{v} \neq \vec{u}$ tem-se $\frac{\partial^p f}{\partial \vec{v}^p}(A) \leq 0$, então A não é minimizante de f (trata-se de um caso inconclusivo em que A é um possível maximizante);
- v. $\exists \vec{u} \in \mathbb{R}^n \setminus \{0\}$ tal que $\frac{\partial^p f}{\partial \vec{u}^p}(A) = 0$, e para $\vec{v} \neq \vec{u}$ tem-se $\frac{\partial^p f}{\partial \vec{v}^p}(A) \geq 0$, então A não é maximizante de f (trata-se de um caso inconclusivo em que A é um possível minimizante). ■

- No caso especial em que $n = 2$, $p = 2$ e que f é de classe C^2 em U , é possível usar o seguinte critério que se baseia no cálculo do *determinante Hessiano*

$$\Delta_f(A) = \begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(A) & \frac{\partial^2 f}{\partial x \partial y}(A) \\ \frac{\partial^2 f}{\partial y \partial x}(A) & \frac{\partial^2 f}{\partial y^2}(A) \end{vmatrix} = \frac{\partial^2 f}{\partial x^2}(a) \frac{\partial^2 f}{\partial y^2}(A) - \left[\frac{\partial^2 f}{\partial x \partial y}(A) \right]^2.$$

Proposição [Critério do Hessiano]

(a) Se $\Delta_f(A) > 0$ então a é ponto extremante de f .

- É um maximizante se $\frac{\partial^2 f}{\partial x^2}(A) < 0$.
- É um minimizante se $\frac{\partial^2 f}{\partial x^2}(A) > 0$.

(b) Se $\Delta_f(A) < 0$ então a não é ponto extremante de f .

(c) Se $\Delta_f(A) = 0$ então o caso é inconclusivo. ■

- O caso inconclusivo referido nas alíneas anteriores, (b) iv.,v. do Teorema, requer um estudo mais aprofundado da função nas direcções ditas *singulares*, definidas pelos vectores $\vec{u}$ que anulam a derivada direcciona de ordem p da função f no ponto A . Neste caso, calculam-se

as derivadas direccionais de ordem $k > p$ em A , apenas segundo as direcções singulares, até encontrar a primeira que não se anule identicamente no ponto A . Suponhamos que esta derivada é de ordem m ($m > p$).

(a) Se m é ímpar então A não é ponto extremante de f .

(b) Se m é par e

- i. Se as derivadas $\frac{\partial^m f}{\partial \vec{u}^m}(A)$ e $\frac{\partial^p f}{\partial \vec{v}^p}(A)$, para $\vec{u}$ definindo uma direcção sigular e $\vec{v}$ não definindo uma direcção sigular, possuem sinais contrários, então A não é extremante.
- ii. Se as derivadas $\frac{\partial^m f}{\partial \vec{u}^m}(A)$ e $\frac{\partial^p f}{\partial \vec{v}^p}(A)$ possuem o mesmo sinal, então a dúvida persiste.

A dúvida só será desfeita (pela negativa) se for possível encontrar uma curva $\mathcal{C}$ no domínio de f , e passando por A , relativamente à qual a restrição $f|_{\mathcal{C}}$ não possua um extremo em A ou possua aí um extremo de natureza diferente daquela que se pretende confirmar.

Para o calculo das derivadas direccionais de f segundo um dado vector $\vec{v}$ é importante ter em atenção o seguinte:

- Sendo f de classe C^p , a derivada direccional de f de ordem k , $k = 1, 2, \dots, p$, pode ser calculada com base nas derivadas parciais de f de ordem k . Assim, para $\vec{v} = (v_1, v_2, \dots, v_n)$,

$$\frac{\partial^k f}{\partial \vec{v}^k}(X) = \sum_{i_1, \dots, i_k=1}^n v_{i_1} \cdots v_{i_k} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(X),$$

mas, pelo teorema de Schwarz e usando um raciocínio de combinatória, conclui-se que no somatório anterior surgem $\frac{k!}{k_1! \cdots k_n!}$ parcelas, onde k_i representa o número de ocorrências do índice i , iguais a

$$v_1^{k_1} \cdots v_n^{k_n} \frac{\partial^k f}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}(X).$$

Consequentemente

$$\frac{\partial^k f}{\partial \vec{v}^k}(X) = \sum_{\substack{k_1, \dots, k_n = 0 \\ k_1 + \dots + k_n = k}}^k (v_1)^{k_1} \dots (v_n)^{k_n} \frac{k!}{k_1! \dots k_n!} \frac{\partial^k f}{\partial x_1^{k_1} \dots \partial x_n^{k_n}}(X), \quad (2)$$

onde os escalares $\frac{k!}{k_1! \dots k_n!}$ são designados por coeficientes multinomiais.

- No caso particular em que $n = 2$, f é uma função de duas variáveis e pondo $X = (x, y)$ e $\vec{v} = (v_1, v_2)$, a expressão (2) assume a forma

$$\frac{\partial^k f}{\partial \vec{v}^k}(x, y) = \sum_{\ell=0}^k (v_1)^\ell (v_2)^{k-\ell} \frac{k!}{\ell!(k-\ell)!} \frac{\partial^k f}{\partial x^\ell \partial y^{k-\ell}}(x, y), \quad (3)$$

onde os coeficientes $\frac{k!}{\ell!(k-\ell)!} = \binom{k}{\ell}$ se designam, em particular, por coeficientes binomiais e podem ser extraídos do triângulo de Pascal,

			1					(ordem $k = 1$)
			1		1			(ordem $k = 2$)
		1		2		1		(ordem $k = 3$)
	1		3		3		1	(ordem $k = 4$)
	1	4		6		4	1	(ordem $k = 5$)
1	5	10		10		5	1	
...	...	...	...	...	...	...	...	

resultando

$$\begin{aligned} \frac{\partial f}{\partial \vec{v}}(X) &= v_1 \frac{\partial f}{\partial x}(X) + v_2 \frac{\partial f}{\partial y}(X) \\ \frac{\partial^2 f}{\partial \vec{v}^2}(X) &= (v_1)^2 \frac{\partial^2 f}{\partial x^2}(X) + 2v_1 v_2 \frac{\partial^2 f}{\partial x \partial y}(X) + (v_2)^2 \frac{\partial^2 f}{\partial y^2}(X) \\ \frac{\partial^3 f}{\partial \vec{v}^3}(X) &= (v_1)^3 \frac{\partial^3 f}{\partial x^3}(X) + 3(v_1)^2 v_2 \frac{\partial^3 f}{\partial x^2 \partial y}(X) + 3v_1 (v_2)^2 \frac{\partial^3 f}{\partial x \partial y^2}(X) \\ &\quad + (v_2)^3 \frac{\partial^3 f}{\partial y^3}(X) \end{aligned}$$

$$\begin{aligned}
\frac{\partial^4 f}{\partial \vec{v}^4}(X) &= (v_1)^4 \frac{\partial^4 f}{\partial x^4}(X) + 4(v_1)^3 v_2 \frac{\partial^4 f}{\partial x^3 \partial y}(X) + 6(v_1)^2 (v_2)^2 \frac{\partial^4 f}{\partial x^2 \partial y^2}(x) \\
&\quad + 4v_1 (v_2)^3 \frac{\partial^4 f}{\partial x \partial y^3}(X) + (v_2)^4 \frac{\partial^4 f}{\partial y^4}(X) \\
&\quad \dots \qquad \dots \qquad \dots \qquad \dots
\end{aligned}$$

Extremos condicionados

O problema da determinação dos extremos condicionados de uma função $f: D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, de classe C^1 num aberto D , consiste na procura dos extremos de f quando sujeita a certas restrições nas variáveis independentes, a que chamamos *equações de ligação* ou *vínculos*, e que indicamos por

$$g_1(X) = 0, \quad g_2(X) = 0, \quad \dots, \quad g_m(X) = 0, \quad (4)$$

onde $g_1, g_2, \dots, g_m$ são funções de classe C^1 em D .

1. Em alguns casos, é possível usar as equações de ligação (4) para explicitar m das variáveis x_i em termos das restantes e transformar o problema de extremos condicionados num problema de extremos “livres”, como os que tratamos anteriormente.
2. Um método que não requer a resolução das equações de ligação é o chamado *método dos multiplicadores de Lagrange*.

Caso 1 A função f está sujeita a uma única condição de ligação, digamos

$$g(X) = 0.$$

Procedemos da seguinte forma.

- (a) Encontramos os pontos $X \in \mathbb{R}^n$ verificando $g(X) = 0$ tais que

$$\nabla g(X) = 0.$$

- (b) Construimos a função de Lagrange, $H : D \times \mathbb{R} \longrightarrow \mathbb{R}$, de $n + 1$ variáveis, definida por

$$H(x_1, \dots, x_n, \lambda) = f(x_1, \dots, x_n) - \lambda g(x_1, \dots, x_n) \quad (5)$$

onde λ é designado por *multiplicador de Lagrange* e funciona como um parâmetro a especificar.

- (c) Procuramos os candidatos a extremos de f condicionados por $g(X) = 0$, que são precisamente os pontos $(x_1, \dots, x_n)$ para os

quais existe $\lambda \in \mathbb{R}$ tal que $(x_1, \dots, x_n, \lambda)$ é um ponto crítico de H e, portanto, uma solução do sistema

$$\begin{cases} \frac{\partial H}{\partial x_1}(x_1, \dots, x_n, \lambda) = 0 \\ \vdots \\ \frac{\partial H}{\partial x_n}(x_1, \dots, x_n, \lambda) = 0 \\ \frac{\partial H}{\partial \lambda}(x_1, \dots, x_n, \lambda) = 0 \end{cases} \iff \begin{cases} \frac{\partial f}{\partial x_1}(x_1, \dots, x_n) = \lambda \frac{\partial g}{\partial x_1}(x_1, \dots, x_n) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_1, \dots, x_n) = \lambda \frac{\partial g}{\partial x_n}(x_1, \dots, x_n) \\ g(x_1, \dots, x_n) = 0 \end{cases} \quad (6)$$

(d) Para cada ponto $(x_1, \dots, x_n)$ encontrado em (a) e cada ponto $(x_1, \dots, x_n, \lambda)$ encontrado em (c), averiguamos se, de facto, se trata de um extremo condicionado de f . Para tal, há dois resultados fundamentais a ter em conta:

- **primeiro**, o Teorema de Weierstrass, que estabelece que uma função contínua num compacto (fechado e limitado) atinge aí os seus extremos,

que é aplicável apenas quando a condição de ligação define um conjunto compacto;

- **segundo**, a proposição que passamos a enunciar e que se baseia no estudo da função H^* definida por

$$H^*(x_1, \dots, x_n) = f(x_1, \dots, x_n) - \lambda g(x_1, \dots, x_n), \quad (7)$$

onde λ assume um valor concreto;

Proposição

Se $(a_1, \dots, a_n, \lambda)$ é um ponto estacionário de H e $(a_1, \dots, a_n)$ é um extremo da correspondente função H^* , então $(a_1, \dots, a_n)$

é também um extremo de f condicionado por $g(X) = 0$ com a mesma natureza da do extremo de H^* .

Esta Proposição é aplicável apenas quando o ponto estacionário $(a_1, \dots, a_n, \lambda)$ da função H de Lagrange é tal que $(a_1, \dots, a_n)$ é um extremo de H^* . Quando $(a_1, \dots, a_n)$ é um ponto de sela de H^* , nada se pode concluir sobre a natureza do candidato $(a_1, \dots, a_n)$ a extremo condicionado.

Caso 2 A função f está sujeita a m condições de ligação, digamos

$$g_1(X) = 0, \dots, g_m(X) = 0.$$

Procedemos de forma perfeitamente semelhante. Primeiro procuramos os pontos $X = (x_1, \dots, x_n)$ verificando as condições de ligação $g_1(X) = 0, \dots, g_m(X) = 0$ tais que

$\nabla g_1(X), \dots, \nabla g_m(X)$ são vectores linearmente dependentes.

Depois, tudo se baseia na construção de uma função de Lagrange que agora envolve as m funções g_i ,

$$\begin{aligned} H(x_1, \dots, x_n, \lambda) &= f(x_1, \dots, x_n) - \lambda_1 g_1(x_1, \dots, x_n) \\ &\quad - \dots - \lambda_m g_m(x_1, \dots, x_n), \end{aligned} \tag{8}$$

com $\lambda_1, \dots, \lambda_m$ *multiplicadores de Lagrange*.