ANÁLISE
Cap. 3 – Cálculo diferencial em $\mathbb{R}^n$

Dep. Matemática UMinho

Março 2020

3. Cálculo diferencial em $\mathbb{R}^n$

Derivada direcional

Sejam $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $a \in D$ e v um vetor de $\mathbb{R}^n$ unitário (isto é, tal que $\|v\| = 1$).

- A derivada direcional de f no ponto a na direção de v define-se como
$$\lim_{t \longrightarrow 0} \frac{f(a + tv) - f(a)}{t}$$
caso este limite exista e denota-se por $\dfrac{\partial f}{\partial v}(a)$.

Observações

1. A derivada direcional na direção do vetor v

$$\frac{\partial f}{\partial v}(a) = \lim_{t \longrightarrow 0} \frac{f(a+tv) - f(a)}{t} \qquad \text{é um número real}$$

   - $t \in \mathbb{R}$;
   - $a + tv = (a_1, \ldots, a_n) + t(v_1, \ldots, v_n) \in \mathbb{R}^n$;
   - $f(a+tv),\ f(a) \in \mathbb{R} \qquad \text{logo} \qquad f(a+tv) - f(a) \in \mathbb{R}$;

2. o símbolo ∂ em $\dfrac{\partial f}{\partial v}(a)$ lê-se “dê curvo”.

3. $\dfrac{\partial f}{\partial v}(a)$ indica a variação de f a partir de a e na direção de v.

4. Utilizamos nesta definição apenas vetores unitários porque dos vetores só nos interessa a direção (e sentido).

Exercício:: 1

1. Sejam $f(x,y) = x^2 + xy$, $a = (3,4)$, $v = (1,1)$ e $w = (1,0)$.

   (a) Indique vetores v_1, w_1, unitários, com a mesma direção e sentido de v, w, respetivamente.

   (b) Calcule $\dfrac{\partial f}{\partial v_1}(a)$ e $\dfrac{\partial f}{\partial w_1}(a)$.

Derivada parcial

- Se e_k designar o $k-$ésimo vetor da base canónica de $\mathbb{R}^n$

$$e_k = (0, \ldots, 0, 1, 0, \ldots, 0)$$

a derivada direcional de f na direção de e_k denomina-se **derivada parcial em ordem a x_k** e denota-se por $\dfrac{\partial f}{\partial x_k}$:

$$\frac{\partial f}{\partial x_k}(a) = \frac{\partial f}{\partial e_k}(a)$$

- O vetor gradiente de f em a é o vetor das derivadas parciais de f em a e denota-se

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \ \ldots, \ \frac{\partial f}{\partial x_n}(a)\right)$$

$$= \frac{\partial f}{\partial x_1}(a)\, e_1 + \ldots + \frac{\partial f}{\partial x_n}(a)\, e_n$$

Caso particular:: $\mathbb{R}^2$

- Se $f : D \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}$, onde D é um conjunto aberto.
  - A derivada parcial de f em ordem a x no ponto (a, b) é o limite, se existir,
  $$\lim_{t \to 0} \frac{f(a+t, b) - f(a, b)}{t}$$
  e denota-se por
  $$\frac{\partial f}{\partial x}(a, b) \qquad \text{ou} \qquad f_x(a, b)$$
  - A derivada parcial de f em ordem a y no ponto (a, b) é o limite, se existir,
  $$\lim_{t \to 0} \frac{f(a, b+t) - f(a, b)}{t}$$
  e denota-se por
  $$\frac{\partial f}{\partial y}(a, b) \qquad \text{ou} \qquad f_y(a, b)\,.$$

Exercício:: 2

2. Sendo $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ a função definida por $f(x, y) = x^2 y$, calcule, usando a definição:

   (a) $\dfrac{\partial f}{\partial x}(0, 0)$

   (b) $\dfrac{\partial f}{\partial y}(1, 2)$

3. Calcule, se existir, $\dfrac{\partial f}{\partial x}(0, 0)$ e $\dfrac{\partial f}{\partial y}(0, 0)$ quando

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{se} \quad (x, y) \neq (0, 0) \\ 0 & \text{se} \quad (x, y) = (0, 0) \end{cases}$$

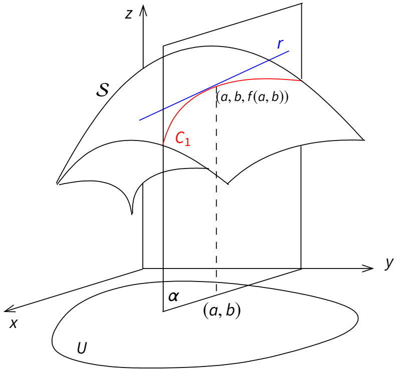
Interpretação geométrica da derivada parcial ($n = 2$)

- Sejam $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ e

$$\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 \ : \ z = f(x, y) \ \wedge \ (x, y) \in D\}$$

- Suponha-se que existe

$$\frac{\partial f}{\partial x}(a, b) = f_x(a, b)$$



- Sejam
  - $\alpha\ :\ y = b$ o plano paralelo ao plano xz e que passa em $(a, b, f(a, b)) \in \mathcal{S}$;
  - C_1 a curva definida pela interseção de α e de $\mathcal{S}$

$$\begin{cases} z = f(x, y) \\ y = b \end{cases} \Longrightarrow z = f(x, b) = \varphi(x);$$

  - $\varphi'(a)$ é o declive da reta tangente a C_1 em $(a, b, f(a, b))$
- Mas

$$\begin{aligned} \frac{\partial f}{\partial x}(a, b) &= \lim_{t \longrightarrow 0} \frac{f((a, b) + t(1, 0)) - f(a, b)}{t} = \lim_{t \longrightarrow 0} \frac{f(a + t, b) - f(a, b)}{t} \\ &= \lim_{t \longrightarrow 0} \frac{\varphi(a + t) - \varphi(a)}{t}, \qquad \text{pois} \qquad f(x, b) = \varphi(x) \\ &= \varphi'(a) \end{aligned}$$

- Conclusão
  - $f_x(a, b)$ é o declive da reta tangente à curva C_1 no ponto $(a, b, f(a, b))$ obtida pela intersecção do plano $y = b$ e da superfície $\mathcal{S}$;
  - De modo análogo, $f_y(a, b)$ é o declive da reta tangente à curva C_2 no ponto $(a, b, f(a, b))$ obtida pela interseção do plano $x = a$ e da superfície $\mathcal{S}$.
- [Caso geral] Se $v = (v_1, v_2)$ é um vetor unitário, é possível considerar o plano vertical que passa no ponto $(a, b, f(a, b))$ e que tem $(v_1, v_2, 0)$ como vetor diretor. Este plano interseta a superfície S segundo uma curva, seja C_3. $\dfrac{\partial f}{\partial v}(a, b)$ é o declive da reta tangente à curva C_3 no ponto $(a, b, f(a, b))$.

Aritmética das derivadas parciais

Sejam $D \subseteq \mathbb{R}^n$ aberto, $\quad a \in D \quad$ e $\quad f, g : D \longrightarrow \mathbb{R}$.

- Se, para $i = 1, \ldots, n$, existirem

$$f_{x_i}(a) \quad \text{e} \quad g_{x_i}(a)$$

tem-se

- $\dfrac{\partial}{\partial x_i}(f+g)(a) = f_{x_i}(a) + g_{x_i}(a)$
- $\dfrac{\partial}{\partial x_i}(fg)(a) = f_{x_i}(a)g(a) + f(a)g_{x_i}(a)$
- $\dfrac{\partial}{\partial x_i}\left(\dfrac{f}{g}\right)(a) = \dfrac{f_{x_i}(a)g(a) - f(a)g_{x_i}(a)}{(g(a))^2}, \quad g(a) \neq 0.$

Observação

- Para derivar f em ordem a x_k considera-se todas as outras variáveis x_i ($i \neq k$) como constantes:

  - [Exemplo] Se $f(x, y) = 2x + y \quad \text{e} \quad g(x, y) = 2\,xy$

    - para derivar em ordem a x considera-se y como uma constante

    $$\frac{\partial f}{\partial x}(x, y) = 2, \qquad \frac{\partial g}{\partial x}(x, y) = 2y.$$

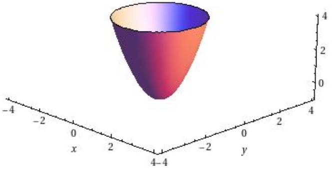
    - para derivar em ordem a y considera-se x como uma constante

    $$\frac{\partial f}{\partial y}(x, y) = 1, \qquad \frac{\partial g}{\partial y}(x, y) = 2x.$$

Exercício:: 3

4. Calcular as derivadas parciais da função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

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



Exercício:: 4

5. Calcule as derivadas parciais das funções seguintes indicando o domínio de validade. Escreva o gradiente.

   (a) $f(x, y) = \text{sen}(x^2 - 3xy)$

   (b) $f(x, y) = x^2 y^2 \, e^{2xy}$

   (c) $f(x, y) = e^x \ln(xy)$

   (d) $f(x, y, z) = x \, e^{x^2+y^2+z^2}$

6. Mostre que sendo $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ a função definida por $f(x, y) = e^{xy}$, então

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

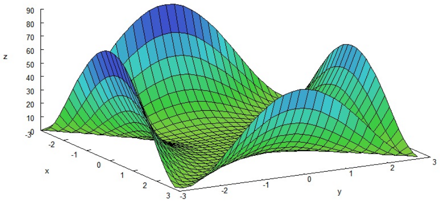
3.2 Funções reais: diferencial

Função diferenciável

- [Revisão de Cálculo] Uma função $f : D \subseteq \mathbb{R} \longrightarrow \mathbb{R}$ diz-se derivável ou diferenciável em $a \in D$ se existe a derivada de f em a: $f'(a)$.
  - Se f é derivável em a existe uma reta tangente ao gráfico de f em $(a, f(a))$ e o seu declive é $f'(a)$:
  $$\ell(x) = f(a) + f'(a)\,(x - a)\,.$$
  - Isto é,

$$\lim_{x \longrightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a) \Leftrightarrow \lim_{x \longrightarrow a} \left[\frac{f(x) - f(a)}{x - a} - f'(a) \right] = 0$$
$$\Leftrightarrow \lim_{x \longrightarrow a} \left[\frac{f(x) - f(a) - f'(a)(x - a)}{x - a} \right] = 0$$
$$\Leftrightarrow \lim_{x \longrightarrow a} \left[\frac{f(x) - \ell(x)}{x - a} \right] = 0$$

- Se $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ qual o significado de f ser diferenciável em a?



A função

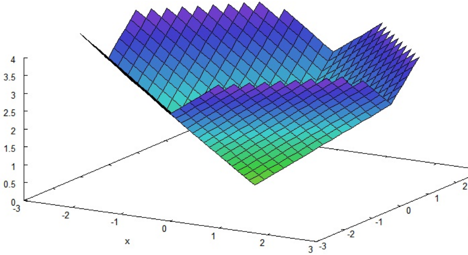
$$g(x, y) = (x^2 - y^2)^2$$

é uma função diferenciável em $\mathbb{R}^2$.

A função

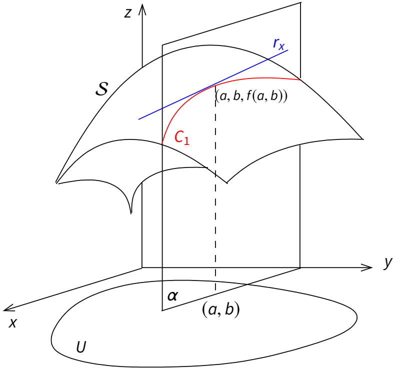
$$f(x, y) = |x| + |y|$$

não é uma função diferenciável em $\mathbb{R}^2$.



Aproximação linear ($n = 2$)

- Sejam
  - r_x a reta tangente ao gráfico de f em $(a, b, f(a, b))$ cujo declive é dado por
    $$f_x(a, b)$$
  - r_y a reta tangente ao gráfico de f em $(a, b, f(a, b))$ cujo declive é dado por
    $$f_y(a, b)$$



- As retas r_x e r_y, concorrentes em $(a, b, f(a, b))$, definem um plano.

- $(1, 0, f_x(a, b))$ e $(0, 1, f_y(a, b))$ são vetores diretores desse plano

- Uma equação desse plano é, para $\lambda, \mu \in \mathbb{R}$

$$(x, y, z) = (a, b, f(a, b)) + \lambda\,(1, 0, f_x(a, b)) + \mu\,(0, 1, f_y(a, b))$$

  ou ainda

$$z = f(a, b) + f_x(a, b)\,(x - a) + f_y(a, b)\,(y - b)$$

- O plano $\pi(x,y) = f(a,b) + f_x(a,b)\,(x-a) + f_y(a,b)\,(y-b)$ será uma boa aproximação linear de f na vizinhança de (a,b) se

$$\lim_{(x,y)\to(a,b)} \frac{|f(x,y) - \pi(x,y)|}{\|(x,y)-(a,b)\|} = 0$$

isto é, se

$$\lim_{(x,y)\to(a,b)} \frac{|f(x,y) - f(a,b) - f_x(a,b)\,(x-a) - f_y(a,b)\,(y-b)|}{\|(x,y)-(a,b)\|} = 0$$

- Considerando a aplicação linear[1] $L : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $L(u_1,u_2) = f_x(a,b)u_1 + f_y(a,b)u_2$, o limite anterior vem

$$\lim_{(x,y)\to(a,b)} \frac{|f(x,y) - f(a,b) - L(x-a, y-b)|}{\|(x,y)-(a,b)\|} = 0$$

[1] L é uma aplicação linear se $L(\alpha\, v + \beta\, w) = \alpha\, L(v) + \beta\, L(w)$

- [Função diferenciável]

  Sejam $f\ :\ D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R},\ a \in D$, D aberto.
  A função f **é diferenciável em** a se existir uma aplicação linear

  $$L\ :\ \mathbb{R}^n \longrightarrow \mathbb{R}$$

  tal que

  $$\lim_{x \longrightarrow a} \frac{|f(x) - f(a) - L(x-a)|}{\|x-a\|} = 0\,.$$

  - Prova-se que a aplicação linear L, se existir, é única.
  - [Diferencial] A aplicação L diz-se diferencial de f em a e denota-se por $df(a)$:

    $$\begin{array}{rccc} df(a): & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & v & \longmapsto & L(v) \end{array}$$

- Sendo $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, D aberto e $a \in D$
  - se f é diferenciável em a então o seu gráfico admite hiperplano tangente em $(a, f(a))$;
  - neste caso, a diferencial de f em a é a aplicação linear dada por

$$df(a)(v) = \frac{\partial f}{\partial x_1}(a)v_1 + \cdots + \frac{\partial f}{\partial x_n}(a)v_n$$

para cada $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$

Exemplo

Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$f(x,y) = \begin{cases} \dfrac{x\,y}{x^2+y^2} & \text{se} \quad (x,y) \neq (0,0) \\ 0 & \text{se} \quad (x,y) = (0,0) \end{cases}$$

- A função será diferenciável em $(0,0)$ se

$$\lim_{(x,y)\to(0,0)} \frac{|f(x,y) - f(0,0) - df(0,0)(x-0, y-0)|}{\|(x,y)-(0,0)\|} = 0$$

  - $f_x(0,0) = 0$ e $f_y(0,0) = 0$ pelo que $df(0,0)(x,y) = 0$.
  - Logo, o limite anterior pode ser escrito como

$$\lim_{(x,y)\to(0,0)} \frac{\left|\dfrac{x\,y}{x^2+y^2} - 0 - 0\right|}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{|x\,y|}{\sqrt{(x^2+y^2)^3}}$$

  - Como este limite não existe, f não é diferenciável em $(0,0)$.

- [Teorema A] Se f é diferenciável em a, então existem todas as derivadas parciais de f em a e

$$\begin{array}{rccl} df(a): & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & v & \mapsto & df(a)(v) = \dfrac{\partial f}{\partial x_1}(a)v_1 + \cdots + \dfrac{\partial f}{\partial x_n}(a)v_n \end{array}$$

  - Se f é diferenciável em a então

$$df(a)(v) = \nabla f(a) \cdot v$$

- [Teorema B] Se todas as derivadas parciais de f existem numa vizinhança de a e são contínuas em a **então** a função f é diferenciável em a .

- [Teorema C] Se f é diferenciável em a, então f é contínua em a .

Exemplo:: Teoremas B+C

Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por

$$f(x,y) = \begin{cases} \dfrac{x^3 y}{x^6 + y^2} & \text{se} \quad (x,y) \neq (0,0) \\ 0 & \text{se} \quad (x,y) = (0,0) \end{cases}$$

- A função f não é contínua em $(0,0)$ pois não existe

$$\lim_{(x,y)\to(0,0)} f(x,y)\,.$$

  - Se f não é contínua em $(0,0)$ pelo Teorema C, não é diferenciável em $(0,0)$.

- Mas a função f tem derivadas direcionais em $(0,0)$ segundo todas as direções .
  - Para um qualquer vetor $u=(u_1,u_2)$ unitário

$$\begin{aligned}\frac{\partial f}{\partial u}(0,0) &= \lim_{t\to 0}\frac{f((0,0)+tu)-f(0,0)}{t} = \lim_{t\to 0}\frac{f(tu_1,tu_2)-0}{t}\\ &= \lim_{t\to 0}\left[\frac{1}{t}\,\frac{t^4u_1^3u_2}{t^6u_1^6+t^2u_2^2}\right] = \lim_{t\to 0}\frac{tu_1^3u_2}{t^4u_1^6+u_2^2} = 0\end{aligned}$$

  - Em particular existem as duas derivadas parciais em $(0,0)$.

  - Como f não é diferenciável, pelo Teorema B, pelo menos uma das derivadas parciais de f não é contínua em $(0,0)$.

Propriedades das funções diferenciáveis

- Se f for diferenciável em $a \in D$ então
  - f é contínua em a;
  - existem todas as derivadas parciais e direcionais de f em a e, para todo o vetor v unitário,
  $$df(a)(v) = \frac{\partial f}{\partial v}(a);$$
  - a diferencial de f em a é a aplicação linear $df(a) : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada para cada $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, por
  $$df(a)(v) = \frac{\partial f}{\partial x_1}(a)v_1 + \cdots + \frac{\partial f}{\partial x_n}(a)v_n$$
  - a diferencial e o gradiente de f em a relacionam-se pela igualdade
  $$df(a)(v) = \nabla f(a) \cdot v$$

Exercícios:: 1

1. Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ tal que

$$f(x,y) = \begin{cases} (x^2+y^2)\,\mathrm{sen}\,\dfrac{1}{\sqrt{x^2+y^2}} & \text{se} \quad (x,y) \neq (0,0) \\ 0 & \text{se} \quad (x,y) = (0,0) \end{cases}$$

   (a) Calcule $f_x(0,0)$ e $f_y(0,0)$.
   (b) Determine f_x e f_y e verifique que não são contínuas em $(0,0)$.
   (c) Mostre que f é diferenciável em $(0,0)$.

Funções de classe C^1

- **[Funções de classe C^1]**

  A função $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, D aberto, diz-se de classe C^1 se as derivadas parciais existem e são contínuas.

  - Se as derivadas parciais existem e são contínuas, a função é diferenciável.
  - Se a função é diferenciável, então é contínua.
  - $f \in C^1$:: f é contínua e todas as suas derivadas parciais são contínuas.

Plano tangente ao gráfico ($n = 2$)

- Se f é diferenciável em (a, b), uma equação do plano tangente ao gráfico de f em $(a, b, f(a, b))$ é
$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b).$$

- Foi visto que o plano tangente ao gráfico de f em $(a, b, f(a, b))$ é uma linearização de f em torno de (a, b).

- Usando o vetor gradiente e o produto escalar de vetores, a equação do plano tangente ao gráfico de f em $(a, b, f(a, b))$ é
$$z = f(a, b) + \nabla f(a, b) \cdot (x - a, y - b).$$

Exercícios:: 2

2. Encontre uma equação do plano tangente ao gráfico da função $f(x,y) = x^2 + y^3$ no ponto de coordenadas $(3, 1, f(3, 1))$.

3.3 Funções reais: derivadas de ordem superior

Derivadas parciais de ordem superior

Sejam $D \subseteq \mathbb{R}^n$ aberto, $a \in D$ e $f : D \longrightarrow \mathbb{R}$ uma função tal que numa vizinhança de $a \in D$ existe $\frac{\partial f}{\partial x_i}$ (i fixo).

- Diz-se que que f tem derivada parcial de segunda ordem em a se $\dfrac{\partial f}{\partial x_i}$ admite derivada parcial em ordem a x_j em a.
  - Esta derivada parcial de segunda ordem em a representa-se por

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \qquad \text{ou} \qquad f_{x_i x_j}(a)\,.$$

  - Se $j = i$ representa-se $\dfrac{\partial^2 f}{\partial x_i^2}(a)$
  - Poderão existir n^2 derivadas parciais de segunda ordem.
  - Define-se a derivada parcial de ordem k como uma derivada parcial de uma derivada parcial de ordem $k-1$.

Exercício:: 1

1. Determine todas as derivadas parciais de segunda ordem da função

$$f(x, y) = \cos(x\, y^2)\,.$$

Função de Classe $\mathcal{C}^k$

Sejam $D \subseteq \mathbb{R}^n$ aberto e $f : D \longrightarrow \mathbb{R}$ uma função.

- [Função de Classe $\mathcal{C}^k$]

  Diz-se que f **é de classe** $\mathcal{C}^k$ quando existem e são contínuas todas as derivadas parciais de f de ordem menor ou igual a k .

  - A função f diz-se de classe $\mathcal{C}^0$ se for contínua.
  - A função f diz-se de classe $\mathcal{C}^\infty$ se for de classe $\mathcal{C}^k$ para todo o k.

Teorema de Schwarz

- [Teorema de Schwarz] Seja $D \subseteq \mathbb{R}^n$ aberto e $f : D \longrightarrow \mathbb{R}$ uma função de classe $\mathcal{C}^2$. Então

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}, \quad i, j = 1, 2, \ldots, n\,.$$

  - Este resultado implica que se f é de classe $\mathcal{C}^k$ então qualquer permutação na ordem de derivação de uma derivada parcial de ordem menor ou igual a k conduz ao mesmo resultado.
  - Por exemplo, em $\mathbb{R}^2$ se f é de classe $\mathcal{C}^3$ tem-se
    - $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$
    - $\dfrac{\partial^3 f}{\partial x^2 \partial y} = \dfrac{\partial^3 f}{\partial y \partial x^2} = \dfrac{\partial^3 f}{\partial x \partial y \partial x}$ e $\dfrac{\partial^3 f}{\partial y^2 \partial x} = \dfrac{\partial^3 f}{\partial x \partial y^2} = \dfrac{\partial^3 f}{\partial y \partial x \partial y}$

Exercício:: 2

2. Verifique o teorema de Schwarz para a função $f(x,y) = xe^y + x^2y$.

3. Mostre que não pode existir uma função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$ cujas derivadas parciais de primeira ordem sejam

$$f_x(x,y) = 2x^3 \qquad \text{e} \qquad f_y(x,y) = yx^2 + x\,.$$

3.4 Funções vetoriais diferenciáveis

Função diferenciável

- [Função diferenciável]

  A função vetorial $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^k$ é diferenciável em $a \in D$, se existir uma aplicação linear

  $$L : \mathbb{R}^n \longrightarrow \mathbb{R}^k$$

  tal que
  $$\lim_{x \longrightarrow a} \frac{\|f(x) - f(a) - L(x-a)\|}{\|x-a\|} = 0\,.$$

  - A esta aplicação linear, que se existir é única, chama-se diferencial de f em a e denota-se por $df(a)$.
  - A função vetorial f diz-se diferenciável em D se for diferenciável em todo o $a \in D$.

- [Teorema] A função vetorial $f = (f_1, \cdots, f_k)$ é diferenciável em $a \in D$ se e só se cada uma das suas funções componentes, f_i, $i = 1, \ldots, k$, for diferenciável em a.

- [Teorema] Se $f = (f_1, \ldots, f_k)$ é diferenciável em $a \in D$ então a sua diferencial é a aplicação linear

$$df(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^k$$

definida para cada $v = (v_1, \cdots, v_n) \in \mathbb{R}^n$ por

$$df(a)(v) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_k}{\partial x_1}(a) & \ldots & \frac{\partial f_k}{\partial x_n}(a) \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

Matriz Jacobiana

- A matriz de $df(a)$ relativamente às bases canónicas de $\mathbb{R}^n$ e $\mathbb{R}^k$ chama-se matriz jacobiana de f em a e denota-se por $Jf(a)$, isto é

$$Jf(a) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial f_k}{\partial x_1}(a) & \dots & \dfrac{\partial f_k}{\partial x_n}(a) \end{bmatrix}$$

  - A matriz Jacobiana de f em a é a matriz $k \times n$ das primeiras derivadas de f;
  - Cada linha de $Jf(a)$ corresponde ao vetor gradiente de f_i em a.

Exercício

1. Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ a função definida por

$$f(x, y) = (y + e^{x+y},\ x + xy^2,\ \mathrm{sen}(x + y))\,.$$

Determine

(a) as funções componentes;

(b) o vetor gradiente de cada uma das funções componentes;

(c) a matriz jacobiana de f;

(d) a lei da transformação linear $df(0, 0)$;

(e) $df(0, 0)(1, 1)$.

2. Determine a matriz jacobiana das funções
   (a) $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ tal que $f(x, y) = (x, y)$;
   (b) $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ tal que $f(x, y) = (xe^y + \cos y,\ x,\ x + e^y)$.

3. Considere a função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por

$$f(x, y) = (2x^2,\ 3y,\ 2xy)\,.$$

   (a) Calcule a matriz jacobiana de f.
   (b) Justifique que a função f é diferenciável e calcule a sua diferencial no ponto $(1, 1)$.
   (c) Determine $df(1, 1)(2, 3)$.

3.5 Propriedades da diferencial e da matriz jacobiana

Aritmética das diferenciais:: funções reais e funções vetoriais

- Sejam $f, g : U \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^k$ funções diferenciáveis em a. Então
  - [Produto por escalar] Sendo $\alpha \in \mathbb{R}$, a função $h(x) = \alpha f(x)$ é diferenciável em a e
$$dh(a) = \alpha df(a)$$
$$\updownarrow$$
$$Jh(a) = \alpha Jf(a)$$
  - [Adição] A função $h(x) = f(x) + g(x)$ é diferenciável em a e
$$dh(a) = df(a) + dg(a)$$
$$\updownarrow$$
$$Jh(a) = Jf(a) + Jg(a)$$

Se $k = 1$, trata-se de funções reais e a matriz Jacobiana, sendo uma matriz $1 \times n$, corresponde a um vetor: o vetor gradiente.

Aritmética das diferenciais:: funções reais

- Sejam $f, g : U \longrightarrow \mathbb{R}$ funções reais diferenciáveis em a. Então
  - [Produto] A função $h(x) = f(x)g(x)$ é diferenciável em a e

$$
\begin{aligned}
dh(a) =& g(a)df(a) + f(a)dg(a) \\
&\Updownarrow \\
\nabla h(a) =& g(a)\nabla f(a) + f(a)\nabla g(a)
\end{aligned}
$$

  - [Quociente] a função $h(x) = \dfrac{f(x)}{g(x)}$ é diferenciável em a e

$$
\begin{aligned}
dh(a) =& \frac{g(a)df(a) - f(a)dg(a)}{g(a)^2} \\
&\Updownarrow \\
\nabla h(a) =& \frac{g(a)\nabla f(a) - f(a)\nabla g(a)}{g(a)^2}
\end{aligned}
$$

Regra da cadeia

- [Regra da cadeia]

  Sejam $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ e $f : \mathbb{R}^p \longrightarrow \mathbb{R}^k$ funções vetoriais tais que g é diferenciável em a e f é diferenciável em $g(a)$. Então a função $h = f \circ g$ é diferenciável em a e

$$dh(a) = d(f \circ g)(a) = df(g(a)) \circ dg(a)$$

$$\updownarrow$$

$$Jh(a) = J(f \circ g)(a) = Jf(g(a))\, Jg(a)$$

[Nota] O produto que ocorre no último membro é um produto de matrizes (não comutativo).

Observação

- A regra da cadeia é a generalização da regra da derivação da função composta estudada em Cálculo.
  - Sejam $f, g : \mathbb{R} \longrightarrow \mathbb{R}$.
    Se g é derivável em a e f é derivável em $g(a)$ então a função $h = f \circ g$ é derivável em a e
    $$h'(a) = f'(g(a))\, g'(a)$$

    (Repare que, como f, g, h são funções $\mathbb{R} \to \mathbb{R}$, Jf, Jg e Jh são matrizes 1×1 que se identificam com f', g' e h', respetivamente.)

[Caso particular:: 1]

Sejam $g : \mathbb{R} \longrightarrow \mathbb{R}^3$ tal que $g(t) = (x(t), y(t), z(t))$ e $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$.
Defina-se $h : \mathbb{R} \longrightarrow \mathbb{R}$ dada por

$$h(t) = (f \circ g)(t) = f(x(t), y(t), z(t)).$$

Então

$$h'(a) = Jh(a) = Jf(g(a))\, Jg(a) = \nabla f(g(a))\, Jg(a)$$

onde

$$\nabla f(g(a)) = \left(\frac{\partial f}{\partial x}(g(a)), \frac{\partial f}{\partial y}(g(a)), \frac{\partial f}{\partial z}(g(a)) \right) \qquad \text{e}$$

$$Jg(a) = \begin{pmatrix} \dfrac{dx}{dt}(a) \\ \dfrac{dy}{dt}(a) \\ \dfrac{dz}{dt}(a) \end{pmatrix}$$

Isto é (omitindo os argumentos)

$$\nabla f\, Jg = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)\begin{pmatrix} \dfrac{dx}{dt} \\ \dfrac{dy}{dt} \\ \dfrac{dz}{dt} \end{pmatrix} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

Assim,

$$h' = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

ou seja

$$h'(a) = \frac{\partial f}{\partial x}(g(a))\frac{dx}{dt}(a) + \frac{\partial f}{\partial y}(g(a))\frac{dy}{dt}(a) + \frac{\partial f}{\partial z}(g(a))\frac{dz}{dt}(a)\,.$$

[Caso particular:: 2]

Sejam $g : \mathbb{R}^3 \to \mathbb{R}^3$ tal que $g(x,y,z) = (u(x,y,z), v(x,y,z), w(x,y,z))$ e $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$.
Defina-se $h : \mathbb{R}^3 \longrightarrow \mathbb{R}$ por

$$h(x,y,z) = (f \circ g)(x,y,z) = f(u(x,y,z), v(x,y,z), w(x,y,z)).$$

Então

$$Jh(x,y,z) = Jf(g(x,y,z))\, Jg(x,y,z) = \nabla f(g(x,y,z))\, Jg(x,y,z)$$

Aqui

$$\nabla f(u,v,w) = \left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w}\right) \qquad \text{e}$$

$$Jg(x,y,z) = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{pmatrix}$$

Isto é (omitindo os argumentos)

$$\nabla h = \nabla f \, Jg =$$

$$= \left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w}\right) \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial z} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} & \dfrac{\partial v}{\partial z} \\ \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} & \dfrac{\partial w}{\partial z} \end{pmatrix}$$

Assim,

$$\begin{cases} \dfrac{\partial h}{\partial x} = \dfrac{\partial f}{\partial u}\dfrac{\partial u}{\partial x} + \dfrac{\partial f}{\partial v}\dfrac{\partial v}{\partial x} + \dfrac{\partial f}{\partial w}\dfrac{\partial w}{\partial x} \\ \dfrac{\partial h}{\partial y} = \dfrac{\partial f}{\partial u}\dfrac{\partial u}{\partial y} + \dfrac{\partial f}{\partial v}\dfrac{\partial v}{\partial y} + \dfrac{\partial f}{\partial w}\dfrac{\partial w}{\partial y} \\ \dfrac{\partial h}{\partial z} = \dfrac{\partial f}{\partial u}\dfrac{\partial u}{\partial z} + \dfrac{\partial f}{\partial v}\dfrac{\partial v}{\partial z} + \dfrac{\partial f}{\partial w}\dfrac{\partial w}{\partial z} \end{cases}$$

[Regra da cadeia:: como fazer]

1. Identificar as funções f, g
2. Verificar se Jf ou Jg é um gradiente
3. Calcular Jf e $Jf(g)$
4. Calcular Jg
5. Efetuar $Jf(g)\, Jg$

Exercícios:: 1

1. Calcular ∇h sendo $h = f \circ g$ onde

$$g: \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x,y,z) & \mapsto & (xz, xy) \end{array} \quad \text{e} \quad f: \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (u,v) & \mapsto & u^2 + v \end{array}$$

2. Considere as funções

$$u: \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & xy \end{array}, \quad v: \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & x+y \end{array},$$

$$w: \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & e^x \end{array}, \quad f: \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ (x,y,z) & \mapsto & x^2y + y^2z \end{array}$$

$$\text{e} \quad h: \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \mapsto & f(u(x,y), v(x,y), w(x,y)) \,. \end{array}$$

Determine $\nabla h(x,y)$.

3. Considere a função $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida por $f(x, y, z) = x^2y - xz$ e $a \in \mathbb{R}$.
   (a) Calcule $df(1, 0, 0)(1, 2, 2)$.
   (b) Usando a regra da cadeia, determine a de modo que a função $g : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $g(t) = f(at^2, at, t^3)$ tenha derivada nula.

4. Calcule $\frac{\partial w}{\partial p}$ e $\frac{\partial w}{\partial q}$, onde $w = r^2 + s^2$ e $r = pq^2, s = p^2 \operatorname{sen} q$.

Interpretação geométrica do vetor gradiente

Sejam $U \subset \mathbb{R}^n$ aberto, $f : U \longrightarrow \mathbb{R}$ diferenciável em $a \in U$.

1. Sendo $v \in \mathbb{R}^n$ um vetor, foi visto que

$$df(a)(v) = \nabla f(a) \cdot v = ||\nabla f(a)||\, ||v|| \cos \alpha$$

(α é a amplitude do ângulo entre $\nabla f(a)$ e v);

e, para $||v|| = 1$,

$$df(a)(v) = \frac{\partial f}{\partial v}(a).$$

Então, se $\nabla f(a) \neq 0$ e $||v|| = 1$

- o maior valor de $\frac{\partial f}{\partial v}(a)$ é igual a $||\nabla f(a)||$ e ocorre quando $\alpha = 0$, isto é
$$v = \frac{\nabla f(a)}{||\nabla f(a)||};$$
- o menor valor de $\frac{\partial f}{\partial v}(a)$ é igual a $-||\nabla f(a)||$ e ocorre quando $\alpha = \pi$, isto é
$$v = -\frac{\nabla f(a)}{||\nabla f(a)||};$$
- $\frac{\partial f}{\partial v}(a)=0$ quando $\alpha = \frac{\pi}{2}$, isto é $v \perp \nabla f(a)$.

- [Consequência 1]

  A derivada direcional $\frac{\partial f}{\partial v}(a)$ mede a variação da função em $a \in U$ na direção do vetor unitário v. Assim, o que atrás foi visto significa que

  - o vetor $\nabla f(a)$ aponta na direção e no sentido de crescimento máximo de f ;

  - o vetor $-\nabla f(a)$ aponta na direção e no sentido de decrescimento máximo de f;

  - $||\nabla f(a)||$ é maior quando as estruturas de nível de f estão mais próximas entre si e menor quando estas estão mais afastadas.

2. Sejam Σ_c a estrutura de nível c de f em $a \in U$ ($c = f(a)$) e v um vetor unitário tangente a Σ_c em a:

Como f é constante em Σ_c tem-se $\frac{\partial f}{\partial v}(a) = 0$ pelo que

$$0 = \frac{\partial f}{\partial v}(a) = \nabla f(a) \cdot v$$

Assim, $\nabla f(a)$ é ortogonal a v.

- [Consequência 2] $\nabla f(a)$ é normal à estrutura de nível de f, Σ_c, em a.

  (Um vetor diz-se normal a uma curva num ponto se for ortogonal à reta tangente à curva nesse ponto, normal a uma superfície num ponto se for ortogonal ao plano tangente à superfície nesse ponto, e assim por diante.)

Reta normal e hiperplano tangente

- Seja $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ diferenciável, $a \in U$ tal que $\nabla f(a) \neq 0$ e Σ_c a estrutura de nível de f em a

$$\Sigma_c = \{x \in U \ : \ f(x) = f(a)\}\ .$$

  - $\nabla f(a)$ é um vetor normal a Σ_c;
  - Uma equação paramétrica da reta normal a Σ_c em a é, então,

$$x = a + \lambda\, \nabla f(a), \qquad \lambda \in \mathbb{R};$$

  - Uma equação cartesiana do hiperplano[2] tangente a Σ_c em a é, então,

$$\nabla f(a) \cdot (x - a) = 0\ .$$

[2] se f é uma função real de variável real o "hiperplano" tangente a Σ_c é uma reta.

Exemplo

- Seja $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ definida por $f(x, y) = x^2 + y^2$ e $P = (2, 3)$.

- curva de nível de f que passa em P:

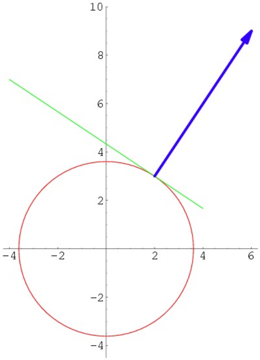
$$\Sigma_c = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 13\}$$

- vetor gradiente de f em P:

$$\nabla f(2, 3) = (4, 6)$$

- reta tangente a Σ_c em P:

$$\nabla f(2, 3) \cdot (x - 2, y - 3) = 0$$
$$\Leftrightarrow 2x + 3y = 13$$



Exercícios:: 2

5. Considere a curva de equação $x^2 + xy + y^2 = 3$.
   (a) Encontre um vetor perpendicular/normal e um vetor tangente à curva no ponto $(-1, -1)$.
   (b) Encontre equações que definam a reta tangente e a reta normal à curva no ponto $(-1, -1)$.

6. Determine o plano tangente ao elipsoide definido por

$$\frac{x^2}{12} + \frac{y^2}{12} + \frac{z^2}{3} = 1$$

no ponto de coordenadas $(2, 2, 1)$.

Extremos Locais

Seja $U \subset \mathbb{R}^n$ um conjunto aberto, $a \in U$ e $f : U \longrightarrow \mathbb{R}$. Diz-se que

- $f(a)$ é um mínimo local, ou a é um minimizante local de f, se existir uma bola $B(a, \varepsilon)$ tal que

$$f(x) \geq f(a), \qquad \forall x \in B(a, \varepsilon);$$

- $f(a)$ é um máximo local, ou a é um maximizante local de f, se existir uma bola $B(a, \varepsilon)$ tal que

$$f(x) \leq f(a), \qquad \forall x \in B(a, \varepsilon);$$

- $f(a)$ é um extremo local, ou a é um extremante local de f, se a for um minimizante ou um maximizante local de f.

Exemplo

- O ponto $a = (0, 0)$ é um minimizante local da função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $f(x, y) = x^2 + y^2$.

  De facto $f(a) = 0$ e, para todo o $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, tem-se

  $$f(x, y) = x^2 + y^2 > 0 = f(a)\,.$$

  Isto é, a é um minimizante de f e $f(a) = 0$ é um mínimo da função.

Teste das 1.as derivadas

Seja $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^1$.

- [Ponto crítico] $a \in U$ é um **ponto crítico** de f se

$$\nabla f(a) = \mathbf{0}$$

  [Teste das 1.as derivadas]

  Se $a \in U$ é um extremante local de f então é um ponto crítico de f.

- [Ponto de sela] $a \in U$ é um **ponto de sela** de f se a é ponto crítico mas não é extremante local de f.

Observações

- O teste das 1.as derivadas estabelece que os únicos candidatos a pontos extremantes de uma função de classe C^1 são os pontos do seu domínio onde se anulam todas as derivadas parciais de primeira ordem da função, simultaneamente.

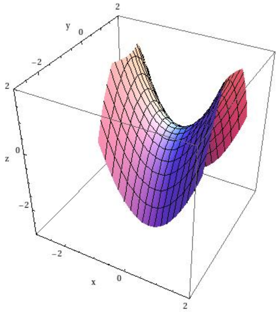
- É possível definir e estudar pontos críticos em funções que não sejam de classe C^1, mas esse estudo está fora do âmbito desta UC.

Exemplo

- A função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $f(x,y) = x^2 - y^2$ tem um ponto crítico mas não tem extremantes.

De facto, $\nabla f(x,y) = (2x, -2y)$ pelo que

$$\nabla f(x,y) = (0,0) \iff (x,y) = (0,0)\,.$$



Seja $a = (0,0)$. Observe-se que $f(a) = 0$. Para os pontos sobre o eixo do x tem-se $y = 0$ e

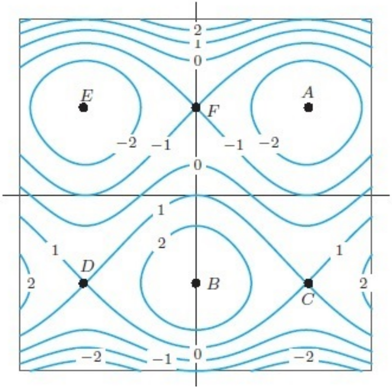
$$f(x,y) = f(x,0) = x^2 > 0 = f(a), \quad x \neq 0\,.$$

Por outro lado, para os pontos sobre o eixo dos y, tem-se $x = 0$ e

$$f(x,y) = f(0,y) = -y^2 < 0 = f(a), \quad y \neq 0\,.$$

Assim, qualquer $B(a,\varepsilon)$, $\varepsilon > 0$, contém pontos onde f assume valores superiores a $f(a)$ e outros pontos onde f assume valores inferiores a $f(a)$.
Logo a, embora seja ponto crítico de f, não é um extremante da função.

Extremos vs curvas de nível



- Curvas de nível fechadas sugerem pontos de extremo (A, B, E);
- Curvas de nível que se cruzam sugerem pontos de sela (C, D, F)

Teste das 2.as derivadas

Seja $U \subset \mathbb{R}^n$ um conjunto aberto e $f : U \longrightarrow \mathbb{R}$ uma função de classe $\mathcal{C}^2$ em $B(a, \varepsilon)$.

- Define-se a matriz Hessiana de f em a por

$$Hf(a) = \begin{pmatrix} f_{x_1x_1}(a) & \dots & f_{x_1x_n}(a) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(a) & \dots & f_{x_nx_n}(a) \end{pmatrix}$$

  onde $f_{x_ix_j} = \dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

- Hf é uma matriz
  - quadrada de dimensão n;
  - simétrica, pois pelo Teorema de Schwarz $f_{x_ix_j}(a) = f_{x_jx_i}(a)$.

- [Teste das 2.^{as} derivadas] (para extremantes locais)

  Seja $U \subset \mathbb{R}^n$ um conjunto aberto, $f : U \longrightarrow \mathbb{R}$ uma função de classe $\mathcal{C}^2$ e $a \in U$ um ponto crítico de f. Então

  - se $Hf(a)$ é definida positiva, a é um minimizante local de f;

  - se $Hf(a)$ é definida negativa, a é um maximizante local de f.

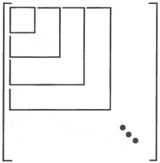
Um pouco de Álgebra Linear

Seja $Q \in \mathcal{M}_{n \times n}$.

1. A matriz Q diz-se
   - definida positiva se $\forall x \in \mathbb{R}^n, x \neq 0,\ x^T Q x > 0$
   - definida negativa se $\forall x \in \mathbb{R}^n, x \neq 0,\ x^T Q x < 0$
2. Se Q é uma matriz real e simétrica então possui n valores próprios reais: $\lambda_1, \ldots, \lambda_n$ (não necessariamente distintos). Neste caso,
   - Q é uma matriz
     - definida positiva se todos os $\lambda_i > 0$;
     - definida negativa se todos os $\lambda_i < 0$;
     - indefinida se existem valores próprios com sinais diferentes.
   - existe uma base de $\mathbb{R}^n$ na qual Q é uma matriz diagonal: $B^{-1}QB = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$ e

$$\det Q = \lambda_1 \ldots \lambda_n$$

3. Sendo Q uma matriz real e simétrica, considerem-se os determinantes das n submatrizes quadradas de Q ao longo da diagonal (menores principais):



- Q é definida positiva se e só se todos estes determinantes forem positivos

- Q é definida negativa se e só se os determinantes de ordem par forem positivos e os de ordem ímpar negativos.

- [Critério dos menores principais]

  Seja $U \subset \mathbb{R}^n$ um conjunto aberto, $f : U \longrightarrow \mathbb{R}$ uma função de classe $\mathcal{C}^3$ e $a \in U$ um ponto crítico de f. Então

  - se todos os menores principais de $Hf(a)$ são positivos, a é um minimizante local de f;
  - se os menores principais de ordem par de $Hf(a)$ são positivos e os de ordem ímpar negativos, a é um maximizante local de f;
  - se todos os menores principais de $Hf(a)$ são não nulos mas a matriz não é definida positiva ou definida negativa, f tem um ponto de sela em a;
  - se algum dos menores principais for nulo, nada se pode concluir sobre a natureza de a.

Exemplo

- A função $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ dada por $f(x, y, z) = x^2 + y^2 + z^2 + xy$ tem um minimizante em $a = (0, 0, 0)$.

  De facto, um único ponto crítico de f é determinado pela solução do sistema de três equações com três incógnitas

  $$\nabla f(x, y, z) = (0, 0, 0) \Leftrightarrow (2x + y, x + 2y, 2z) = (0, 0, 0)$$

  cuja solução é $a = (0, 0, 0)$.
  Por outro lado, a matriz Hessiana de f é a matriz (constante)

  $$Hf(x, y, z) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = Hf(a)$$

  e todos os seus menores principais são positivos, já que

  $$M_1 = 2 > 0, \quad M2 = \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 3 > 0, \qquad M_3 = \det Hf(a) = 6 > 0\,.$$

  Então, pelo critério dos menores principais, conclui-se que f tem um mínimo em $f(a)$.

Teste das 2.as derivadas:: caso $n = 2$

Seja $U \subset \mathbb{R}^2$ um conjunto aberto e $f : U \longrightarrow \mathbb{R}$ uma função de classe $\mathcal{C}^3$ em $B(a, \varepsilon)$.

- A matriz Hessiana de f em a é

$$Hf(a) = \begin{pmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{yx}(a) & f_{yy}(a) \end{pmatrix}$$

- Há dois menores principais:
  - $M_2 = \det Hf(a) = f_{xx}(a) f_{yy}(a) - \left[f_{xy}(a)\right]^2$
  - $M_1 = f_{xx}(a)$

- [Critério dos menores principais] (caso $n = 2$)

  Seja $U \subset \mathbb{R}^2$ um conjunto aberto, $f : U \longrightarrow \mathbb{R}$ uma função de classe $\mathcal{C}^3$ numa vizinhança de $a \in U$ e

$$\det Hf(a) = f_{xx}(a)f_{yy}(a) - \left[f_{xy}(a)\right]^2.$$

  Suponhamos que $a \in U$ é um ponto crítico de f.

  - se $\det Hf(a) > 0$ e
    - $f_{xx}(a) > 0$ então a é um minimizante local de f;
    - $f_{xx}(a) < 0$ então a é um maximizante local de f;
  - se $\det Hf(a) < 0$ então f tem um ponto de sela em a;
  - se $\det Hf(a) = 0$ nada se pode concluir.

Exemplo

- Determine e classifique os pontos críticos da função $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $f(x,y) = 12x^2 + 8y^3 - 24xy$.

Aqui $\nabla f(x,y) = (24x - 24y, 24y^2 - 24x)$ e

$$\nabla f(x,y) = (0,0) \Leftrightarrow \left\{ \begin{array}{l} x - y = 0 \\ y^2 - x = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x = y \\ x^2 - x = 0 \end{array} \right. \Rightarrow$$

$$\Rightarrow \left\{ \begin{array}{l} y = x \\ x(x-1) = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} y = 0 \\ x = 0 \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} y = 1 \\ x = 1 \end{array} \right. ,$$

isto é, f tem dois pontos críticos, sejam

$$A = (0,0) \text{ e } B = (1,1)\,.$$

Agora há que estudar a natureza de cada um destes pontos críticos recorrendo à matriz hessiana de f.
A matriz hessiana de f é

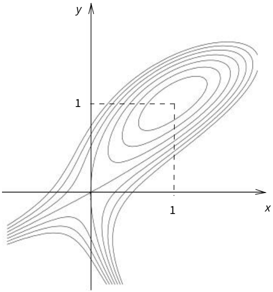
$$Hf(x,y) = \begin{pmatrix} 24 & -24 \\ -24 & 48\,y \end{pmatrix}$$

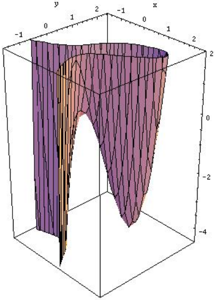
pelo que

$$Hf(A) = \begin{pmatrix} 24 & -24 \\ -24 & 0 \end{pmatrix} \quad \text{e} \quad Hf(B) = \begin{pmatrix} 24 & -24 \\ -24 & 48 \end{pmatrix}.$$

Uma vez que $\det Hf(A) = -24^2 < 0$ conclui-se que A é um ponto de sela.

Por outro lado, $\det Hf(B) = 576 > 0$ e $f_{xx}(B) = 24 > 0$ pelo que B é um minimizante de f .





- O ponto $(0, 0)$ é ponto de sela.
- O ponto $(1, 1)$ é ponto minimizante local.

Extremos condicionados

- [Problema] Pretende-se determinar os extremos da função

$$f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$$

quando as variáveis independentes estão sujeitas a algumas restrições.

Diz-se

- Extremos condicionados ou
- Extremos relativos

Métodos:

- Redução de dimensão
- Multiplicadores de Lagrange

Vocabulário

Sejam $B, U \subset \mathbb{R}^n$ conjuntos abertos, $f : U \longrightarrow \mathbb{R}$ e $g : B \longrightarrow \mathbb{R}$.
Considere-se a estrutura de nível $k = 0$ da função[3] g:

$$\Sigma = \{x \in B \ : \ g(x) = 0\} \, .$$

- [Extremante relativo]

  Um ponto $a \in (U \cap \Sigma)$ diz-se um **extremante de f relativo**, ou condicionado, à condição $g(x) = 0$ se é um extremante da restrição de f ao conjunto Σ, $\left. f \right|_{\Sigma}$;
  e, nesse caso, $f(a)$ diz-se um extremo de f relativo à condição $g(x) = 0$.

[3] Caso se tenha $g(x) = k$ pode-se definir $G(x) = g(x) - k$ e considerar $G(x) = 0$.

Observação

- Um ponto a em que f tem um extremante relativo a uma condição $g(x) = 0$ não é necessariamente um extremante local da função f e também não é necessariamente um ponto crítico de f.

- [Teorema de Weierstrass]

  Se f é uma função contínua e Σ é um conjunto limitado e fechado então $f\big|_{\Sigma}$ atinge um máximo e um mínimo.

Redução de dimensão: exemplo

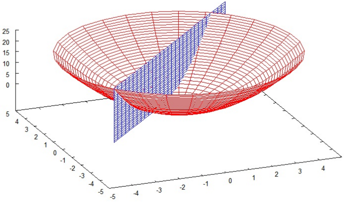
- Determinar os extremantes de $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $f(x,y) = x^2 + y^2$ sujeita à restrição
$$y = x + 1.$$

  - Sejam $g(x,y) = y - x - 1$ e
  $$\Sigma = \{(x,y) \in \mathbb{R}^2 : g(x,y) = 0\} = \{(x,y) \in \mathbb{R}^2 : y = x + 1\}$$

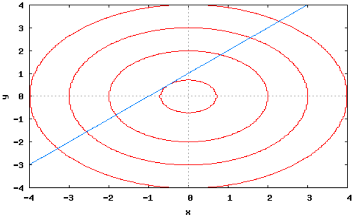
  - Então
  $$f\big|_{\Sigma}(x,y) = f(x,x+1) = x^2 + (x+1)^2 = 2x^2 + 2x + 1 = h(x).$$

$$\ldots$$

- $x = -\dfrac{1}{2}$ é ponto crítico de h.
- $\left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$ é ponto crítico de $f\big|_{\Sigma}$.



Gráficos de $z = x^2 + y^2$ e $y = x + 1$.



Algumas curvas de nível de f e restrição $y = x + 1$.

Multiplicadores de Lagrange

Sejam $B, U \subset \mathbb{R}^n$ conjuntos abertos, $f : U \longrightarrow \mathbb{R}$ e $g : B \longrightarrow \mathbb{R}$ duas funções de classe $\mathcal{C}^1$.

- Considere-se a estrutura de nível da função g:

$$\Sigma = \{x \in B \ : \ g(x) = 0\}.$$

- Suponha-se que $\nabla g(x) \neq \mathbf{0}$, $x \in \Sigma$.
  - Se $a \in (U \cap \Sigma)$ é um extremante local de f relativo à condição $g(x) = 0$ então existe um número real $\lambda \in \mathbb{R}$ tal que

$$\nabla f(a) = \lambda \, \nabla g(a).$$

  - O número λ é chamado multiplicador de Lagrange.
  - O ponto $a \in \Sigma$ diz-se ponto crítico de $f\big|_{\Sigma}$.
- Se, para algum $x \in \Sigma$, $\nabla g(x) = \mathbf{0}$, então este x é também um possível extremante.

[Método dos multiplicadores de Lagrange:: algoritmo]

Para determinar os extremantes da função $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ sujeitos à restrição $\Sigma : g(x) = 0$, há que

1. determinar os possíveis valores de $x \in \Sigma$ para os quais $\nabla g(x) = \mathbf{0}$
2. determinar x (e $\lambda \in \mathbb{R}$) resolvendo o sistema de $n + 1$ equações

$$\begin{cases} \nabla f(x) = \lambda \, \nabla g(x) \\ g(x) = 0 \end{cases}$$

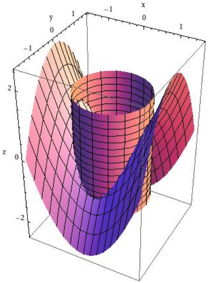
3. calcular o valor de f em todos os valores encontrados nos passos anteriores.

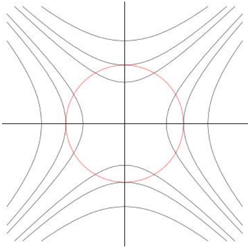
Observação

- Tal como no teste da 1.ª derivada para a determinação de extremos sem restrições, as soluções do sistema $\nabla f(x) = \lambda \, \nabla g(x), \quad g(x) = 0$ são candidatas a pontos extremantes.
    - Se f é contínua e Σ é limitado e fechado há a garantia de que existem extremos (pelo menos um maximizante e pelo menos um minimizante) — Teorema de Weierstrass.
    - Caso f não seja contínua ou Σ não limitado ou não fechado há que estudar o comportamento da função em torno de cada ponto crítico.
- A condição $\nabla f(x) = \lambda \, \nabla g(x)$ significa que os vetores gradientes são paralelos.

Exemplo

- Determinar os extremantes de $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $f(x,y) = x^2 - y^2$ sujeita à condição
$$x^2 + y^2 = 1.$$





Considere-se uma função auxiliar g definida em $\mathbb{R}^2$ por $g(x,y) = x^2 + y^2 - 1$ e a sua curva de nível zero
$$\Sigma = \left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 - 1 = 0\right\}.$$
As funções f e g são funções de classe $C^1(\mathbb{R}^2)$.

Note-se que $\nabla g(x,y) = (2x, 2y)$ nunca se anula em Σ (pois só se anula em $(0,0)$). Assim, os extremantes de $f\big|_\Sigma$ estarão entre as soluções (x,y) do sistema de 3 equações a 3 incógnitas

$$\begin{cases} \nabla f(x,y) = \lambda \nabla g(x,y) \\ g(x,y) = 0 \end{cases} \Rightarrow \begin{cases} 2x = 2\lambda x \\ -2y = 2\lambda y \\ x^2 + y^2 = 1 \end{cases} \Rightarrow \begin{cases} x(1-\lambda) = 0 \\ y(1+\lambda) = 0 \\ x^2 + y^2 = 1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x = 0 \\ y(1+\lambda) = 0 \\ y^2 = 1 \end{cases} \text{ ou } \begin{cases} \lambda = 1 \\ y(1+\lambda) = 0 \\ x^2 + y^2 = 1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x = 0 \\ y(1+\lambda) = 0 \\ y^2 = 1 \end{cases} \text{ ou } \begin{cases} \lambda = 1 \\ y = 0 \\ x^2 = 1 \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} x = 0 \\ \lambda = -1 \\ y = 1 \end{cases} \text{ ou } \begin{cases} x = 0 \\ \lambda = -1 \\ y = -1 \end{cases} \text{ ou } \begin{cases} \lambda = 1 \\ y = 0 \\ x = 1 \end{cases} \text{ ou } \begin{cases} \lambda = 1 \\ y = 0 \\ x = -1 \end{cases}$$

Isto é, os candidatos a extremantes são os pontos

$$A = (0, 1), \quad B = (0, -1), \quad C = (1, 0), \quad D = (-1, 0)\,.$$

Observe-se que, como f é contínua e Σ é um conjunto limitado e fechado, pelo Teorema de Weierstrass há a garantia de que f tem máximo e mínimo em Σ . Podemos então encontrá-los por simples comparação do valor da função nestes quatro pontos. Ora

$$f(A) = f(B) = -1 \quad \text{e} \quad f(C) = f(D) = 1\,.$$

Assim, os minimizantes de $f\big|_\Sigma$ são $A = (0, 1)$ e $B = (0, -1)$ enquanto os maximizantes são $C = (1, 0)$ e $D = (-1, 0)$.

Caso de k restrições

- Havendo várias restrições

$$g_i(x) = 0, \qquad i = 1, \ldots, k$$

há que

1. determinar os possíveis valores de x, pertencentes simultaneamente a todas as restrições, para os quais $\nabla g_1(x), \ldots, \nabla g_k(x)$ não são linearmente independentes;
2. resolver o sistema de $n + k$ equações

$$\begin{cases} \nabla f(x) = \lambda_1 \nabla g_1(x) + \cdots + \lambda_k \nabla g_k(x) \\ g_1(x) = 0 \\ \qquad \vdots \\ g_k(x) = 0 \end{cases}$$

Os valores obtidos nestes dois passos são candidatos a pontos extremantes.

Observação

- O método dos multiplicadores de Lagrange é equivalente ao teste das 1.$^{\text{as}}$ derivadas para extremos livres.
- Tal como no caso de extremos livres, existe um teste das 2.$^{\text{as}}$ derivadas para extremos condicionados.

Teste da hessiana aumentada

- Seja $U \subset \mathbb{R}^n$ um conjunto aberto, $f, g : U \longrightarrow \mathbb{R}$ funções de classe $\mathcal{C}^2$.
  Seja $a \in U, g(a) = 0$ e Σ a superfície de nível 0 de g.
  Suponha-se que $\nabla g(a) \neq 0$ e que existe $\lambda^* \in \mathbb{R}$ tal que $\nabla f(a) = \lambda^* \, \nabla g(a)$.
  Seja $\overline{H}$ a matriz hessiana[4] da função auxiliar $\mathcal{L}(\lambda, x) = f(x) - \lambda \, g(x)$ e designe-se por $\overline{M}_k$ o menor principal de ordem k da matriz $\overline{H}(\lambda^*, a)$.
  Se, para $k \geq 3$,
  - todos os $\overline{M}_k < 0$, então a é um minimizante local de $f\big|_\Sigma$;
  - $\overline{M}_3 > 0, \overline{M}_4 < 0, \ldots$ (isto é, os restantes $\overline{M}_k$ alternam de sinal), então a é um maximizante de $f\big|_\Sigma$.
  - todos os $\overline{M}_k \neq 0$ mas não estão nas situações anteriores, então a é ponto de sela.

[4] note a ordem das variáveis

Extremos globais

Seja $D \subset \mathbb{R}^n$, $a = (a_1, \ldots, a_n) \in D$ e $f : D \longrightarrow \mathbb{R}$. Diz-se que

- $f(a)$ é um mínimo global, ou a é um minimizante global de f, se

$$f(x) \geq f(a), \qquad \forall x \in D;$$

- $f(a)$ é um máximo global, ou a é um maximizante global de f, se

$$f(x) \leq f(a), \qquad \forall x \in D;$$

- $f(a)$ é um extremo global, ou a é um extremante global de f, se a for um minimizante ou um maximizante global de f.

- Relembremos o Teorema de Weierstrass:

  Sejam D um conjunto limitado e fechado de $\mathbb{R}^n$ e $f : D \longrightarrow \mathbb{R}$ uma função contínua.

  Então f assume um valor máximo absoluto e um valor mínimo absoluto em D.

- Sendo D fechado, $D = \text{int}\, D \cup \text{fr}\, D$

  - [int D] é um conjunto aberto: os pontos críticos de f são tais que $\nabla f(a) = 0$.

  - [fr D] os pontos críticos de f podem ser determinados usando o método dos multiplicadores de Lagrange.

Estratégia

- Sendo D um conjunto fechado e limitado, para determinar os extremantes globais de f em D há que

  1. determinar todos os pontos críticos de f em int D: $\nabla f(x) = 0$;
  2. se fr D for a reunião de várias curvas (ou superfícies, ou…), determinar as suas interseções;
  3. determinar os extremantes de f em cada uma das curvas (ou superfícies, ou…) que compõem fr D: método de redução de ordem ou método dos multiplicadores de Lagrange;
  4. calcular o valor de f em todos os pontos encontrados;
  5. classificar os pontos extremantes (maximizante/minimizante) por comparação dos valores de f.

Exemplo

- Pretende-se conhecer os extremos da função definida por $f(x, y) = x^2 - 2xy + 2y$ no retângulo

$$R = \left\{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 3,\ 0 \leq y \leq 2\right\}.$$

O conjunto R é limitado e fechado e a função f é contínua em R. Então, pelo teorema de Weierstrass, a função f assume um valor máximo e um valor mínimo em R.

1. O gradiente de f é $\nabla f(x, y) = (2x - 2y, -2x + 2)$ e os pontos críticos da função são as soluções do sistema $\nabla f(x, y) = (0, 0)$ ou seja, é o ponto $A = (1, 1)$. Tem-se que $A \in R$ e $f(A) = 1$.

2. Para determinar os extremos de f sobre a fronteira de R usa-se o método de redução de ordem pelo que o estudo se reduz ao problema de determinação de extremos de funções reais de uma variável real.

   A fronteira de R é constituída por 4 segmentos de reta

$$
\begin{aligned}
L_1 &: \ y = 0, x \in [0,3] \quad \text{e} \quad f|_{L_1}(x,y) = f(x,0) = x^2 \\
L_2 &: \ x = 3, y \in [0,2] \quad \text{e} \quad f|_{L_2}(x,y) = f(3,y) = 9 - 4y \\
L_3 &: \ y = 2, x \in [0,3] \quad \text{e} \quad f|_{L_3}(x,y) = f(x,2) = x^2 - 4x + 4 \\
L_4 &: \ x = 0, y \in [0,2] \quad \text{e} \quad f|_{L_4}(x,y) = f(0,y) = 2y
\end{aligned}
$$

Sobre L_1, $0 \leq x \leq 3$ e $f(x,0) = x^2$ é uma função monótona que tem valor mínimo em $x = 0$, $f(0,0) = 0$, e valor máximo em $x = 3$, $f(3,0) = 9$. Os candidatos a extremantes de f são $B = (0,0)$, $C = (3,0)$.

Sobre L_2, $0 \leq y \leq 2$ e $f(3,y) = 9 - 4y$ é uma função decrescente pelo que tem o seu máximo em $y = 0$, $f(3,0) = 9$ e o seu mínimo em $y = 2$, $f(3,2) = 1$. Os candidatos a extremantes de f são $C = (3,0)$, $D = (3,2)$.

Sobre L_3, $0 \leq x \leq 3$ e $f(x,2) = x^2 - 4x + 4 = (x-2)^2$ que tem o seu valor mínimo em $x = 2$, $f(2,2) = 0$ e o seu valor máximo em $x = 0$, $f(0,2) = 4$. Os candidatos a extremantes de f são, então, $E = (2,2)$, $F = (0,2)$.

Por fim, sobre L_4, $0 \leq y \leq 2$ e $f(0,y) = 2y$ é uma função monótona com valor mínimo em $y = 0$, $f(0,0) = 0$ e valor máximo em $y = 2$, $f(0,2) = 4$. Os candidatos a extremantes de f são, então, $B = (0,0)$, $F = (0,2)$.

3. Os valores assumidos por f nos pontos anteriores são, então,

$$\begin{aligned} &f(A) = f(1,1) = 1, \qquad && f(B) = f(0,0) = 0, \quad && f(C) = f(3,0) = 9, \\ &f(D) = f(3,2) = 1, \qquad && f(E) = f(2,2) = 0, \quad && f(F) = f(0,2) = 4\,. \end{aligned}$$

4. Observe-se que os pontos de intersecção dos diferentes segmentos de reta que definem a fronteira de R são os pontos B, C, D e F encontrados no Passo 2.

5. Atendendo ao Passo 3., conclui-se que o máximo absoluto de f em R é $f(3,0) = 9$ e que o mínimo absoluto é $f(0,0) = f(2,2) = 0\,$.