

Resumo de Aulas Teóricas de Análise Matemática II

Rui Albuquerque
Universidade de Évora
2011/2012

Aula 1

O espaço euclidiano \mathbb{R}^n :

Espaço vectorial, espaço de pontos, vectores $a = (a_1, \dots, a_n)$,
 $x = (x_1, \dots, x_n)$, etc...

Admite soma de vectores e multiplicação por escalares $\lambda \in \mathbb{R}$. Por exemplo,

$$\lambda v = \lambda(v_1, \dots, v_n) = (\lambda v_1, \dots, \lambda v_n) .$$

Norma euclideana = comprimento de um vector = distância do ponto à origem:

$$\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2} .$$

Propriedades:

$\forall u, v \in \mathbb{R}^n, \lambda \in \mathbb{R},$

$$\|u + v\| \leq \|u\| + \|v\| \quad (\text{desigualdade triangular}),$$

$$\|\lambda v\| = |\lambda| \|v\| ,$$

$$\|v\| = 0 \Leftrightarrow v = 0 .$$

Noções topológicas em \mathbb{R}^n , noção de vizinhança de um ponto.

Bola aberta de centro em $a \in \mathbb{R}^n$ e raio $\epsilon \in \mathbb{R}^+$:

$$B_\epsilon(a) = \{x \in \mathbb{R}^n : \|x - a\| < \epsilon\} .$$

As vizinhanças podem ser dadas por n -intervalos: dado o mesmo ponto $a \in \mathbb{R}^n$, os “cubos”

$$]a_1 - \epsilon, a_1 + \epsilon[\times \cdots \times]a_n - \epsilon, a_n + \epsilon[$$

geram a mesma topologia em \mathbb{R}^n . Claramente, pois, sempre com o mesmo centro, toda a bola contém um cubo e todo o cubo contém uma bola.

Classificação dos pontos e subconjuntos

Dado um subconjunto $S \subset \mathbb{R}^n$ e dado $a \in \mathbb{R}^n$. Este ponto diz-se:

Ponto interior a S : quando a e toda uma sua vizinhança estão contidos em S .

Ponto exterior a S : quando a e toda uma sua vizinhança estão contidos em $\mathbb{R}^n \setminus S$ (complementar de S).

Ponto fronteiro a S : quando qualquer vizinhança de a tem pontos em S e fora de S .

Definem-se em seguida de forma óbvia os subconjuntos de \mathbb{R}^n :

$$\text{int}(S) \qquad \text{ext}(S) \qquad \text{fr}(S)$$

como o interior, o exterior e a fronteira de S .

Por exemplo, $\text{int}(S) = \{\text{pontos interiores a } S\}$.

Por definição, cada par de entre estes três subconjuntos tem intersecção vazia um com o outro.

Claramente,

$$\text{int}(\mathbb{R}^n \setminus S) = \text{ext}(S) \quad \text{e} \quad \text{fr}(\mathbb{R}^n \setminus S) = \text{fr}(S) .$$

S diz-se aberto se $\text{int}(S) = S$.

S diz-se fechado se $\mathbb{R}^n \setminus S$ é aberto. Ou seja, se $S \supseteq \text{fr}(S)$. Ou ainda, o que é o mesmo, $S = S \cup \text{fr}(S)$.

S diz-se limitado se existir um número positivo $L > 0$ tal que $\|x\| \leq L, \forall x \in S$.

S diz-se compacto se for fechado e limitado.

Exercício: Prove que:

- 1) toda a bola aberta é um subconjunto aberto.
- 2) Toda a bola fechada $\overline{B}_\epsilon(a) = B_\epsilon(a) \cup \text{fr}(B_\epsilon(a))$ é um subconjunto compacto de \mathbb{R}^n .

Aula 2

Funções, limites, continuidade

Seja $S \subset \mathbb{R}^n$ e f uma função em S com valores em \mathbb{R}^m :

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

$$\begin{aligned}x = (x_1, \dots, x_n) \quad \longmapsto \quad f(x) = f(x_1, \dots, x_n) &= (y_1, \dots, y_m) \\(y_1, \dots, y_m) &= (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)).\end{aligned}$$

Claramente, f tem *componentes* $f_i : S \rightarrow \mathbb{R}$, $f = (f_1, \dots, f_m)$.

Recordamos que S é o domínio ou conjunto de partida de f .

Recordamos que $\text{Im } f = f(S) = \{f(x) : x \in S\}$ é a imagem de f .

f diz-se injectiva se ... e sobrejectiva se ... (recordar!).

f diz-se limitada se $f(S)$ é um subconjunto limitado.

Seja $a \in \mathbb{R}^n$.

Dizemos que f tem limite $b \in \mathbb{R}^m$ em a se:

$$\forall \delta > 0, \exists \epsilon > 0 : \forall x \in S \cap B_\epsilon(a) \implies f(x) \in B_\delta(b) .$$

Ou seja,

$$\forall \delta > 0, \exists \epsilon > 0 : \forall x \in S \text{ e } \|x - a\| < \epsilon \implies \|f(x) - b\| < \delta .$$

Neste caso escrevemos

$$\lim_{x \rightarrow a} f(x) = b .$$

Proposição

O limite, se existe, é único.

Prova-se, por definição, que, sendo $f, g : S \rightarrow \mathbb{R}^m$ e $\lambda : S \rightarrow \mathbb{R}$ funções em S com limites em a :

$$\lim_{x \rightarrow a} f(x) = b, \quad \lim_{x \rightarrow a} g(x) = c, \quad \lim_{x \rightarrow a} \lambda(x) = l,$$

então existem os limites:

$$\lim_{x \rightarrow a} (f + g)(x) = \lim_{x \rightarrow a} (f(x) + g(x)) = b + c,$$

$$\lim_{x \rightarrow a} (\lambda \cdot f)(x) = \lim_{x \rightarrow a} \lambda(x)f(x) = l \cdot b.$$

Mais ainda, se $l \neq 0$, então existe o limite $\lim_{x \rightarrow a} \frac{1}{\lambda}(x) = \frac{1}{l}$.

Sendo $a \in S$ dizemos que f é contínua em a se existe o limite em a e este é igual a $f(a)$, ou seja,

$$\lim_{x \rightarrow a} f(x) = f(a) .$$

Consequência imediata do anterior:

Proposição

A soma de duas funções contínuas em a é uma função contínua em a .

O produto de uma função contínua em a por uma função escalar contínua em a é uma função contínua em a .

A inversa de uma função contínua e não nula em a é uma função contínua em a .

Função composta e mais sobre limites

Suponhamos agora que $f : S \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma função com limite b em a e que $h : U \subset \mathbb{R}^m \rightarrow \mathbb{R}^k$ é uma função com limite c em b .

(Repare bem nos pontos e nos conjuntos)

Poderemos então falar da função composta:

$$h \circ f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad h \circ f(x) = h(f(x)) .$$

O seu domínio, D , será obviamente o conjunto de pontos de S cuja imagem por f está no domínio de h (isto é mesmo o que faz sentido!).

Agora, é simples demonstrar que c é o limite de $h \circ f$ em a :

$$\lim_{x \rightarrow a} (h \circ f)(x) = c .$$

Proposição

A composição de duas funções como f e h acima, a primeira contínua em a , a segunda contínua em $f(a)$, é uma função contínua em a .

Aula 3

Uma função diz-se contínua (no seu domínio) se for contínua em todos os pontos.

Exercício: Mostre que toda a função racional em pontos de \mathbb{R}^3 , $f(x, y, z) = \frac{p(x, y, z)}{q(x, y, z)}$, com p, q polinómios, é uma função contínua no seu domínio.

Proposição

Se existir um dos limites, têm-se as seguintes equivalências:

$$\lim_{x \rightarrow a} f(x) = b \Leftrightarrow \lim_{x \rightarrow 0} f(a + x) = b \Leftrightarrow \lim_{x \rightarrow a} \|f(x) - b\| = 0 .$$

Poderemos também falar de limites direccionais no ponto a e na direcção de um vector v de uma dada função f :

$$\lim_{t \rightarrow 0} f(a + tv) = b \quad \text{com } t \text{ a variar em } \mathbb{R} .$$

Nota. Há funções que têm todos os limites direccionais (para todo o v) e no entanto não têm limite em a !

É o caso de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definida por

$$f(x, y) = \begin{cases} 1 & \text{se } y \geq x^2 \text{ ou } y \leq 0 \\ 0 & \text{se } y < x^2 \text{ e } y > 0 \end{cases} .$$

Limites laterais ou parciais são os limites na direcção dos eixos canónicos, ou seja na direcção de um vector da base canónica de \mathbb{R}^n .

Nota. Nem sempre existem os limites laterais, podendo existir o limite no ponto. Por exemplo,

$\lim_{x \rightarrow 0} (\lim_{y \rightarrow 0} x \operatorname{sen} y)$ não existe, mas existe o $\lim_{(x,y) \rightarrow (0,0)} x \operatorname{sen} y = 0$.

Derivação e diferenciabilidade I

Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função definida num aberto D e seja $a \in D$.

Define-se a derivada parcial de f no ponto a em ordem a x_i como o limite, se existir,

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a) &= \lim_{x_i \rightarrow a_i} \frac{f(a_1, \dots, x_i, \dots, a_n) - f(a_1, \dots, a_n)}{x_i - a_i} \\ &= \lim_{h \rightarrow 0} \frac{f(a_1, \dots, a_i + h, \dots, a_n) - f(a_1, \dots, a_n)}{h}\end{aligned}$$

(do lado esquerdo está a notação para o limite).

Mais geralmente, falamos de derivada direccional de f no ponto $a \in D$ e na direcção de $v \in \mathbb{R}^n$ quando existe o

$$df(a)(v) = f'(a)(v) = \lim_{h \rightarrow 0} \frac{f(a + hv) - f(a)}{h}$$

(do lado esquerdo estão duas notações usuais para o limite, se existir).

É de facto noção mais geral, pois é claro que

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

onde $e_j = (0, \dots, 0, 1, 0, \dots)$.

Note-se que, fixando $v \in \mathbb{R}^n$, se existirem as derivadas em todos os pontos, voltamos a ter uma nova função $x \mapsto df(x)(v)$.

Podemos então falar em nova derivada ou derivada de segunda ordem de f em x na direcção de $w \in \mathbb{R}^n$ e depois de v :

$$d^2f(x)(w, v) = \lim_{h \rightarrow 0} \frac{df(x + hw)(v) - df(x)(v)}{h}.$$

E, assim por diante, definimos as derivadas de terceira ordem, quarta ordem, ordem p , etc.

No caso das direcções principais, temos as notações:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \quad \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}, \quad \text{etc...}$$

Por agora não vamos dizer mais do que isto sobre as derivadas de segunda ou maiores ordens.

Aula 4

Aplicações lineares

Recordemos agora a noção de *aplicação* linear: uma função

$$A : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

satisfazendo as propriedades, $\forall v, w \in \mathbb{R}^n, \forall \lambda \in \mathbb{R}$,

$$A(v + w) = A(v) + A(w) , \quad A(\lambda v) = \lambda A(v) .$$

Fixemos a base canónica já antes referida.

Por exemplo para \mathbb{R}^n trata-se do sistema de vectores e_1, \dots, e_n onde

$$e_i = (0, \dots, 0, 1, 0, \dots, 0) \quad \text{com 1 no lugar } i.$$

Em \mathbb{R}^m tratar-se-á do sistema de vectores f_1, \dots, f_m onde

$$f_j = (0, \dots, 0, 1, 0, \dots, 0) \quad \text{com 1 no lugar } j,$$

mas este vector, bem entendido, com m entradas.

Como bem se sabe, fixadas as bases, uma aplicação linear A fica *determinada* pela sua matriz $[a_{ij}]$.

Cada vector $A(e_i) \in \mathbb{R}^m$ escreve-se de novo à custa da base canónica f_1, \dots, f_m de \mathbb{R}^m

$$A(e_i) = \sum_{j=1}^m a_{ij} f_j .$$

A matriz *determina* A pois, se $v \in \mathbb{R}^n$ se escreve como $v = (v_1, \dots, v_n) = \sum_{i=1}^n v_i e_i$, então

$$A(v) = A\left(\sum v_i e_i\right) = \sum v_i A(e_i) = \sum_{i,j} v_i a_{ij} f_j .$$

Isto prova em particular o seguinte

Corolário

Toda a aplicação linear é contínua.

Derivação e diferenciabilidade II

Voltemos às derivadas.

É possível dar exemplos de funções $f(x, y)$, $(x, y) \in \mathbb{R}^2$, que têm as duas derivadas parciais $\frac{\partial f}{\partial x}$ e $\frac{\partial f}{\partial y}$ num ponto a mas nesse mesmo ponto a função não é contínua.

Ora, tal situação não é nada conveniente.

Nota. Um exemplo a estudar:

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \quad \text{no ponto } a = (0, 0).$$

Recordemos brevemente a fórmula de Taylor (a que voltaremos mais tarde) para funções reais de variável real: se existir, vale

$$f(x + h) = f(x) + f'(x)h + f''(x)\frac{h^2}{2} + f'''(x)\frac{h^3}{6} + \dots .$$

Repare-se então que o chamado *resto de segunda ordem* tende para 0 com h mesmo quando dividido por h .

Daqui se abstrai a noção de diferenciabilidade conveniente para funções definidas em abertos de \mathbb{R}^n e com valores noutro espaço vectorial \mathbb{R}^m .

Dizemos que uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ definida num aberto D é diferenciável no ponto $x \in D$ se existir uma aplicação linear $A_x : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tal que, escrevendo para um vector v

$$f(x + v) = f(x) + A_x(v) + o(v) ,$$

resulta

$$\lim_{v \rightarrow 0} \frac{o(v)}{\|v\|} = 0 .$$

Se f for diferenciável em x , denotamos

$$A_x(v) = df(x)(v) = f'(x)(v) .$$

A aplicação linear $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ é chamada diferencial de f
no ponto x .

Propriedades imediatas da diferenciabilidade:

- ▶ A notação não é ambígua nem está repetida: se for diferenciável, a função f tem mesmo derivada parcial na direcção de v no ponto x e qualquer que seja v .
- ▶ Por razões anteriores, só existe uma aplicação linear A_x nas condições da definição acima.
- ▶ Vemos agora que $df(x)$ é mesmo uma aplicação linear (isto não é imediato da fórmula dada muito atrás).
- ▶ Uma função diferenciável em x é contínua em x .

Exercício: Prove estas afirmações (como viu durante a aula teórica).

Recordemos o teorema do valor médio ou de Lagrange: para uma função f definida num intervalo $[\alpha, \beta] \subset \mathbb{R}$ e diferenciável nesse intervalo, existe um ponto $x_0 \in]\alpha, \beta[$ tal que

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} = f'(x_0) .$$

É claro, temos o mesmo para a restrição de uma função real e diferenciável num aberto $D \subset \mathbb{R}^n$ que contenha 'intervalos'.

Teorema (do valor médio ou dos acréscimos finitos)

Sejam $a, v \in \mathbb{R}^n$ e $f : D \rightarrow \mathbb{R}$ função diferenciável. Suponhamos que o segmento $\llbracket a, a + v \rrbracket = \{x_t = a + tv \in \mathbb{R}^n : t \in [0, 1]\}$ está contido em D . Então existe um ponto x_{t_0} nesse segmento tal que

$$f(a + v) - f(a) = df(x_{t_0})(v) .$$

Aula 5

Jacobiano

Chama-se matriz jacobiana de $f = (f_1, \dots, f_m)$ no ponto x à matriz de $f'(x) = df(x)$ escrita à custa das bases canónicas de \mathbb{R}^n e \mathbb{R}^m :

$$\text{Jac } f = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} .$$

Havendo necessidade, dizemos qual o ponto em questão assim:
 $\text{Jac}(f)(x)$.

Por exemplo, se $g : \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável em $x = (x_1, \dots, x_n)$ e $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ é um vector qualquer, então temos as identidades:

$$dg_x(v) = \sum_{i=1}^n v_i dg_x(e_i) = \sum_{i=1}^n v_i \frac{\partial g}{\partial x_i}(x)$$

correspondendo a multiplicação de matrizes

$$\text{Jac } g(x) v = \begin{bmatrix} \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = dg_x(v).$$

Regras da diferenciação

Prova-se, por definição, que, sendo $f, g : D \rightarrow \mathbb{R}^m$ e $\lambda : D \rightarrow \mathbb{R}$ funções definidas num aberto $D \subset \mathbb{R}^n$ e diferenciáveis em $a \in D$, então também são diferenciáveis em a as funções $f + g$ e $\lambda.f$.

Tem-se para qualquer $v \in \mathbb{R}^n$

$$(f + g)'(a) = f'(a) + g'(a)$$

e a **regra de Leibniz** (note bem as aplicações lineares em causa)

$$(\lambda.f)'(a)(v) = \lambda'(a)(v).f(a) + \lambda(a).f'(a)(v) .$$

Outras versões do resultado anterior:

$$d(f + g) = df + dg, \quad d(\lambda f) = (d\lambda)f + \lambda df,$$

$$\frac{\partial}{\partial x_i}(f + g) = \frac{\partial f}{\partial x_i} + \frac{\partial g}{\partial x_i}, \quad \frac{\partial(\lambda f)}{\partial x_i} = \frac{\partial \lambda}{\partial x_i} f + \lambda \frac{\partial f}{\partial x_i},$$

ou ainda

$$\text{Jac}(f + g) = \text{Jac } f + \text{Jac } g, \quad \text{Jac}(\lambda f) = \begin{bmatrix} f_1 \\ \vdots \\ f_m \end{bmatrix} \text{Jac } \lambda + \lambda \text{Jac } f.$$

Regra da derivação da função composta:

Suponhamos agora que $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ é uma função diferenciável em a , que $f(a)$ pertence a um aberto $\tilde{D} \subset \mathbb{R}^m$ e que $h : \tilde{D} \rightarrow \mathbb{R}^k$ é uma função diferenciável em $f(a)$.

Então prova-se que $h \circ f$ é diferenciável em a e que, $\forall v \in \mathbb{R}^n$,

$$d(h \circ f)(a)(v) = dh(f(a))(df(a)(v)) .$$

Por outras palavras,

$$(h \circ f)'(a)(v) = h'(f(a))(f'(a)(v)) .$$

O diferencial $(h \circ f)'(a) : \mathbb{R}^n \rightarrow \mathbb{R}^k$ verifica portanto

$$(h \circ f)'(a) = h'(f(a)) \circ f'(a) .$$

E em termos matriciais, recordando que a composição de aplicações lineares dá lugar à multiplicação das respectivas matrizes que as representam, tem-se

$$\text{Jac}(h \circ f)(a) = \text{Jac}(h)(f(a)) \text{Jac} f(a) .$$

Uma função diz-se diferenciável (no seu domínio D , aberto) se for diferenciável em cada um dos pontos de D .

Exemplo 1:

Seja $g : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ uma função diferenciável. Como habitual, usamos as coordenadas $(x, y, z) \in \mathbb{R}^3$.

Seja $f : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$ uma função diferenciável num intervalo I , uma curva no espaço que admitimos passando em D . Claro que $f(t) = (f_1(t), f_2(t), f_3(t))$.

Então $g \circ f : I \rightarrow \mathbb{R}$ é diferenciável e teremos,

pela regra da derivação da função composta:

$$\frac{d}{dt}(g \circ f) = \frac{\partial g}{\partial x} \frac{df_1}{dt} + \frac{\partial g}{\partial y} \frac{df_2}{dt} + \frac{\partial g}{\partial z} \frac{df_3}{dt}.$$

Se quisermos mais pormenor, o que se tem é:

$$\begin{aligned}(g \circ f)'(t) &= \\ &= \frac{\partial g}{\partial x}(f(t)) \frac{df_1}{dt}(t) + \frac{\partial g}{\partial y}(f(t)) \frac{df_2}{dt}(t) + \frac{\partial g}{\partial z}(f(t)) \frac{df_3}{dt}(t).\end{aligned}$$

Repare-se na multiplicação de matrizes:

$$\text{Jac}(g \circ f)_t = \text{Jac}(g)_{f(t)} \text{Jac}(f)_t.$$

Exemplo 2:

Seja $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ e $g : \mathbb{R}^2 \rightarrow \mathbb{R}^5$.

Suponhamos $h(x, y)$ função de (x, y) e $g(a_1, a_2)$ função de (a_1, a_2) .

Então, por exemplo, a derivada em ordem a x da quarta componente de g composta com h : $g_4(h_1(x, y), h_2(x, y))$:

$$\frac{\partial g_4 \circ h}{\partial x} = \frac{\partial g_4}{\partial a_1} \frac{\partial h_1}{\partial x} + \frac{\partial g_4}{\partial a_2} \frac{\partial h_2}{\partial x} .$$

Aula 6

Derivação e diferenciabilidade III

O seguinte critério permite decidir quando é que uma dada função é diferenciável:

Teorema

Uma função $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é diferenciável em D se admitir as n derivadas parciais $\frac{\partial f}{\partial x_i}$ e $n - 1$ delas forem contínuas em D .

Fazemos aqui a demonstração pois esta não foi feita na aula teórica:

Usamos o método de indução em n . Para $n = 1$, está provado porque a diferenciabilidade é o mesmo que a existência de derivada.

Suponhamos por hipótese que o resultado foi provado até dimensão $n - 1$ e tratemos de provar para dimensão n . Seja $a \in D$ e $v \in \mathbb{R}^n$. Como existem as derivadas parciais, definimos o diferencial de f :

$$df(a)(v) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}.$$

Vejamos agora a condição de diferenciabilidade: $f(a + v) - f(a) - df(a)(v) = o(v)$ é um infinitésimo com v , ou seja $\lim o(v)/\|v\| = 0$ quando $v \rightarrow 0$.

Façamos a decomposição $v = (v_1, \dots, v_n) = u + v_n e_n$ onde $u = (v_1, \dots, v_{n-1}, 0)$. Podemos supôr, sem perda de generalidade, que é $\frac{\partial f}{\partial x_n}$ que é contínua. Então a restrição de f a cada um dos hiperplanos em que a_n está fixo é uma função diferenciável pela hipótese de indução (tem $n - 1 - 1$ derivadas parciais contínuas). Agora, somando e subtraindo $f(a + u)$ e usando a linearidade do diferencial, temos

$$o(v) = f(a + u + v_n e_n) - f(a + u) - v_n \frac{\partial f}{\partial x_n}(a) + f(a + u) - f(a) - df(a)(u) .$$

Pelo teorema dos acréscimos finitos no ponto $a + u$ e na direcção de e_n , conclui-se que existe um $\theta \in]0, 1[$ tal que o anterior fica igual a

$$\begin{aligned} &= v_n \frac{\partial f}{\partial x_n}(a + u + \theta v_n e_n) - v_n \frac{\partial f}{\partial x_n}(a) + f(a + u) - f(a) - df(a)(u) \\ &= v_n \left(\frac{\partial f}{\partial x_n}(a + u + \theta v_n e_n) - \frac{\partial f}{\partial x_n}(a) \right) + o(u) \end{aligned}$$

Reparando que $\|v\| = \sqrt{\|u\|^2 + v_n^2}$, lembrando que $\partial_n f$ é contínua e a hipótese de indução, torna-se fácil

concluir que $\lim_{v \rightarrow 0} \frac{o(v)}{\|v\|} = 0$.

Pelo teorema, tornou-se muito fácil decidir se uma função é diferenciável ou não: basta reparar nas derivadas parciais!!!

Já falámos de derivadas de ordem superior a 1 (Aula 3).
Uma utilização criteriosa do teorema dos acréscimos finitos (como a que se usou na demonstração acima) permite mostrar o teorema seguinte.

Primeiro, o diferencial de ordem p no ponto x define-se como:

$$d^p f(x)(v_1, \dots, v_p) = d\left(d^{p-1}f(x)(v_1, \dots, v_{p-1})\right)(v_p) ,$$

onde $v_1, \dots, v_{p-1}, v_p \in \mathbb{R}^n$ são as várias direcções - vectores - em que repetidamente se diferencia, na variável x apenas, a função f . (nota: do lado esquerdo está a notação para a derivada da derivada de ordem $p - 1$ do lado direito, se existir.)

Surge então o seguinte resultado de grande economia de recurso.

Teorema (da igualdade das derivadas mistas)

Se $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é duas vezes diferenciável, então para quaisquer direcções u, v , tem-se

$$d^2f(x)(u, v) = d^2f(x)(v, u) .$$

Exercício: Estude a simetria do diferencial de 2ª ordem pela definição e repare que a demonstração do teorema não é imediata.

Tomando $u = e_i$ e $v = e_j$ vectores da base canónica, o resultado acima traduz-se em:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

ou ainda

$$\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} = \frac{\partial^3 f}{\partial x_k \partial x_j \partial x_i} .$$

Exercício: 1) Explique a segunda igualdade acima.

2) Experimente a primeira com $f(x, y) = x \cos(xy)$.

Aula 7

Eis finalmente os conjuntos de funções num aberto D que mais vamos estudar (notação do lado esquerdo):

$$C_D^0 = \{f : D \rightarrow \mathbb{R} : f \text{ é contínua em } D\} .$$

$$C_D^1 = \{f : \exists \text{ derivadas parciais } \frac{\partial f}{\partial x_i} \text{ e são contínuas em } D\} .$$

\vdots

$$C_D^k = \{f : \exists \text{ derivadas parciais } \frac{\partial f^k}{\partial x_{i_1} \cdots \partial x_{i_k}} \text{ contínuas em } D\} .$$

Função de classe C^k em D é a que pertence a C_D^k .

Função de classe C^∞ é a que é de classe C^k para todo o $k \in \mathbb{N}$.

Claro que $C^\infty \subset \dots \subset C^k \subset \dots \subset C^2 \subset C^1 \subset C^0$.

Por teoremas vistos acima, deduzimos

Teorema

C_D^k é um espaço vectorial sobre \mathbb{R} .

Exercício: Utilize indução matemática para provar que a composição de aplicações de classe C^k é de classe C^k .

Produto interno euclidiano, gradiente

A norma euclidiana vista na Aula 1 surge de um produto interno, de uma *métrica*.

Recordemos da Álgebra Linear que se define o produto interno euclidiano entre dois vectores $u = (x_1, \dots, x_n)$ e $v = (y_1, \dots, y_n)$ de \mathbb{R}^n como o valor real

$$\langle u, v \rangle = \langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = x_1 y_1 + \dots + x_n y_n .$$

Trata-se de uma aplicação $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ bilinear, simétrica e não-degenerada, ou seja, $\forall u, v, w \in \mathbb{R}^n, \lambda \in \mathbb{R}$:

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$$

$$\langle \lambda u, v \rangle = \langle u, \lambda v \rangle = \lambda \langle u, v \rangle$$

$$\langle u, v \rangle = \langle v, u \rangle$$

$$\langle u, u \rangle \geq 0 \text{ com igualdade se e só se } u = 0 .$$

Exercícios: 1) Provar as propriedades de $\langle \cdot, \cdot \rangle$.

2) Para qualquer $v \in \mathbb{R}^n$ fixo, é linear a aplicação $h : \mathbb{R}^n \rightarrow \mathbb{R}$ definida por $h(x) = \langle v, x \rangle$.

É claro que

$$\langle u, u \rangle = \|u\|^2 .$$

Dito de outra forma, $\|u\| = \sqrt{\langle u, u \rangle}$.

Vectores $u, v \in \mathbb{R}^n$ tais que $\langle u, v \rangle = 0$ dizem-se ortogonais.

Temos sempre a chamada desigualdade de Cauchy-Bunyakovski-Schwarz: $|\langle u, v \rangle| \leq \|u\| \|v\|$, com igualdade se e só se os dois vectores são linearmente dependentes. A demonstração resulta do estudo da função em λ definida por $\|u + \lambda v\|^2$ sabendo-se de antemão que é sempre positiva.

Temos assim que o número $\cos \angle(u, v) = \frac{\langle u, v \rangle}{\|u\| \|v\|}$ está entre -1 e 1 , o qual de facto coincide com o coseno.

Nota. Estas propriedades são provadas em curso de Álgebra Linear e Geometria Analítica.

Fórmula muito prática:

$$\langle u, v \rangle = \|u\| \|v\| \cos \angle(u, v).$$

Tal como a seguinte, voltando ao início, em multiplicação de matrizes:

$$\langle u, v \rangle = [x_1 \quad \cdots \quad x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = u^T v$$

Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função diferenciável num aberto D . Chamamos gradiente de f no ponto x ao vector

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

Um ponto $a \in D$ diz-se um ponto crítico de f se $df(a) = 0$, ou seja, se $\frac{\partial f}{\partial x_i}(a) = 0, \forall 1 \leq i \leq n$. O mesmo é dizer

$$a \text{ é um ponto crítico de } f \iff \nabla f(a) = 0 .$$

Como se vê, o gradiente determina o diferencial de f :

$$df(x)(v) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(x) = \langle \nabla f(x), v \rangle .$$

Nota. O vector gradiente dá-nos outra forma de olhar para df , isto é, para a matriz-linha $\text{Jac } f$.

Exercício: Suponha f diferenciável numa bola e que $df(x) = 0$ em todos os pontos. Mostre que f é uma constante.

Sugestão: teorema de Lagrange.

Aula 8

Interpretação geométrica, função implícita

Começemos com um exemplo em dimensão 1:

Seja $f(x)$ uma função C^1 real de variável real num intervalo $D \subset \mathbb{R}$.

O seu gráfico é dado como o conjunto de pontos de $\mathbb{R}^{1+1} = \mathbb{R}^2$
 $G_f = \{(\overline{x, f(x)}) : x \in D\}$. A recta tangente ao gráfico, em cada ponto, é gerada por $T = (1, f'(x))$.

O gradiente da função $(x, w) \xrightarrow{\Phi} w - f(x)$ é o vector
 $\nabla\Phi = (-f'(x), 1)$. Tem-se

$$\langle T, \nabla\Phi \rangle = \langle (1, f'(x)), (-f'(x), 1) \rangle = -f'(x) + f'(x) = 0 .$$

Em conclusão, o gradiente de Φ é ortogonal ao gráfico de f .

Novo exemplo, em dimensão 3:

Seja $f(x, y, z)$ uma função C^1 real e definida num aberto $D \subset \mathbb{R}^3$.

O seu gráfico é igual ao conjunto de pontos de \mathbb{R}^{3+1} dado por

$$G_f = \{(x, y, z, f(x, y, z)) : (x, y, z) \in D\}.$$

O *espaço tangente* ao gráfico, em cada ponto, é gerado por três vectores:

$$T_1 = \left(1, 0, 0, \frac{\partial f}{\partial x}\right), \quad T_2 = \left(0, 1, 0, \frac{\partial f}{\partial y}\right), \quad T_3 = \left(0, 0, 1, \frac{\partial f}{\partial z}\right).$$

Para se perceber este espaço, tomem-se curvas no gráfico e suas velocidades: T_1 surge como a *velocidade* da curva

$$x \mapsto (x, 0, 0, f(x, y, z)),$$

T_2 surge como a velocidade da curva $y \mapsto (0, y, 0, f(x, y, z))$, etc.

O gradiente da função $\Phi(x, y, z, w) = w - f(x, y, z)$ é o vector $\nabla\Phi = (-\nabla f, 1)$. Tem-se então (exercício), $\forall i = 1, 2, 3$,

$$\langle T_i, \nabla\Phi \rangle = 0 .$$

Dito de outra forma, $d\Phi_{(x,y,z,w)}(T_i) = 0$.

Em conclusão, o gradiente de Φ é ortogonal ao gráfico de f .

Nos dois exemplos anteriores repare-se que G_f é sempre o conjunto dos zeros de Φ : $\{(x, w) \in D \times \mathbb{R} : \Phi(x, w) = 0\}$, um conjunto de soluções de $\Phi(x, w) = 0$.

Em geral, suponhamos que é dada $\Phi : V \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ de classe C^1 . Um conjunto de nível c (chamado superfície de nível caso $n = 2$ ou curva de nível caso $n = 1$) é definido como

$$M_c = \{(x_1, \dots, x_{n+1}) : \Phi(x_1, \dots, x_{n+1}) = c\} .$$

Suponhamos que alguma derivada parcial $\frac{\partial \Phi}{\partial x_{j_0}} \neq 0$. Então, pelos exemplos acima, podemos afirmar que o espaço tangente a M_c em cada ponto é dado pelos vectores T_j tais que $\langle T_j, \nabla \Phi \rangle = 0$.

Nota. Em linguagem da Geometria Analítica: o espaço tangente é o núcleo do diferencial, $T_x(M_c) = \text{Nuc } d\Phi_x$.

Como conclusão, de novo: o gradiente de Φ é ortogonal a M_c .

É também possível recuperar a função implícita.

Teorema (da função implícita em dimensão $n + 1$)

Suponhamos que é dada Φ como acima. Suponhamos, sem perda de generalidade, que num certo ponto (a_1, \dots, a_n, b) do domínio de Φ temos

$$\Phi(a_1, \dots, a_n, b) = c \text{ e } \frac{\partial \Phi}{\partial x_{n+1}}(a_1, \dots, a_n, b) \neq 0 .$$

Então existe uma vizinhança $D \subset \mathbb{R}^n$ de (a_1, \dots, a_n) e uma função $f : D \rightarrow \mathbb{R}$ de classe C^1 tal que

$$f(a_1, \dots, a_n) = b \text{ e } \Phi(x_1, \dots, x_n, f(x_1, \dots, x_n)) = c .$$

Mais ainda, pela regra da derivação da função composta, temos

$$0 = \frac{\partial \Phi}{\partial x_i} + \frac{\partial \Phi}{\partial x_{n+1}} \frac{\partial f}{\partial x_i} .$$

Ou seja, $\forall i \leq n$,

$$\frac{\partial f}{\partial x_i} = -\frac{\partial \Phi}{\partial x_i} / \frac{\partial \Phi}{\partial x_{n+1}} .$$

Exemplo: $\Phi(x, y, z) = x^2 + y^2 + z^2$. Então $\frac{\partial \Phi}{\partial z} = 2z \neq 0$ para $z \neq 0$.

Claro que a função implícita só existe para $\Phi = c \geq 0$, aí sendo igual a $z = \sqrt{c - x^2 - y^2}$ em torno da solução $(0, 0, \sqrt{c})$.

Aula 9

Teoremas da função implícita e da função inversa

Continuando a aula anterior, suponhamos que em vez de uma temos duas equações

$$\phi_1(x_1, \dots, x_{n-1}, x_n, x_{n+1}) = 0, \quad \phi_2(x_1, \dots, x_{n-1}, x_n, x_{n+1}) = 0$$

e que temos uma solução de ambas: $\phi(a_1, \dots, a_{n-1}, b_1, b_2) = 0$. Poderemos então tentar escrever x_{n+1} à custa das n primeiras variáveis, usando ϕ_1 ou ϕ_2 , e depois talvez escrever x_n à custa das variáveis que sobram.

É para tal necessário que as duas equações sejam *mesmo* funcionalmente dependentes de x_n e x_{n+1} , e independentes uma da outra. Prova-se - não o faremos neste curso - que tal condição é garantida por:

$$\det \begin{bmatrix} \frac{\partial \phi_1}{\partial x_n} & \frac{\partial \phi_1}{\partial x_{n+1}} \\ \frac{\partial \phi_2}{\partial x_n} & \frac{\partial \phi_2}{\partial x_{n+1}} \end{bmatrix} \neq 0 .$$

Continuando o exemplo anterior, suponhamos que além de $x^2 + y^2 + z^2 = c \geq 0$ temos $2xy - z^2 = 0$, então vem $(x + y)^2 = c$ e daí resulta, por exemplo, $y = \sqrt{c} - x$ e $z = \sqrt{2x(\sqrt{c} - x)}$ em torno da solução $(\frac{\sqrt{c}}{2}, \frac{\sqrt{c}}{2}, \sqrt{\frac{c}{2}})$.

Generalizando, temos o

Teorema (da função implícita em dimensão $n + m$)

Suponhamos que é dada $\Phi : V \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ de classe C^1 num aberto V e que num certo ponto $(a_1, \dots, a_n, b_1, \dots, b_m)$ temos

$$\Phi(a_1, \dots, a_n, b_1, \dots, b_m) = 0 \quad \text{e} \quad \det \left[\frac{\partial \Phi_i}{\partial x_{j+n}} \right]_{i,j=1, \dots, m} \neq 0$$

(determinante calculado só no ponto em questão). Então existe uma vizinhança $D \subset \mathbb{R}^n$ de (a_1, \dots, a_n) e uma função $f : D \rightarrow \mathbb{R}^m$ de classe C^1 tal que

$$f(a_1, \dots, a_n) = (b_1, \dots, b_m) \quad \text{e} \quad \Phi(x_1, \dots, x_n, f(x_1, \dots, x_n)) = 0 .$$

Do teorema anterior aplicado a $\phi(y, x) = y - f(x)$ resulta logo o

Teorema (da função inversa)

Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ uma função de classe C^1 tal que num certo ponto a do aberto D verifica:

$$\det(\text{Jac } f(a)) \neq 0 .$$

Então existe uma vizinhança de a onde f é invertível com inversa f^{-1} de classe C^1 .

Derivando a igualdade $f^{-1} \circ f(x) = x$ pela regra da derivada da função composta do lado esquerdo e por definição do lado direito, temos, para todo o vector v ,

$$df^{-1}(y) \circ df(x)(v) = v ,$$

onde $y = f(x)$, ou seja

$$df^{-1}(y) = (df(x))^{-1} .$$

Para a representação matricial, vem imediatamente

$$\text{Jac } f^{-1}(y) = (\text{Jac } f(x))^{-1} .$$

Exemplo

Recorde a função exponencial $e^x : \mathbb{R} \rightarrow \mathbb{R}^+$ e sua inversa, que se sabe bem por que existe. A inversa é \log , verificando assim $e^{\log y} = y$ e $\log e^x = x$.

A derivada da exponencial é conhecida: $d(e^x)(v) = e^x v = w$.
Então a derivada da inversa é

$$d(\log y)(w) = v = \frac{1}{e^x} w = \frac{1}{y} w .$$

Exercício: Encontre um intervalo onde $y = \sin x$ é invertível e prove que $(\arcsin y)' = \frac{1}{\sqrt{1-y^2}}$.

Aula 10

Extremos relativos

Quase sempre daqui em diante $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ será uma função de classe C^∞ com valores reais.

$a \in D$ diz-se um ponto onde f tem um máximo (local) ou, por abuso de linguagem, diz-se um máximo (local) de f se

$$\exists \epsilon > 0 : f(x) \leq f(a), \quad \forall x \in B_\epsilon(a) .$$

Note-se que o valor $f(a)$ é que é um máximo do subconjunto $f(B_\epsilon) \subset \mathbb{R}$.

Se a condição for $f(x) \leq f(a), \quad \forall x \in D$, então o ponto a passa por ser um máximo global de f .

Da mesma forma se define mínimo (local) de f : um ponto $a \in D$ para o qual $\exists \epsilon > 0 : f(x) \geq f(a), \forall x \in B_\epsilon(a)$.

Mas é o valor $f(a)$ que é mínimo!

a é mínimo global se a condição for $f(a) \leq f(x), \forall x \in D$.

Os quatro tipos acima são chamados de extremos relativos.

Falando de mínimos e máximos, em geral, falamos de extremos locais.

Nota. falamos de máximos ou mínimos estritos se eles forem isolados. Ou seja, para o máximo, por exemplo, na definição acima tem-se $f(x) < f(a), \forall x \neq a$.

Da definição de diferenciabilidade e do facto de uma aplicação linear não nula de \mathbb{R}^n em \mathbb{R} tomar sempre, perto da origem 0, valores positivos e negativos (porquê?), resulta logo o

Teorema

Se a é um extremo relativo, então $df(a) = 0$.

Assim, os extremos estão entre os pontos críticos de f (cf. Aula 7, recordemos que $df(a) = 0 \Leftrightarrow \nabla f(a) = 0 \Leftrightarrow f'(a) = 0$).

Por f ser de classe C^2 , podemos falar da fórmula de Taylor de ordem 2.

Eis como se obtém de forma rápida:

Primeiro, fixados $a \in D$, $v \in \mathbb{R}^n$, repare-se que derivando em ordem a $t \in \mathbb{R}$ a seguinte função temos logo (pela regra da derivada da função composta e por ser $\frac{d}{dt}(a + tv) = v$):

$$\frac{d^2}{dt^2}(f(a + tv)) = \frac{d}{dt}(f'(a + tv)(v)) = f''(a + tv)(v)(v)$$

Em particular, em $t = 0$, a primeira derivada em ordem a t é mesmo $f'(a)(v)$ e a segunda derivada é $f''(a)(v)(v)$. Posto isto, o resto $o_2(a, v, t)$ em

$$f(a + tv) = f(a) + tf'(a)(v) + \frac{t^2}{2}f''(a)(v)(v) + o_2$$

há-de verificar $\lim_{t \rightarrow 0} \frac{o_2(a, v, t)}{t^2} = 0$.

(Para $t = 1$ e v suficientemente pequeno resulta a próxima identidade)

Como dizíamos,

$$f(a + v) = f(a) + f'(a)(v) + \frac{1}{2}f''(a)(v, v) + o_2$$

com $\frac{o_2(v)}{\|v\|^2} \rightarrow 0$ quando $v \rightarrow 0$.

Notas. É muito fácil provar que o diferencial de 2^a ordem $f''(x)(u, v)$ é linear em cada uma das entradas u ou v (cf. Aula 6 donde aliás se prova o mesmo em geral para o diferencial de ordem p).

Se virmos $v = (v_1, \dots, v_n) \in \mathbb{R}^n$, a expressão de $f''(x)(v, v)$ é um polinómio nos v_i de grau 2, pelo que $f''(x)(v, v)$ é uma *forma quadrática*.

Situemo-nos agora numa vizinhança $B_\epsilon(a)$ dum ponto crítico a .
Como $f'(a) = 0$, temos

$$f(a + v) = f(a) + \frac{1}{2}f''(a)(v, v) + o_2$$

e assim, para ϵ suficientemente pequeno, o que determina se $f(a + v) \geq f(a)$ ou $f(a + v) \leq f(a)$ numa vizinhança de a , isto é com $\|v\| < \epsilon$, é mesmo o sinal de $f''(a)(v, v)$ — desde que este número não se anule (mas há excepções!).

Suponhamos agora que nalguma base $\tilde{e}_1, \dots, \tilde{e}_n$ de \mathbb{R}^n , a matriz de $f''(a)$ era diagonal:

$$\tilde{H} = [f''(a)(\tilde{e}_i, \tilde{e}_j)] = \begin{bmatrix} \lambda_1 & 0 & & 0 \\ 0 & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

É claro que, $\forall v \in \mathbb{R}^n$, $v = x_1 \tilde{e}_1 + \dots + x_n \tilde{e}_n$ para certos escalares $x_i \in \mathbb{R}$.

Então

$$f''(a)(v, v) = \sum_{i,j=1}^n x_i x_j f''(a)(\tilde{e}_i, \tilde{e}_j) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2 .$$

e, da fórmula de Taylor,

$$f(a + v) = f(a) + \frac{1}{2}(\lambda_1 x_1^2 + \cdots + \lambda_n x_n^2) + o_2$$

pelo que

- se $\lambda_i > 0$, $\forall i = 1, \dots, n$, conclui-se que $f(a)$ é um mínimo estrito
- se $\lambda_i < 0$, $\forall i = 1, \dots, n$, conclui-se que $f(a)$ é um máximo estrito

- se, $\forall i = 1, \dots, n$, $\lambda_i \neq 0$, alguns de sinal $+$ e outros de sinal $-$, então o ponto a é chamado de ponto-sela.

Hessiana, o teorema dos extremos relativos

Chamamos matriz hessiana de f à matriz das segundas derivadas:

$$H = H(f)(a) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right].$$

Considerem-se dois vectores de \mathbb{R}^n

$$u = (u_1, \dots, u_n) = \sum_i u_i e_i, \quad v = (v_1, \dots, v_n) = \sum_i v_i e_i.$$

Por lineariedade,

$$d^2 f(a)(u, v) = \sum_{i,j=1}^n u_i v_j f''(a)(e_i, e_j) = \sum_{i,j=1}^n u_i v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

Em termos matriciais,

$$\begin{aligned} d^2f(a)(u, v) &= \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ & \ddots & \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\ &= u^T(Hv) \end{aligned}$$

Como H é simétrica (pela igualdade das derivadas mistas), denotando por \cdot^T a transposta, temos $H = H^T$ e logo

$$u^T(Hv) = \langle u, Hv \rangle = \langle Hu, v \rangle = (Hu)^T v.$$

Vamos chamar determinantes principais aos números reais Δ_k que resultam de tomarmos os determinantes das matrizes quadradas de H que vão crescendo com a diagonal. Ou seja,

$$\Delta_k = \det \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & & \frac{\partial^2 f}{\partial x_k \partial x_1} \\ & \ddots & \\ \frac{\partial^2 f}{\partial x_1 \partial x_k} & & \frac{\partial^2 f}{\partial x_k^2} \end{bmatrix}$$

com $k = 1, \dots, n$.

Por exemplo, $\Delta_1 = \frac{\partial^2 f}{\partial x_1^2}$ e $\Delta_n = \det H$.

O sinal destes números é invariante por condensação simétrica!

Teorema

Seja $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função C^2 num aberto D e seja $a \in D$ um ponto crítico de f .

- Se na sucessão $1, \Delta_1, \Delta_2, \dots, \dots, \Delta_n$ todos têm sinal positivo, então a é um mínimo de f .
- Se na sucessão $1, \Delta_1, \Delta_2, \dots, \dots, \Delta_n$ os sinais estão sempre a alterar (+ - + - ...), então a é um máximo de f .
- Se na sucessão $1, \Delta_1, \Delta_2, \dots, \dots, \Delta_n$ aparecem ambas repetição e alternância de sinal, então a é um ponto-sela de f .
- Se algum $\Delta_i = 0$, não devemos concluir nada.

Demonstração.

Queremos primeiro ver que H é diagonalizável. Isto é, conseguimos encontrar matriz $n \times n$ de mudança de base, denotada E , tal que $E^T H E = \tilde{H}$ como suposemos atrás, ou seja \tilde{H} diagonal. Com efeito, nesse caso encontraremos uma base de vectores próprios associados aos valores próprios λ_j e conseguiremos decidir dos extremos.

Confirme-se, como condição necessária, que H tem valores próprios reais. Para isso lembremos que todo o polinómio tem uma raiz complexa. Logo o polinómio característico de H tem uma raiz $\alpha + \sqrt{-1}\beta \in \mathbb{C}$. Seja $u + \sqrt{-1}v$ um vector próprio de H no espaço vectorial $\mathbb{R}^n \oplus \sqrt{-1}\mathbb{R}^n$. Significa então que $H(u + \sqrt{-1}v) = (\alpha + \sqrt{-1}\beta)(u + \sqrt{-1}v)$. Como H é real, as partes reais e imaginária equacionam-se:

$$Hu = \alpha u - \beta v \quad \text{e} \quad Hv = \alpha v + \beta u .$$

Da igualdade $\langle Hu, v \rangle = \langle u, Hv \rangle$ vem também $\alpha \langle u, v \rangle - \beta \|v\|^2 = \beta \|u\|^2 + \alpha \langle u, v \rangle$. Logo $\beta = 0$. Porque u e v não são ambos nulos.

Assim descobrimos pelo menos um, eventualmente dois vectores próprios reais associados a $\alpha \in \mathbb{R}$ (podem u e v ser o mesmo)!

$$Hu = \alpha u \quad \text{e} \quad Hv = \alpha v .$$

Agora recordemos o processo de condensação. Na triangularização de uma matriz simétrica como H , para chegar à diagonalização podemos fazê-lo a par de mais um processo que não lhe quebra a simetria. A cada passo do processo de operações elementares sobre linhas fazêmo-lo repetir sobre colunas respectivas. E se algum elemento da diagonal principal for nulo, então ora se troca com a linha em baixo que não tenha na mesma coluna essa entrada nula, fazendo também a troca das colunas respectivas, ora, sendo todas essas entradas nulas, passa-se à entrada principal seguinte, à direita e em baixo, onde poderá já haver entrada não nula. Assim temos o processo

$$H \rightsquigarrow E_1 H E_1^T \rightsquigarrow E_2 (E_1 H E_1^T) E_2^T \rightsquigarrow \dots \rightsquigarrow \tilde{H} = \text{matriz diagonal} .$$

A cada passo, não só se mantém a simetria como se mantêm os sinais dos determinantes principais Δ_j . Vejamos: é claro que somar linhas ou colunas não altera os determinantes. Para além disso, a cada passo, os determinantes principais de por exemplo $E_2HE_2^T$, que altera por suposto a linha e coluna 2, verificam

$$|E_2HE_2^T| = |E_2||H||E_2^T| = |E_2|^2|H|.$$

Isto porque o determinante de um produto de matrizes é o produto dos determinantes e o determinante da matriz transposta é igual ao da matriz dada. Como $|E_2|^2 > 0$ e E_1 não altera nem a linha 1 nem as que estão abaixo da linha 2, está explicado por que se matêm os sinais de Δ_j . Para outros processos a prova é análoga. Por exemplo, os sinais de $\Delta_1, \Delta_2, \Delta_3$ são sempre +, -, - na seguinte cadeia:

$$\begin{bmatrix} 2 & 4 & -1 \\ 4 & 0 & 3 \\ -1 & 3 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 2 & 0 & -1 \\ 0 & -8 & 5 \\ -1 & 5 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & -8 & 5 \\ 0 & 5 & -\frac{1}{2} \end{bmatrix} \rightsquigarrow \begin{bmatrix} 2 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & \frac{21}{8} \end{bmatrix}.$$

Por último lembremos que ao condensar a matriz para encontrar a base de vectores próprios, só podemos utilizar as transformações sobre linhas (para não fazer o erro do tipo $2v + 3w = 5*$). Então temos uma base $\beta = \{u_1, \dots, u_n\}$ de \mathbb{R}^n tal que $EHu_i = \sum_{j \leq i} x_{ji}u_j$. Ou seja, $EH = X$ com X nula por baixo da diagonal principal (matriz triangular superior). Mas note-se que E , pelas definições, é de dois tipos: ou triangular inferior ou simétrica. No primeiro caso, a base $E^T\beta$ é uma base de vectores próprios, pois repare-se: temos $EHE^T = XE^T$ simétrica, por outro lado o produto de duas triangulares superiores é triangular superior. Pelo que XE^T só pode ser matriz diagonal! O segundo caso, que é quando a linha e coluna i se trocam respectivamente pelas linha e coluna $j > i$, corresponde sem outro perigo (!) a trocar o vector u_i pelo u_j na base. Tal transformação E tem $\det < 0$ mas a sua inversa é $E^{-1} = E^T = E$, ela própria, e portanto também não influi nos Δ_j como se viu acima. No final da condensação à esquerda estamos sempre na situação do primeiro caso. E logo a respectiva mudança de base dá-nos a base de vectores próprios que queríamos.

Finalmente, de termos uma base de vectores próprios de H com a qual f se escreve no ponto a essencialmente como $f(x_1, \dots, x_n) = f(a) + \lambda_1 x_1^2 + \dots + \lambda_n x_n^2$, como vimos atrás, resulta que:

- $f(a)$ é um máximo se todos os $\lambda_i < 0$, ou seja, se os determinantes principais $\Delta_1, \Delta_2, \dots$ da matriz H , a hessiana de f , a matriz inicial, vão alternando o sinal.
- $f(a)$ é um mínimo se todos os $\lambda_i > 0$, isto é, todos os $\Delta_i > 0$.
- a é um ponto-sela se nenhum dos λ_i é nulo, algum $\lambda_i < 0$ e outro $\lambda_j > 0$; tal que significa que na cadeia dos Δ_i vão aparecer repetições e alterações de sinal.
- finalmente se v é um vector próprio associado a um $\lambda_i = 0$, então é o resto de ordem 2 na fórmula de Taylor na direcção de v que há-de dizer da natureza do ponto crítico.

As condições recíprocas das acima também valem, porque o polinómio característico de H é invariante por transformação de bases. Mais precisamente, dadas duas bases de \mathbb{R}^n notemos Q a matriz de mudança de base. Então $\det(H - \lambda 1_n) = \det(Q(H - \lambda 1_n)Q^{-1}) = \det(QHQ^{-1} - \lambda 1_n)$. Pelo que λ é valor próprio de H se e só se é valor próprio de QHQ^{-1} .
(Em Álgebra Linear diríamos: um valor próprio é um invariante da representação matricial de uma dada aplicação linear.)

Exemplo: Seja $f(x, y, z) = x^2 + 4xy - xz + 3yz$. O único ponto crítico é $(0, 0, 0)$. A hessiana é a matriz que aparece na demonstração acima. A origem é assim um ponto-sela pois na cadeia dos determinantes principais $1, 2, -16, -42$ temos repetição e alternância de sinal.

Aula 11

Extremos condicionados, multiplicadores de Lagrange

Suponhamos que temos duas funções reais de classe C^1 definidas num aberto $D \subset \mathbb{R}^n$.

$$f, \phi : D \longrightarrow \mathbb{R}$$

Seja M a superfície de nível 0 de ϕ : $M = \{x \in D : \phi(x) = 0\}$.

Suponhamos que queremos encontrar os extremos de f sobre M ; essencialmente interessam-nos os máximos e mínimos que f toma quando restringida a M . Esse é um problema de extremos condicionados: extremos de f na condição de $\phi = 0$.

Recorde: $\nabla\phi(x)$ é um vector ortogonal a T_xM em cada ponto $x \in M$ (cf. Aula 8).

Vendo agora f a variar sobre M , interessam-nos os pontos críticos no sentido em que restringimos o diferencial de f a TM . Ou melhor, um ponto extremo a de f sobre M fará anular $df(a)(v)$ em todas as direcções v sobre M , isto é, nas direcções v tangentes a M . Em suma,

a é ponto critico de f condicionado a ϕ

$$\iff df(a)(v) = 0, \forall v \in T_aM .$$

Escrevendo $df(a)(v) = \langle \nabla f(a), v \rangle$, vemos que a é crítico condicionado se $\nabla f(a) \perp T_a M$. Supondo que a é ponto regular de ϕ , isto é, $\nabla \phi(a) \neq 0$, acabamos de provar que os dois gradientes têm de ser paralelos, colineares!

Teorema

Seja a um ponto regular de ϕ . Então a é extremo condicionado de f sobre $M = \{\phi = 0\}$ se $\exists \lambda \in \mathbb{R}$ tal que

$$\nabla f(a) + \lambda \nabla \phi(a) = 0 .$$

A mesma equação escreve-se $\nabla(f + \lambda\phi)(a) = 0$ — então procurar os extremos condicionados acima é o mesmo que procurar os pontos críticos de $F(x, \lambda) = f(x) + \lambda\phi(x)$ com a função F definida em $D \times \mathbb{R}$.

Com efeito, $\frac{\partial F}{\partial \lambda}(x, \lambda) = 0$ se e só se $\phi(x) = 0$, que é a condição inicial, a não esquecer.

Este método de resolução chama-se **método dos multiplicadores de Lagrange**.

Exemplo: Supomos o problema de minimizar a distância de $(0, 0)$ à curva M dada por $x^2 + y^3 = 3$.

Todos os pontos de M são regulares.

A função “distância” mais apropriada é $f(x, y) = x^2 + y^2$.

$F(x, y, \lambda) = x^2 + y^2 + \lambda(x^2 + y^3 - 3)$ tem 3 pontos críticos:

$(0, \sqrt[3]{3}, \lambda_0)$, $(\sqrt{3}, 0, -1)$ e $(\sqrt{\frac{73}{27}}, \frac{2}{3}, -1)$. Por inspecção directa, o mínimo procurado está em $(0, \sqrt[3]{3})$.

Nota. Os valores de λ não nos interessam sumamente, são auxiliares de cálculo.

Aula 12

Multiplicadores de Lagrange com duas condições

Suponhamos que estamos no problema anterior de extremos de f , mas que em vez de uma temos duas condições ϕ_1, ϕ_2 . Portanto $f, \phi_1, \phi_2 : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ são funções de classe pelo menos C^1 . Cada uma das superfícies $M_i = \{x \in D : \phi_i(x) = 0\}$ oferece uma condição a f ; porém queremos extremos sobre $M = M_1 \cap M_2$.

Usando os mesmos argumentos acima temos de analisar TM .

Tem-se sempre $T(M_1 \cap M_2) \subseteq TM_1 \cap TM_2$. Teremos igualdade se de facto o ortogonal a $\nabla\phi_1$ e o ortogonal a $\nabla\phi_2$ não coincidirem, ou seja, se os dois vectores não forem linearmente dependentes, o que vamos supôr - e o que consiste na regularidade de (ϕ_1, ϕ_2) .

Novamente, f tem extremo em $a \in M_1 \cap M_2$ se o seu gradiente for ortogonal a $T(M_1 \cap M_2)$ nesse ponto. Ou seja, $\nabla f(a)$ está no plano gerado por $\nabla\phi_1(a)$ e $\nabla\phi_2(a)$. Ou seja, o primeiro gradiente é combinação linear dos dois últimos.

Teorema

Seja a um ponto regular de (ϕ_1, ϕ_2) . Então a é extremo condicionado de f sobre $M = \{\phi_1 = 0, \phi_2 = 0\}$ se $\exists \lambda_1, \lambda_2 \in \mathbb{R}$ tais que

$$\nabla f(a) + \lambda_1 \nabla \phi_1(a) + \lambda_2 \nabla \phi_2 = 0 .$$

De novo, torna-se então mais prático considerar a função

$$F(x_1, \dots, x_n, \lambda_1, \lambda_2) = f(x) + \lambda_1 \phi_1(x) + \lambda_2 \phi_2(x)$$

e estudar os seus pontos críticos na totalidade.

Exemplo: Encontrar o mínimo de $f(x, y, z) = |z|$ sobre o corte da esfera de raio 3 e centro na origem pelo plano $2x + 3y - z = 0$.

Então

$$F(x, y, z, \lambda, \mu) = z + \lambda(x^2 + y^2 + z^2 - 9) + \mu(2x + 3y - z) = 0 .$$

O ponto onde se dá igual extremo é $\pm \sqrt{\frac{18}{91}} \left(1, \frac{3}{2}, \frac{13}{2}\right)$.

Exercício: Generalizar a definição de regularidade e o teorema anterior para m condições ϕ_1, \dots, ϕ_m .

Aula 13

Noções sobre curvas em \mathbb{R}^3

Neste curso chamamos caminho, curva, linha ou trajectória, denotada por γ , a uma função contínua $r : [a, b] \rightarrow \mathbb{R}^3$.

Uma reparametrização da curva γ é dada pela composição de r com uma função $\phi : [c, d] \rightarrow [a, b]$ bijectiva e crescente.

Estamos interessados na imagem de r e não *muito* na forma como γ é percorrida.

$r_1 : [c, d] \rightarrow \mathbb{R}^3$ e $r_2 : [a, b] \rightarrow \mathbb{R}^3$ são parametrizações equivalentes de γ (ou dizem-se curvas equivalentes) se existe uma reparametrização da primeira que se transforme na segunda, ou seja,

$$r_1 = r_2 \circ \phi$$

A curva inversa de $\gamma \equiv r(t)$, $a \leq t \leq b$, é definida por
 $-\gamma \equiv r^-(t) = r(a + b - t)$.

Trata-se da mesma curva, percorrida de trás para a frente.

A justaposição de duas curvas, a primeira terminando onde a segunda começa, é a curva que resulta das duas funções parametrizada num só intervalo.

Exemplo. As curvas $r(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$, e $r_1(t) = (\cos 5t, \sin 5t)$, $t \in [0, \frac{2}{5}\pi]$ são equivalentes. Mas não são equivalentes a $r_2(t) = (\cos t^3\pi, \sin t^3\pi)$, $t \in [0, 1]$.

Chamamos comprimento de $\gamma \equiv r(t)$, $t \in [a, b]$, ao número real

$$L_\gamma = \sup_{\mathcal{P}} \sum_{i=0}^{n-1} \|r(t_{i+1}) - r(t_i)\|$$

onde \mathcal{P} denota *partições* do intervalo fechado $[a, b]$ descritas por $t_0 = a < t_1 < t_2 < \dots < t_n = b$.

Teorema

Se r é de classe C^1 , então o comprimento é dado por

$$L_\gamma = \int_a^b \left\| \frac{dr}{dt} \right\| dt.$$

Fazemos a justificação de acordo com a teoria do integral de Riemann estudada em Análise Matemática de uma variável.

Note-se primeiro que

$$\sum_{i=0}^{n-1} \|r(t_{i+1}) - r(t_i)\| = \sum_{i=0}^{n-1} \left\| \int_{t_i}^{t_{i+1}} r'(t) dt \right\| \leq \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \|r'(t)\| dt = \int_a^b \|r'(t)\| dt.$$

Agora seja $s(t)$ o comprimento da curva restringida a $[a, t]$ (a curva que vai do ponto $r(a)$ a $r(t)$).

Para $t > t_0$, temos pela desigualdade triangular e pela desigualdade acima que

$$\|r(t) - r(t_0)\| \leq s(t) - s(t_0) \leq \int_{t_0}^t \|r'\|.$$

Dividindo tudo por $t - t_0$ e tomando o limite quando $t \rightarrow t_0$, resulta $\|r'(t_0)\| \leq s'(t_0) \leq \|r'(t_0)\|$, donde se conclui (para $t_0 < t$ seria o mesmo) que $s'(t_0) = \|r'(t_0)\|$.

A função $s(t) = \int_a^t \|r'(\tilde{t})\| d\tilde{t}$ é conhecida como a função comprimento de arco.

s é claramente crescente e bijectiva, do intervalo $[a, b]$ em $[0, L_\gamma]$.
Deve-se notar bem:

$$ds = \|r'(t)\| dt .$$

Seja agora $\phi : [0, L_\gamma] \rightarrow [a, b]$ a função inversa de $s(t)$, isto é, $\phi(s) = t$. É claro que ϕ também é crescente.

Então $r \circ \phi$ dá-nos uma parametrização equivalente de γ .

Eis a parametrização por comprimento de arco de γ .

Integrais de linha

Seja agora D um aberto de \mathbb{R}^n e $f : D \rightarrow \mathbb{R}$ uma função C^0 .

Seja $\gamma = r(t)$, $a \leq t \leq b$, uma curva cuja imagem está contida em D .

Prova-se que o integral

$$\int_{\gamma} f \, ds = \int_a^b f(r(t)) \|r'(t)\| \, dt$$

“não depende da (escolha da) parametrização”; mais precisamente, não depende da escolha de r entre todas as parametrizações equivalentes de γ .

$\int_{\gamma} f \, ds$ chama-se integral de f ao longo de γ relativo ao comprimento de arco.

Suponhamos que γ é dada por $r(t) = (x(t), y(t), z(t))$ (só para simplificar, escolhemos \mathbb{R}^3).

Podemos falar do integral

$$\int_{\gamma} f \, dx = \int_a^b f(r(t)) x'(t) \, dt$$

o qual também não depende da escolha da parametrização

$\int_{\gamma} f \, dx$ chama-se integral de f ao longo de γ relativo a x .

Da mesma forma define-se o $\int_{\gamma} f \, dy$ ou $\int_{\gamma} f \, dz$.

Aula 14

Trabalho de um campo vectorial e campo conservativo; potencial

Seja agora $F : D \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ uma função classe C^0 — descrita então por

$$F(x, y, z) = f_1(x, y, z)e_1 + f_2(x, y, z)e_2 + f_3(x, y, z)e_3$$

o que também se designa de campo vectorial sobre D .

Seja $\gamma = r(t)$ uma curva como acima e $\vec{t} = \frac{r'}{\|r'\|}$ o *versor* da velocidade de r (supondo que não se anule).

Tem-se a igualdade

$$\begin{aligned} \int_{\gamma} \langle F, \vec{t} \rangle ds &= \int_a^b \langle F(r(t)), r'(t) \rangle dt = \\ &= \int_a^b \left(f_1(r(t)) x'(t) + f_2(r(t)) y'(t) + f_3(r(t)) z'(t) \right) dt = \\ &= \int_{\gamma} f_1 dx + f_2 dy + f_3 dz . \end{aligned}$$

Assim se comporta o integral de linha: o trabalho de F sobre γ , conceito da Física, é o integral do campo vectorial F ao longo do caminho na direcção da tangente.

Suponhamos agora que F é um gradiente, ou seja, existe $\psi : D \rightarrow \mathbb{R}$ tal que $F = \nabla\psi$. Neste caso diz-se que F é um campo conservativo ou campo vectorial gradiente.

À função escalar ψ dá-se o nome de potencial.

Proposição

Condição necessária para um campo vectorial $F = (f_1, \dots, f_n)$ de classe C^1 ser conservativo é que

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} .$$

Repare-se que dizer que $F = (f_1, \dots, f_n)$ é um campo conservativo, ou seja $(f_1, \dots, f_n) = \left(\frac{\partial \psi}{\partial x_1}, \dots, \frac{\partial \psi}{\partial x_n} \right)$, é o mesmo que dizer que $f_1 dx_1 + \dots + f_n dx_n = d\psi$.

Neste caso, temos:

Proposição

Se F é um campo vectorial gradiente, com potencial ψ , então

$$\int_{\gamma} f_1 dx_1 + \dots + f_n dx_n = \psi(B) - \psi(A)$$

para qualquer curva γ em D com início em A e fim em B .

Em particular, é nulo o trabalho de F ao longo de qualquer curva γ fechada:

$$\int_{\gamma} \langle F, \vec{t} \rangle ds = \psi(A) - \psi(A) = 0 .$$

A condição anterior permite *integrar* o campo vectorial F . Ou seja, encontrar o resultado recíproco.

Com efeito, pode-se provar que sendo o integral de circulação de $\langle F, \vec{t} \rangle$ nulo ao longo de toda a curva fechada, então F é conservativo, isto é, é um campo vectorial gradiente.

Aula 15

Integral de Riemann em dimensão > 1

Temos agora por objecto de estudo um dos mais incisivos instrumentos do cálculo em várias variáveis. Nem todos os teoremas podem ser provados pelos meios simples que temos procurado mostrar. Essas outras demonstrações deixaremos para curso mais especializado.

Seja I um intervalo de \mathbb{R}^n . Existem reais $a_i \leq b_i$ tais que $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$.

Chamamos volume de I ao número real

$$V(I) = (b_1 - a_1) \cdots (b_n - a_n).$$

Chamamos diâmetro de I ao número real

$$d(I) = \sqrt{(b_1 - a_1)^2 + \cdots + (b_n - a_n)^2}.$$

Chamamos partição do intervalo I a um conjunto $P = \{I_1, I_2, \dots, I_k\}$ de subintervalos de I tais que

$$\text{int } I_i \cap \text{int } I_j = \emptyset \quad \forall i \neq j \quad \text{e} \quad I_1 \cup I_2 \cup \dots \cup I_k = I .$$

Dadas duas partições P_1, P_2 do mesmo intervalo I , dizemos que P_1 é mais fina que P_2 (o que se denota por $P_1 \preceq P_2$) se cada intervalo de P_1 está contido num intervalo de P_2 .

É claro que dadas duas partições, há sempre uma terceira mais fina que qualquer uma delas (porquê?).

Chamamos diâmetro da partição P ao máximo dos diâmetros dos intervalos que a compõem.

Seja $\xi = \{X_1, X_2, \dots, X_k\}$ uma sucessão de pontos do intervalo I . A sucessão diz-se compatível com a partição $P = \{I_1, I_2, \dots, I_k\}$, se $X_i \in I_i, \forall 1 \leq i \leq k$.

Agora seja $f : I \subset \mathbb{R}^n \rightarrow \mathbb{R}$ uma função real definida no intervalo I , onde se tem também uma partição P e uma sucessão ξ compatível com P . Ao número real

$$S(f, P, \xi) = \sum_{i=1}^k f(X_i) V(I_i)$$

damos o nome de soma de Riemann de f relativa a P e a ξ .

Pode-se compreender a soma de Riemann como a soma dos volumes dos paralelepípedos de base I_i e altura $f(X_i)$.

Dizemos que f é integrável à Riemann, integrável-R ou, simplesmente, que f é integrável sobre o intervalo I se existir um número real A tal que

$$\forall \delta > 0, \exists \epsilon > 0 \text{ tal que } \forall P \text{ partição de } I$$
$$d(P) < \epsilon \implies |S(f, P, \xi) - A| < \delta, \forall \xi \text{ compatível com } P .$$

Outra forma de dizer é que, em vez de um ϵ , existe uma partição P_0 tal que, sempre que $P \preceq P_0$, temos a mesma desigualdade entre as respectivas somas de Riemann e o número A .

Proposição

Se f é integrável, então A é único.

Denota-se

$$A = \int_I f .$$

O conhecimento da construção dos números reais como limites de sucessões de Cauchy, permite ainda escrever a mesma condição sem sequer invocar A :

Teorema

$f : I \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é integrável à Riemann se e só se

$\forall \delta > 0, \exists P_0$ partição de I , tal que $\forall P_1, P_2 \preceq P_0 \implies$

$$|S(f, P_1, \xi_1) - S(f, P_2, \xi_2)| < \delta, \forall \xi_1, \xi_2 .$$

(Subentende-se, ξ_1, ξ_2 compatíveis com P_1, P_2 respectivamente).

O primeiro propósito do integral de Riemann é o cálculo de volumes.

Seja M um subconjunto qualquer do intervalo I limitado:
 $M \subset I \subset \mathbb{R}^n$. Seja $\chi_M : I \rightarrow \mathbb{R}$ a função definida por

$$\chi_M(x) = \begin{cases} 1 & x \in M \\ 0 & x \notin M \end{cases} .$$

Dizemos que M é mensurável (à Jordan) se χ_M é integrável.
Se tal for o caso, chamamos volume de M ao número real
 $V(M) = \int_I \chi_M$.

Propriedades do integral de Riemann

Teorema

Sejam $f, g : I \rightarrow \mathbb{R}$ funções integráveis num intervalo I . Então:

(i) $f + g$, fg e $|f|$ são integráveis

(ii) se $|f(x)| \geq c > 0$, $\forall x \in I$, então $\frac{1}{f}$ é integrável.

O seguinte teorema acenta na noção de continuidade uniforme de f , função contínua sobre um compacto, que nos garante em termos estreitos o controle global da variação de f no intervalo limitado e fechado I (compacto).

Aula 16

Teorema

Toda a função contínua em I é integrável à Riemann.

Mais geralmente:

Teorema

Toda a função contínua num conjunto mensurável $M \subset I$ é integrável, no sentido que é integrável a função F dada por $F(x) = f(x)$, se $x \in M$, e $F(x) = 0$ se $x \in I \setminus M$.

Claro que se escreve

$$\int_M f = \int_I F .$$

Teorema (do valor médio)

Seja $f : M \rightarrow \mathbb{R}$ contínua num conjunto conexo e compacto. Então existe $x_0 \in M$ tal que $\int_M f = f(x_0) \int_M 1 = f(x_0) V(M)$.

Nota. Um conjunto diz-se conexo se não for a união de dois abertos de \mathbb{R}^n disjuntos e não-vazios. Intuitivamente exprime-se a ideia de que M é constituído de uma só *peça*.

Os dois próximos teoremas permitem calcular muitos integrais. O primeiro é fácil de provar.

Teorema (fundamental do cálculo integral)

Seja $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ integrável. Então, sendo $\phi(x) = \int_a^x f(t) dt$, $x \in [a, b]$, tem-se

$$\phi' = f .$$

Em particular,

$$\int_x^b f(t) dt = \phi(b) - \phi(x) = [\phi]_x^b .$$

Teorema (Fubini)

Seja $f : I \times J \rightarrow \mathbb{R}$ contínua num intervalo de \mathbb{R}^{n+p} , com $I \subset \mathbb{R}^n$, $J \subset \mathbb{R}^p$ intervalos fechados e limitados. Então:

(i) $y \mapsto f(x, y)$ é integrável em J , $\forall x \in I$

(ii) $x \mapsto \int_J f(x, y) dy$ é contínua e por isso integrável em I

(iii) tem-se

$$\iint_{I \times J} f = \int_I \left(\int_J f(x, y) dy \right) dx .$$

Do teorema de Fubini conclui-se que é indiferente a ordem de integração!

Como corolários dos resultados acima temos que se f é integrável sobre um subconjunto $M \subset \mathbb{R}^{n+1}$, descrito pela condição $M = \{(x, y) \in \mathbb{R}^{n+1} : x \in I \subset \mathbb{R}^n, \varphi_1(x) \leq y \leq \varphi_2(x), \forall x \in I\}$ para certas funções contínuas $\varphi_1, \varphi_2 : I \rightarrow \mathbb{R}$, então

$$\iint_M f(x, y) \, dx \, dy = \int_I \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) \, dy \, dx .$$

Exemplo 1: Procura-se calcular a área entre a parábola x^2 e a cúbica x^3 para $x \in [-2, 2]$.

Note-se que à esquerda de 1 temos $x^3 < x^2$, enquanto à direita temos o contrário. Assim, a área (volume em 2 dimensões) é calculada por

$$\begin{aligned} \left(\int_{-2}^1 \int_{x^3}^{x^2} + \int_1^2 \int_{x^2}^{x^3} \right) 1 \, dy \, dx &= \int_{-2}^1 (x^2 - x^3) \, dx + \int_1^2 (x^3 - x^2) \, dx = \\ &= \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_{-2}^1 + \left[\frac{x^4}{4} - \frac{x^3}{3} \right]_1^2 = \\ &= \frac{1}{3} - \frac{1}{4} + \frac{8}{3} + \frac{16}{4} + \frac{16}{4} - \frac{8}{3} - \frac{1}{4} + \frac{1}{3} = \frac{102}{12} = \frac{17}{2}. \end{aligned}$$

Exemplo 2: Calculamos agora o volume da região M em \mathbb{R}^3 entre o parabolóide $z = 3 - x^2 - y^2$ e o plano $z = x + y$.

É claro que a intersecção das duas superfícies é dada por

$$3 - x^2 - y^2 = x + y \iff \left(x + \frac{1}{2}\right)^2 + \left(y + \frac{1}{2}\right)^2 = \frac{7}{2}$$

portanto uma circunferência. É fácil compreender que a região fica descrita por

$$M = \left\{ (x, y, z) : x + y \leq z \leq 3 - x^2 - y^2, \left(x + \frac{1}{2}\right)^2 + \left(y + \frac{1}{2}\right)^2 \leq \frac{7}{2} \right\}.$$

A partir daqui o problema resolve-se facilmente recorrendo a coordenadas polares, que se introduzem a seguir.

Aula 17

Teorema geral de mudança de variáveis em integração

O teorema que aqui vem em título é sem dúvida o mais difícil de provar neste nosso curso, o qual não se deseja sobrecarregar com mais definições e instrumentos que seriam apenas necessários para tal prova. Mas é também um daqueles resultados que o próprio engenho e a necessidade nos iluminam. Aliás o teorema foi utilizado em diversas formas aplicadas, por grandes matemáticos como Euler ou Lagrange, mais de cem anos antes da primeira demonstração rigorosa.

Vejamos as hipóteses.

É dado um subconjunto mensurável $E \subset U \subset \mathbb{R}^n$ contido num aberto U , por sua vez contido num intervalo limitado de \mathbb{R}^n .

É dada uma função $\phi : U \rightarrow V$ bijectiva sobre um outro aberto V , de classe C^1 e com inversa contínua.

Denotamos ainda por $M = \phi(E)$ o subconjunto imagem de E por ϕ .

Finalmente, é dada função $f : M \rightarrow \mathbb{R}$ integrável à Riemann.

Teorema (de mudança de variável em integração)

Nas condições descritas anteriormente, temos que $M = \phi(E)$ é mensurável e que

$$\int_M f(y) \, dy = \int_E f \circ \phi(x) |\det \text{Jac } \phi(x)| \, dx .$$

Note-se que a função integranda do lado direito contém o *módulo* do determinante da matriz jacobiana.

Um pouco mais devemos observar sobre a demonstração.

Recordemos que uma qualquer base $\{u_1, \dots, u_n\}$ de \mathbb{R}^n define um volume ou sólido, a saber, o *paralelepípedo* P de arestas u_i e um mesmo vértice a origem do referencial. P é o sólido com pontos $\sum_i x_i u_i$ onde $(x_1, \dots, x_n) \in [0, 1]^n$. Ora demonstra-se em curso próprio de Álgebra Linear que

$$V(P) = \det \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} .$$

Este valor é obviamente igual ao determinante da aplicação linear que transforma a base canónica $\{e_i\}$ na base dos $\{u_j\}$.

Assim, se $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ é uma aplicação linear qualquer (em particular transforma subespaços lineares em subespaços lineares), então o determinante de L é o quociente entre os volumes de um paralelepípedo $L(P)$, imagem por L de outro paralelepípedo P do qual partimos:

$$V(L(P)) = V(P) \det L$$

(porquê?).

Finalmente note-se que, infinitesimalmente, o aparecimento do determinante da jacobiana de ϕ no teorema geral de mudança de variável se deve ao que acabamos de observar.

Coordenadas polares e cilíndricas

Somente para \mathbb{R}^2 e \mathbb{R}^3 .

A transformação ϕ de coordenadas cartesianas (x, y) de \mathbb{R}^2 em coordenadas polares

$$(x, y) = (\rho \cos \theta, \rho \sin \theta)$$

onde $\rho = \sqrt{x^2 + y^2} \in]0, +\infty[$ é o raio e $\theta \in]0, 2\pi[$ o ângulo que o ponto-vector faz com o semi-eixo positivo dos x , é muito importante.

Feitas as contas, temos

$$\det \text{Jac } \phi = \det \begin{bmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{bmatrix} = \rho > 0 .$$

Exemplo 1: Podemos agora calcular o volume que queríamos atrás (Aula 16):

Escrevendo $D = \{(x, y) \in \mathbb{R}^2 : (x + \frac{1}{2})^2 + (y + \frac{1}{2})^2 \leq \frac{7}{2}\}$.

Tínhamos parado no cálculo de

$$\begin{aligned} \iint_D \int_{x+y}^{3-x^2-y^2} 1 \, dz \, dx \, dy &= \iint_D 3 - x^2 - y^2 - x - y \, dz \, dx \, dy \\ &= \iint_D \frac{7}{2} - (x + \frac{1}{2})^2 - (y + \frac{1}{2})^2 \, dx \, dy . \end{aligned}$$

Porém! Note-se que a transformação que convém não é exactamente aquela das coordenadas de centro em $(0, 0)$ mas sim outras de centro em $(\frac{1}{2}, \frac{1}{2})$.

É claro que o jacobiano de $(x, y) = (-\frac{1}{2} + \rho \cos \theta, -\frac{1}{2} + \rho \sin \theta)$ também resulta em ρ .

O integral anterior continua como

$$= \int_0^{2\pi} \int_0^{\sqrt{\frac{7}{2}}} \left(\frac{7}{2} - \rho^2\right) \rho \, d\rho \, d\theta = 2\pi \left[\frac{7}{4}\rho^2 - \frac{\rho^4}{4} \right]_0^{\sqrt{\frac{7}{2}}} = \frac{49}{8}\pi .$$

Exemplo 2: Para calcular a área do círculo $B_r(0)$ de raio r , temos

$$\iint_{B_r(0)} 1 \, dy \, dx = \int_0^{2\pi} \int_0^r \rho \, d\rho \, d\theta = 2\pi \left[\frac{\rho^2}{2} \right]_0^r = \pi r^2$$

tal como se esperava.

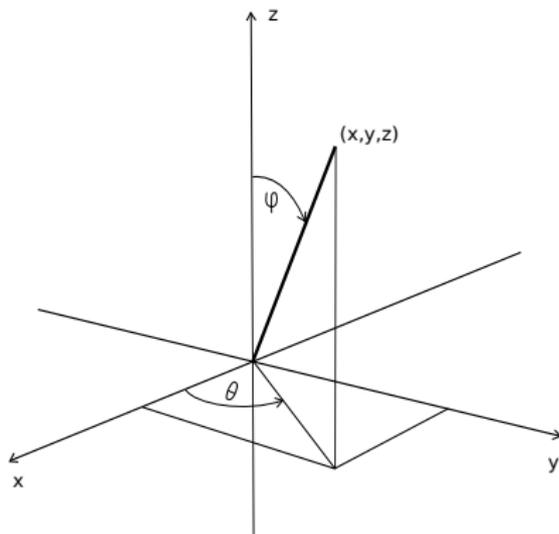
Note-se que no exemplo 1 considerámos inicialmente uma região em \mathbb{R}^3 . De facto o que fizemos foi utilizar coordenadas cilíndricas, que mais não são do que as polares em conjunto com uma terceira, cartesiana, z :

$$(x, y, z) = (\rho \cos \theta, \rho \sin \theta, z) .$$

Verifica-se facilmente que o determinante da matriz jacobiana é o mesmo: ρ .

Coordenadas esféricas

São apenas para \mathbb{R}^3 .



Necessitando descrever os pontos em três dimensões por meio de coordenadas esféricas, como se viu na figura, temos:

$$(x, y, z) = (\rho \operatorname{sen} \varphi \cos \theta, \rho \operatorname{sen} \varphi \operatorname{sen} \theta, \rho \cos \varphi) = \phi(\rho, \theta, \varphi)$$

com $0 \leq \rho < +\infty$, $0 \leq \theta < 2\pi$, $0 < \varphi < \pi$.

Note-se que $\rho = \sqrt{x^2 + y^2 + z^2}$.

O factor a ter em conta pelo teorema da mudança de variável em integração é

$$\begin{aligned} \det \operatorname{Jac} \phi &= \begin{vmatrix} \operatorname{sen} \varphi \cos \theta & -\rho \operatorname{sen} \varphi \operatorname{sen} \theta & \rho \cos \varphi \cos \theta \\ \operatorname{sen} \varphi \operatorname{sen} \theta & \rho \operatorname{sen} \varphi \cos \theta & \rho \cos \varphi \operatorname{sen} \theta \\ \cos \varphi & 0 & -\rho \operatorname{sen} \varphi \end{vmatrix} \\ &= \rho^2 \operatorname{sen} \varphi . \end{aligned}$$

Eis um exemplo fundamental, o cálculo do volume da bola $B = B_r(0)$ de raio r :

$$V(B) = \int_B 1 = \int_0^\pi \int_0^{2\pi} \int_0^r \rho^2 \operatorname{sen} \varphi \, d\rho \, d\theta \, d\varphi = \frac{4}{3}\pi r^3$$

fórmula que era conhecida na antiguidade clássica.

Aula 18

Teorema de Green

Eis um primeiro teorema que relaciona a integração sobre um domínio com aquela sobre a sua fronteira.

Suponhamos que $\gamma \subset \mathbb{R}^2$ é uma curva *seccionalmente* de classe C^1 , simples e fechada (simples significa que só passa uma vez em cada ponto, para o que não conta o inicial — que neste caso se supõe igual ao final por a curva ser fechada).

Nas condições anteriores, γ define um domínio do plano, denotado $\text{int } \gamma \subset \mathbb{R}^2$, a saber, o maior aberto limitado que tem γ como fronteira.

Suponhamos ainda que γ é percorrida no sentido positivo ou trigonométrico, ou seja, no sentido contrário ao dos ponteiros do relógio, no plano descrito em coordenadas (x, y) .

(O que está em causa é a escolha de uma *orientação*, outro conceito matemático que não dependerá do *observador*).

Teorema (de Green)

Sejam $f, g : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ duas funções de classe C^1 num aberto.

Seja γ uma curva nas condições acima, tal que a região do plano $R = \gamma \cup \text{int}(\gamma) \subset D$. Então:

$$\oint f \, dx + g \, dy = \iint_R \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \, dy .$$

Note-se que se $F = (f, g)$ é um campo vectorial conservativo em D , então necessariamente $\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} = 0$ (veja-se a Aula 14).

Tendo em conta o recíproco deste resultado, analisado no fim da referida aula, podemos afirmar ainda que, se para qualquer $\gamma \subset D$, curva simples e fechada, o interior de γ também está contido em D (por exemplo se D é *convexo*), então a condição acima é também suficiente. Ou seja, dado $F = (f, g)$

$$\exists \psi \text{ tal que } d\psi = f dx + g dy \iff \frac{\partial g}{\partial x} = \frac{\partial f}{\partial y} .$$

A condição geométrica sobre D é identificada como a noção de D ser simplesmente conexo. É esse, com efeito, o caso dos subconjuntos convexos (por definição, os que contêm os segmentos de recta entre os seus pontos).

Observe-se ainda no teorema de Green a coerência dos sinais se mudarmos o sentido à curva.

Outra conclusão imediata é a fórmula

$$\frac{1}{2} \oint x \, dy - y \, dx = \iint_R 1 \, dx \, dy = \text{Área}(R) .$$

Integrais de superfície

Não se pode conceber a noção de área de uma dada superfície de acordo com a intuição comum, mas sim como uma nova definição; pois uma superfície, em geral, possui aquilo que é conhecido como *curvatura*, dependente em cada ponto e de forma intrínseca da estrutura métrica do espaço.

Chamamos superfície à imagem \mathcal{S} de, ou a, uma função injectiva e contínua

$$\begin{aligned} f : D \subset \mathbb{R}^2 &\longrightarrow \mathbb{R}^3 \\ (s, t) &\longmapsto (x(s, t), y(s, t), z(s, t)) . \end{aligned}$$

Mais geralmente, uma superfície poderá ser descrita por várias expressões funcionais, como seja o caso da descrição da fronteira de um *sólido*.

Suponhamos f de classe C^1 .

Tendo em conta a diferenciabilidade em cada ponto $(s_0, t_0) \in D$,
requere-se que infinitesimalmente a área de \mathcal{S} sobre o rectângulo
 $R_{s,t}$ de arestas geradas pelos pontos (s_0, t_0) , $(s_0 + s, t_0)$, $(s_0, t_0 + t)$
(logo de área st) coincida *no limite* com a área do paralelogramo
gerado pelas imagens por meio de f dos vértices de $R_{s,t}$.

Por outras palavras, requere-se o *elemento de área* igual a

$$d\sigma = \frac{1}{st} \text{Área}(df(R_{s,t})) ds dt .$$

Como se viu na Aula 17, o determinante é uma operação algébrica que nos dá a área em \mathbb{R}^2 ou o volume em \mathbb{R}^3 . Podemos calcular a área do paralelogramo $df(R_{s,t})$ em \mathbb{R}^3 utilizando o determinante, se aos dois vectores $\frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}$ juntarmos um terceiro N de norma 1, que lhes seja perpendicular (por isso perpendicular a TS). A área desejada será o volume do paralelepípedo com aquela base e com altura 1.

Queremos então encontrar um vector $N_1 = (a, b, c) \perp \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t}$:

$$\begin{cases} ax'_s + by'_s + cz'_s = 0 \\ ax'_t + by'_t + cz'_t = 0 \end{cases} .$$

Uma solução é

$$a = y'_s z'_t - y'_t z'_s, \quad b = z'_s x'_t - z'_t x'_s, \quad c = x'_s y'_t - x'_t y'_s .$$

Então, sendo $N = \frac{N_1}{\|N_1\|} = \frac{(a,b,c)}{\sqrt{a^2+b^2+c^2}}$,

$$\begin{aligned} \frac{1}{st} \text{Área}(df(R_{s,t})) &= \det \left(N, \frac{\partial f}{\partial s}, \frac{\partial f}{\partial t} \right) = \frac{1}{\|N_1\|} \begin{vmatrix} a & b & c \\ x'_s & y'_s & z'_s \\ x'_t & y'_t & z'_t \end{vmatrix} \\ &= \frac{1}{\|N_1\|} (a^2 + b^2 + c^2) = \sqrt{a^2 + b^2 + c^2} = \|N_1\|. \end{aligned}$$

Nota. $-N_1$ também é solução do sistema que dá $N_1 = (a, b, c)$. Obviamente escolhemos o sinal em N de acordo com o que faz o determinante positivo! Isto corresponde a *uma escolha de orientação* pois há duas normais, $\pm N$, à superfície dada.

Eis por que se define a área de uma superfície \mathcal{S} por

$$A(\mathcal{S}) = \iint_{\mathcal{S}} 1 \, d\sigma = \iint_D \sqrt{a^2 + b^2 + c^2} \, ds \, dt$$

Podemos escrever a regra mnemónica (correspondendo a uma escolha de orientação em \mathbb{R}^3):

$$N_1 = \begin{vmatrix} e_1 & e_2 & e_3 \\ x'_s & y'_s & z'_s \\ x'_t & y'_t & z'_t \end{vmatrix}$$

e logo

$$d\sigma = \sqrt{a^2 + b^2 + c^2} \, ds \, dt = \|N_1\| \, ds \, dt .$$

Definimos ainda o integral de superfície para uma dada função contínua $g : \mathcal{S} \rightarrow \mathbb{R}$ por

$$\iint_{\mathcal{S}} g \, d\sigma = \int \int_D g(x(s, t), y(s, t), z(s, t)) \|N_1\| \, ds \, dt .$$

Nota. Pode-se definir muito naturalmente uma *equivalência de superfícies*, tal como fizemos para as curvas, essencialmente considerando que se teriam outras coordenadas para o mesmo subconjunto \mathcal{S} de \mathbb{R}^3 , e pode-se depois demonstrar que o integral de superfície não depende da escolha dessas *parametrizações* de \mathcal{S} .

Exemplo: o mais habitual é S ser dada por uma equação
 $z = \phi(x, y)$, $(x, y) \in D$; nos termos anteriores trata-se da função,
injectiva,

$$f(x, y) = (x, y, \phi(x, y)) .$$

Então encontramos facilmente a expressão

$$\iint_S g \, d\sigma = \iint_D g(x, y, z(x, y)) \sqrt{1 + (\phi'_x)^2 + (\phi'_y)^2} \, dx \, dy .$$

Exercício: mostre que a área da esfera $S_r^2 = \{\|(x, y, z)\| = r\}$ de raio r é $4\pi r^2$.

Aula 19

Bordo de uma superfície e orientação induzida no bordo

Para melhor compreender um dos teoremas cimeiros deste curso devemos observar algumas situações gerais da teoria das superfícies.

Uma superfície \mathcal{S} em \mathbb{R}^3 poderá ser dada não apenas por uma, mas por várias funções injectivas; nesse caso \mathcal{S} coincide com a reunião das imagens dessas várias funções, ou seja, $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \dots$.

É o caso da superfície de um cubo, a união de seis *faces*.

É também o caso de uma superfície dada implicitamente por

$\phi(x, y, z) = 0$, como é exemplo a esfera $x^2 + y^2 + z^2 = 1$, união de $z = \sqrt{1 - x^2 - y^2}$ com $z = -\sqrt{1 - x^2 - y^2}$.

Em geral, vamos admitir que as imagens das funções das \mathcal{S}_i não se intersectam, senão, eventualmente, sobre as suas *arestas* ou *bordos*. O resultado poderá não ter bordo nenhum (como é o exemplo da esfera). Também existem superfícies descritas por uma só expressão e com bordos disconexos.

Sendo $\varphi : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^3$, definida num aberto D , dando origem a uma superfície \mathcal{S} , chamamos bordo de \mathcal{S} , ao conjunto

$$\partial\mathcal{S} = \varphi(\text{fr}(D)) .$$

Admitimos φ como estando também definida na fronteira de D , por isso chamamos orientação induzida no bordo de \mathcal{S} ao sentido da curva $\varphi \circ \gamma$ quando γ é o caminho que descreve a $\text{fr}(D)$ no sentido trigonométrico do plano \mathbb{R}^2 .

Suponhamos que \mathcal{S} é dada por

$\varphi(s, t) = (x(s, t), y(s, t), z(s, t))$, $(s, t) \in D$, função de classe C^1 .

Referimos já a existência de uma normal N a \mathcal{S} (estamos também sempre a admitir que $\frac{\partial \varphi}{\partial s}, \frac{\partial \varphi}{\partial t} \in \mathbb{R}^3$ são linearmente independentes).

Como se sabe, $N \perp TS$.

\mathcal{S} é naturalmente *orientada* pela base de \mathbb{R}^3 $N, \frac{\partial \varphi}{\partial s}, \frac{\partial \varphi}{\partial t}$ ou, se se preferir, $\frac{\partial \varphi}{\partial s}, \frac{\partial \varphi}{\partial t}, N$
(se não se fizerem mais assumpções - atenção - cf. banda de Möebius).

Por exemplo um disco $D = B_1(0) \subset \mathbb{R}^2$ coincide com uma superfície em \mathbb{R}^3 , por meio de $(x, y) \mapsto (x, y, 0)$. É fácil ver que $N = e_3 = (0, 0, 1)$ e que a orientação induzida no bordo é dada pelo sentido óbvio $\vec{t}_{(x,y,0)} = (-y, x, 0)$ para pontos no bordo $x^2 + y^2 = 1$. Por exemplo no ponto $(1, 0, 0)$ temos $\vec{t} = e_2$.

Este exemplo é paradigmático.

Suponhamos que nos dão a superfície \mathcal{S} e a normal N em cada ponto, sem nos darem as funções coordenadas. Põe-se então a questão:

Para um ponto $(x, y, z) \in \partial\mathcal{S}$, como adivinhar a orientação induzida, ou seja, como escolher \vec{t} ?

A resposta é: tomar o plano tangente TS nesse ponto, imaginar nesse plano um vector unitário \vec{u} que aponta para fora de S e depois escolher \vec{t} de tal forma que as bases

\vec{u}, \vec{t}, N e e_1, e_2, e_3 têm a mesma orientação.

(Mais precisamente, de forma que o determinante de transformação de uma na outra seja positivo).

Exemplo. Seja $S = \{x^2 + y^2 + z^2 = 2, -1 \leq z \leq 1\}$ e N a normal exterior à bola.

Então o bordo ∂S tem duas componentes. E onde $z = 1$, temos $\vec{t}_{(x,y,z)} = (y, -x, 0)$. Onde $z = -1$, temos $\vec{t}_{(x,y,z)} = (-y, x, 0)$.

O rotacional e o teorema de Stokes

Seja $F : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ um campo vectorial de classe C^1 definido num aberto de \mathbb{R}^3 . Na habitual base canónica, existem funções $f, g, h : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ tais que

$$F = (f, g, h) = f e_1 + g e_2 + h e_3 .$$

Chamamos rotacional de F ao novo campo vectorial $\text{rot } F$ definido, também sobre U , por

$$\begin{aligned} \text{rot } F &= \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f & g & h \end{vmatrix} \\ &= \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z} \right) e_1 + \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x} \right) e_2 + \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) e_3 . \end{aligned}$$

Finalmente podemos enunciar:

Teorema (de Stokes)

*É dada uma superfície S seccionalmente de classe C^2 em \mathbb{R}^3 .
Supõe-se que a superfície está incluída no domínio de um campo
vectorial F de classe C^1 . É também dada uma normal N sobre S .
Então*

$$\iint_S \langle \text{rot } F, N \rangle d\sigma = \oint_{\partial S} \langle F, \vec{t} \rangle ds$$

*onde se toma ∂S com a orientação induzida, \vec{t} como o versor da
tangente ao bordo e s o comprimento de arco.*

Uma versão local do teorema, admitindo que são conhecidas as funções coordenadas $\varphi(\tilde{s}, t)$, com $(\tilde{s}, t) \in D$, obriga-nos a recordar

$$d\sigma = \|N_1\| d\tilde{s} dt, \quad N = \frac{N_1}{\|N_1\|} = \frac{(a, b, c)}{\sqrt{a^2 + b^2 + c^2}}$$

e

$$\int \langle F, \vec{t} \rangle ds = \int f dx + g dy + h dz .$$

Diz-nos então claramente o teorema de Stokes que

$$\begin{aligned} \iint_D \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z} \right) a + \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x} \right) b + \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) c d\tilde{s} dt \\ = \oint_{\partial S} f dx + g dy + h dz . \end{aligned}$$

Note-se que, se $h = 0$ e f, g não dependem de z , então temos uma versão do teorema de Green *no espaço*.

Claro que a versão inicial do teorema de Stokes resulta da versão local, da decomposição $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \dots$ e da linearidade dos integrais de linha e de superfície.

Note-se que, uma vez N fixado, o mais frequente é a orientação induzida numa parte do bordo comum a duas faces $\mathcal{S}_1, \mathcal{S}_2$ ser deduzida pela primeira, simétrica da que se deduz pela segunda.

Exemplo 1. Seja \mathcal{Q} a superfície de um cubo em $[0, 1]^3 \subset \mathbb{R}^3$ *aberto* em cima, ou seja, faltando-lhe a face superior. Seja N a normal exterior (perpendicular a cada face e apontando para fora).
(cont.)

Seja $F = xy^2 e_3$.

Então a circulação nas arestas de lado e em baixo cancelam duas a duas; só sobrando a circulação das arestas da face que falta.

Note-se que a circulação, em cima, se faz toda no mesmo sentido!

Porém, a circulação de F no bordo é nula porque $e_3 \perp \partial Q$. Por outro lado, $\text{rot } F = 2xy e_1 - y^2 e_2$. Então

$$\begin{aligned} \iint_S \langle \text{rot } F, N \rangle d\sigma &= \left(\iint_{\{x=1\}} - \iint_{\{x=0\}} \right) 2xy d\sigma \\ &- \left(\iint_{\{y=1\}} - \iint_{\{y=0\}} \right) y^2 d\sigma = \int_0^1 \int_0^1 2y dy dx - 1 dx dy = 0 \end{aligned}$$

como esperávamos.

Exemplo 2: Seja F um campo vectorial de classe C^1 qualquer, numa região contendo a superfície esférica $S \equiv x^2 + y^2 + z^2 = r^2$ de raio r . O bordo é vazio. Então

$$\iint_S \langle \text{rot } F, N \rangle d\sigma = 0 .$$

O integral $\iint_S \langle G, N \rangle d\sigma$ chama-se fluxo de G através de S .

Aula 20

Divergência e o teorema de Gauss ou da divergência

Suponhamos que é dado um campo vectorial

$F = (f, g, h) : U \longrightarrow \mathbb{R}^3$ de classe C^1 num aberto $U \subset \mathbb{R}^3$.

Define-se a divergência de F como a função escalar

$\operatorname{div} F : U \longrightarrow \mathbb{R}$,

$$\operatorname{div} F = \frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} .$$

Este conceito, por estranho que pareça, surge de forma natural em geometria diferencial. Tem importância por si mesmo na medição da variação instantânea de F , aliás como mostra o próximo teorema.

Suponhamos que M é uma região de \mathbb{R}^3 com volume não nulo; como exemplo, podemos ter M um aberto não vazio.

Um tal *sólido* dá lugar a uma superfície $\mathcal{S} = \partial M$, o bordo de M , a qual vamos admitir coincidir com a fronteira de M em \mathbb{R}^3 .

Admitimos também de modo ilustrativo, como já ocorreu e ocorre geralmente, que \mathcal{S} tem algum plano tangente $T\mathcal{S}$ em cada ponto.

Chamamos normal exterior a $\mathcal{S} = \partial M$ ao vector unitário N_e , ortogonal a $T\mathcal{S}$ e que aponta para fora de M .

Portanto, tendo em conta as hipóteses, N_e só está definido nos pontos de \mathcal{S} onde não haja ambiguidade de todas as noções anteriores.

Teorema (de Gauss ou da divergência)

Suponhamos que é dada uma região M , com bordo seccionalmente de classe C^2 , contida no domínio de um campo vectorial F como acima. Então

$$\iiint_M \operatorname{div} F = \iint_{\partial M} \langle F, N_e \rangle d\sigma .$$

Note-se que o integral triplo é sobre o volume M .

O resultado é linear sobre uma união $M = M_1 \cup M_2 \cup \dots$.

Por exemplo, se se fizer uma decomposição *cirúrgica* de M , repare-se que as normais exteriores vão aparecer em bordos contíguos em direcções iguais mas sentido oposto.

A demonstração do teorema é um interessante e fácil exercício, desde que admitamos que M se possa decompor em regiões mais simples do tipo $M_i = I \times D$, com $I \subset \mathbb{R}$ e $D \subset \mathbb{R}^2$.

Exemplo. Seja $M = B_r(0) \subset \mathbb{R}^3$ e suponhamos $F = (x + y)e_1$. Então $\mathcal{S} = \partial M$ é a superfície esférica de raio 1.

Como $\operatorname{div} F = 1$, temos $\iiint_M \operatorname{div} F = V(M) = \frac{4}{3}\pi r^3$.

Por outro lado, é óbvio que $N_e = \frac{1}{r}(x, y, z)$ em cada ponto $(x, y, z) \in \mathcal{S}$. Então $\langle F, N_e \rangle = \frac{1}{r}(x^2 + xy)$.

Usando a parametrização de S usual, $z = \pm\sqrt{r^2 - x^2 - y^2}$, temos

$$d\sigma = \|N_1\| dx dy = \frac{r}{\sqrt{r^2 - x^2 - y^2}}$$

e logo

$$\iint_{\partial M} \langle F, N_e \rangle d\sigma = 2 \iint_D \frac{x^2 + xy}{\sqrt{r^2 - x^2 - y^2}} dx dy .$$

Em coordenadas polares chegamos também a $\frac{4}{3}\pi r^3$.

Bibliografia

- “Calculus 2”, Tom Apostol, McGraw Hill,
- “Análise Real”, F. R. Dias Agudo, Volume I e volume II, Escolar Editora.
- “Cálculo Diferencial e Integral”, volume II, N. Piskounov, Lopes da Silva Editora
- “Physics for scientists and engineers”, Fishbane and alt., Prentice Hall