Notas de Apoio
à disciplina de Matemática
do Mestrado Integrado em Ciências Farmacêuticas

Gonçalo Gutierres

Departamento de Matemática
Faculdade de Ciências e Tecnologia
Universidade de Coimbra
2022

Este texto contém um resumo da matéria da disciplina de Matemática do Mestrado Integrado em Ciências Farmacêuticas da Universidade de Coimbra. Ele é uma expansão dos diapositivos apresentados nas aula teóricas e contém, entre outras coisas, explicações mais detalhadas e a resolução de alguns dos exercícios propostos. Alguns dos assuntos poderão não estar apresentados pela mesma ordem que foram lecionados nas aulas.

Esta é a versão disponibilizada para os alunos do ano lectivo de 2022/2023.

Indíce

1 Funções Elementares

Genericamente vamos classificar como funções elementares todas as funções reais que nos são familiares, ou construídas a partir destas. Estas são as funções que têm as propriedades mais "agradáveis". Vamos portanto excluir todas as funções que precisam de ser definidas por ramos.

Definição 1.1. Uma função real diz-se uma *função elementar* se for uma função constante, polinomial, exponencial, logarítmica, trigonométrica, ou se resultar destas através de adições, produtos, quocientes ou radiciações. A composição e a inversa de funções elementares também são funções elementares. Vamos considerar como domínio de uma função elementar o conjunto dos pontos onde ela tem derivada finita.

Exemplos 1.2.

1. $f(x) = x^3 - 5x + 7$; (função polinomial)

2. $g(x) = \dfrac{3x-2}{\cos x}$;

3. $h(x) = \sqrt{\sin^2 x - x}$, no conjunto $D = \{x \in \mathbb{R} : \sin^2 x - x > 0\}$.

Advém diretamente da definição que todas as funções elementares são continuamente deriváveis. Aliás, mesmo sem a restrição adicional de que a derivada tem que ser finita todas as funções elementares são contínuas. Esta restrição adicional serve apenas para excluir casos patológicos como a função módulo, $f : \mathbb{R} \to \mathbb{R}$ com $f(x) = |x| = \sqrt{x^2}$, de serem funções elementares.

1.1 Funções Invertíveis

Nem todas as funções são invertíveis, para isso é preciso haver uma correspondência unívoca entre o domínio e o conjunto de chegada. Quando isso não acontece, podemos considerar restrições desses conjuntos.

Definições 1.3.

1. Uma função $f : A \to B$ é *injetiva* se $\forall x, y \in A,\ f(x) = f(y) \Rightarrow x = y$.

2. Uma função $f : A \to B$ é *sobrejetiva* $\forall y \in B, \exists x \in A : f(x) = y$.

3. Uma função é *bijetiva* se é injetiva e sobrejetiva.

Se $f : A \to B$ é bijetiva, então existe uma função $g : B \to A$ tal que

$$\forall x \in A,\ g(f(x)) = x \text{ e } \forall y \in B,\ f(g(y)) = y.$$

A função g é a *função inversa* de f, e representa-se por f^{-1}.

A função inversa, se existir, é única e pode igualmente ser definida pela condição $f(x) = y \Leftrightarrow x = f^{-1}(y)$, para todo $x \in A$ e para todo $y \in B$. Esta é a condição que habitualmente usamos para definir a inversa de funções injetivas, considerando como domínio da função inversa o contradomínio da função

Exemplo 1.4. A função exponencial é bijetiva se considerarmos como conjunto de chegada, o conjunto $]0,+\infty[$. A função logaritmo (de base e) é definida precisamente como sendo a função inversa da exponencial, ou seja $e^x = y \Leftrightarrow x = \ln y$.

$$\begin{array}{rccc} f: & \mathbb{R} & \to &]0,+\infty[\\ & x & \mapsto & e^x \end{array} \qquad \begin{array}{rccc} f^{-1}: &]0,+\infty[& \to & \mathbb{R} \\ & y & \mapsto & \ln y \end{array}$$

Todas as funções injetivas podem ser facilmente transformadas em funções bijetivas, restringindo o conjunto de chegada. Se uma função não é injetiva, por vezes considera-se uma restrição injetiva da função. Ou seja, uma restrição do domínio a um conjunto onde a função é injetiva.

Provavelmente o exemplo mais conhecido desta situação é a função quadrado $f(x) = x^2$. Esta função não é injetiva em $\mathbb{R}$, mas é injetiva em $[0,+\infty[$. A função raiz quadrada é definida como a função inversa da restrição de f aos reais não negativos.

$$\begin{array}{rccc} g: & [0,+\infty[& \to & [0,+\infty[\\ & y & \mapsto & \sqrt{y} \end{array}$$

Temos então que $x^2 = y \wedge x \geq 0 \Leftrightarrow x = \sqrt{y}$. Em linguagem corrente, dizemos que a raiz quadrada de y é o número não negativo x que elevado ao quadrado dá y.

Funções trigonométricas inversas

As funções trigonométricas seno, cosseno, tangente e cotangente não são injetivas. No entanto, como vimos anteriormente, podemos considerar a função inversa de restrições destas funções. As restrições principais das funções seno, cosseno, tangente e cotangente têm domínio $[-\frac{\pi}{2}, \frac{\pi}{2}]$, $[0,\pi]$, $]-\frac{\pi}{2}, \frac{\pi}{2}[$ e $]0,\pi[$, respectivamente. As suas inversas são designadas por arco seno, arco cosseno, arco tangente e arco cotangente.

Definições 1.5. As funções *arco seno, arco cosseno, arco tangente e arco cotangente* são definidas como as funções inversas das restrições principais das funções seno, cosseno, tangente e cotangente.

- $x = \arcsin y \Leftrightarrow y = \sin x \ \wedge \ x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$

$$\begin{array}{rccc} \sin: & [-\frac{\pi}{2}, \frac{\pi}{2}] & \rightarrow & [-1,1] \\ & x & \mapsto & \sin x \end{array} \qquad \begin{array}{rccc} \arcsin: & [-1,1] & \rightarrow & [-\frac{\pi}{2}, \frac{\pi}{2}] \\ & y & \mapsto & \arcsin y \end{array}$$

- $x = \arccos y \Leftrightarrow y = \cos x \ \wedge \ x \in [0, \pi]$

$$\begin{array}{rccc} \cos: & [0,\pi] & \rightarrow & [-1,1] \\ & x & \mapsto & \cos x \end{array} \qquad \begin{array}{rccc} \arccos: & [-1,1] & \rightarrow & [0,\pi] \\ & y & \mapsto & \arccos y \end{array}$$

- $x = \arctan y \Leftrightarrow y = \tan x \ \wedge \ x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$

$$\begin{array}{rccc} \tan: &]-\frac{\pi}{2}, \frac{\pi}{2}[& \rightarrow & \mathbb{R} \\ & x & \mapsto & \tan x \end{array} \qquad \begin{array}{rccc} \arctan: & \mathbb{R} & \rightarrow &]-\frac{\pi}{2}, \frac{\pi}{2}[\\ & y & \mapsto & \arctan y \end{array}$$

- $x = \operatorname{arccot} y \Leftrightarrow y = \cot x \ \wedge \ x \in]0, \pi[$

$$\begin{array}{rccc} \cot: &]0,\pi[& \rightarrow & \mathbb{R} \\ & x & \mapsto & \cot x \end{array} \qquad \begin{array}{rccc} \operatorname{arccot}: & \mathbb{R} & \rightarrow &]0,\pi[\\ & y & \mapsto & \operatorname{arccot} y \end{array}$$

Nota. Temos que $\sin(\arcsin y) = y$, uma vez que esta expressão só tem significado para $y \in [-1,1]$, mas em geral não podemos escrever que $\arcsin(\sin x) = x$, uma vez que esta igualdade só é válida para $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$. Por exemplo $\arcsin(\sin \pi) = \arcsin(0) = 0 \neq \pi$.

Funções hiperbólicas

As funções hiperbólicas têm algumas propriedades semelhantes às funções trigonométricas, e por isso "herdaram" o nome das funções trigonométricas. Vamos então definir as funções seno hiperbólico, cosseno hiperbólico, tangente hiperbólica e cotangente hiperbólica. Esta família de funções pode ser definida, de modo equivalente, geométrica ou algebricamente. Neste texto vamos apenas usar a definição algébrica. Fazemos apenas uma pequena observação geométrica, o nome de funções *hiperbólicas* advém do facto da hipérbole $x^2 - y^2 = 1$ representar para estas funções um papel similar ao do círculo trigonométrico $x^2 + y^2 = 1$ para as funções trigonométricas.

Definições 1.6. As funções *seno hiperbólico, cosseno hiperbólico, tangente hiperbólica e cotangente hiperbólica* são definidas pelas expressões:

1. $\sinh x = \frac{1}{2}\left(e^x - e^{-x}\right),$ [seno hiperbólico]

2. $\cosh x = \frac{1}{2}\left(e^{x} + e^{-x}\right),$ [cosseno hiperbólico]

3. $\tanh x = \dfrac{\sinh x}{\cosh h} = \dfrac{e^{x} - e^{-x}}{e^{x} + e^{-x}},$ [tangente hiperbólica]

4. $\coth x = \dfrac{\cosh x}{\sinh h} = \dfrac{e^{x} + e^{-x}}{e^{x} - e^{-x}}.$ [cotangente hiperbólica]

As funções sinh, cosh, tanh têm domínio $\mathbb{R}$ e a função coth tem domínio $\mathbb{R} \setminus \{0\}$.

As principais fórmulas envolvendo funções trigonométricas têm fórmulas correspondentes para as funções hiperbólicas. Vamos enunciar apenas algumas das mais importantes. A sua demonstração não é muito difícil uma vez que pode ser feita a partir das propriedades da função exponencial.

1. $\cosh^2 x - \sinh^2 x = 1$ [fórmula fundamental]

2. $\cosh(2x) = \cosh^2 x + \sinh^2 x$

3. $\sinh(2x) = 2\sinh x\, \cosh x$

4. $(\sinh x)' = \cosh x$

5. $(\cosh x)' = \sinh x$

As funções seno, tangente e cotangente hiperbólicas são injetivas, e portanto invertíveis. Por isso apenas é preciso considerar uma restrição de domínio para definir a inversa de cosseno hiperbólico. As inversas destas funções são designadas por argumento seno hiperbólico, argumento cosseno hiperbólico, argumento tangente hiperbólica e argumento cotangente hiperbólica.

Definições 1.7. As funções argumento seno hiperbólico, argumento cosseno hiperbólico, argumento tangente hiperbólica e argumento cotangente hiperbólica são definidas como as funções inversas das funções seno, tangente e cotangente hiperbólicas e da restrição principal do cosseno hiperbólico.

- $x = \arg\sinh y \Leftrightarrow y = \sinh x$

$$\begin{array}{rccc} \sinh : & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & \sinh x \end{array} \qquad \begin{array}{rccc} \arg\sinh : & \mathbb{R} & \to & \mathbb{R} \\ & y & \mapsto & \arg\sinh y \end{array}$$

- $x = \arg\cosh y \Leftrightarrow y = \cosh x \,\wedge\, x \geq 0$

$$\begin{array}{rccc} \cosh : & [0,+\infty] & \to & [1,+\infty] \\ & x & \mapsto & \cosh x \end{array} \qquad \begin{array}{rccc} \arg\cosh : & [1,+\infty] & \to & [0,+\infty] \\ & y & \mapsto & \arg\cosh y \end{array}$$

- $x = \arg\tanh y \Leftrightarrow y = \tanh x$

$$\begin{array}{rccc} \tanh: & \mathbb{R} & \to &]-1,1[\\ & x & \mapsto & \tanh x \end{array} \qquad \begin{array}{rccc} \arg\tanh: &]-1,1[& \to & \mathbb{R} \\ & y & \mapsto & \arg\tanh y \end{array}$$

- $x = \arg\coth y \Leftrightarrow y = \coth x$

$$\begin{array}{rccc} \coth: & \mathbb{R}\setminus\{0\} & \to & \mathbb{R}\setminus[-1,1] \\ & x & \mapsto & \coth x \end{array} \qquad \begin{array}{rccc} arg\coth: & \mathbb{R}\setminus[-1,1] & \to & \mathbb{R}\setminus\{0\} \\ & y & \mapsto & \arg\coth y \end{array}$$

Nota. A única destas funções hiperbólicas que não é injetiva é o cosseno hiperbólico. Por esse motivo temos que considerar um restrição no domínio para que seja invertível.

2 Cálculo Diferencial

O Cálculo Diferencial é a área da Matemática onde se estudam taxas de variação, nomeadamente o declive de retas ao gráfico de funções de uma ou mais variáveis. Nesta disciplina vamos fazer uma breve revisão de alguns dos conceitos conhecidos para funções reais de uma variável, e generalizá-los para funções de duas ou mais variáveis.

2.0 Limites e Continuidade de funções reais de variável real

A noção de limite de uma função dá-nos informação sobre o que acontece a essa função na proximidade de um ponto dado.

Intuitivamente, dizemos que o *limite* da função f, quando x tende para a é o valor L se f puder tomar valores tão próximos de L quanto queiramos, escolhendo x numa vizinhança do ponto a, mas diferente de a. Vamos agora definir limite em linguagem matemática.

Definição 2.1. Seja $f : D \to \mathbb{R}$ e a um ponto de acumulação[1] de D.
$\lim\limits_{x\to a} f(x) = L$ se $\forall \epsilon > 0\, \exists \delta > 0,\ 0 < |x-a| < \delta \Rightarrow |f(x) - L| < \epsilon$.

Exemplos 2.2.

1. $\lim\limits_{x\to 1} 2x + 3 = 5$ [regra da substituição]

2. $f(x) = \begin{cases} x^3 - 3 & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases}$ $\lim\limits_{x\to 0} f(x) = -3 \neq 1 = f(0)$

[1] a é um ponto de acumulação de D se $\forall \delta > 0,\ ]a-\delta, a+\delta[\setminus\{a\} \cap D \neq \emptyset$.

3. Vamos verificar que $\lim\limits_{x\to 0} \sin\frac{\pi}{x}$ não existe. Para tal vamos usar um resultado que nos diz que se o limite existe, então a imagem de qualquer sucessão que tenda para 0, tem que tender para esse limite.

   Consideremos as sucessões, convergentes para 0, de termo geral $u_n = \frac{1}{n}$ e $v_n = \frac{2}{4n+1}$. Se os limites das sucessões $f(u_n)$ e $f(v_n)$ forem diferentes ou um deles não existir, então $\lim\limits_{x\to 0} \sin\frac{\pi}{x}$ não existe.

   $\lim f(u_n) = \lim \sin\frac{\pi}{1/n} = \lim \sin n\pi = 0$

   $\lim f(v_n) = \lim \sin\frac{\pi}{2/(4n+1)} = \lim \sin(\frac{\pi}{2} + 2n\pi) = 1$

   Podemos então concluir que o limite não existe.

Limites laterais

Se na definição de limite considerarmos que os valores de x, pelos quais nos aproximamos de a, são sempre maiores (menores) do que a, chamamos a esse limite *limite à direita (esquerda)*.

- $\lim\limits_{x\to a^+} f(x) = L$ se $\forall\epsilon > 0\, \exists\delta > 0,\ 0 < x - a < \delta \Rightarrow |f(x) - L| < \epsilon.$ [limite à direita]
- $\lim\limits_{x\to a^-} f(x) = L$ se $\forall\epsilon > 0\, \exists\delta > 0,\ 0 < a - x < \delta \Rightarrow |f(x) - L| < \epsilon.$ [limite à esquerda]

O limite existe se ambos os limites laterais existirem e forem iguais. Os limites laterais só estão definidos se a for um ponto de acumulação do domínio à direita ou à esquerda, respectivamente.

Exemplo 2.3.

$$f(x) = \begin{cases} x+2 & \text{se } x < 0 \\ -5 & \text{se } x \geq 0 \end{cases} \qquad \lim_{x\to 0^-} f(x) = 2 \neq -5 = \lim_{x\to 0^+} f(x)$$

Limites no infinito

Tal como no caso finito, o limite de uma função quando esta se aproxima de um dado ponto é infinito se puder tomar valores tão grandes quanto queiramos, escolhendo x numa vizinhança do ponto a.

Por outro lado dizemos que o limite de uma função no infinito existe, se à medida que tomarmos valores cada vez maiores nos aproximamos tanto quanto queiramos do valor limite.

Estes dois tipos de limites podem ser escritos rigorosamente em linguagem matemática do seguinte modo.

- $\lim\limits_{x\to a} f(x) = +\infty$ se $\forall M > 0\, \exists\delta > 0,\ 0 < |x - a| < \delta \Rightarrow f(x) > M.$
- $\lim\limits_{x\to a} f(x) = -\infty$ se $\forall M < 0\, \exists\delta > 0,\ 0 < |x - a| < \delta \Rightarrow f(x) < M.$

- $\lim\limits_{x \to +\infty} f(x) = L$ se $\forall \epsilon > 0 \, \exists N > 0, \ x > N \Rightarrow |f(x) - L| < \epsilon$.
- $\lim\limits_{x \to -\infty} f(x) = L$ se $\forall \epsilon > 0 \, \exists N < 0, \ x < N \Rightarrow |f(x) - L| < \epsilon$.

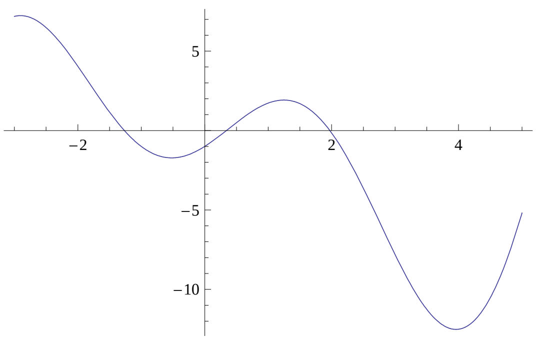
Além destes casos, podemos facilmente definir quando é que um limite no infinito é infinito, assim como os casos em que os limites laterais são infinitos.

O limite da soma, diferença, produto, quociente e potência de duas funções, no mesmo ponto, é a soma, diferença, produto e quociente dos limites, sempre que ambos estiverem definidos e o seu resultado não for uma indeterminação. As indeterminações serão estudadas mais à frente.

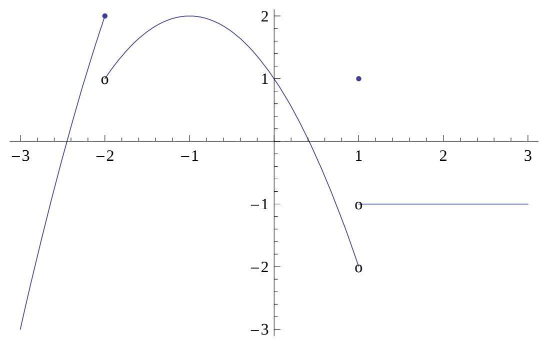
Continuidade

Definição 2.4. Uma função $f : D \to \mathbb{R}$ é *contínua* em $a \in D$ se $\lim\limits_{x \to a} f(x) = f(a)$. Caso contrário, diz-se que a função é descontínua em a. Da mesma forma se pode definir continuidade lateral.

Uma função é contínua no conjunto D se é contínua em todos os pontos do seu domínio.



Função contínua no intervalo $[-3, 5]$.



Função descontínua em -2 e 1. No entanto, é contínua à esquerda em -2.

Uma continuidade diz-se removível se $\lim\limits_{x \to a} f(x)$ existe e é finito. Isto significa que se existe uma função g tal que $f(x) = g(x)$, para $x \neq a$, mas que é contínua em a. Ou seja, a descontinuidade de f em a foi removida.

Exemplo 2.5.

$$f(x) = \begin{cases} x \sin \frac{\pi}{x} & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases} \qquad \lim_{x \to 0} x \sin \frac{\pi}{x} = 0 \text{ (infinitésimo } \times \text{ f.limitada)}$$

A descontinuidade de f é removível. Se removermos essa descontinuidade, obtemos a função g.

$$g(x) = \begin{cases} x \sin \frac{\pi}{x} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

A função $g : \mathbb{R} \to \mathbb{R}$ é um exemplo de uma função contínua em $\mathbb{R}$, mas que não pode ser escrita numa única "expressão".

Tal como para os limites, a soma, diferença, produto, quociente e potência de duas funções contínuas é uma função contínua no seu domínio. Também a composição de duas funções contínuas e a inversa de uma função contínua são funções contínuas, nos respectivos domínios.

Podemos assim afirmar que todas as funções elementares são funções contínuas. As *funções elementares* são as que resultam das funções polinomiais, logarítmicas, trigonométricas e trigonométricas inversas através de um número finito de somas, diferenças, produtos, quocientes e potências. Por exemplo as funções de expressão analítica e^x, $\frac{2x^3+1}{\cos x}$ e $\sqrt{\ln x + 5}$ são funções elementares. De facto, com excepção da função módulo, todas as funções estudadas até aqui, que se possam escrever numa expressão única, são funções elementares.

2.1 Derivadas

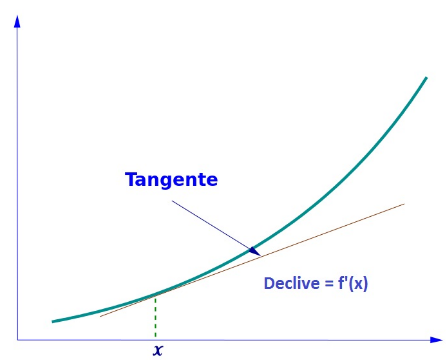
Definição 2.6. Seja f uma função definida no intervalo $]a-\epsilon, a+\epsilon[$, com $\epsilon > 0$. A *derivada* da função f no ponto $a \in \mathbb{R}$ é

$$f'(a) = \frac{df}{dx}(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}, \text{ se este limite existir.}$$

A última igualdade resulta simplesmente da mudança de variável $x - a = h$.

Notas.

1. Se só existir limite à esquerda ou à direita (ou forem diferentes), chamam-se derivada à direita $f'_d(a)$ e derivada à esquerda $f'_e(a)$, respectivamente, de f no ponto a.



2. Geometricamente, a derivada no ponto x representa o declive da reta tangente nesse ponto, como indicado na figura.
3. Em geral, $f'(a)$ representa a taxa de variação instantânea de y em função de x.
4. À notação de derivada do tipo $\dfrac{df}{dx}$ chama-se notação de Leibniz, por ter isso introduzida pelo matemático alemão Gottfried W. Leibniz.

Derivada da função composta

Se f é diferenciável em x e g é diferenciável em $g(x)$, então $g \cdot f$ é diferenciável em x e

$$(g \cdot f)'(x) = g'(f(x))\, f'(x).$$

Se escrevermos $y = g \cdot f(x)$ e $u = f(x)$, então a derivada da função composta pode ser escrita em termos simbólicos da seguinte maneira.

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} \qquad \text{[regra da cadeia]}$$

Exemplo 2.7. Muitas das regras de derivadas mais usadas baseiam-se na regra da derivada da função composta, ou regra da cadeia.

Sejam $u = f(x) = 3x$ e $y = g(u) = \sin u$. Vamos agora ver qual é a derivada da função $g \cdot f$.

$(g \cdot f)'(x) = g'(f(x))\ f'(x) = g'(u)\, f'(x) = \cos u \times 3 = \cos(3x) \times 3$.

Usando a notação de Leibniz $\dfrac{dy}{dx} = \dfrac{dy}{du} \times \dfrac{du}{dx} = \cos u \times 3 = \cos(3x) \times 3$.

Claro que normalmente usamos esta regra diretamente, sem detalhar os passos intermédios, ou seja $(g \cdot f)'(x) = (\sin(3x))' = \cos(3x) \times 3$.

Exercício

Num recipiente cónico com raio igual a $2dm$ e altura igual a $4dm$ é colocada água à taxa constante de $2dm^3/min$. A que velocidade está a subir a água, quando está a $3dm$ de altura?

Designemos por V o volume de água num dado instante, por h a altura da água e por r o raio da superfície de água. Como a altura do cone é de $4dm$ e o raio do cone é de $2dm$, usando uma regra de três simples temos que $r = \frac{h}{2}$.

Usando a fórmula do volume do cone deduzimos que o volume de água no recipiente quando a água está à altura h é

$$V = \frac{1}{3}\pi h r^2 = \frac{1}{3}\pi h \left(\frac{h}{2}\right)^2 = \frac{\pi}{12}h^3 \ .$$

Por outro lado, do enunciado, sabemos que $\dfrac{dV}{dt} = 2$.

Finalmente vamos usar a regra da derivada da composta para calcular o valor de $h'(t)$.

$$\frac{dV}{dt} = \frac{dV}{dt} \times \frac{dh}{dt} \Leftrightarrow 2 = \left(\frac{\pi}{4}h^2\right) \times h'(t) \Leftrightarrow h'(t) = \frac{8}{\pi h^2}$$

Quando $h = 3$, a água está a subir à velocidade à velocidade de $h'(t) = \dfrac{8}{9\pi}\ dm/min$.

Derivada da função inversa

A partir da derivada da função composta, deduz-se uma regra para a derivada da função inversa.

Seja g a função inversa de f, ou seja $f \cdot g(x) = x$. Vamos supor que $f'(g(x))$ existe e é diferente de zero. Usando a regra da função composta.

$(f \cdot g)'(x) = x' \Leftrightarrow f'(g(x))\ g'(x) = 1 \Leftrightarrow g'(x) = \dfrac{1}{f'(g(x))}$

Teorema 2.8. *Se f é diferenciável e invertível em x, com $f'(x) \neq 0$, então a sua função inversa f^{-1} é diferenciável em $y = f(x)$ e*

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'\left(f^{-1}(y)\right)}.$$

Exemplo 2.9. Vamos aproveitar a regra da derivada da função inversa, para deduzir a derivar da função arcsin.

Sejam $f(x) = \sin x = y$ e $f^{-1}(y) = \arcsin y = x$, com $x \in]-\frac{\pi}{2}, \frac{\pi}{2}[$ e $y \in]-1, 1[$. Os intervalos são abertos para que $f'(x) = \cos x$ seja diferente de zero. Então

$$(\arcsin y)' = \frac{1}{(\sin x)'} = \frac{1}{\cos x} = \frac{1}{\cos(\arcsin y)} = \frac{1}{\sqrt{1-y^2}}.$$

De maneira idêntica podemos deduzir as derivadas das outras funções trigonométricas inversas.

Seja u uma função real derivável, então:

1. $(\arcsin u)' = \dfrac{u'}{\sqrt{1-u^2}}$;
2. $(\arccos u)' = -\dfrac{u'}{\sqrt{1-u^2}}$;
3. $(\arctan u)' = \dfrac{u'}{1+u^2}$;
4. $(\operatorname{arccot} u)' = -\dfrac{u'}{1+u^2}$.

Regra de l'Hôpital

A regra de l'Hôpital é uma regra muito prática que serve para levantar indeterminações do tipo $\frac{0}{0}$ ou $\frac{\infty}{\infty}$. No entanto, qualquer outro tipo de indeterminação pode ser transformada numa indeterminação deste tipo, e portanto a regra pode ser aplicada.

Teorema 2.10. *Sejam f e g funções diferenciáveis num intervalo que contém a (a pode ser igual a $\pm\infty$).*

Se $\lim_{x\to a} f(x) = 0 = \lim_{x\to a} g(x)$ ou $\lim_{x\to a} f(x) = \pm\infty = \lim_{x\to a} g(x)$, e $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ existe, então

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Demonstração. (caso particular) Vamos provar apenas o caso da indeterminação $\frac{0}{0}$ em que as funções f' e g' são contínuas em $a \in \mathbb{R}$, o que implica que f e g também sejam contínuas em a e portanto $f(a) = g(a) = 0$. Consideramos ainda $g'(a) \neq 0$.

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = \frac{f'(a)}{g'(a)} = \frac{\lim_{x\to a} \frac{f(x)-f(a)}{x-a}}{\lim_{x\to a} \frac{g(x)-g(a)}{x-a}} = \lim_{x\to a} \frac{f(x)-f(a)}{g(x)-g(a)} = \lim_{x\to a} \frac{f(x)}{g(x)}$$

□

Exemplos 2.11.

1. $\lim\limits_{x\to 1} \dfrac{\ln x}{x-1} \left(\dfrac{0}{0}\right) = \lim\limits_{x\to 1} \dfrac{1/x}{1} = 1$

2. $\lim\limits_{x\to +\infty} \dfrac{e^x}{x^2} \left(\dfrac{\infty}{\infty}\right) = \lim\limits_{x\to +\infty} \dfrac{e^x}{2x} \left(\dfrac{\infty}{\infty}\right) = \lim\limits_{x\to +\infty} \dfrac{e^x}{2} = +\infty$

3. $\lim\limits_{x\to 0^+} \dfrac{\cos x}{\sin x} = +\infty \neq \lim\limits_{x\to 0^+} \dfrac{-\sin x}{\cos x} = 0$

   Neste caso não podemos aplicar a regra de l'Hôpital porque este limite não é uma indeterminação.

4. $\lim\limits_{x\to +\infty} \dfrac{x+\cos x}{x} \left(\dfrac{\infty}{\infty}\right) \neq \lim\limits_{x\to +\infty} \dfrac{1-\sin x}{1}$

   O limite da direita não existe, no entanto $\lim\limits_{x\to +\infty} \dfrac{x\sin x+\cos x}{x^2} = \lim\limits_{x\to +\infty} 1 + \dfrac{\cos x}{x} = 1.$

Nota. Uma consequência importante da regra de l'Hôpital é o seguinte resultado:

se uma função f está definida no intervalo $]a-\delta, a+\delta[$ e tem derivada em $]a-\delta, a+\delta[\setminus\{a\}$ tal que $\lim\limits_{x\to a} f'(x)$ existe, então $f'(a)$ também existe e é igual a $\lim\limits_{x\to a} f'(x)$.

Indeterminações do tipo 0^0, ∞^0 e 1^∞

Estas indeterminações transformam-se numa indeterminação do tipo $0 \times \infty$ usando as propriedades dos limites e do logaritmo. A partir daí podemos usar um dos métodos disponíveis para levantar esta indeterminação, como a regra de l'Hôpital por exemplo. De seguida vamos ver como se faz essa transformação.

$$\lim_{x\to a} f(x)^{g(x)} = \qquad (f(x) > 0 \text{ numa vizinhança de } a)$$

$$e^{\ln\left(\lim_{x\to a} f(x)^{g(x)}\right)} = e^{\lim_{x\to a} \ln\left(f(x)^{g(x)}\right)} = \qquad (\ln \text{ é contínua})$$

$$e^{\lim_{x\to a} g(x)\times \ln f(x)} \qquad (\text{propriedade do logaritmo})$$

O limite $\lim\limits_{x\to a} g(x) \times \ln f(x)$ é uma indeterminação do tipo $0 \times \infty$ (ou $\infty \times 0$ conforme o caso).

Exemplo 2.12. $\lim\limits_{x\to 0^+} x^x \left(0^0\right) = e^{\ln\left(\lim_{x\to 0^+} x^x\right)} = e^{\lim_{x\to 0^+} \ln(x^x)} = e^{\lim_{x\to 0^+} (x \times \ln x)} \left(e^{0\times(-\infty)}\right)$

Cálculo Auxiliar:

$\lim\limits_{x\to 0^+} x \times \ln x \, (0 \times \infty) = \lim\limits_{x\to 0^+} \dfrac{\ln x}{1/x} = \lim\limits_{x\to 0^+} \dfrac{(\ln x)'}{(1/x)'} = \lim\limits_{x\to 0^+} \dfrac{1/x}{-1/x^2} = \lim\limits_{x\to 0^+} -x = 0$

Finalmente concluímos que: $\lim\limits_{x\to 0^+} x^x = e^0 = 1$

2.2 Funções de Várias variáveis reais

No nosso dia a dia deparamo-nos frequentemente com situações em que precisamos de fazer cálculos com mais do que uma variável. Na prática estamos a calcular o valor de uma função com mais do que uma variável num determinado ponto. Comecemos por ver um exemplo simples de uma função de três variáveis que talvez já tenham usado.

Exemplo 2.13. Quantos custa o combustível gasto, em média, num quilómetro por um automóvel, sabendo que gastou x litros de combustível para percorrer y quilómetros, e o preço do combustível é de z euros por litro? Não é difícil descobrir que o custo por quilómetro é dada pela função

$$C(x,y,z) = \frac{x\,z}{y} \text{ euro}/km\ .$$

Ou seja, o custo médio depende de três variáveis, x, y e z.

De seguida vamos generalizar alguns dos conceitos conhecidos para funções reais de uma variável a funções reais de várias variáveis. As definições serão feitas, quase sempre, para funções de duas variáveis, ou seja em $\mathbb{R}^2 = \{(x,y)\ :\ x \in \mathbb{R} \text{ e } y \in \mathbb{R}\}$, ainda que na maioria dos casos seja fácil generalizá-las.

Definição 2.14. Uma função real de duas variáveis reais é uma função

$$\begin{array}{rccc} f: & D \subseteq \mathbb{R}^2 & \to & \mathbb{R} \\ & (x,y) & \mapsto & f(x,y) \end{array},$$

onde o *domínio* é um subconjunto de $\mathbb{R}^2$ (o conjunto dos pontos do plano) e o *contradomínio* está contido em $\mathbb{R}$.

Definição 2.15. O *gráfico de uma função real de duas variáveis reais* é uma superfície no espaço, ou seja em $\mathbb{R}^3$,

$$\text{Gráfico de } f = \{(x,y,z) \in \mathbb{R}^3\ :\ z = f(x,y) \text{ e } (x,y) \in D_f\}\,.$$

O gráfico de uma função de duas variáveis pode ser representado geometricamente, no entanto neste caso não é possível generalizar o conceito a funções de mais variáveis.

Exemplo 2.16. Consideremos a função de duas variáveis com expressão analítica $f(x,y) = \sqrt{9 - x^2 - y^2}$. Para calcular o domínio desta função procedemos como habitualmente calculando o conjunto de pontos do plano onde a expressão $f(x,y)$ está definida. O domínio é então

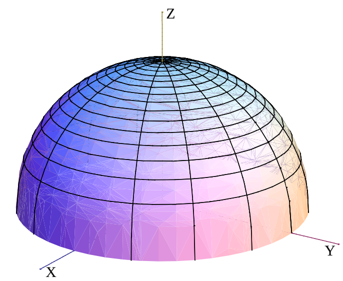
$$D_f = \{(x,y)\ :\ 9 - x^2 - y^2 \geq 0\} = \{(x,y)\ :\ x^2 + y^2 \leq 9\},$$

ou seja o domíno desta função é o círculo de raio 3 e centro na origem.

Para calcular o contradomínio de f basta observar que
$0 \leq x^2 + y^2 \leq 9 \Leftrightarrow 9 - 0 \geq 9 - (x^2 + y^2) \geq 9 - 9 \Rightarrow \sqrt{9} \geq \sqrt{9 - x^2 - y^2} \geq 0$. Temos então que o contradomínio é o conjunto:

$$D'_f = \{f(x,y)\ :\ (x,y) \in D_f\} = \{\sqrt{9 - x^2 - y^2}\ :\ x^2 + y^2 \leq 9\} = [0,3]\,.$$

O gráfico de f é o conjunto dos pontos do espaço que verificam a equação $z = \sqrt{9 - x^2 - y^2}$ o que implica que $z^2 = 9 - x^2 - y^2 \Leftrightarrow x^2 + y^2 + z^2 = 9$. Como $z = \sqrt{9 - x^2 - y^2} \geq 0$, o gráfico de f é a semi-esfera superior da esfera de centro na origem e raio 3.



O gráfico de f é a parte superior da esfera $x^2 + y^2 + z^2 = 9$.

2.3 Derivadas Parciais

Generalizar o conceito de derivada é um dos assuntos mais sensíveis no estudo de funções de várias variáveis. Algumas propriedades de função derivável em $\mathbb{R}$ só são obtidas em $\mathbb{R}^n$ com a noção de diferenciabilidade (da qual vamos falar apenas muito levemente), no entanto existe uma noção mais fraca mas mais fácil de utilizar e à qual podemos aplicar muitos dos conhecimentos adquiridos para funções de uma variável, a derivada parcial.

Antes da definição formal de derivada parcial, vamos ver como é que podemos usar as regras de derivação para calcular as derivadas parciais de funções elementares. Uma função de duas variáveis tem duas derivadas parciais, uma em ordem a cada uma das duas variáveis – usualmente x e y. A derivada parcial tem duas notações diferentes, que aproveitamos para introduzir.

Exemplos 2.17.

1. $f(x, y) = 3x + xy$

   A derivada parcial de f em ordem a x é $\dfrac{\partial f}{\partial x}(x, y) = f_x(x, y) = 3 + y$

   [y funciona como uma constante.]

   A derivada parcial de f em ordem a y é $\dfrac{\partial f}{\partial y}(x, y) = f_y(x, y) = x$

   [x funciona como uma constante.]

2. $f(x, y) = \sin(x^2 y) + e^x - y^2$

   $$\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = 2xy\cos(x^2 y) + e^x$$

   $$\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = x^2\cos(x^2 y) - 2y$$

Vamos de seguida apresentar a definição de derivada parcial num ponto (a,b), para funções de duas variáveis. Para funções de três ou mais variáveis a definição é em tudo semelhante. Esta definição é feita a partir da definição de derivada para funções de uma variável, fixando uma variável de cada vez.

Consideremos uma função f de duas variáveis e (a,b) um ponto do seu domínio. Define-se a função, de uma variável, $g(x) = f(x,b)$, ou seja, fixa-se a segunda variável $y = b$.

Se $g'(a)$ existe, então a esse valor chama-se *derivada parcial de f em ordem a x no ponto* (a,b). De igual modo podemos definir derivada parcial de f em ordem a y no ponto (a,b).

Definição 2.18. Sejam f uma função de duas variáveis e $(a,b) \in D_f$.

1. A *derivada parcial de f em ordem a x* no ponto (a,b) é o valor do limite

$$\frac{\partial f}{\partial x}(a,b) := \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h} \quad (= g'(a))\,.$$

2. A *derivada parcial de f em ordem a y* no ponto (a,b) é o valor do limite

$$\frac{\partial f}{\partial y}(a,b) := \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}\,.$$

Naturalmente as derivadas parciais podem não existir em todos os pontos. De seguida vamos ver um exemplo onde num dos pontos existe uma derivada parcial, mas não a outra.

Exemplo 2.19.

Consideremos a função de expressão analítica $f(x,y) = \begin{cases} \frac{x^3+y^2}{x^2+y^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = 0 \end{cases}$.

Nos pontos que não estão na fronteira entre os dois ramos, neste caso todos excepto o ponto $(0,0)$, podemos usar as regras de derivação para calcular as derivadas parciais. Ou seja, o único ponto "anguloso"é o ponto $(0,0)$. Vamos então verificar se a função tem derivadas parciais neste ponto.

$$\frac{\partial f}{\partial x}(0,0) := \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{h^3/h^2}{h} = \lim_{h\to 0} \frac{h^3}{h^3} = 1$$

$$\frac{\partial f}{\partial y}(0,0) := \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{h^2/h^2}{h} = \lim_{h\to 0} \frac{1}{h^3}$$

Ou seja $\frac{\partial f}{\partial x}(0,0) = 1$, mas $\frac{\partial f}{\partial y}(0,0)$ não existe.

A derivada parcial, sendo definida à custa da derivada de uma função de uma variável, tem essencialmente o mesmo significado que esta. A derivada parcial representa assim uma taxa de variação. Voltando ao Exemplo 2.13 do início da secção 2.2, se calcularmos a derivada de C em ordem a z, estamos a calcular a variação de C (€$/km$) em relação à variação de z

($€/litro$), ou seja o resultado será dado em $\dfrac{€/km}{€/litro} = litros/km$. Se derivarmos em ordem a y, por exemplo, o resultado será em $€/km^2$.

Interpretação Geométrica

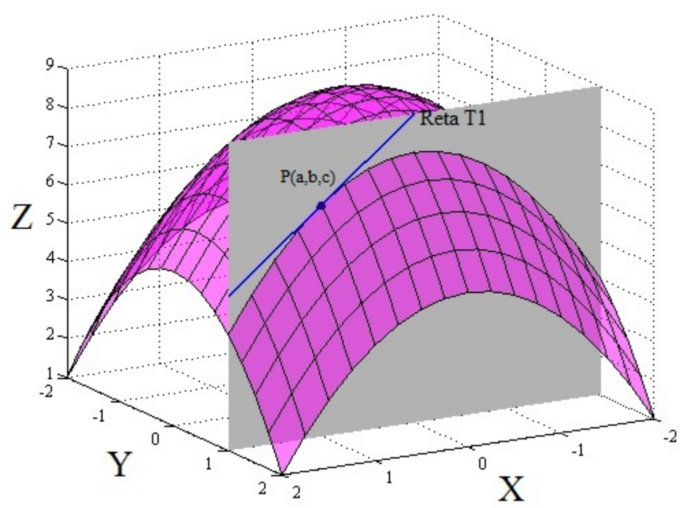
Para as funções de duas variáveis as derivadas parciais representam o declive de uma reta tangente. Neste caso a reta tangente ao gráfico de uma função num ponto não é única, existem uma infinidade delas. As derivadas parciais representam os declives de duas dessas retas, as retas paralelas ao plano XOZ e ao plano YOZ, respectivamente.

Sejam $z = f(x, y)$ o gráfico de uma função $f : \mathbb{R}^2 \to \mathbb{R}$ com $f(a, b) = c$, T_1 a reta tangente ao gráfico em $P = (a, b, c)$ e paralela ao plano XOZ; e T_2 a reta tangente ao gráfico em P e paralela ao plano YOZ. Temos que:

$$\text{declive de } T_1 = \frac{\partial f}{\partial x}(a, b) \quad \text{e} \quad \text{declive de } T_2 = \frac{\partial f}{\partial y}(a, b)\,.$$

Por exemplo, a equação $z = 8 - x^2 - y^2$ define o gráfico da função $f(x, y) = 8 - x^2 - y^2$. Como $f(1, 1) = 6$, a reta tangente ao gráfico em $P = (1, 1, 6)$ que é paralela ao plano XOZ é a reta T_1 que tem declive igual a $\dfrac{\partial f}{\partial x}(1, 1) = 6$.

A sua equação vetorial é $(x, y, z) = (1, 1, 6) + t(1, 0, -2), \;\; t \in \mathbb{R}$.



Reta tangente ao gráfico na direção do eixo dos XX.

Derivadas Parciais de Segunda Ordem

Tal como nas funções de uma variável, é possível derivar a mesma função várias vezes. Neste caso é preciso ter em atenção que tal como podemos derivar a primeira vez em relação em cada uma das variáveis, o mesmo acontece quando derivamos pela segunda vez. Por exemplo, para uma função de três variáveis existem três derivadas parciais (de primeira ordem) e $3 \times 3 = 9$ derivadas parciais de segunda ordem, 3^3 de terceira ordem e assim sucessivamente. Por facilidade vamos definir as derivadas de segunda ordem de uma função de duas variáveis, sendo que é imediato perceber como se procede para funções de mais variáveis.

Definição 2.20. Seja $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$, então as (quatro) derivadas parciais de segunda ordem de f são:

- $\dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial x}\right) = \dfrac{\partial^2 f}{\partial x^2} = f_{xx}$;
- $\dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) = \dfrac{\partial^2 f}{\partial x \partial y} = f_{yx}$; [deriva-se primeiro em ordem a y e depois em ordem a x]
- $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right) = \dfrac{\partial^2 f}{\partial y \partial x} = f_{xy}$; [deriva-se primeiro em ordem a x e depois em ordem a y]
- $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial y}\right) = \dfrac{\partial^2 f}{\partial^2 y} = f_{yy}$.

Por iteração deste processo definem-se as derivadas parciais de ordens superiores.

O próximo teorema diz-nos que, em geral, não é preciso calcular f_{yx} se já conhecermos o valor de f_{xy}.

Teorema 2.21 (de Clairaut)**.** *Todas as funções elementares satisfazem a igualdade* $f_{yx}(a, b) = f_{xy}(a, b)$.

De modo prático esta igualdade aplica-se sempre que pudermos usar as regras de derivação, ou seja sempre que a função não estiver definida por ramos, que é o caso de quase todas as funções que iremos usar nesta disciplina.

O Teorema de Clairaut também se aplica a derivadas de ordem superior, por exemplo para todas as funções elementares temos que $f_{xxy} = f_{xyx} = f_{yxx}$.

Exemplo 2.22. Seja $f(x, y) = x^2 y - 3y$. As derivadas parciais de primeira ordem são $\dfrac{\partial f}{\partial x} = 2xy$ e $\dfrac{\partial f}{\partial y} = x^2 - 3$. As derivadas de segunda ordem são:

- $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial x}\right) = \dfrac{\partial}{\partial x}(2xy) = 2y$;

- $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right) = \dfrac{\partial}{\partial x}\left(x^2 - 3\right) = 2x;$
- $\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{\partial}{\partial y}\left(2xy\right) = 2x;$
- $\dfrac{\partial^2 f}{\partial^2 y} = \dfrac{\partial}{\partial y}\left(x^2 - 3\right) = 0\,.$

Vetor Gradiente

O vetor gradiente é um vetor que condensa a informação sobre as derivadas parciais num vetor, assim o *vetor gradiente* de uma função de n variáveis $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ é

$$\nabla f(x_1, x_2, \ldots, x_n) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)\ ;$$

onde as derivadas parciais são calculadas no ponto $(x_1, x_2, \ldots, x_n)$. O cálculo das derivadas parciais pode ser feito por definição ou usando as regras de derivação quando tal for possível. O vetor gradiente num ponto só está definido quando existem todas as derivadas parciais nesse ponto.

Exemplo 2.23. Para $f(x, y, z) = x^2 y + 2yz$, o vetor gradiente em (x, y, z) é

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = \left(2xy, x^2 + 2z, 2y\right)\ .$$

Se considerarmos um ponto concreto do domínio, o vetor gradiente é um vetor concreto (neste caso no espaço). Por exemplo,

$$\nabla f(1, 2, 0) = (4, 1, 4)\ .$$

2.4 Derivada Direcional

Como já foi referido anteriormente, as derivadas parciais representam apenas o declive de duas das retas tangentes ao gráfico de uma função de duas variáveis (ou mais geralmente a taxa de variação da função em n direções distintas para funções de n variáveis). Mas não há nenhum motivo para considerarmos apenas duas retas tangentes e não todas as outras. Acontece que, para funções "bem comportadas", conhecer as derivadas parciais é suficiente para calcular o declive de todas as retas tangentes ao gráfico de uma função de duas variáveis. Ao declive dessas retas chama-se derivada direcional. Ao contrário do que fizemos para as derivadas parciais, vamos começar por ver a definição formal de derivada direcional.

Definição 2.24. A *derivada direcional* de $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ no ponto (a, b) na direção, e sentido, do vetor unitário $\hat{u} = (u_1, u_2)$, isto é, $||\hat{u}|| = 1$, é

$$D_{\hat{u}} f(a, b) := \lim_{h \to 0} \frac{f(a+h\, u_1, b+h\, u_2) - f(a, b)}{h}\,.$$

Notas.

1. $D_{\hat{u}} f(a, b)$ mede a variação de f na direção, e sentido, de $\hat{u}$. Se o vetor não for unitário, considerarmos o vetor unitário com a mesma direção e sentido. Ou seja, para $\vec{u} \neq \vec{0}$, considera-se o vetor $\hat{u} = \dfrac{\vec{u}}{||\vec{u}||}$.

2. As derivadas parciais são casos particulares das derivadas direcionais. Considerando os versores $\hat{\imath} = (1, 0)$ e $\hat{\jmath} = (0, 1)$,
$$D_{\hat{\imath}} f(a, b) = f_x(a, b) \quad \text{e} \quad D_{\hat{\jmath}} f(a, b) = f_y(a, b)\,.$$

3. Se definirmos $g(h) = f(a + h\, u_1, b + h\, u_2)$, então $D_{\hat{u}} f(a, b) = g'(0)$, o que mostra que a derivada direcional representa uma taxa de variação. Neste caso a taxa de variação na direção e sentido do versor $\hat{u}$.

4. O sentido é importante, uma vez que se consideramos dois vetores simétricos, as respetivas derivadas direcionais também são simétricas, $D_{-\hat{u}} f(a, b) = -D_{\hat{u}} f(a, b)$.

5. De modo semelhante se definem as derivadas direcionais de funções com mais variáveis.

A noção de função diferenciável é uma noção complexa e que não temos tempo para abordar com detalhe neste curso. Intuitivamente dizemos que uma função de duas variáveis é *diferenciável* num ponto se o gráfico da função tem um **único** plano tangente (não vertical) ao gráfico nesse ponto. As funções diferenciáveis (de duas ou mais variáveis) num ponto têm todas as derivadas parciais nesse ponto, mas o contrário pode não se verificar.

É de notar que para funções de uma variável o conceito de diferenciabilidade coincide com o de derivabilidade. Intuitivamente podemos igualmente dizer que uma função de uma variável é diferenciável num ponto, se o gráfico da função tem uma **única** reta tangente (não vertical) ao gráfico nesse ponto.

As funções que vamos usar são usualmente diferenciáveis, uma que todas as funções elementares são diferenciáveis no interior do seu domínio. Sempre que tivermos uma função elementar, vamos usar o teorema que se segue para calcular a derivada direcional. Este resultado, na prática, generaliza o uso das regras de derivação a todas as derivadas direcionais. O teorema e a proposição seguintes são válidos para funções de n variáveis.

Teorema 2.25. *Se f é diferenciável em $(a,b) \in D_f$, então*

$$D_{\hat{u}} f(a,b) = \nabla f(a,b) \cdot \hat{u} = f_x(a,b)u_1 + f_y(a,b)u_2\,.$$

Sempre que estamos nas condições deste teorema é fácil descobrir qual é a direção em que derivada a direcional toma o maior (ou o menor) valor possível.

Proposição 2.26. *Se f é diferenciável em $(a,b) \in D_f$, então:*

1. *A derivada direcional de f em (a,b) é* **máxima** *na direção e sentido do vetor gradiente, ou seja para $\hat{u} = \dfrac{\nabla f(a,b)}{||\nabla f(a,b)||}$.*
2. *A derivada direcional de f em (a,b) é* **mínima** *na direção do vetor gradiente, mas no sentido oposto, ou seja para $\hat{u} = -\dfrac{\nabla f(a,b)}{||\nabla f(a,b)||}$.*
3. *A derivada direcional de f em (a,b) é* **nula** *na direção perpendicular ao vetor gradiente.*

Demonstração. Como f é diferenciável,

$D_{\hat{u}} f(a,b) = \nabla f(a,b) \cdot \hat{u} = ||\nabla f(a,b)||\,||\hat{u}||\cos\theta = ||\nabla f(a,b)||\cos\theta$, onde θ é o ângulo formado pelos vetores $\nabla f(a,b)$ e $\hat{u}$.

A função f e o ponto (a,b) são fixos, portanto este valor depende apenas do cosseno do ângulo formado pelos dois vetores. Portanto, o valor é máximo quando $\cos\theta = 1 \Leftrightarrow \theta = 0$, ou seja quando os vetores $\nabla f(a,b)$ e $\hat{u}$ têm a mesma direção e sentido. Neste caso o valor da derivada direcional é igual a $||\nabla f(a,b)||$. De igual modo, o valor da derivada direcional é mínimo quando $\cos\theta = -1 \Leftrightarrow \theta = \pi$, ou seja quando os vetores $\nabla f(a,b)$ e $\hat{u}$ têm a mesma direção mas sentidos opostos; e neste caso o valor da derivada direcional é igual a $-||\nabla f(a,b)||$.

A derivada direcional é nula quando $\cos\theta = 0 \Leftrightarrow \theta = \frac{\pi}{2}$, ou seja quando os vetores $\nabla f(a,b)$ e $\hat{u}$ são perpendiculares. $\square$

Exercício

Suponhamos que $h(x,y) = 2 - xy + x^2 + y$ representa a altitude, medida em km, numa dada região. Uma pessoa na origem do referencial caminha na direção SO. Está a subir ou a descer? Em que direção é que o declive é maior?

A função h indica a altitude de um determinado ponto, quando são conhecidas as suas coordenadas. (As coordenadas usualmente são a latitude e a longitude.) Consideremos o eixo dos XX no sentido $Oeste - Este$ e o eixo dos YY no sentido $Sul - Norte$.

Portanto a direção SO corresponde ao versor $\hat{u} = \left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$. Queremos descobrir se o valor $D_{\hat{u}} h(0,0)$ é positivo ou negativo.

O gradiente de h é $\nabla h(x,y) = (-y+2x, -x+1)$ e

$$D_{\hat{u}}h(0,0) = \nabla h(0,0)\cdot \hat{u} = (0,1)\cdot\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right) = -\frac{\sqrt{2}}{2} < 0\,.$$

O declive é máximo na direção do vetor $\nabla h(0,0) = (0,1)$, ou seja na direção $Sul - Norte$ e o seu valor é igual a $||(0,1)|| = 1$. Se todas os valores forem na mesma unidade de medida (km), então este declive corresponderia a uma inclinação de 45^{o}.

2.5 Aproximação Linear

Nem sempre é possível conhecer a maneira como uma função se comporta em todo o seu domínio, por isso por vezes é necessário recorrer a resultados aproximados. Uma das maneiras mais simples de obter a aproximação a uma função é usando uma aproximação linear. Para funções reais de uma variável, a aproximação linear de uma função f é a função de expressão $mx+b$ tal que $y = mx+b$ é a reta tangente ao gráfico da função f num determinado ponto. Portanto para calcularmos uma aproximação linear basta conhecer o valor da função e da(s) sua(s) derivada(s) num ponto.

Sejam $f : D \subseteq \mathbb{R} \to \mathbb{R}$ uma função real de variável real e a um ponto interior do domínio. A *aproximação linear* de f em torno de $a \in D$ é a função:

$$L(x) = f(a) + f'(a)\,(x-a)\,.$$

Quando $|x-a|$ é um valor muito pequeno, e f tem derivada contínua num intervalo que contém a e x, então $L(x)$ é uma boa aproximação de $f(x)$, ou seja

$$f(x) \approx L(x) = f(a) + f'(a)\,(x-a)\,.$$

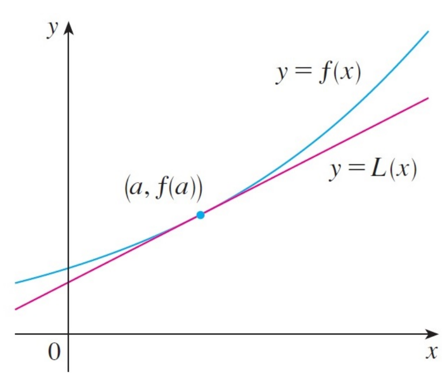
Notas.

1. A aproximação linear serve para *estimar* o valor de uma função, quando apenas sabemos como se comporta num ponto.

2. Cada função tem várias aproximações lineares, dependendo do ponto a escolhido.

3. O gráfico da aproximação linear $y = L(x)$ é a reta tangente ao gráfico da função $y = f(x)$ no ponto $(a, f(a))$.

4. A aproximação linear consiste em substituir uma função, pela função afim que tem por gráfico uma reta tangente ao gráfico da função.



Exercício

Calcule o valor aproximado de $\sqrt{4.01}$?

Considera-se a função real de expressão analítica $f(x) = \sqrt{x}$. Uma vez conhecida a função deduzimos que $f'(x) = \dfrac{1}{2\sqrt{x}}$.

Neste caso vamos usar a aproximação linear da função f em torno do ponto $a = 4$ para calcular um valor aproximado para $f(4.01) = \sqrt{4.01}$, ou seja vamos usar $x = 4.01$ e $\Delta x := x - a = 0.01$:

$$L(x) = f(4) + f'(4)\,(x-a) = 2 + \frac{1}{4}\,(x-4)\,.$$

Finalmente, concluímos que $f(4.01) \approx L(4.01) = f(4) + f'(4) \times 0.01 = 2 + \dfrac{1}{4} \times 0.01 = 2.0025$. Note-se que o valor real de $\sqrt{4.01}$ é $2.0024984\ldots$, o que significa que obtivemos uma aproximação com um erro inferior 2×10^{-6} quando o valor do *erro* inicial era $\Delta x = 0.01 = 10^{-2}$ mostrando que neste caso a aproximação linear é uma boa ferramenta.

A aproximação linear pode também ser útil quando sabemos como a função se comporta num ponto, ou num intervalo pequeno, e queremos deduzir qual o seu comportamento em zonas próximas.

De seguida vamos generalizar este processo para funções de duas variáveis. Tudo o vai ser feito para funções de duas variáveis pode facilmente ser adaptado para funções de mais variáveis. Em toda a construção que faremos a seguir vamos considerar que as funções são funções elementares, e que os pontos considerados são pontos interiores do seu domínio.

Aproximação Linear - funções de duas variáveis

Sejam $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ uma função real de duas variáveis e (a,b) um ponto interior do domínio D.

A *aproximação linear* de f em torno de $(a,b) \in \mathbb{R}^2$ é a função linear

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

Tal como no caso de uma variável, quando os valores $|x-a|$ e $|y-b|$ são muito pequenos, e f é diferenciável numa vizinhança de (a,b) que contém (x,y), então $L(x,y)$ é uma boa aproximação de $f(x,y)$, ou seja

$$f(x,y) \approx L(x,y) = f(a,b) + f_x(a,b)\,(x-a) + f_y(a,b)\,(y-b)\,.$$

O gráfico da aproximação linear $z = L(x,y)$ é o plano tangente ao gráfico da função $z = f(x,y)$ no ponto $P = (a,b,f(a,b))$. Ou seja, a aproximação linear consiste em substituir uma função, pela função linear que tem por gráfico um plano tangente ao gráfico da função.

Deste modo concluímos ainda que a equação do plano tangente ao gráfico da função $z = f(x, y)$ no ponto $P = (a, b, f(a, b))$ é o plano de equação

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

Diferencial

A noção de diferencial está intimamente ligada com a noção de aproximação linear e as duas podem ser facilmente confundidas. Em alguns dos exercícios será indiferente usar qualquer uma das notações, podendo no entanto existir vantagens de cálculo que tornem preferível usar uma delas. Optámos por introduzir a noção formal de diferencial, uma vez que será útil nos capítulos seguintes.

Definição 2.27.

1. Para uma função de uma variável $y = f(x)$, o *diferencial* de y é

$$dy = f'(x)\,dx = \frac{dy}{dx}\,dx\,.$$

2. Para uma função de duas variáveis $z = f(x, y)$, o diferencial de z ou *diferencial total* é

$$dz = f_x(a, b)\,dx + f_y(a, b)\,dy = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy\,.$$

A noção de diferencial é uma noção simbólica e na definição de diferencial total, dx e dy representam variáveis independentes.

Lendo dx e dy como variáveis independentes, podemos fazer as substituições $dx = \Delta x = x - a$, $dy = \Delta y = y - b$ e definindo $\Delta z = f(x, y) - f(a, b)$, temos que

$$\Delta z \approx dz\,, \text{ para valores pequenos de } dx \text{ e } dy.$$

Na notação dos diferenciais, a aproximação linear pode assim ser escrita

$$f(x, y) \approx f(a, b) + dz\,.$$

De igual modo para funções de uma variável podemos escrever $f(x) \approx f(a) + dy\,.$

Exercício

A função $T(x, y)$ mede a temperatura em graus Celsius no ponto (x, y) de uma superfície plana. Um rato é largado na superfície e encontra-se na posição $P = (2, 1)$, medida em cm. Sabendo que $T_x(2, 1) = 3$, $T_y(2, 1) = 5$ e a posição do rato foi determinada com um erro inferior a $1mm$, indique um valor aproximado para o erro cometido na temperatura a que está o local onde se encontra o rato.

O valor que queremos calcular é o valor absoluto de $\Delta T = T(2+dx, 1+dy) - T(2,1)$, com $|dx| \leq 0.1$ e $|dy| \leq 0.1$.

Vamos usar a abordagem dos diferenciais, ou seja:

$$|\Delta T| \approx |dT| = |T_x(2,1)\,dx + T_y(2,1)\,dy| \leq |T_x(2,1)\,dx| + |T_y(2,1)\,dy|$$

$$= |T_x(2,1)|\,|dx| + |T_y(2,1)|\,|dy| \leq 3 \times 0.1 + 5 \times 0.1 = 0.8\ .$$

O erro cometido é inferior a 0.8 graus Celsius.

De modo semelhante às aproximações lineares se definem as aproximações quadráticas, que consistem em aproximar funções através de funções de segundo grau.

- $Q(x) = L(x) + \dfrac{f''(a)}{2}\,(x-a)^2$ [Função de um variável.]
- $Q(x,y) = L(x,y) + \dfrac{f_{xx}(a,b)}{2}\,(x-a)^2 + f_{xy}(a,b)\,(x-a)(y-b) + \dfrac{f_{yy}(a,b)}{2}\,(y-a)^2$
[Função de duas variáveis.]

2.6 Derivada da Função Composta (regra da cadeia)

Tal como nos caso das funções de uma variável, é possível determinar a derivada (parcial) de uma função composta a partir das derivadas de cada uma das funções. Antes de enunciarmos o teorema para um caso mais geral, começamos por o caso mais elementar considerando uma função composta: $\mathbb{R} \to \mathbb{R}^2 \to \mathbb{R}$.

Teorema 2.28 (caso 1)**.** *Sejam $z = f(x,y)$ uma função com derivadas parciais e $x(t)$, $y(t)$ funções deriváveis, então*

$$\frac{dz}{dt} = \frac{\partial f}{\partial x} \times \frac{dx}{dt} + \frac{\partial f}{\partial y} \times \frac{dy}{dt}\ .$$

De notar que as variáveis x e y funcionam simultaneamente com variáveis independentes em relação a z e variáveis dependentes em a relação a t. Esta relações de dependência podem ser traduzidas mum diagrama a que costuma chamar *diagrama de árvore.*

$$\begin{tikzcd}[column sep=small] & z \arrow[dl, dash] \arrow[dr, dash] & \\ x \arrow[d, dash] & & y \arrow[d, dash] \\ t & & t \end{tikzcd}$$

Exemplos 2.29.

1. Consideremos $z = \sin(2x+y)$, com $x(t) = t^2$ e $y(t) = e^t$. A derivada de z em ordem a t é:

$$\frac{dz}{dt}(t) = \frac{\partial z}{\partial x}(x,y) \times \frac{dx}{dt}(t) + \frac{\partial z}{\partial y}(x,y) \times \frac{dy}{dt}(t) =$$

$$= 2\cos(2x+y) \times 2t + \cos(2x+y) \times e^t \ .$$

Para valores concretos de t podemos calcular a derivada desta função. Por exemplo quando $t = 1$, os valores de x e y são $x = 1^2 = 1$ e $y = e^1 = e$, e assim temos que:

$$\frac{dz}{dt}(1) = 4\cos(e+2) + e\cos(e+2) = (4+e)\cos(e+2)\,.$$

2. A regra da derivada da composta para funções de duas variáveis serve também para deduzir regras de derivação que utilizamos habitualmente. Fazendo $z = u^v$ com $u(x)$ e $v(x)$ funções de uma variável podemos deduzir uma regra de derivação para potências com variável tanto na base com no expoente.

$$(u^v)' = \frac{dz}{dx} = \frac{\partial z}{\partial u} \times \frac{du}{dx} + \frac{\partial z}{\partial v} \times \frac{dv}{dx} = u^{v-1}\,u' + u^v\,(\ln u)\,v'$$

Exercício

Num determinado gás existe a relação $PV = 8.31T$. Determine a taxa de variação da pressão(P), quando a temperatura(T) é $300k$ e está a aumentar à taxa de 0.1k/s; e o volume(V) é de $100l$ e aumenta $0.2l/s$.

Vamos resolver este exercício usando a derivada da composta. Vamos primeiro analisar os dados do problema. Consideramos t_0 o instante em que o problema é colocado. Nesse instante é-nos dito que $T_0 = 300$, $V_0 = 100$, $\dfrac{dT}{dt}(t_0) = 0.1$ e $\dfrac{dV}{dt}(t_0) = 0.2$.

Por outro lado $PV = 8.31T \Leftrightarrow P = 8.31\dfrac{T}{V}$ e consequentemente $\dfrac{\partial P}{\partial T} = \dfrac{8.31}{V}$ e $\dfrac{\partial P}{\partial V} = -\dfrac{8.31T}{V^2}$.

Finalmente estamos em condições de determinar a taxa de variação.

$$\frac{dP}{dt} = \frac{\partial P}{\partial V} \times \frac{dV}{dt} + \frac{\partial P}{\partial T} \times \frac{dT}{dt}$$

$$\frac{dP}{dt}(t_0) = \frac{8.31}{100} \times 0.1 - \frac{8.31 \times 300}{100^2} \times 0.2 = -0.04155\,KPa/s$$

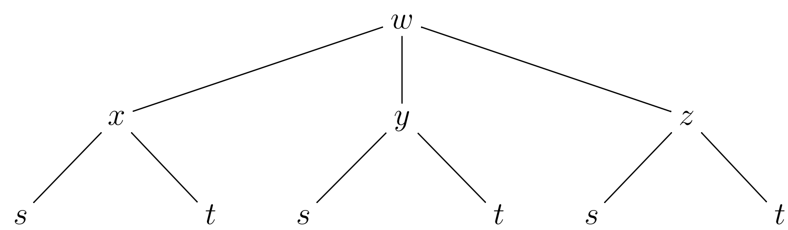
O Teorema 2.28 é o caso mais elementar possível com $z = f(x,y)$ uma função de duas ou mais variáveis. De seguida vamos expor uma situação um pouco mais geral de onde facilmente se deduzem os restantes casos.

Teorema 2.30 (caso 2). *Sejam $w = f(x,y,z)$, $x(s,t)$, $y(s,t)$ e $z(s,t)$ funções com derivadas parciais, então*

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x} \times \frac{\partial x}{\partial s} + \frac{\partial w}{\partial y} \times \frac{\partial y}{\partial s} + \frac{\partial w}{\partial z} \times \frac{\partial z}{\partial s}\,;$$

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x} \times \frac{\partial x}{\partial t} + \frac{\partial w}{\partial y} \times \frac{\partial y}{\partial t} + \frac{\partial w}{\partial z} \times \frac{\partial z}{\partial t}\,.$$

Diagrama de árvore



Exemplo 2.31. Consideremos $u = 3x^2y - xz$, como $x = s^2 + t^2$, $y = 2s$ e $z = t + 2$.

As derivadas parciais de w em ordem a s e a t são, respeticamente:

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \times \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \times \frac{dy}{ds} = (6xy - z) \times 2s + 3x^2 \times 2\,;$$

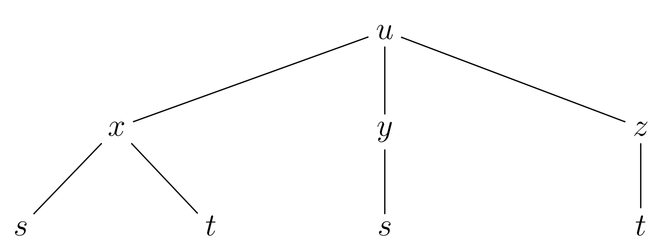
$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \times \frac{\partial x}{\partial t} + \frac{\partial u}{\partial z} \times \frac{dz}{dt} = (6xy - z) \times 2t - x \times 1\,;$$

Quando $(s,t) = (1,2)$, temos que $(x,y,z) = (5,2,4)$ e portanto:

$$\frac{\partial u}{\partial s}(1,2) = 56 \times 2 + 75 \times 2 = 262\,;$$

$$\frac{\partial u}{\partial t}(1,2) = 56 \times 4 - 5 \times 1 = 219\,.$$

Neste exemplo o diagrama de árvore toma um aspeto um pouco diferente, uma vez que a variável x depende das duas variáveis s e t, enquanto y e z são funções de apenas uma variável. Esta diferença justifica também a diferença de notação, visto que as derivadas de y e de z são derivadas totais e não derivadas parciais.



Uma iteração deste processo permite usar a regra da cadeia para as derivadas da composta de três ou mais funções.

2.7 Extremos de Funções de 2 Variáveis

As funções que temos estado a estudar têm $\mathbb{R}$ como conjunto de chegada, e por isso é possível definir os extremos absolutos, e locais, de modo idêntico ao que é feito para funções de uma variável. Estas definições são válidas para funções de duas ou mais variáveis. Por simplicidade elas são enunciadas para funções de duas variáveis.

Definições 2.32. Sejam $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ e $(a, b) \in D$.

1. $f(a, b)$ é um máximo (mínimo) absoluto de f se

$$(\forall (x, y) \in D)\ f(a, b) \geq f(x, y)\ (f(a, b) \leq f(x, y)).$$

2. A função f tem um máximo (mínimo) local em (a, b) se

$$(\exists \epsilon > 0)\ ||(x, y) - (a, b)|| < \epsilon \Rightarrow f(a, b) \geq f(x, y)\ (f(a, b) \leq f(x, y)).$$

Ao ponto (a, b) chama-se ponto maximizante (minimizante) local, ou absoluto, conforme o caso.

Para funções deriváveis de uma variável temos como condição necessária, para um ponto ser extremante local, a derivada ser nula nesse ponto. Esse resultado é facilmente generalizável para funções com mais variáveis.

Proposição 2.33. *Sejam $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ e $(a, b) \in D$. Se f tem um extremo local em (a, b), e as derivadas parciais existem nesse ponto, então $f_x(a, b) = 0 = f_y(a, b)$.*

Geometricamente isto significa que o plano tangente ao gráfico da função é horizontal. (Ver descrição do plano tangente na secção **Aproximação Linear**).

Aos pontos (a, b) tais que $f_x(a, b) = 0 = f_y(a, b)$ chamam-se *pontos críticos* de f.

Depois de calcularmos os pontos críticos, temos que identificar quais deles são pontos extremantes. Para funções de duas variáveis não é possível fazer um quadro de sinal. No entanto existe um resultado, semelhante ao que a seguir se enuncia para funções reais de variável real, que pode ser usado determinar se um ponto crítico é ou não extremante.

Teorema 2.34. *Sejam $f : D \subseteq \mathbb{R} \to \mathbb{R}$ e $a \in D$ tais que f tem derivadas de segunda ordem, e $f'(a) = 0$.*

1. *Se $f''(a) > 0$, então $f(a)$ é um mínimo local.*

2. *Se $f''(a) < 0$, então $f(a)$ é um máximo local.*

Teorema 2.35. *Sejam $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ e $(a,b) \in D$ tais que f tem derivadas parciais de segunda ordem e (a,b) é um ponto crítico de f.*

Define-se $D(a,b) = f_{xx}(a,b)\, f_{yy}(a,b) - \left(f_{xy}(a,b)\right)^2$.

1. *Se $D(a,b) > 0$ e $f_{xx}(a,b) > 0$, então $f(a,b)$ é um mínimo local.*
2. *Se $D(a,b) > 0$ e $f_{xx}(a,b) < 0$, então $f(a,b)$ é um máximo local.*
3. *Se $D(a,b) < 0$, então $f(a,b)$ não é um extremo local.*
   [(a,b) é um ponto sela].

Se $D(a,b) = 0$, então seria necessário estudar as derivadas (parciais) de ordem superior para saber se um ponto é extremante local. Por vezes é possível analisar esses casos usando diretamente a definição, como veremos mais à frente no Exemplo 2.38.

Exemplo 2.36. Consideremos a função $f(x,y) = x^4 + y^4 - 4xy$.

1. Determinar os pontos críticos de f.

$$\begin{cases} f_x = 0 \\ f_y = 0 \end{cases} \Leftrightarrow \begin{cases} 4x^3 - 4y = 0 \\ 4y^3 - 4x = 0 \end{cases} \Leftrightarrow \begin{cases} y = x^3 \\ x = y^3 \end{cases} \Leftrightarrow \begin{cases} -- \\ x^9 = x \end{cases} \Leftrightarrow$$

$$\Leftrightarrow (x,y) = (0,0) \vee (x,y) = (-1,-1) \vee (x,y) = (1,1)$$

2. Verificar se os pontos críticos são extremos locais.

   $D(x,y) = f_{xx}(x,y)\, f_{yy}(x,y) - \left(f_{xy}(x,y)\right)^2 = (12x^2)(12y^2) - (-4)^2 = 144x^2y^2 - 16 = 16(9x^2y^2 - 1)$.

   Como $D(0,0) = -16 < 0$, concluímos que $(0,0)$ é um ponto sela.

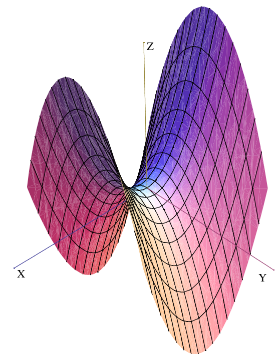
   Como $D(1,1) = D(-1,-1) = 16 \times 8 > 0$ e $f_{xx}(1,1) = f_{xx}(-1,-1) = 12 > 0$, concluímos que f tem dois mínimos locais $f(1,1) = -2$ e $f(-1,-1) = -2$.

Este exemplo mostra que é possível uma função de duas variáveis ter dois mínimos locais, e nenhum máximo local.

Neste caso é possível provar que -2 é também o mínimo absoluto. Para tal reescreve-se a função $f(x) = \left(x^2 - 1\right)^2 + \left(y^2 - 1\right)^2 + 2\left(x - y\right)^2 - 2 \geq -2$, provando assim que -2 é um minorante da função.

Exemplo 2.37. Ponto Sela

A função $f(x, y) = x^2 - y^2$ não tem nenhum extremo local, mas tem um ponto sela $(0, 0)$. (Verificar!)
O gráfico desta função deu origem à utilização do nome *ponto sela* para um ponto crítico não extremante.



Exemplo 2.38. Consideremos a função $f(x, y) = x^2 - 2xy^2 + y^4 - y^5$. O único ponto crítico desta função é $P = (0, 0)$ com $D(0, 0) = 0$. Portanto não podemos usar o Teorema 2.35 para verificar se P é extremante. Para isso vamos usar a definição. Sejam $P_1 = (a^2, a)$ e $P_2 = (a^2, -a)$, com $a > 0$, dois pontos infinitamente próximos de P, é fácil verificar que $f(P_1) = -a^5 < 0 = f(0, 0) < a^5 = f(P_2)$. Ou seja P é um ponto sela. (A escolha dos pontos P_1 e P_2 não foi aleatória, ela surge naturalmente dos cálculos efetuados para determinar os pontos críticos. Devem verificar.)

3 Cálculo Integral

O Cálculo Integral, de modo genérico, é área da Matemática onde se estudam as Anti-Derivadas ou Primitivas e a utilização que se faz delas. Os integrais podem estar definidos para uma ou mais variáveis, ainda que nesta disciplina apenas abordaremos o caso de uma variável.

3.1 Primitivas, Anti-Derivadas ou Integral Indefinido

Na primeira secção deste capítulo começamos por ver a definição de primitiva, ou anti-derivada. O nome anti-derivada indica-nos claramente o que estamos a calcular, embora a primitiva seja de utilização mais comum. De seguida estudaremos algumas das principais técnicas de primitivação

Definição 3.1. Uma função F é uma *primitiva* de f se $F' = f$.

A primitiva de uma função não é única. A mesma função pode ter várias primitivas. Vamos ver alguns exemplos desse facto.

Exemplos 3.2.

1. $F_1(x) = x$ e $F_2(x) = x+3$ são duas primitivas da função $f(x) = 1$. $(F_1'(x) = F_2'(x) = 1)$

2. $G_1(x) = \sin^2 x$ e $G_2(x) = -\cos^2 x$ são duas primitivas da função $g(x) = 2\sin x \cos x$. $(G_1'(x) = G_2'(x) = 2\sin x \cos x)$

   As funções G_1 e G_2 estão relacionadas da seguinte forma $G_1(x) = G_2(x) + 1$, ou seja $\sin^2 x = -\cos^2 x + 1$. Este é um resultado geral, como veremos a seguir.

Teorema 3.3.

1. *Se F é uma primitiva de f, então $F + C$, com $C \in \mathbb{R}$, é uma primitiva de f.*

2. *Se F_1 e F_2 são duas primitivas da mesma função, num determinado intervalo, então existe $C \in \mathbb{R}$ tal que $F_1 = F_2 + C$.*

Notação: O conjunto de todas as primitivas de uma função representa-se por $\displaystyle\int f(x)\,dx$:

$$\int f(x)\,dx = F(x) + C, \text{ onde } F' = f \text{ e } C \in \mathbb{R}.$$

- F é uma primitiva de f.
- $\int f(x)\,dx$ é a família de (todas as) primitivas de f, ou *integral indefinido*.

Os integrais herdam algumas propriedades das derivadas. Os resultados do próximo teorema são imediatos a partir dos resultados análogos para as derivadas.

Teorema 3.4. *Sejam $\alpha \in \mathbb{R}$, f e g funções primitiváveis.*

1. $\int \alpha f(x)\,dx = \alpha \int f(x)\,dx$
2. $\int f(x) + g(x)\,dx = \int f(x)\,dx + \int g(x)\,dx$

Ao contrário das derivadas, não existe uma fórmula para calcular primitivas. A primitivação é simplesmente definida como o "inverso" da derivação. Assim, se invertermos uma tabela de derivadas obtemos uma *tabela de primitivas*. As primitivas que se obtêm através da inversão das derivadas conhecidas designam-se por *primitivas imediatas*.

A seguir vamos ver exemplos de primitivas imediatas, fórmulas mais gerais que aparecem na tabela de primitivas e de aplicação do Teorema 3.4.

Exemplos 3.5.

1. $\int 2x\,dx = x^2 + C,\ C \in \mathbb{R}$
2. $\int x^n\,dx = \dfrac{x^{n+1}}{n+1} + C,\ n \neq -1$
3. $\int f'(x)\cos\left(f(x)\right)\,dx = \sin\left(f(x)\right) + C$
4. $\int x^3 - 2x^2 + 7\,dx = \int x^3\,dx - 2\int x^2\,dx + \int 7\,dx = \dfrac{x^4}{4} - 2\dfrac{x^3}{3} + 7x + C$
5. $\int \dfrac{10}{x(10-x)}\,dx = \int \dfrac{1}{x} + \dfrac{1}{10-x}\,dx = \int \dfrac{1}{x} - \dfrac{-1}{10-x}\,dx = \ln x - ln(10-x) + C$
   $= \ln\dfrac{x}{10-x} + C,\ x \in]0,10[$
   Esta primitiva foi resolvida apenas no intervalo $]0,10[$. De modo idêntico podemos calcular a primitiva no intervalo $]-\infty,0[$ ou no intervalo $]10,+\infty[$.

Primitivação por Partes

A regra da *primitivação por partes* é deduzida a partir da regra da derivada do produto. Esta regra é o mais parecido que temos de uma regra de primitivação do produto. Na bibliografia, por vezes, ela surge descrita usando a notação dos diferenciais.

Sejam g e h duas funções deriváveis num determinado intervalo. A regra da derivada do produto diz-nos que

$$(h\,g)' = h'\,g + h\,g'\,.$$

Primitivando ambos os membros da igualdade, e isolando uma das parcelas do segundo membro, obtemos a igualdade pretendida.

$$\int (h\,g)'\,(x)\,dx = \int h'(x)\,g(x)\,dx + \int h(x)\,g'(x)\,dx \Leftrightarrow$$

$$\int h'(x)\,g(x)\,dx = \int (h\,g)'\,(x)\,dx - \int h(x)\,g'(x)\,dx\,.$$

Como a primitiva da derivada é, a menos de constante, a própria função.

$$\int h'(x)\,g(x)\,dx = h(x)\,g(x) - \int h(x)\,g'(x)\,dx$$

Finalmente, fazendo $h' = f$ e $h = F$, obtemos a fórmula da *primitivação por partes*.

$$\int f(x)\,g(x)\,dx = F(x)\,g(x) - \int F(x)\,g'(x)\,dx,$$

onde F é uma primitiva de f.

Nos próximos exemplos identificaremos em cada caso qual a função por onde devemos começar a primitivar (f) e qual a função que começamos por derivar (g).

Exemplos 3.6.

1. $\displaystyle\int \underbrace{x}_{g}\,\underbrace{e^x}_{f}\,dx = x\,\underbrace{e^x}_{F} - \int e^x \times \underbrace{1}_{g'}\,dx = x\,e^x - e^x + C$

2. $\displaystyle\int \ln x\,dx = \int \underbrace{\ln x}_{g} \times \underbrace{1}_{f} = (\ln x) \times x - \int \frac{1}{x}\,x = x\,\ln x - x + C$

3. $\displaystyle\underbrace{x}_{f}\,\underbrace{\arctan x}_{g}\,dx = \frac{x^2}{2}\,\arctan x - \int \frac{x^2}{2}\,\frac{1}{x^2+1}dx = \frac{x^2}{2}\,\arctan x - \frac{1}{2}\int \frac{x^2+1-1}{x^2+1}dx$

$$= \frac{x^2}{2}\,\arctan x - \frac{1}{2}\int 1 - \frac{1}{x^2+1}dx = \frac{x^2}{2}\,\arctan x - \frac{1}{2}(x - \arctan x) + C$$

$$= \frac{1}{2}\left(\arctan x(x^2+1) - x\right) + C$$

Primitivas por Substituição

Tal como a regra da primitivação por partes, a regra das *primitivas por substituição* é deduzida a partir de uma regra conhecida para as derivadas. Neste caso a regra da derivada da composta ou regra da cadeia (para funções de uma variável). Ela também pode ser vista como uma mudança de variável. Comecemos por um exemplo que já conhecemos, e que nos indica que uma grande quantidade de regras de primitivação imediata faz uso de uma mudança de variável.

É fácil verificar que $\displaystyle\int \cos(t^2)\,2t\,dt = \sin(t^2) + C$. Por outro lado se fizermos $x = t^2$, e usando que $dx = \frac{dx}{dt}\,dt = 2t\,dt$, chegamos à igualdade

$$\int \cos(t^2)\,\underbrace{2t\,dt}_{dx} = \int \cos(x)\,dx = \sin x + C = \sin(t^2) + C\,.$$

Claro que vamos continuar a calcular esta primitiva diretamente, sem escrever a substituição explicitamente. Este exemplo serve apenas para ilustrar como funciona a substituição para o cálculo de primitivas.

Sejam g e h duas funções deriváveis. A regra da derivada da composta diz-nos que

$$(h \circ g)'\,(t) = h'(g(t))\,g'(t)\,.$$

Mais uma vez, primitivando ambos os membros da igualdade obtemos a fórmula pretendida.

$$\int h'(g(t))\,g'(t)\,dt = h(g(t)) + C$$

Fazendo $h' = f$ e $h = F$, temos que

$$\int f(g(t))\,g'(t)\,dt = F(g(t)) + C, \quad \text{onde } F \text{ é uma primitiva de } f.$$

Tal como vimos no exemplo anterior se $x = g(t)$, então $dx = g'(t)\,dt$ e a mudança de variável, ou substituição $x = g(t)$ implica que,

$$\int f(g(t))\,g'(t)\,dt = \int f(x)\,dx = F(x) + C = F(g(t)) + C.$$

A notação dos diferenciais $x = g(t) \Rightarrow dx = g'(t)dt = \dfrac{dx}{dt}dt$, torna mais fácil recordar a regra da substituição.

A fórmula tal como a deduzimos, e vimos no exemplo, é lida da *"esquerda para a direita"*. Ela pode, no entanto, ser usada no sentido inverso. Nesse caso temos de procurar qual a substituição mais adequada em cada caso. A tabela de primitivas dá-nos algumas sugestões de mudanças de variável.

Teorema 3.7 (Primitiva por substituição)**.** *Seja f contínua num intervalo I. Se $x = g(t)$ é uma função diferenciável em I tal que $g'(t) \neq 0$ no interior do intervalo, então g é invertível, $t = g^{-1}(x)$, e*

$$\int f(x)\,dx = \left(\int f(g(t))\,g'(t)\,dt\right)_{t=g^{-1}(x)}.$$

Veremos de seguida mais um exemplo de primitiva calculada por substituição. Para melhor compreensão usaremos a notação dos diferenciais. É preciso notar que o resultado final deverá ser uma função na mesma variável que a função dada inicialmente.

Exemplo 3.8.

$$\int \sqrt{1-x^2}\,dx$$

Faz-se a substituição $x = \sin t \Rightarrow dx = (\sin t)'dt = \cos t\,dt$.

Para $x \in [-1, 1]$ (o domínio da função) e $t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, x é uma função bijetiva.

$$\int \sqrt{1-x^2}\,dx = \int \sqrt{1-\sin^2 t}\ \cos t\,dt = \int \sqrt{cos^2 t}\ \cos t\,dt$$

$$= \int |\cos t|\ \cos t\,dt = \int \cos^2 t\,dt \qquad \left(\cos t \geq 0 \text{ para } t \in \left[-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right]\right).$$

$$(*) = \int \frac{1}{2}\,(1+\cos 2t)\ dt = \frac{1}{2}\left(t + \frac{\sin 2t}{2}\right) + C$$

$$= \frac{1}{2}(t + \frac{2\sin t\cos t}{2}) + C = \frac{1}{2}\left(\arcsin x + x\sqrt{1-x^2}\right) + C$$

Notas: $\cos t = \sqrt{1-x^2}$, uma vez que $\cos^2 t + \sin^2 t = 1$, $\sin t = x$ e $\cos t \geq 0$.

(*) Na tabela de primitivas existe uma secção sobre como primitivar potências de funções trigonométricas. Neste caso usámos a igualdade $\cos^2 t = \dfrac{1}{2}\,(1+\cos 2t)$

Pode ser necessário usar uma combinação destas técnicas de primitivação e/ou usar repetidamente a mesma técnica para calcular a primitiva desejada.

Frequentemente é mais fácil calcular uma derivada do que calcular a respetiva primitiva. Por isso, uma maneira eficiente de verificar que uma primitiva está bem calculada é calcular a derivada do resultado obtido.

Como vimos atrás é preciso ter ainda em atenção que a primitiva não é única, e que algumas funções podem ser escritas de várias maneiras.

3.2 Integral Definido

Antes de passarmos à definição formal de integral definido, e posteriormente à maneira como se calcula, começamos com um exemplo muito simples que já estudaram em Física e que serve para ilustrar a relação entre primitiva e integral definido, que veremos a seguir.

Problema

Uma mota vai dos 0 aos $180km/h = 50m/s$ em $10s$ com aceleração uniforme. Quantos m percorre nos últimos $4s$?

Resolver este problema consiste em saber qual é a velocidade média da mota no intervalo $[6, 10]$ e multiplicar pelo tempo. Como o movimento é uniforme a distância percorrida é:

$$dist = v_{\text{média}} \times (10 - 6) = \frac{50 + 30}{2} \times 4 = 160m.$$

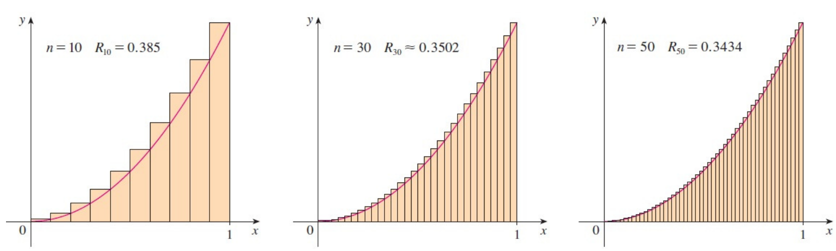
Geometricamente, calculámos a área dum trapézio com bases 50 e 30, e altura 4.

E se o movimento não for uniforme? Se a velocidade for modelada por uma qualquer função contínua $v(t)$? O problema contínua a consistir em calcular a área da região limitada pelo eixo dos XX (onde é medido o tempo) e pelo gráfico da função $v(t)$. Podemos calcular essa área usando aproximações sucessivas.

Definição 3.9. Seja f uma função contínua no intervalo fechado $[a, b]$ (e portanto limitada).

O *integral definido* de f no intervalo $[a, b]$ é:

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f\left(a + \frac{b-a}{n} i\right) \times \frac{b-a}{n}.$$



Iterações quando $n = 10, 30, 50$ para a função $f(x) = x^2$.

Voltando ao problema anterior, a distância percorrida pela mota é escrita em termos de integrais como $\int_6^{10} 5t\,dt$, ou em geral a distância percorrida pela mota entre os instantes $t = a$ e $t = b$ quando a velocidade instantânea em cada momento t é $v(t)$ vai ser calculada pelo integral:

$$\int_a^b v(t)\,dt\,.$$

Notas.

1. À função f chamamos *função integranda* e aos valores a e b *extremos de integração.*

2. Sejam $a < c < b$, $\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$.

3. A definição de integral definido pode ser generalizada para funções limitadas com um número finito de descontinuidades.

4. O *Valor Médio* de f no intervalo $[a, b]$ é calculado usando a fórmula

$$\mathsf{Média}(f) = \frac{1}{b-a} \int_a^b f(x)\, dx\,.$$

Proposição 3.10. *Se $f(x) \geq 0$, para $x \in [a, b]$, então $\int_a^b f(x)\, dx$ é o valor da área da região limitada pelas retas $x = a$, $x = b$, $y = 0$ e $y = f(x)$.*

Como vimos anteriormente $\displaystyle\int_6^{10} 5x\, dx = \frac{(50+30) \times (10-6)}{2} = 160$ é área dum trapézio com bases 50 e 30, e altura 4.

Teorema 3.11 (Teorema Fundamental do Cálculo Integral)**.**

1. Seja f contínua em $[a, b]$.

A função $F(x) = \int_a^x f(t)\, dt$ é uma primitiva de $f(x)$ no intervalo $[a, b]$. Ou seja, todas as funções contínuas são primitiváveis.

2. Se f é contínua em $[a, b]$ e F uma primitiva de f, então:

$$\int_a^b f(x)\, dx = F(b) - F(a)\,.$$

Apesar de todas as funções contínuas serem primitiváveis, nem todas são primitiváveis como uma soma finita de funções elementares.

O ponto *2.* do *Teorema Fundamental do Cálculo Integral* dá-nos uma maneira prática de calcular um integral definido. Voltando ao problema que nos tem servido de guia, ele diz que para calcular o deslocamento basta conhecer a posição final e a posição inicial.

Exemplos 3.12.

1. $\displaystyle\int_6^{10} 5x\, dx = F(10) - F(6) = \frac{5}{2}10^2 - \frac{5}{2}6^2 = 250 - 90 = 160,$

   onde $F(x) = \frac{5}{2}x^2$ é uma das primitivas de F.

   Em geral não vamos escrever a primitiva F separadamente. Nos próximos exemplos vamos introduzir a notação que vai evitar fazê-lo.

2. $\displaystyle\int_0^1 x^2\,dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1^3}{3} - \frac{0^3}{3} = \frac{1}{3}$

   Este é o exemplo da figura que acompanha a Definição 3.9 (definição de Integral Definido).

3. Consideremos a função definida por ramos $f(x) = \begin{cases} \cos x & \text{se } x > 0 \\ x & \text{se } x \leq 0 \end{cases}$.

   $\displaystyle\int_{-2}^{\pi/2} f(x)\,dx = \int_{-2}^{0} x\,dx + \int_0^{\pi/2} \cos x\,dx =$

   $\displaystyle\left[\frac{x^2}{2}\right]_{-2}^{0} + [\sin x]_0^{\pi/2} = (0-2) + \sin\frac{\pi}{2} - \sin 0 = -1.$

   A função deste exemplo tem uma descontinuidade no ponto 0, no entanto a integração em cada dos dois ramos da função pode ser feita usando o Teorema Fundamental do Cálculo Integral como habitualmente (ver Nota 2. atrás). Reparemos que neste caso existem os dois limites laterais em 0, ainda que sejam diferentes.

4. $\displaystyle\int_{-2}^{1} -\frac{1}{x^2} \neq \left[\frac{1}{x}\right]_{-2}^{1}$

   A função $f(x) = -\frac{1}{x^2}$ é ilimitada no intervalo $[-2, 1]$ (e não está definida em 0), o que torna este exemplo diferente do anterior. Esta função não é integrável no intervalo $[-2, 1]$ mas é integrável em qualquer intervalo que não contenha o 0.

De seguida vamos ver como é que a primitivação por partes e por substituição podem ser usadas para calcular integrais definidos. As regras que enunciamos a seguir são apenas uma reescrita adaptada a este novo contexto. Os integrais podem ser sempre calculados usando o caso geral.

Integração por Partes

Seja F uma primitiva de f.

$$\int_a^b f(x)\,g(x)\,dx = F(b)g(b) - F(a)g(a) - \int_a^b F(x)\,g'(x)\,dx$$

A vantagem de usar esta fórmula diretamente, em vez de calcular a primitiva de $f(x)\,g(x)$ primeiro e fazer as substituições depois, é principalmente porque nos permite simplificar a escrita.

Exemplo 3.13.

$$\int_0^2 xe^x\,dx = [e^x x]_0^2 - \int_0^2 e^x 1\,dx = (2e^2 - 0) - [e^x]_0^2 = 2e^2 - (e^2 - e^0) = e^2 + 1.$$

Integração por Substituição

Sejam f uma função contínua em $[a,b]$ e $g : [c,d] \to [a,b]$ uma função bijetiva e com derivada contínua, então

$$\int_a^b f(x)\,dx = \int_c^d f(g(t))\,|g'(t)|\,dt.$$

Usar esta fórmula diretamente tem a grande de vantagem de não ser preciso voltar à variável inicial x para calcular o valor do integral. (Comparar com o Teorema 3.7.)

Exemplos 3.14.

1. Neste exemplo vamos usar as mesmas função e substituição do que no Exemplo 3.8, e por isso vamos omitir alguns passos que já foram feitos.

   Usámos a substituição $x = \sin t$, e portanto

   $-1 \leq x \leq 1 \Leftrightarrow -1 \leq \sin t \leq 1 \Rightarrow -\dfrac{\pi}{2} \leq t \leq \dfrac{\pi}{2}.$

   Explicitamente a função mudança de variável é $g : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1,1]$ com $g(t) = \sin t$.

   Já tínhamos calculado que $dx = \sin t\,dt$.

   $$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2 t}|\cos t|\,dt = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t\,dt = \ldots =$$
   $$= \frac{1}{2}\left[t + \frac{\sin 2t}{2}\right]_{-\pi/2}^{\pi/2} = \frac{1}{2}\left(\frac{\pi}{2} + \frac{\sin \pi}{2} - \left(-\frac{\pi}{2}\right) - \frac{\sin(-\pi)}{2}\right) = \frac{\pi}{2}$$

   Nota: O que acabámos de calcular foi a área da metade superior do círculo de raio igual a 1 e centro na origem.

2. Calcular $\displaystyle\int_1^{\sqrt[3]{2}} \frac{\sqrt{x^3-1}}{x}\,dx$, fazendo a mudança de variável $u^2 = x^3 - 1$.

   Faz-se a substituição $u^2 = x^3 - 1 \Leftrightarrow x = \sqrt[3]{u^2+1} = \left(u^2+1\right)^{1/3}$.

   $$\frac{dx}{du} = \frac{1}{3}(2u)\left(u^2+1\right)^{-2/3} \Rightarrow dx = \frac{2}{3}\frac{u}{\sqrt[3]{(u^2+1)^2}}du = \frac{2u}{3x^2}du$$

   $$1 \leq x \leq \sqrt[3]{2} \Leftrightarrow 1 \leq x^3 \leq 2 \Leftrightarrow 0 \leq x^3 - 1 \leq 1 \Rightarrow 0 \leq u \leq 1$$

   Temos assim que

   $$\int_1^{\sqrt[3]{2}} \frac{\sqrt{x^3-1}}{x}\,dx = \int_0^1 \frac{u}{x}\frac{2u}{3x^2}\,du = \int_0^1 \frac{2}{3}\frac{u^2}{x^3}\,du = \int_0^1 \frac{2}{3}\frac{u^2}{u^2+1}\,du =$$
   $$\frac{2}{3}\int_0^1 1 - \frac{1}{u^2+1}\,du = \frac{2}{3}\left[u - \arctan u\right]_0^1 = \frac{2}{3}(1 - 0 - \arctan 1 + \arctan 0) = \frac{2}{3}\left(1 - \frac{\pi}{4}\right).$$

Quando fazemos uma substituição temos que mudar a função integranda, mas também os extremos de integração.

3.3 Aplicações do Cálculo Integral

As aplicações do cálculo integral definido resultam todas, direta ou indiretamente, da fórmula que vimos anteriormente para o cálculo de média de uma função:

$$\mathsf{Média}(f) = \frac{1}{b-a}\int_a^b f(x)\,dx\,.$$

De modo análogo se a função a integrar representar uma taxa de variação, ou seja a função é uma derivada g', então $\displaystyle\int_a^b g'(x)\,dx$ é o acréscimo total da função g no intervalo $[a, b]$. Este é a situação que descrevemos no problema inicial da Secção 3.2.

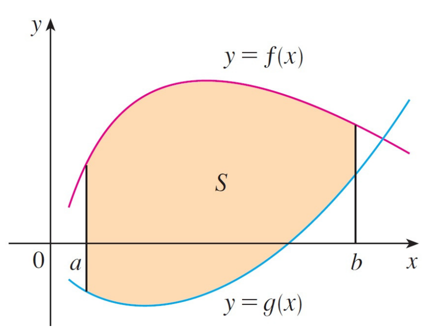
De seguida veremos algumas aplicações geométricas do cálculo integral.

Áreas

O integral definido de funções não negativas foi introduzido como o cálculo de uma área. O resultado que enunciamos a seguir generaliza esse facto.

Proposição 3.15. *Sejam f e g funções contínuas em $[a, b]$, com $g(x) \leq f(x)$ para $x \in [a, b]$. A área da região S limitada pelos gráficos de f e g e pelas retas verticais $x = a$ e $x = b$ é*

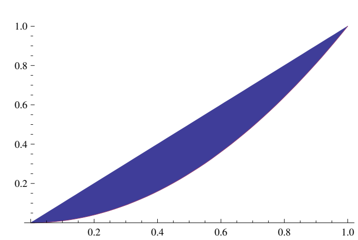
$$\mathsf{Área}(S) = \int_a^b f(x) - g(x)\,dx \ \ [=(\text{altura média})\ \times(b-a)]\,.$$



Exemplos 3.16.

1. Calcular a área da região $A = \{(x, y)\ :\ x^2 \leq y \leq x\}$.

   Neste exemplo os valores dos extremos de integração não são dados explicitamente, por isso temos que os deduzir. As curvas $y = x$ e $y = x^2$ intersectam-se quando $x = 0$ ou $x = 1$. Portanto os extremos de integração são 0 e 1. A figura ajuda a compreender melhor o que fizemos.



$$\text{Área}(A) = \int_0^1 x - x^2\,dx = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{6}$$

2. No Exemplo 3.14(1) calculámos a área do semi-círculo:

$$\{(x,y)\,:\,0 \le y \le \sqrt{1-x^2}\} = \{(x,y)\,:\,x^2+y^2 \le 1 \wedge y \ge 0\}\,.$$

Sólidos de Revolução

Os integrais definidos podem também ser usados para calcular o volume de certo tipo de sólidos, os sólidos de revolução.

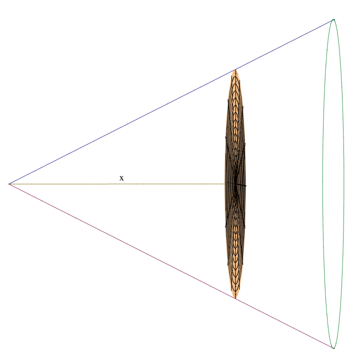
Definição 3.17. Um *sólido de revolução* é um sólido que resulta da rotação de uma figura plana em torno de um eixo, chamado *eixo de revolução*.

O cilindro, o cone e a esfera são exemplos de sólidos de revolução que resultam da rotação de um retângulo, de um triângulo e de um semi-círculo, respectivamente.

Vamos usar o exemplo do cone (do qual já sabemos calcular o volume) para ilustrar como podemos, em geral, calcular o volume de um sólido de resolução se conheceremos a região e o eixo de revolução.

Consideremos um cone de raio r e altura h. O volume do cone é igual ao produto da altura pelo valor médio da área de cada secção circular do cone. Seja x a distância do centro desse círculo ao vértice do cone (representado na figura). Então o raio do círculo é igual a $\frac{rx}{h}$ e portanto a área desse círculo é $A(x) = \pi\left(\frac{rx}{h}\right)^2$. O volume do cone é então igual ao produto de h pelo valor médio de $A(x)$ e portanto:

$$\text{Volume}(cone) = \text{média de } A(x) \times h = \int_0^h A(x)\,dx = \int_0^h \pi\left(\frac{rx}{h}\right)^2\,dx = \frac{1}{3}\pi r^2 h\,.$$



$A(x)$ é a área da região sombreada.

Rotação em torno do eixo dos XX

No caso geral podemos considerar que o raio de cada círculo é dado por uma função $f(x)$. Neste caso a função que nos dá a área é $A(x) = \pi f(x)^2$.

Proposição 3.18. *Seja f contínua em $[a, b]$. O volume do sólido de revolução gerado pela rotação em torno do eixo dos XX da figura limitada pelo gráfico de f, pelo eixo dos XX e pelas retas $x = a$ e $x = b$ é*

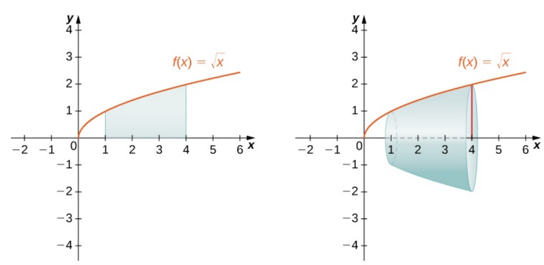
$$\text{volume} = \pi \int_a^b f(x)^2\, dx\,.$$

Mais geralmente, se fizermos a rotação em torno de uma reta horizontal $y = k$, então o volume é calculado pela expressão

$$\text{volume} = \pi \int_a^b (f(x) - k)^2\, dx\,.$$

Exemplo 3.19. Calcular o volume do sólido de revolução gerado pela rotação em torno do eixo dos XX da região $\{(x, y) : 0 \leq y \leq \sqrt{x},\ 1 \leq x \leq 4\}$.

$$\text{volume} = \pi \int_1^4 \sqrt{x}^2\, dx = \pi \int_1^4 x\, dx = \pi \left[\frac{x^2}{2}\right]_1^4 = \frac{15\pi}{2}$$



Sólido de revolução gerado pela rotação da região limitada por $f(x) = \sqrt{x}$.

Rotação em Torno do Eixo dos YY

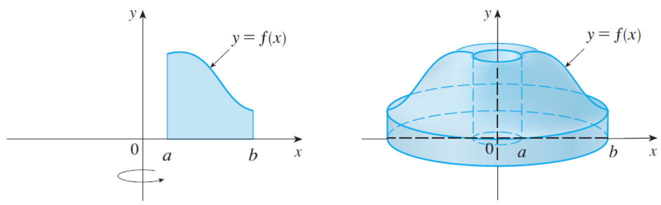
Se a rotação da figura for feita em torno do eixo dos YY, ou de qualquer outra reta vertical, o cálculo do volume é feito usando uma outra fórmula. É preciso ter em atenção que trocando o papel das variáveis x e y podemos sempre transformar esta "situção" na anterior, e usar a respetiva fórmula.

Proposição 3.20. *Sejam $0 \leq a \leq b$, f contínua em $[a, b]$ tal que $f(x) \geq 0$. O volume do sólido de revolução gerado pela rotação em torno do eixo dos YY da figura limitada pelo gráfico de f, pelo eixo dos XX e pelas retas $x = a$ e $x = b$ é*

$$\text{volume} = 2\pi \int_a^b x\, f(x)\, dx\,.$$

Neste caso $f(x)$ representa a altura de um cilindro de raio x. A área da superfície desse cilindro é $S(x) = 2\pi x\, f(x)$, e o volume do sólido pode ser visto como:

$$(\text{média de } S(x)) \times (b - a).$$



Exemplo 3.21. Consideremos a região do plano $A = \{(x, y) \,|\, 0 \leq y \leq x^2,\ x \leq 1\}$.

1. Calcular o volume do sólido de revolução gerado pela rotação em torno do eixo dos XX da região A.

$$\text{Volume}_X = \pi \int_0^1 \left(x^2\right)^2\, dx = \pi \left[\frac{x^5}{5}\right]_0^1 = \frac{\pi}{5}$$

2. Calcular o volume do sólido de revolução gerado pela rotação em torno do eixo dos YY da região A.

$$\text{Volume}_Y = 2\pi \int_0^1 x\, x^2\, dx = 2\pi \left[\frac{x^4}{4}\right]_0^1 = \frac{\pi}{2}$$

Comprimento de Curvas

Uma outra aplicação geométrica do cálculo integral é o cálculo do comprimento do gráfico de uma função derivável num intervalo limitado.

Proposição 3.22. *Seja f uma função com derivada contínua em $[a, b]$. O comprimento do gráfico de f entre os pontos $(a, f(a))$ e $(b, f(b))$ é*

$$C_a^b(f) = \int_a^b \sqrt{1 + f'(x)^2}\, dx\,.$$

Exemplo 3.23. Calcular o perímetro de uma circunferência de raio r.

$$x^2 + y^2 = r^2 \Leftrightarrow y = \pm\sqrt{r^2 - x^2}$$

Para calcular o perímetro de meia circunferência, faz-se $f(x) = \sqrt{r^2 - x^2}$ e $f'(x) = \dfrac{-x}{\sqrt{r^2 - x^2}}$ e portanto o perímetro da circunferência é:

$$\text{Perímetro} = 2\int_{-r}^{r} \sqrt{1 + \frac{x^2}{r^2 - x^2}}\, dx = 2r \int_{-r}^{r} \frac{\frac{1}{r}}{\sqrt{1 - \frac{x^2}{r^2}}}\, dx =$$

$$2r \left[\arcsin\left(\frac{x}{r}\right)\right]_{-r}^{r} = 2r\left(\arcsin 1 - \arcsin(-1)\right) = 2\pi r\ .$$

3.4 Integrais Impróprios

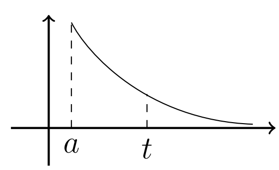
Os integrais impróprios são uma generalização dos integrais definidos, e resultam de considerarmos as hipóteses de o intervalo de integração ou a função integranda serem ilimitados (ou eventualmente ambos). Como exemplo básico podemos pensar que é possível uma região ilimitada ter uma área finita. Apesar de ser um pouco contra-intuitivo, este caso não é muito diferente do que acontece quando somamos todos os termos de uma progressão geométrica. Claro que já sabemos que nem todas as somas com um número infinito de parcelas têm um valor finito, e também nem todos os *integrais impróprios* têm um valor finito. Aos integrais impróprios que realmente representam um valor finito chamaremos *convergentes*. Tal como referido anteriormente, os integrais impróprios podem ser de dois tipos.

Intervalos Ilimitados

Seja f uma função contínua e não negativa em $[a, +\infty)$ tal que $\lim\limits_{x\to+\infty} f(x) = 0$.

Se $t > a$, então a área $A(t)$ da região limitada pelo gráfico de $f(t)$, o eixo dos XX e as retas verticais $x = a$ e $x = t$ é dada por:

$$A(t) = \int_a^t f(x)dx$$



Se $\lim\limits_{t\to+\infty} A(t)$ existe (e é finito), então este limite pode ser interpretado como a área da região limitada pelo gráfico de f, pelo eixo dos XX e pela reta $x = a$.

Definições 3.24.

(a) Se $\displaystyle\int_a^t f(x)\, dx$ existe para $t \geq a$, então

$$\int_a^{+\infty} f(x)\, dx = \lim_{t\to+\infty} \int_a^t f(x)\, dx, \text{ se este limite existir.}$$

(b) Se $\displaystyle\int_t^b f(x)\,dx$ existe para $t \leq b$, então

$$\int_{-\infty}^{b} f(x)\,dx = \lim_{t\to-\infty} \int_t^b f(x)\,dx, \text{ se este limite existir.}$$

(c)

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{+\infty} f(x)\,dx, \text{ para qualquer } c \in \mathbb{R}$$

Exemplos 3.25.

1. $\displaystyle\int_1^{+\infty} \frac{1}{x}\,dx = \lim_{t\to+\infty} \int_1^t \frac{1}{x}\,dx = \lim_{t\to+\infty} [\ln x]_1^t = \lim_{t\to+\infty} \ln t - \ln 1 = +\infty$

2. $\displaystyle\int_{-\infty}^{0} e^x\,dx = \lim_{t\to-\infty} \int_t^0 e^x\,dx = [e^x]_t^0 = \lim_{t\to-\infty} e^0 - e^t = 1$

3. $\displaystyle\int_{-\infty}^{+\infty} x\,dx = \int_{-\infty}^{0} x\,dx + \int_0^{+\infty} x\,dx = \infty - \infty$

   Este integral não está definido. Este tipo de indeterminações não pode ser levantada.

Funções Ilimitadas

Definição 3.26.

(a) Seja f contínua no intervalo $[a, b[$. Se $\displaystyle\lim_{x\to b^-} f(x)$ for infinito, então

$$\int_a^b f(x)\,dx = \lim_{t\to b^-} \int_a^t f(x)\,dx, \text{ se este limite existir.}$$

(b) Seja f contínua no intervalo $]a, b]$. Se $\displaystyle\lim_{x\to a^+} f(x)$ for infinito, então

$$\int_a^b f(x)\,dx = \lim_{t\to a^+} \int_t^b f(x)\,dx, \text{ se este limite existir.}$$

Nota. $\displaystyle\int_0^1 x \sin\left(\frac{1}{x}\right) dx$ não é um integral impróprio, pois $\displaystyle\lim_{x\to 0^+} x \sin\left(\frac{1}{x}\right) = 0$.

Exemplos 3.27.

1. $\displaystyle\int_0^1 \frac{1}{\sqrt[3]{x}}\,dx = \lim_{t\to 0^+} \int_t^1 x^{-1/3}\,dx = \lim_{t\to 0^+} \left[\frac{x^{2/3}}{2/3}\right]_t^1 = \lim_{t\to 0^+} \frac{1}{2/3} - \frac{t^{2/3}}{2/3} = \frac{3}{2}$

2. $\displaystyle\int_0^1 \frac{1}{x-1}\,dx = \lim_{t\to 1^-}\int_0^t \frac{1}{x-1}\,dx = \lim_{t\to 1^-}\int_0^t \frac{-1}{1-x}\,dx = \lim_{t\to 1^-}\left[\ln(1-x)\right]_0^t =$

   $\displaystyle\lim_{t\to 1^-} \ln(1-t) - \ln 1 = -\infty$

Definição 3.28. Se o limite de um integral impróprio existe, e é finto, o integral diz-se *convergente*. Se um integral impróprio não é convergente, diz-se *divergente*.

Critério de Comparação

Como vimos anteriormente nem sempre é possível calcular com exatidão o valor dum integral, mas por vezes é possível calcular um valor aproximado ou um intervalo de possíveis resultados. Para os integrais impróprios vamos usar um critério simples de comparação que nos vais ajudar a verificar se um integral impróprio é convergente ou divergente.

Proposição 3.29 (Critério de Comparação)**.** *Sejam f e g contínuas no intervalo $]a,b[$, com $0 \leq f(x) \leq g(x)$ para x nesse intervalo.*

1. *Se $\displaystyle\int_a^b g(x)\,dx$ é convergente, então $\displaystyle\int_a^b f(x)\,dx$ também é convergente.*
2. *Se $\displaystyle\int_a^b f(x)\,dx$ é divergente, então $\displaystyle\int_a^b g(x)\,dx$ também é divergente.*

Os valores a e b podem ser $\pm\infty$.

A proposição anterior é uma consequência do Teorema do Enquadramento, que é válido para todo o tipo de integrais e que vamos enunciar de seguida.

Teorema 3.30 (Teorema do Enquadramento)**.** *Sejam f, g e h funções contínuas no intervalo $]a,b[$, com $f(x) \leq g(x) \leq h(x)$ para x nesse intervalo.*

Se $\displaystyle\int_a^b f(x)\,dx$ e $\displaystyle\int_a^b h(x)\,dx$ são convergentes, então $\displaystyle\int_a^b g(x)\,dx$ também é convergente,

e

$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx \leq \int_a^b h(x)\,dx.$$

Exemplo 3.31. Um tipo de funções importante em Estatística são as funções de distribuição normal ou Gaussiana. A menos de parâmetros constantes, elas resultam de modificações da função $f(x) = e^{-x^2}$. Apesar da função $f(x)$ não ser integrável como soma de funções elementares, é possível determinar a natureza do integral $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2}\,dx$.

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = 2\int_{0}^{+\infty} e^{-x^2}\,dx \qquad (e^{-x^2} \text{ é uma função par.})$$

$$\int_{0}^{+\infty} e^{-x^2} = \underbrace{\int_{0}^{1} e^{-x^2}\,dx}_{\text{finito}} + \int_{1}^{+\infty} e^{-x^2}\,dx$$

Para $x \geq 1$,

$$0 \leq e^{-x^2} \leq e^{-x} \Rightarrow 0 \leq \int_{1}^{+\infty} e^{-x^2}\,dx \leq \int_{1}^{+\infty} e^{-x}\,dx = \frac{1}{e}.$$

Conclui-se então que o integral $\displaystyle\int_{-\infty}^{+\infty} e^{-x^2}\,dx$ é convergente.

4 Equações Diferenciais

Uma *equação diferencial* é uma equação que envolve uma ou mais derivadas de uma função incógnita $y = f(x)$ $(y', y'', ...)$. A equação pode conter a própria função y, outras funções conhecidas de x e constantes. A solução de uma equação diferencial é uma função, ou mais geralmente uma família de funções, tal como a solução de uma equação algébrica é um valor real ou um conjunto de valores. Tal como acontece com as equações algébricas, existem equações diferenciais que não têm solução.

Não existe uma regra geral para resolver equações diferenciais, apenas algumas regras parciais para resolver certos tipos de equações. Neste curso vamos aprender a resolver dois tipos de equações diferenciais: *equações de variáveis separáveis* e *equações lineares de primeira ordem*, mas antes disso vamos introduzir alguns conceitos sobre equações diferenciais um pouco mais globais.

De referir que nesta área da Matemática existe muita investigação feita atualmente, tanto teórica como aplicada.

- A variável da função incógnita, normalmente x, é a *variável independente* e a função incógnita, normalmente y, é a *variável dependente*.
- A *ordem de uma equação* é a ordem da derivada de maior grau que aparece na equação.

Exemplos de Equações Diferenciais

1. $y' = x$ ordem 1
2. $y'' + y'y = x^2 + 1$ ordem 2
3. $y'''x + \sin x = y'$ ordem 3

Definições 4.1.

1. Ao conjunto de todas as soluções de uma equação diferencial chama-se *Solução Geral* ou *Integral Geral*.
2. A uma solução da equação diferencial chama-se *Solução Particular*.

Nota. As equações diferenciais estão definidas num intervalo de $\mathbb{R}$, e cada uma das suas soluções deve ser considerada apenas num intervalo determinado.

Exemplo 4.2. A equação $y' = x$ tem como solução geral $y(x) = \dfrac{x^2}{2} + C$, onde C é uma constante pertencente a $\mathbb{R}$.

As funções $y_1 = \frac{x^2}{2}$ ou $y_2 = \frac{x^2}{2} + 5$ são soluções particulares da equação diferencial $y' = x$.

Exercício

(a) Mostre que $y = C_1 e^{5x} + C_2 e^{-3x}$, $C_1, C_2 \in \mathbb{R}$ é a solução geral da equação diferencial $y'' - 2y' - 15y = 0$.

Começamos por calcular as duas primeiras derivadas da função dada, e substituímos na equação para verificar se a igualdade é válida.

$y = C_1 e^{5x} + C_2 e^{-3x}$

$y' = 5C_1 e^{5x} - 3C_2 e^{-3x}$

$y'' = 25C_1 e^{5x} + 9C_2 e^{-3x}$

$$y'' - 2y' - 15y = 25C_1 e^{5x} + 9C_2 e^{-3x} - 2\left(5C_1 e^{5x} - 3C_2 e^{-3x}\right) - 15\left(C_1 e^{5x} + C_2 e^{-3x}\right) =$$

$$(25C_1 - 10C_1 - 15C_1)e^{5x} + (9C_2 + 6C_2 - 15C_2)e^{-3x} = 0$$

(b) Determinar a solução particular que verifica as condições $y(0) = 0$ e $y'(0) = 1$.

Vamos usar as duas *condições iniciais* que nos são dadas para determinar o valor das constantes C_1 e C_2.

$$\begin{cases} y(0) = 0 \\ y'(0) = 1 \end{cases} \Leftrightarrow \begin{cases} C_1 + C_2 = 0 \\ 5C_1 - 3C_2 = 1 \end{cases} \Leftrightarrow \begin{cases} C_1 = \frac{1}{8} \\ C_2 = -\frac{1}{8} \end{cases}$$

$y = \dfrac{1}{8}e^{5x} - \dfrac{1}{8}e^{-3x}$ é a solução particular pretendida.

4.1 Equações Diferenciais de Primeira Ordem

Nesta secção vamos aprender a resolver dois tipos específicos de equações diferenciais, mas antes disso vamos enunciar um resultado que nos garante que todas as equações de 1ª ordem têm solução.

Teorema 4.3 (Teorema de Existência e Unicidade de soluções)**.** *Sejam $\varphi(x, y)$ e $\frac{\partial \varphi}{\partial y}(x, y)$ funções contínuas e (a, b) um ponto do domínio de φ.*

Então existe uma única solução da equação $y' = \varphi(x, y)$ que satisfaz a condição $y(a) = b$.

Todas as equações de primeira ordem se podem escrever na forma $y' = \varphi(x, y)$, para alguma função φ. Portanto, este teorema diz-nos que qualquer equação diferencial de primeira ordem tem uma solução única que verifica a condição $y(a) = b$, quaisquer que sejam os valores a e b (desde que a condição esteja bem definida). Tal como já foi referido anteriormente, essa solução pode não estar definida em $\mathbb{R}$ mas apenas num intervalo que contenha a. De notar que este resultado apenas indica que tal solução existe, ainda que nem sempre exista uma maneira "fácil" de a calcular.

Equações de Variáveis Separáveis

Uma equação diferencial é de *variáveis separadas* se puder ser escrita de forma a que os termos que envolvem a variável independente fiquem num lado da equação, e os termos que envolvem a variável dependente fiquem do outro lado. Ou seja, pode escrita na forma:

$$M(x) = N(y)y',$$

onde M é função de x e N é função de y.

Uma equação pode não ser de variáveis separadas, mas ser equivalente a uma equação de variáveis separadas. A uma dessas equações chamamos equação diferencial de *variáveis separáveis*. A seguinte equação é um desses exemplos.

$$(4y + yx^2)y' = (2x + xy^2) \Leftrightarrow \frac{y}{2 + y^2}y' = \frac{x}{4 + x^2}.$$

Solução das equações de variáveis separáveis

$$M(x) = N(y)y' \Leftrightarrow M(x) = N(y)\frac{dy}{dx}$$

Integra-se ambos os membros da equação.

$$\int M(x)\, dx = \int N(y)\frac{dy}{dx}\, dx \Leftrightarrow \int M(x)\, dx = \int N(y)\, dy$$

A última equivalência é obtida usando a relação existente entre dy e dx tal como referido aquando da integração por substituição.

Desta maneira obtemos a solução implícita da equação de variáveis separáveis.

Para se obter a solução explícita é, por vezes, necessário fazer uma restrição do domínio da função $y(x)$.

Nota. A *solução implícita* de uma equação diferencial é uma relação entre as variáveis dependente e independente (y e x), enquanto a *solução explícita* resulta de isolar a variável dependente num dos membros da igualdade. Nem sempre é necessário, ou sequer possível, calcular a solução explícita de uma equação diferencial.

Exemplos 4.4.

1. Determine a solução geral da equação diferencial $\dfrac{y}{x}y' = -1$.

   $\dfrac{y}{x}y' = -1 \overset{x\neq 0}{\Longrightarrow} yy' = -x \Leftrightarrow y\dfrac{dy}{dx} = -x \Longrightarrow$

   $\displaystyle\int y\,dy = \int -x\,dx \Leftrightarrow \frac{y^2}{2} = -\frac{x^2}{2} + k \overset{2k=c}{\Longleftrightarrow} x^2 + y^2 = c$

   A solução implícita da equação é $x^2 + y^2 = c$, com $c > 0$ e $x \neq 0$.

   Para $y > 0$ e $x \in]0, \sqrt{c}[$, a solução explícita é $y(x) = \sqrt{c - x^2}$.

2. Determine a solução particular da equação diferencial $y^2e^x + y' = 0$, que verifica a condição inicial $y(0) = 1$.

   $y^2e^x + y' = 0 \overset{y\neq 0}{\Longrightarrow} e^x + \dfrac{y'}{y^2} = 0 \Leftrightarrow \dfrac{1}{y^2}\dfrac{dy}{dx} = -e^x \Leftrightarrow$

   $\displaystyle\int -\frac{1}{y^2}\,dy = \int e^x\,dx \Leftrightarrow \frac{1}{y} = e^x + k, \quad k \in \mathbb{R}.$

   Para determinar a solução particular que verifica $y(0) = 1$ faz-se na solução geral $x = 0$ e $y = 1$.

   $\frac{1}{1} = e^0 + k \Leftrightarrow k = 0$

   A solução pretendida obtém-se substituindo k por 0 na solução geral, ou seja $y = e^{-x}$.

Problema de Misturas

Um contentor contém $20Kg$ de sal dissolvidos em $5000l$ de água. Uma solução de água salgada com $0.03Kg$ de sal por litro entra no contentor a uma velocidade de $25l/min$. A solução é misturada completamente e sai do contentor à mesma velocidade. Que quantidade de sal existe no contentor ao fim de meia hora?

Considerações iniciais:

- Considera-se para efeitos formulação do problema que a mistura da água que entra com a que está no contentor é feita instantaneamente.
- A concentração inicial é de $\dfrac{20}{5000} = 0.004Kg/l$.
- A quantidade máxima de sal que o contentor poderá ter é de $5000 \times 0.03 = 150Kg$.

Notação:

- $y(t)$ - quantidade de sal em Kg.
- $y(0) = 20Kg$.
- $y'(t)$ =taxa de entrada - taxa de saída= $E(t) - S(t)$ (medido em Kg/min)
- $E(t) = \underbrace{0.03Kg/l}_{\text{concentração}} \times \underbrace{25l/min}_{\text{velocidade}} = 0.75Kg/min$
- $S(t) = \dfrac{y(t)}{5000} \times 25 = \dfrac{y(t)}{200}$
- $y'(t) = E(t) - S(t) = \dfrac{150 - y}{200}$

Resolução da Equação Diferencial

$$\frac{dy}{dt} = \frac{150 - y}{200} \Leftrightarrow \frac{1}{150 - y}\frac{dy}{dt} = \frac{1}{200} \Leftrightarrow$$

$$-\int \frac{-1}{150 - y}dy = \int \frac{1}{200}dt \overset{y<150}{\Longleftrightarrow} -\ln(150 - y) = \frac{t}{200} + c, \quad c \in \mathbb{R} \Leftrightarrow$$

$$\ln(150 - y) = -\frac{t}{200} - c \Leftrightarrow 150 - y = e^{-\frac{t}{200}}e^{-c} \overset{k=e^{-c}}{\Longleftrightarrow}$$

[Quando c é uma constante real, e^{-c} é uma constante real positiva.]

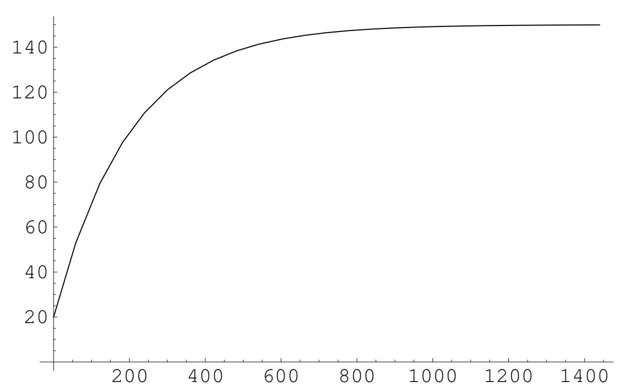
$150 - y = ke^{-\frac{t}{200}} \Leftrightarrow y = 150 - ke^{-\frac{t}{200}}, \quad k > 0$

Falta determinar a solução que verifica $y(0) = 20$.

$y(0) = 20 \Leftrightarrow 20 = 150 - k \Leftrightarrow k = 130$

A solução é $y(t) = 150 - 130e^{-\frac{t}{200}}$ e portanto

$y(30) = 150 - 130e^{-\frac{3}{20}} \approx 38,1Kg$.



$t \in [0, 1440] \qquad 1440min = 24h$

O gráfico mostra a evolução da quantidade de sal segundo este modelo, durante as primeiras $24h$. É fácil verificar, analítica e graficamente, que $\lim_{t \to +\infty} y(t) = 150$, ou seja a concentração tende para $0.03 Kg/l$.

Equações Diferenciais Lineares

As equações diferenciais lineares são um tipo particular de equações diferenciais, geralmente mais fáceis de resolver. Antes de vermos como se resolvem as equações lineares de 1ª ordem, vamos definir equações diferenciais lineares e dar alguns exemplos. No que segue $y^{(n)}$ representa a derivada de ordem n de y.

Uma *equação diferencial linear* de ordem n é uma equação diferencial que pode ser escrita na seguinte forma:

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = f(x)\,,$$

onde $a_0(x)$, $a_1(x), \cdots$, $a_n(x)$, $f(x)$ são funções reais com $a_n(x) \neq 0$. Supondo que $a_n(x)$ não se anula num determinado intervalo, podemos dividir toda a expressão por $a_n(x)$ e considerar sempre $a_n(x) = 1$.

Alguns exemplos de equações diferenciais lineares, e a respetiva ordem.

1. $y' + 2y = x$ — 1ª ordem
2. $y''' + \frac{1}{x}y' - y = \sin x$ — 3ª ordem
3. $xy'' - y' = 0$ — 2ª ordem
4. $x^2y' + y = x^2 \overset{x\neq 0}{\Longleftrightarrow} y' + \frac{1}{x^2}y = 1$ — 1ª ordem

Uma *equação diferencial linear de primeira ordem* é uma equação que pode ser escrita na forma

$$y' + p(x)y = q(x)\,,$$

onde $p(x)$ e $q(x)$ são funções contínuas num dado intervalo. Uma equação escrita nesta forma diz-se escrita na *forma canónica.*

Se $q(x) = 0$, então $y' + p(x)y = 0$ é uma equação de variáveis separáveis[2], e portanto
$y' + p(x)y = 0 \overset{y\neq 0}{\Longleftrightarrow} \dfrac{y'}{y} = -p(x) \overset{y>0}{\Longleftrightarrow} \ln y = -\displaystyle\int p(x)\,dx \Leftrightarrow y = e^{-\int p(x)\,dx}$.

Motivado pela solução deste caso particular, no caso geral vamos usar o fator integrante $\mu(x) = e^{\int p(x)\,dx}$. Um fator integrante é uma função que multiplicada pela equação permite

[2]Uma equação diferencial pode ser simultaneamente de variáveis separáveis e linear de primeira ordem.

transformá-la numa equação de solução conhecida, ou mais facilmente resolúvel. O objetivo de usar um fator integrante é semelhante ao que temos quando usamos uma mudança de variável noutros contextos (limites, integração, ...).

Solução Geral das Equações Lineares de Primeira Ordem

A partir da equação escrita na forma canónica $y' + p(x)y = q(x)$, multiplica-se ambos os membros da equação pelo fator integrante $\mu(x) = e^{\int p(x)\,dx}$.

$$y' + p(x)y = q(x) \Rightarrow \mu(x)y' + \overbrace{p(x)\mu(x)}^{\mu'(x)}\, y = \mu(x)q(x)$$

Depois de multiplicarmos pelo fator integrante, obtemos do lado esquerdo da equação a derivada do produto das funções $y(x)$ e $\mu(x)$.

$$\frac{d}{dx}\left(\mu(x)y(x)\right) = \mu(x)q(x) \Leftrightarrow \mu(x)y(x) = \int \mu(x)q(x)\,dx \Leftrightarrow$$

$$y(x) = \frac{1}{\mu(x)} \int \mu(x)q(x)\,dx$$

$$y = e^{-\int p(x)\,dx}\left(\int e^{\int p(x)\,dx} q(x)\,dx + c\right)$$

Deduzimos assim uma fórmula resolvente para as equações lineares de 1ª ordem. Apesar disso, nos exemplos que apresentamos a seguir as equações diferenciais vão ser resolvidas usando o fator integrante em cada caso.

Exemplo 4.5.

1. Determinar a solução geral da equação diferencial $y' + y = e^x$.

   O fator integrante desta equação é $\mu(x) = e^{\int 1 dx} = e^x$.

   $$y' + y = e^x \Leftrightarrow e^x y' + e^x y = e^{2x} \Leftrightarrow \frac{d}{dx}\left(e^x y\right) = e^{2x} \Leftrightarrow$$

   $$e^x y = \int e^{2x} dx \Leftrightarrow y = e^{-x}\left(\frac{e^{2x}}{2} + c\right) \Leftrightarrow$$

   $$y = \frac{1}{2}e^x + c\,e^{-x}, \quad c \in \mathbb{R}$$

2. Determinar a solução da equação diferencial $x^2 y' + xy = 1$, para $x > 0$.

   $$x^2 y' + xy = 1 \Leftrightarrow y' + \frac{1}{x}y = \frac{1}{x^2}$$

   O fator integrante é $\mu(x) = e^{\int \frac{1}{x} dx} = e^{\ln x} = x$.

(Para $x < 0$, o fator integrante seria $-x$. Esta diferença de sinal não teria nenhuma implicação prática.)

Multiplica-se ambos os membros da equação pelo fator integrante.

$$y' + \frac{1}{x}y = \frac{1}{x^2} \Leftrightarrow xy' + y = \frac{1}{x} \Leftrightarrow \frac{d}{dx}(xy) = \frac{1}{x} \Leftrightarrow$$

$$xy = \int \frac{1}{x}dx \Leftrightarrow y = \frac{1}{x}(\ln x + c) \Leftrightarrow y = \frac{\ln x}{x} + \frac{c}{x}, \quad c \in \mathbb{R}$$

Note-se que não existe nenhuma solução desta equação cujo o domínio contenha o valor $x = 0$.

Problema de Misturas – segunda versão

Um contentor contém $20Kg$ de sal dissolvidos em $5000l$ de água. Uma solução de água salgada com $0.03Kg$ de sal por litro entra no contentor a uma velocidade de 25l/min. A solução é misturada completamente e sai do contentor à velocidade de $50l/min$.

A principal alteração nesta versão em relação à anterior é que a velocidade de saída é superior à velocidade de entrada da água. Isso resulta no facto de passados $\frac{5000}{50-25} = 200$ minutos o tanque ficará vazio. A equação resultante desta modelação é uma equação linear de 1ª ordem.

- $y'(t)$ =taxa de entrada - taxa de saída= $E(t) - S(t)$ (medido em Kg/min)
- $E(t) = 0.03Kg/l \times 25l/min = 0.75Kg/min$
- $S(t) = \dfrac{y(t)}{5000-25t} \times 50 = \dfrac{2y(t)}{200-t}$

Resolução da Equação Diferencial

$$y' = \frac{3}{4} - \frac{2y}{200-t}, \; t \in [0, 200[$$

$$y' = \frac{3}{4} - \frac{2y}{200-t} \Leftrightarrow y' + \frac{2y}{200-t} = \frac{3}{4}$$

O fator integrante é $\mu(t) = e^{\int \frac{2}{200-t}dt} = e^{-2\ln(200-t)} = \dfrac{1}{(200-t)^2}$.

$$\frac{1}{(200-t)^2}y' + \frac{2}{(200-t)^3}y = \frac{3/4}{(200-t)^2} \Leftrightarrow$$

$$\frac{1}{(200-t)^2}y = \frac{-3/4}{200-t} + c \Leftrightarrow y(t) = \frac{3}{4}(200-t) + c\,(200-t)^2$$

Falta determinar o valor de c para achar a solução particular pretendida.

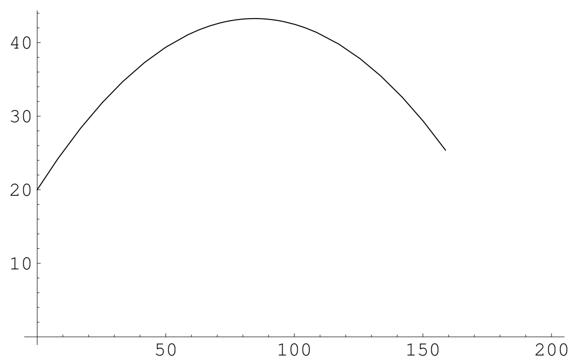
$$y(0) = 20 \Leftrightarrow 20 = \frac{3}{4}200 + c200^2 \Leftrightarrow c = -\frac{13}{4000}$$

A solução do problema é

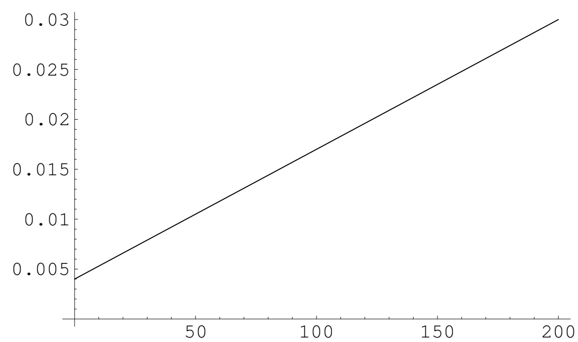
$$y(t) = \frac{3}{4}(200 - t) - \frac{13}{4000}(200 - t)^2$$

Neste caso, como a quantidade de água é variável, é interessante verificar qual a evolução da concentração de sal na água ao longo tempo. A concentração ao longo do tempo é dada pela expressão

$$c(t) = \frac{y(t)}{5000 - 25t} = \frac{3}{100} - \frac{13}{100000}(200 - t).$$



Evolução da quantidade de sal.



Evolução da concentração.

4.2 Equações Diferenciais como Modelos Matemáticos

A matemática é a linguagem que usamos para descrever, compreender ou prever fenómenos físico-químicos do mundo real, tendo as equações diferenciais um papel de relevo neste processo.

A representação ou descrição de leis naturais, físicas ou de fenómenos do mundo real usando conceitos matemáticos é conhecida como *Modelo Matemático*. Ao longo do semestre temos estudado vários modelos para exemplificar a utilização em contexto real dos conceitos matemáticos que estudámos, mas neste capítulo a utilidade da modelação matemática é mais facilmente demonstrada, e por isso dedicamos esta secção a estudar alguns modelos em particular.

Vamos começar por dois modelos que são normalmente usados para estudar o crescimento populacional: o modelo exponencial e o modelo logístico. O primeiro é um modelo mais básico e o segundo mais complexo.

Quando falamos de crescimento populacional, podemos falar do crescimento da população de animais ou plantas num determinado espaço, ou de uma cultura de bactérias, por exemplo. Estes modelos também podem ser usados para estudar a disseminação de doenças.

Em ambos os casos, quanto mais fechado ao exterior o sistema for mais fiável será o modelo. Os estudos reais são feitos com modelos mais complexos, mas em muitas situações estes modelos dão-nos alguma indicação sobre a situação real.

Modelo de Crescimento Exponencial

Neste modelo a população de uma comunidade (animais, vegetais, bactérias, ...) cresce a uma taxa proporcional ao tamanho dessa mesma população.

Em termos matemáticos, se $y(t)$ é o número de indivíduos da população no instante t, então $y'(t)$ é a taxa de crescimento da população nesse mesmo instante t.

O modelo de crescimento exponencial é descrito pela equação

$$y' = k\,y\,, \quad k \neq 0\ .$$

Esta equação é simultaneamente linear e de variáveis separáveis.

A solução da equação $y' = k\,y$ é

$$y(t) = y_0\,e^{kt}\ ,$$

onde $y_0 = y(0)$ é a população inicial.

A equação resolve-se da mesma maneira para valores de k positivos ou negativos, no entanto o sinal da derivada $y'(t)$ é diferente e portanto correspondem a situações distintas.

- $k > 0$ – *Modelos de crescimento exponencial.*
- $k < 0$ – *Modelos de decaimento exponencial.*

O modelo de crescimento exponencial é frequentemente bastante exato no início do crescimento de uma população, mas torna-se desadequado à medida que o tempo passa, por isso houve necessidade de criar modelos mais adaptados a evoluções mais longas.

Modelo Logístico

Neste modelo a população cresce exponencialmente quando a população é pequena, e tende a estabilizar quando a população se aproxima da chamada *capacidade de suporte*. A capacidade de suporte é a população que o ambiente consegue sustentar a longo prazo. No caso da disseminação de certo tipo de doenças pode-se considerar como *capacidade de suporte* o percentagem da população necessária para se atingir a imunidade de grupo, altura em que a doença tenderá naturalmente a estabilizar. Na situação pandémica da Covid19, esse número é obviamente muito alto e por isso foi necessário tomar medidas para sustê-la antes de atingir esse valor.

Em termos matemáticos, a *equação logística* é:

$$y' = k\,y\,(L-y)\,, \quad k>0\,,$$

onde $L>0$ é a capacidade de suporte.

A taxa de crescimento é proporcional ao produto do tamanho da população e da capacidade que a população ainda tem para aumentar.

Solução da Equação Logística

$$\frac{dy}{dt} = k\,y\,(L-y) \Leftrightarrow \frac{1}{y\,(L-y)}\,dy = k\,dt \Leftrightarrow$$

$$\int \frac{1}{y\,(L-y)}dy = \int k\,dt \Leftrightarrow \frac{1}{L}\int \frac{1}{y} + \frac{1}{L-y}dy = \int k\,dt \Leftrightarrow$$

$$\ln|y| - \ln|L-y| = kLt + c \Leftrightarrow \ln\left|\frac{L-y}{y}\right| = -kLt - c \Leftrightarrow$$

$$\left|\frac{L-y}{y}\right| = e^{-kLt-c} \Longleftrightarrow \frac{L-y}{y} = Ce^{-kLt},\ C = \pm e^{-c} \neq 0$$

Temos assim que a solução da equação logística é:

$$y = \frac{L}{C\,e^{-kLt}+1} \Leftrightarrow y = \frac{L}{C\,D^t+1},$$ onde $D = e^{-kL} > 0$ é uma constante real positiva.

Fazendo $y_0 = y(0)$, a população no instante inicial, obtemos que $C = \dfrac{L-y_0}{y_0}$.

Não é difícil verificar que neste modelo a população tende para o valor da capacidade de suporte, $\lim\limits_{t\to\infty} y(t) = \lim\limits_{t\to\infty} \dfrac{L}{Ce^{-kLt}+1} = L$, com $k>0$.

- Se $y_0 \in]0, L[$ a população vai aumentar.
- Se $y_0 \in]L, +\infty[$ a população vai diminuir.

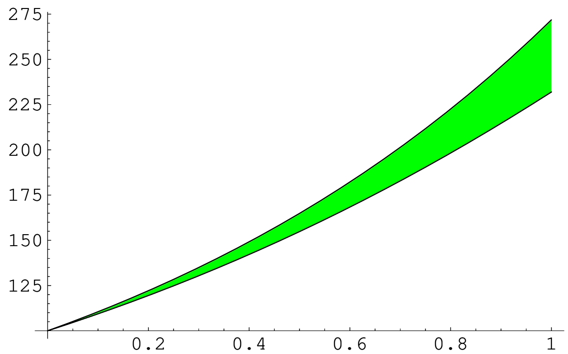
Teoricamente também é possível considerar o modelo logístico com $k<0$. Neste caso a população tenderia para 0.

Exponencial vs Logística

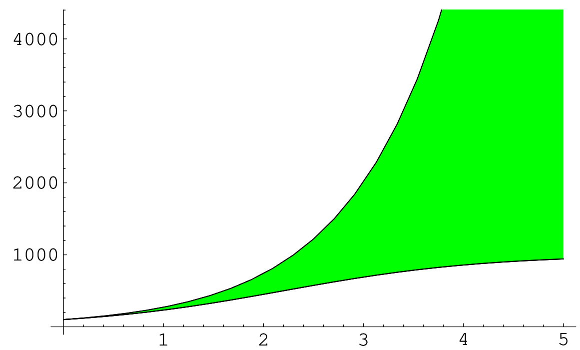
Os próximos gráficos servam para comparar a evolução de uma população de crescimento exponencial e de crescimento logístico. Consideramos os mesmos valores iniciais para a população e para a taxa de crescimento, 100 em ambos casos ($y'(0) = k\,y(0)$ e $y'(0) = k\,L\,y(0)$, respetivamente). A taxa de crescimento considerada é propositadamente muito alta para que sejam facilmente visíveis as diferenças entre os dois modelos

Do lado esquerdo vemos que nos instantes iniciais não há grandes diferenças entre as duas curvas, mas do lado direito a diferença entre os dois modelos já é bem visível, com o modelo logístico a estabilizar enquanto o exponencial contínua a crescer.

$y_0 = 100 \qquad L = 1000 \qquad k = 1 \text{ (exponencial)} \qquad kL = 1 \text{ (logística)}$

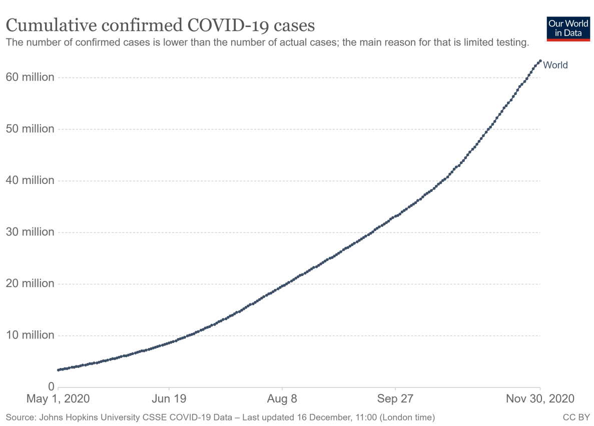


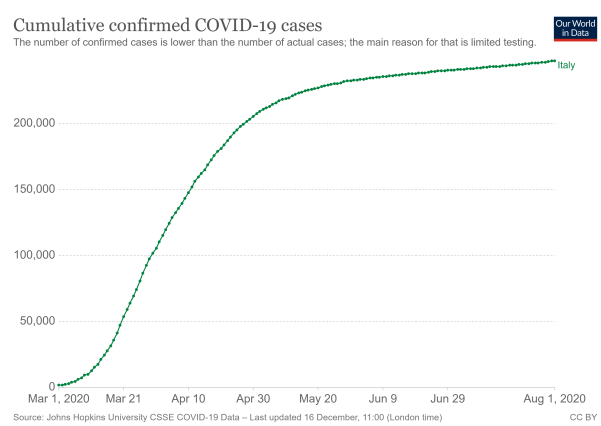
$t \in [0, 1]$



$t \in [0, 5]$

A diferença entre o crescimento duma pandemia, no seu início ou atingindo uma maior percentagem da população, pode também ser observado nos gráficos do número de casos de Covid19 nos primeiros meses em Itália e no Mundo. Essencialmente, durante esse período a pandemia cresceu a um ritmo *quase* exponencial em todo mundo, onde a proporção de casos registados em relação ao total da população era ainda bastante baixa. Em Itália depois de um início onde a doença era desconhecida, e cresceu naturalmente, a curva começou a ser *achatada* em virtude de medidas restritivas muito severas. Neste caso o modelo comportou-se de modo semelhante ao da curva logística, o que na nossa linguagem significa que a *capacidade de suporte* do sistema foi drasticamente reduzida.

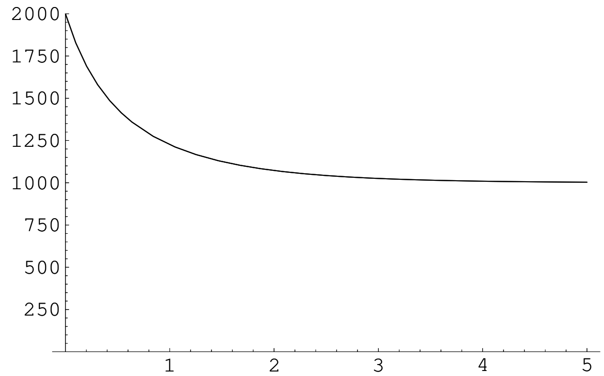




Logística de Crescimento Negativo

Para que o estudo fique completo mostramos também o gráfico de um função logística onde o valor inicial é maior do que a capacidade de suporte, $y_0 > L$.

$y_0 = 2000 \qquad L = 1000 \qquad kL = 1$



$t \in [0, 5]$

Exercício

Um estudante portador do vírus da gripe regressa a um colégio com 1000 alunos. Suponha que o colégio está isolado e que o vírus se propaga com uma taxa de variação proporcional não apenas ao número y de alunos já infectados mas também ao número de alunos não infectados. Determine o número de alunos infectados após 6 dias, sabendo que passados 4 dias eles são já 50.

Estamos nas condições no modelo logístico com $L = 1000$ e $y_0 = 1$. Da solução da equação logística, sabemos que:

$$y = \frac{L}{C\,D^t + 1} \text{ com } C = \frac{L - y_0}{y_0} = \frac{1000 - 1}{1} = 999\,.$$

Por outro lado, sabemos também que

$$y(4) = 50 \Leftrightarrow 50 = \frac{1000}{999\,D^4 + 1} \Leftrightarrow 999\,D^4 + 1 = \frac{1000}{50} \Leftrightarrow$$

$$D^4 = \frac{19}{999} \Leftrightarrow D = \left(\frac{19}{999}\right)^{1/4}.$$

Finalmente concluímos que

$$y = \frac{1000}{999\,\left(\frac{19}{999}\right)^{t/4} + 1}\,,$$

e portanto

$$y(6) = \frac{1000}{999\,\left(\frac{19}{999}\right)^{3/2} + 1} \approx 276\,.$$

De seguida vamos descrever dois modelos práticos de aplicação das equações diferenciais, o modelo de *desintegração radioativa* e a *lei de arrefecimento de Newton*. Em ambos os casos, eles são essencialmente modelos de decaimento exponencial.

Desintegração Radioativa e Datação por Carbono

Neste modelo matemático, que descreve a desintegração radioativa de uma dada substância, é assumido que a taxa de desintegração dos átomos é proporcional ao número de átomos presentes em cada instante.

Em termos matemáticos, se $m(t)$ representa a massa da substância radioativa no instante t, então $m'(t)$ é a taxa de (de)crescimento da massa nesse instante.

Este modelo de decaimento exponencial é descrito pela equação

$$m' = k\,m\,, \quad k < 0\,.$$

A solução da equação $m' = k\,m$ é

$$m(t) = m_0\,e^{kt}\,,$$

onde $m_0 = m(0)$ é a massa inicial.

A constante k é usualmente determinada em termos da meia-vida do material, isto é o tempo necessário para uma quantidade inicial dessa substância, m_0, por desintegração se reduzir a metade. Designemos por $V > 0$ a meia-vida do material, logo

$$\frac{m(V)}{m_0} = \frac{1}{2} \Leftrightarrow \frac{m_0\,e^{kV}}{m_0} = \frac{1}{2} \Leftrightarrow k = -\frac{\ln 2}{V}\,.$$

Temos assim que a solução da equação $m' = k\,m$ é

$$m(t) = m_0\,e^{-\frac{\ln 2}{V}t} = m_0\,2^{-\frac{t}{V}}\,,$$

onde V é a meia-vida do material.

Exemplo 4.6. Um pedaço de madeira foi encontrado com $1/500$ da quantidade original de Carbono 14. Determine a sua idade, sabendo que a meia-vida de Carbono 14 é 5600.

Pelos cálculos que fizemos anteriormente, sabemos que:

$$m(t) = m_0\,2^{-\frac{t}{5600}}\,.$$

Queremos descobrir o valor de t quando $m(t) = \frac{m_0}{500}$.

$$\frac{m_0}{500} = m_0\,2^{-\frac{t}{5600}} \Leftrightarrow \frac{1}{500} = 2^{-\frac{t}{5600}} \Leftrightarrow 500 = 2^{\frac{t}{5600}} \Leftrightarrow$$

$$\frac{t}{5600} = \log_2 500 \Leftrightarrow t = 5600 \times \log_2 500 \approx 50200$$

O pedaço de madeira tem aproximadamente 50200 anos.

Lei de Arrefecimento de Newton

A lei do arrefecimento de Newton estabelece que a taxa a que um corpo arrefece é proporcional à diferença entre a temperatura do corpo e a temperatura do meio que circunda o corpo, o chamado meio ambiente.

Em termos matemáticos, se $T(t)$ representa a temperatura do corpo no instante t, então $T'(t)$ é a taxa de arrefecimento (ou de aquecimento, conforme o caso). Designamos por T_a a temperatura do meio ambiente.

Este modelo é descrito pela equação

$$T' = k\,(T - T_a)\,, \quad k < 0 \text{ assumindo que o corpo está a arrefecer .}$$

O modelo é de decaimento exponencial em relação a $T - T_a$, e não em relação à variável T.

A solução da equação do arrefecimento é

$T' = k(T - T_a) \Leftrightarrow \dfrac{1}{T - T_a}T' = k \Leftrightarrow$

$\displaystyle\int \frac{1}{T - T_a}dT = \int k\,dt \Leftrightarrow \ln(T - T_a) = kt + c,\ c \in \mathbb{R},\, T > T_a \Leftrightarrow$

$T = T_a + C\,e^{kt}$, $C > 0$.

Além disso, fazendo $t = 0$ obtemos $C = T_0 - T_a$, (T_0 é a temperatura inicial) pelo que a solução da equação é

$$T(t) = T_a + (T_0 - T_a)e^{kt}.$$

Por exemplo, para determinar a hora de um óbito usa-se $T_0 = 37^{\mathrm{o}}C$.

5 Álgebra Linear

A Álgebra Linear é uma área muito vasta da matemática da qual iremos estudar apenas uma pequena parte. O nosso principal objetivo no estudo que vamos realizar é mostrar como calcular a reta de regressão linear, e estudar outros problemas semelhantes. Para isso vamos introduzir uma ferramenta que nos vai ser útil para o fazer, as Matrizes.

5.1 Matrizes

Definição 5.1. Uma *matriz* A do tipo $m \times n$ é um quadro que se obtém dispondo $m \times n$ números reais em m linhas e n colunas.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

De forma abreviada, escrevemos $A = [a_{ij}]_{m \times n}$, onde a_{ij} denota o elemento (ou entrada ou componente) genérico da matriz situado na linha i, coluna j de A.

- Se $m = n$ (nºlinhas=nºcolunas), a matriz A diz-se *quadrada* de ordem n.
- Uma matriz do tipo $1 \times n$ chama-se *vetor linha*.
- Uma matriz do tipo $m \times 1$ chama-se *vetor coluna*.

Exemplos 5.2.

$$A = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix} \quad B = \begin{bmatrix} -2 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 5 \\ 1 \\ 4 \\ 6 \end{bmatrix} \quad E = \begin{bmatrix} 3 & 0 & 2 \end{bmatrix}$$

1. A matriz A é do tipo 2×2 (matriz quadrada).
2. A matriz B é do tipo 3×2.
3. A matriz C é do tipo 2×3.
4. A matriz D é do tipo 4×1 (vetor coluna).
5. A matriz E é do tipo 1×3 (vetor linha).

Matrizes Especiais

Existem algumas matrizes que desempenham um papel especial na aritmética de matrizes. Nomeadamente a matriz a nula e a matriz identidade que desempenham o papel de elemento neutro para a adição e para a multiplicação de matrizes, respetivamente, como veremos mais à frente.

1. *Matriz Nula*
   (todas as entradas são 0)

$$0_{m\times n} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

2. *Matriz Diagonal*
   (matriz quadrada que tem zeros fora da diagonal principal)

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

3. *Matriz Identidade*
   (matriz diagonal com $d_1 = d_2 = \cdots = d_n = 1$)

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

A matriz nula e a matriz identidade não são uma única matriz, mas sim uma família de matrizes. Para cada tipo de matrizes existe uma matriz nula, e para cada tipo de matriz quadrada existe uma matriz identidade. (Uma matriz diagonal é um caso ainda mais geral.)

De seguida vamos estudar as principais operações de matrizes. O nosso modelo serão as operações com vetores, que nos vai servir da base para a definição das operações com matrizes.

Adição de Matrizes

Só é possível somar matrizes do mesmo tipo. A *soma de duas matrizes* é feita somando os elementos que estão na mesma posição, tal como é feito com vetores do mesmo "tamanho".

Exemplo 5.3.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -4 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & -1 & -2 \\ 2 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix}$$

$$A+B=\begin{bmatrix}1+0 & 0-1 & 0-2\\ -4+2 & 0+0 & 1+2\\ 0+0 & 2+1 & 0+1\end{bmatrix}=\begin{bmatrix}1 & -1 & -2\\ -2 & 0 & 3\\ 0 & 3 & 1\end{bmatrix}$$

Multiplicação por um Escalar

O *produto de uma matriz por um escalar* é uma matriz do mesmo tipo da primeira, que se obtém multiplicando todas as entradas pelo valor escalar. Mais uma vez, este caso é idêntico ao que acontece quando multiplicamos um escalar por um vetor.

Exemplo 5.4.

O produto de 2 pela matriz $A=\begin{bmatrix}1 & 0 & 0\\ -4 & 0 & 1\\ 0 & 2 & 0\end{bmatrix}$ é a matriz $2A=\begin{bmatrix}2 & 0 & 0\\ -8 & 0 & 2\\ 0 & 4 & 0\end{bmatrix}$.

Teorema 5.5. *Sejam A, B e C matrizes do tipo $m \times n$ e $\alpha, \beta \in \mathbb{R}$. São válidas as seguintes propriedades.*

1. $A+B=B+A$ *(comutatividade)*

2. $A+(B+C)=(A+B)+C$ *(associatividade)*

3. *Existe uma única matriz* $\mathbf{0}$ *do tipo* $m \times n$ *tal que* $A+\mathbf{0}=A$. *(elemento neutro da soma)*

4. *Para cada matriz* A *existe uma única matriz* $-A=(-1)A$ *tal que* $A+(-A)=\mathbf{0}$. *(elemento simétrico)*

5. $\alpha(\beta A)=(\alpha\beta)A$ *(associatividade do produto por escalares)*

6. $(\alpha+\beta)A=\alpha A+\beta A$ *(distributividade)*

7. $\alpha(A+B)=\alpha A+\alpha B$ *(distributividade)*

Nota. A diferença entre duas matrizes do mesmo tipo pode definir-se à custa da adição e da multiplicação por -1:

$$A-B:=A+(-1)B.$$

Antes de definirmos o produto de matrizes, vamos relembrar como se calcula o produto escalar de vetores. O produto de matrizes de um vetor linha por um vetor coluna é uma matriz 1×1 em que a única entrada tem exatamente o valor do produto escalar desses dois vetores.

Produto Escalar de Vetores

O produto escalar do vetor $a = (a_1, a_2, \cdots, a_n)$ pelo vetor $b = (b_1, b_2, \cdots, b_n)$ é

$$< a, b >= a_1b_1 + a_2b_2 + \cdots + a_nb_n = \sum_{i=1}^{n} a_ib_i \,.$$

Se considerarmos a como vetor linha, e b como vetor coluna, definimos o produto das matrizes

$$a = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \text{ e } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

como sendo a matriz 1×1 em que a única entrada é o valor $< a, b >$.

$$a\,b = [a_1b_1 + a_2b_2 + \cdots + a_nb_n] = \begin{bmatrix} \sum_{i=1}^{n} a_ib_i \end{bmatrix}$$

Multiplicação de uma Matriz por um Vetor Coluna

Sejam agora A uma matriz $m \times n$; $L_1,\ L_2, \cdots L_m$ as suas linhas e b um vetor coluna $n \times 1$.

A multiplicação da matriz A pelo vector coluna b é um vetor coluna do tipo $m \times 1$, onde o elemento na linha i é o valor do produto interno da linha i, L_i, pelo vetor b:

$$Ab = \begin{bmatrix} < L_1, b > \\ < L_2, b > \\ \vdots \\ < L_m, b > \end{bmatrix} .$$

Exemplo 5.6.

$$A = \begin{bmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{bmatrix} \qquad b = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$$

$$Ab = \begin{bmatrix} 2 \times (-1) + 4 \times 2 + (-2) \times 2 \\ 4 \times (-1) + 9 \times 2 + (-3) \times 2 \\ (-2) \times (-1) + (-3) \times 2 + 7 \times 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \\ 10 \end{bmatrix}$$

Multiplicação de Matrizes

Sejam agora A uma matriz $m \times n$ como anteriormente, B uma matriz $n \times p$ e C_1, $C_2, \cdots C_p$ as suas colunas. O produto AB é uma matriz $m \times p$, onde o elemento na posição (i, j), i.e. na linha i e coluna j, é igual ao produto escalar da linha i de A com a coluna j de B, $< L_i, C_j >$.

Se designarmos por a_{ij} e por b_{ij} os elementos na posição (i, j) de A e B, respectivamente, então $< L_i, C_j >= \displaystyle\sum_{k=1}^{n} a_{ik} b_{kj}$.

O produto de duas matrizes só está definido quando o número de colunas da primeira matriz é igual ao número de linhas da segunda.

Exemplo 5.7. Consideremos $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, do tipo 3×3, e $B = \begin{bmatrix} 3 & 2 \\ -5 & 1 \\ 8 & 0 \end{bmatrix}$, do tipo 3×2.

Então o produto de A por B está definido e

$$AB = \begin{bmatrix} 1 \times 3 + 2 \times (-5) + 3 \times 8 & 1 \times 2 + 2 \times 1 + 3 \times 0 \\ 4 \times 3 + 5 \times (-5) + 6 \times 8 & 4 \times 2 + 5 \times 1 + 6 \times 0 \\ 7 \times 3 + 8 \times (-5) + 9 \times 8 & 7 \times 2 + 8 \times 1 + 9 \times 0 \end{bmatrix} = \begin{bmatrix} 17 & 4 \\ 35 & 13 \\ 52 & 22 \end{bmatrix}.$$

O produto BA não está definido, pois o número de colunas de B é diferente do número de linhas de A.

Teorema 5.8. *Sejam A, B, C e D matrizes, $\alpha \in \mathbb{R}$. São válidas as seguintes igualdades sempre que as multiplicações estiverem definidas.*

1. $A(BC) = (AB)C$ *(associatividade)*

2. $A(B + C) = AB + AC$ *(distributividade à esquerda)*

3. $(B + C)D = BD + CD$ *(distributividade à direita)*

4. $\alpha(AB) = (\alpha A)B = A(\alpha B)$

5. *Se A é uma matriz $m \times n$, então*

$$AI_n = A \text{ e } I_m A = A. \quad \textit{(elemento neutro)}$$

Nem todas as propriedades das operações com números reais são válidas para as operações com matrizes, em particular **o produto de matrizes não é comutativo**.

Matrizes Invertíveis

Definição 5.9. Uma matriz quadrada A, de ordem n, é *invertível* se existe uma matriz B (quadrada de ordem n) tal que
$$AB = I_n = BA.$$
Se existir, essa matriz é única, designa-se por inversa de A e denota-se por A^{-1}.

Nem toda a matriz quadrada tem inversa, no entanto se uma matriz quadrada tiver inversa de um dos lados, então é invertível. O próximo teorema diz exatamente isso.

Teorema 5.10. *Seja A uma matriz quadrada.*

1. *Se B é uma matriz quadrada tal que $BA = I$, então $B = A^{-1}$.*

2. *Se B é uma matriz quadrada tal que $AB = I$, então $B = A^{-1}$.*

Matriz Invertível de ordem 2

Uma matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ é invertível se e só se $ad - bc \neq 0$.

Neste caso a sua inversa é
$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Exemplo 5.11.
$$\begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}^{-1} = \frac{1}{5 \times 3 - 2 \times 7} \begin{bmatrix} 3 & -2 \\ -7 & 5 \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -7 & 5 \end{bmatrix}$$

Nota. Ao valor $ad - bc$ chama-se *determinante* da matriz A. O determinante para matrizes quadradas de ordem superior também se pode definir, no entanto não o faremos no nosso curso.

Teorema 5.12. *Sejam A e B matrizes invertíveis e $\alpha \neq 0$.*

1. *A matriz AB é invertível e $(AB)^{-1} = B^{-1}A^{-1}$.*

2. *A matriz αA é invertível e $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$.*

Transposta de uma Matriz

Definição 5.13. A *transposta* da matriz A é a matriz que resulta de A, escrevendo as linhas como colunas (ou as colunas como linhas) pela mesma ordem. Esse matriz designa-se por A^t.

Se A for do tipo $m \times n$, então a matriz transposta de A, A^t, é do tipo $n \times m$.

Exemplo 5.14. $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}$ $\qquad A^t = \begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$

Definição 5.15. Uma matriz quadrada é *simétrica* se $A^t = A$.

Proposição 5.16. *Sejam A e B duas matrizes, $\alpha \in \mathbb{R}$. São válidas as seguintes propriedades:*

1. $(A^t)^t = A$.
2. $(A + B)^t = A^t + B^t$.
3. $(\alpha A)^t = \alpha A^t$.
4. $(AB)^t = B^t A^t$.
5. *Se a matriz A é invertível, então A^t é invertível e $(A^t)^{-1} = (A^{-1})^t$.*
6. *Se a A é invertível e simétrica, então A^{-1} também é simétrica.*

Sistemas de Equações Lineares

Um *sistema de equações lineares* do tipo $m \times n$, isto é com m equações e n incógnitas, é um conjunto de equações lineares da forma

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = b_2 \\ & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = b_m \end{cases},$$

onde os valores a_{ij} e b_i são constantes e as variáveis x_j são as incógnitas do sistema.

Forma Matricial

Usando o produto de matrizes, um sistema de equações lineares pode ser escrito como uma equação matricial

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Ou seja, o sistema pode ser escrito na forma $Ax = b$, onde:

- a matriz A é uma matriz $m \times n$, e é designada por *matriz do sistema* ou matriz dos coeficientes;
- o vetor coluna x é uma matriz $n \times 1$ e é o *vector das incógnitas*;
- o vetor coluna b é uma matriz $m \times 1$ e é o *vector dos termos independentes.*

O conhecimento da inversa de uma matriz, quando existe, permita resolver um sistema de equações lineares através de uma simples multiplicação de matrizes.

Teorema 5.17. *Se A é uma matriz invertível de ordem n, então para cada vector b do tipo $n \times 1$, o sistema $Ax = b$ tem uma e uma única solução, nomeadamente $x = A^{-1}b$.*

Mas nem todos os sistemas têm uma solução única.

Definição 5.18. Os sistemas lineares podem ser classificados como:

(i) impossíveis (os que não têm soluções);

(ii) possíveis e determinados (têm uma única solução);

(iii) possíveis e indeterminados (têm uma quantidade infinita de soluções).

Já vimos uma maneira simples de resolver os sistemas possíveis e determinados quando a matriz do sistema é invertível.

De seguida vamos ver como achar uma solução para um sistema impossível. Claro que sendo um sistema impossível não tem solução, o problema que vamos estudar a seguir consiste em achar a melhor solução possível. Existem várias alternativas para definir o que é a "melhor" solução, na linguagem de funções o melhor é o mínimo de uma função (de várias variáveis). O método que vamos usar para determinar a melhor solução baseia-se na forma matricial do sistema.

5.2 Método dos Mínimos Quadrados

Sejam A uma matriz do tipo $m \times n$ e b um vector $m \times 1$.

Se um sistema $Ax = b$ for impossível, não existe nenhum vector x $n \times 1$ que satisfaça a igualdade $Ax = b$. No entanto, podemos procurar os vetores $\overline{x}$ que tornem mínima a distância entre $A\overline{x}$ e b, isto é, a norma $\|A\overline{x} - b\|$.

Definição 5.19. Um vector $\overline{x}$ do tipo $n \times 1$ é *solução no sentido dos mínimos quadrados* do sistema $Ax = b$ se

$$\|A\overline{x} - b\| = \min\{\|Ax - b\| : x \text{ é um vetor } n \times 1\}.$$

Para determinar a solução no sentido dos mínimos quadrados temos que achar o mínimo da função de n variáveis $f(x_1, \cdots, x_n) = \|Ax - b\|$, que é equivalente a determinar o mínimo da função $g(x_1, \cdots, x_n) = \|Ax - b\|^2$. No entanto como já foi dito anteriormente, vamos usar a forma matricial do sistema para calcular essa solução.

Teorema 5.20. *Sejam A uma matriz do tipo $m \times n$ e b um vetor $m \times 1$. Então o sistema $Ax = b$ tem sempre solução no sentido dos mínimos quadrados.*

O vetor $\overline{x}$ é solução no sentido dos mínimos quadrados de $Ax = b$ se e só se é solução do sistema $(A^t A)x = A^t b$.

Se o sistema $Ax = b$ é possível, então a solução o sistema $(A^t A)x = A^t b$ tem exatamente a(s) mesma(s) solução(ões).

Portanto, para calcularmos uma solução no sentido dos mínimos quadrados de $Ax = b$ apenas temos de resolver o sistema $A^t A\overline{x} = A^t b$. Notemos que a matriz $A^t A$ é simétrica.

Definição 5.21. O erro da solução aproximada, no sentido dos mínimos quadrados, é o valor de $\|A\overline{x} - b\|$, onde $\overline{x}$ é (um)a solução no sentido dos mínimos quadrados de $Ax = b$.

Exemplo 5.22. Sejam $A = \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix}$ e $b = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$.

O sistema $Ax = b$ é impossível. Para achar a solução no sentido dos mínimos quadrados resolvemos o sistema:

$$A^t A\overline{x} = A^t b \Leftrightarrow \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \overline{x} = \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \Leftrightarrow$$

$$\begin{bmatrix} 6 & 0 \\ 0 & 30 \end{bmatrix} \overline{x} = \begin{bmatrix} 2 \\ 16 \end{bmatrix} \Leftrightarrow \overline{x} = \begin{bmatrix} 1/3 \\ 8/15 \end{bmatrix}$$

O erro desta solução é

$$\|A\overline{x} - b\| = \| \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/3 \\ 8/15 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \| =$$

$$\| \begin{bmatrix} 3 \\ -2/5 \\ 1/5 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} \| = \| \begin{bmatrix} 0 \\ -12/5 \\ -24/5 \end{bmatrix} \| = \sqrt{(12/5)^2 + (24/5)^2} \approx 5.36$$

Isto significa que o mínimo da função
$g(x_1, x_2) = \|Ax - b\|^2 = (x_1 + 5x_2 - 3)^2 + (2x_1 - 2x_2 - 2)^2 + (-x_1 + x_2 - 5)^2$ *é atingido quando* $x_1 = 1/3$ *e* $x_2 = 8/15$.

Há vários problemas em que o objectivo passa pelo determinação de uma função $y = f(x)$ que se ajuste a um conjunto de dados experimentais da melhor maneira possível. Normalmente pretende-se que o gráfico da função seja tão simples quanto possível, como uma reta ou uma parábola.

Vamos começar por estudar como se calcula a reta de regressão linear.

Regressão Linear

Queremos determinar a recta $y = mx + b$ que melhor se ajusta a um conjunto de pontos do plano: $(x_1, y_1); (x_2, y_2); \cdots ; (x_n, y_n)$.

Substituindo os valores dos pontos na equação da reta, obtemos um sistema linear com n equações (uma por cada ponto) e duas incógnitas, m e b.

$$\left\{ \begin{array}{rcl} m\,x_1 + b & = & y_1 \\ m\,x_2 + b & = & y_2 \\ & \cdots & \\ m\,x_n + b & = & y_n \end{array} \right. \Leftrightarrow \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Se os pontos $(x_1, y_1); (x_2, y_2); \cdots ; (x_n, y_n)$ não forem colineares (ou estiverem sobre uma reta vertical), o sistema é impossível. Neste caso podemos procurar a reta que melhor se aproxima dos pontos no sentido dos mínimos quadrados.

Exemplo 5.23. Determinemos a reta $y = mx + b$ que melhor se ajusta, no sentido dos mínimos quadrados, aos pontos $(1,2), (2,2)$ e $(3,4)$.

$$\begin{cases} m + b &= 2 \\ 2m + b &= 2 \\ 3m + b &= 4 \end{cases} \Leftrightarrow \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

O sistema é impossível. A solução no sentido dos mínimos quadrados é a solução do sistema

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 18 \\ 8 \end{bmatrix}.$$

Resolvendo o sistema obtém-se que $m = 1$ e $b = 2/3$, e portanto a reta que melhor se ajusta, no sentido dos mínimos quadrados, aos pontos dados tem equação $y = x + 2/3$.

Erro na Regressão Linear

O cálculo do erro faz-se, como no caso geral, calculando a norma da diferença entre os valores estimados e os valores reais, ou seja o erro é

$$\| \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} m\,x_1 + b \\ m\,x_2 + b \\ \vdots \\ m\,x_n + b \end{bmatrix} \|,$$

onde m e b são os valores obtidos na solução no sentido dos mínimos quadrados.

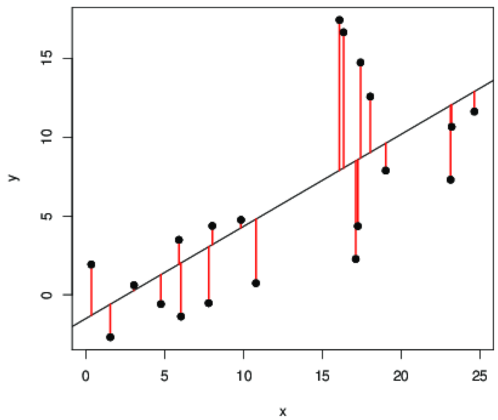
No exemplo anterior, o erro é

$$\| \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 + 2/3 \\ 2 + 2/3 \\ 3 + 2/3 \end{bmatrix} \| = \sqrt{6}/3 \approx 0.82$$

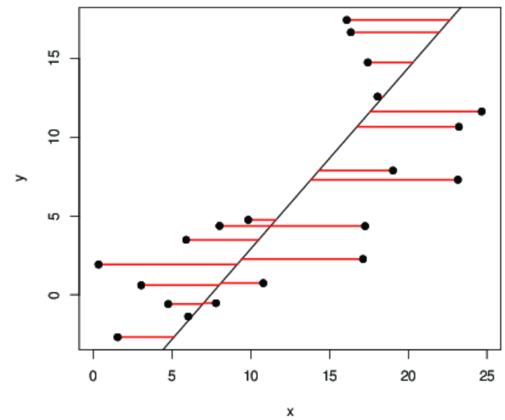
Graficamente, o erro é a norma do vetor em que as coordenadas são o módulo da diferença entre a ordenada de cada ponto (x_i, y_i) e a ordenada do ponto com a mesma abscissa da reta de regressão linear, $|y_i - (mx_i + b)|$.

Nota. A reta $y = mx+b$ serve para estimar o valor de y em função de x, mas não o contrário uma vez que o erro é medido na variável y.

Para estimar o valor de x em função de y é preciso calcular a reta $x = ny + a$, que só por coincidência será a mesma.



Erro medido na variável y.



Erro medido na variável x.

Regressão Quadrática

O ajuste de uma curva a um conjunto de pontos não precisa de ser feito por uma reta. Podemos, por exemplo, usar uma parábola de equação $y = ax^2 + bx + c$ (regressão quadrática). Neste caso as incógnitas que é preciso determinar são a, b, e c.

Exemplo 5.24. Determinar a parábola $y = ax^2 + bx + c$ que melhor se ajusta, no sentido dos mínimos quadrados, ao conjunto de pontos $\{(0,0),(1,1),(2,3),(-1,-1)\}$. Vamos ter um sistema com 4 equações e 3 incógnitas, substituindo na equação da parábola x e y pela ordenada e pela abscissa de cada um dos pontos para obter as quatro equações.

$$\left\{\begin{array}{rrrcr} a\,0^2+ & b\,0+ & c & = & 0 \\ a\,1^2+ & b\,1+ & c & = & 1 \\ a\,2^2+ & b\,2+ & c & = & 3 \\ a\,(-1)^2+ & b\,(-1)+ & c & = & -1 \end{array}\right. \Leftrightarrow \left\{\begin{array}{rcr} c & = & 0 \\ a+b+c & = & 1 \\ 4a+2b+c & = & 3 \\ a-b+c & = & -1 \end{array}\right. \Leftrightarrow \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ -1 \end{bmatrix}$$

A solução no sentido dos mínimos quadrados deste sistema é a solução do sistema:

$$\begin{bmatrix} 0 & 1 & 4 & 1 \\ 0 & 1 & 2 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & 1 & 4 & 1 \\ 0 & 1 & 2 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \\ -1 \end{bmatrix} \Leftrightarrow \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 12 \\ 8 \\ 3 \end{bmatrix}$$

Resolvendo o sistema obtém-se que $a = \frac{1}{4}$, $b = \frac{21}{20}$ e $c = -\frac{3}{20}$, portanto a parábola que melhor se ajusta, no sentido dos mínimos quadrados, aos pontos dados tem equação $y = \dfrac{x^2}{4} + \dfrac{21x}{20} - \dfrac{3}{20}$.

Referências

[1] J. Stewart, *Cálulo* – Volumes I e II, 5ªed., Thomson Learning, São Paulo, 2007.

[2] A. P. Santana, J. F. Queiró, *Introdução à Álgebra Linear*, Gradiva, Lisboa, 2010.

[3] Dennis G. Zill, *Equações Diferenciais com aplicações em Modelagem*, Cengage Learning Editores, 2003.