Capítulo 1

Preliminares

Vou apresentar/recordar apenas as noções básicas, com alguns exemplos, sobre espaços métricos, espaços vectoriais normados e espaços com produto interno. A ideia é simplesmente a de fixar notações e preencher alguma lacuna que os alunos tenham sobre estes tópicos

1.1 Espaços métricos

Sejam X um conjunto não vazio e $d : X \times X \to \mathbb{R}$ uma função. Dizemos que d é uma *distância* ou *métrica sobre* X se, para todo $x, y, z \in X$, se verifica:

- $d(x, y) \geq 0$;
- $d(x, y) = 0 \Leftrightarrow x = y$;
- $d(x, y) = d(y, x)$;
- $d(x, z) \leq d(x, y) + d(y, z)$ (chamada desigualdade triangular).

Ao par (X, d) chamamos *espaço métrico*. Omitiremos, frequentemente, a referência à métrica de um espaço métrico, designando-o simplesmente por X. Chama-se *métrica discreta* sobre X à função d definida por $d(x, x) = 0$ e $d(x, y) = 1$, se $x \neq y$. Se d for uma métrica sobre X e satisfizer a condições $d(x, z) \leq \max\{d(x, y), d(y, z)\}$ quaisquer que sejam $x, y, z \in X$, dizemos que d é uma *ultra-métrica*.

Sobre o conjunto $\mathbb{R}^n$ há duas distâncias que são bem conhecidas: d_2 e d_∞. Elas são definidas por (sendo $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{R}^n$)

$$d_2\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2},$$
$$d_2\big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}.$$

Um exemplo (mais estranho) é o da métrica discreta aplicada a um qualquer conjunto X não vazio: $d(x, y) = 1$, se x, y forem elementos distintos de X (é claro que $d(x, x)$ tem de ser igual a 0).

Dados um espaço métrico (X, d), $x_0 \in X$ e $r > 0$, chamamos aos conjuntos

- $B_d(x_0, r) = \{x \in X : d(x, x_0) < r\}$, *bola de centro em* x_0 *e raio* r;
- $D_d(x_0, r) = \{x \in X : d(x, x_0) \leq r\}$, *disco de centro em* x_0 *e raio* r;
- $S_d(x_0, r) = \{x \in X : d(x, x_0) = r\}$, *esfera de centro em* x_0 *e raio* r.

Por exemplo, em $\mathbb{R}^2$, $D_{d2}(x_0, r)$ é um círculo centrado em x_0 e raio r enquanto que $D_{d_\infty}(x_0, r)$ é uma quadrado centrado em x_0 e lado $2r$. Se d for a métrica discreta então, por exemplo $D_d(x_0, r)$ é igual a $\{x_0\}$, se $r < 1$, e igual a X, se $r \geq 1$.

Não havendo dúvidas de qual a métrica em questão, omitiremos o d que aparece subscrito.

Definição 1.1. *Seja (X, d) um espaço métrico e $A \subseteq X$. Dizemos que A é:*

- aberto *de X se para todo o ponto de A existir uma bola centrada nesse ponto que está contida em A;*
- *fechado de X se para qualquer ponto de $X \setminus A$ existir uma bola centrada nesse ponto que não intersecta A, isto é, está contida em $X \setminus A$.*

Note-se que um conjunto é aberto se e só se o seu complementar é fechado. Note-se ainda que $\emptyset$ e X são conjuntos que são simultaneamente abertos e fechados.

Definição 1.2. *Dizemos que uma* sucessão $(a_n)_n$ *de elementos de X é:*

- convergente *se existe $a \in X$ tal que*
$$\forall \varepsilon > 0 \; \exists p \in \mathbb{N} \; \forall n \geq p \quad d(a_n, a) \leq \varepsilon;$$
- de Cauchy *se*
$$\forall \varepsilon > 0 \; \exists p \in \mathbb{N} \; \forall n, m \geq p \quad d(a_n, a_m) \leq \varepsilon.$$

Nestas definições podemos substituir $\leq$ por $<$ que a noção não se altera. Isto é verdade porque qualquer bola centrada num ponto contem sempre um disco contido nesse ponto.

É simples de ver que qualquer sucessão convergente cujos elementos pertencem a um fechado de X, converge para um elemento desse fechado. E isto caracteriza os espaços fechados em espaços métricos.

Um espaço métrico em que toda a sucessão de Cauchy é convergente é designado por *espaço métrico completo.*

Definição 1.3. *Sejam (X, d) e (Y, D) dois espaços métricos e $f : X \to Y$ uma função. Dado $x_0 \in X$, dizemos que f é contínua em x_0 se*
$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in X \quad d(x, x_0) \leq \delta \Rightarrow D(f(x), f(x_0)) \leq \varepsilon.$$
Dizemos que a função f *é* contínua *se f for contínua em todos os pontos do seu domínio.*

Note-se que a implicação na definição de continuidade significa $f(D(x_0), \delta)) \subseteq D(f(x_0), \epsilon)$. À imagem do que foi dito acima para as sucessões, podemos também na definição substituir $\leq$ por $<$ que a noção não se altera. Se fizermos essa alteração isso significa $f(B(x_0), \delta)) \subseteq B(f(x_0), \epsilon)$.

Definição 1.4. *Sejam (X, d) e (Y, D) e.m, e $f : X \to Y$ uma função. Dizemos que f é:*

- <u>uniformemente contínua</u> *se*
$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, y \in X \quad d(x, y) \leq \delta \Rightarrow D(f(x), f(y)) \leq \varepsilon.$$
Essencialmente, o δ pode ser escolhido independentemente dos pontos.
- *um* <u>homeomorfismo</u> *se é contínua, bijectiva e f^{-1} é contínua;*
- *uma* <u>isometria</u> *se*
$$\forall x, y \in X \quad D(f(x), f(y)) = d(x, y).$$

Dizemos que dois *espaços métricos* são *homeomorfos* se existe um homeomorfismo entre eles e *isométricos* se existe uma isometria sobrejectiva entre eles (note-se que uma isometria é sempre injectiva). No primeiro caso, os dois espaços são "o mesmo" a nível topológico" e no segundo caso, eles são ´' 'o mesmo" a nível métrico (que implica, topológico).

Sejam X um conjunto e d_1, d_2 duas métricas em X. Dizemos que a *métrica* d_1 é *mais fina* que a métrica d_2 se a função identidade de (X, d_1) em (X, d_2) for contínua. Se esta função for um homeomorfismo dizemos que as *métricas* d_1 e d_2 são *equivalentes.*

Vou agora introduzir 3 noções topológica, sendo que a terceira é a mais importante no contexto da disciplina. A noção de compacto que apresento não é a definição usual mas é uma definição equivalente para espaços métricos.

Definição 1.5. *Um espaço métrico X diz-se:*

- conexo *se os únicos subconjuntos abertos e fechados de X são $\emptyset$ e X;*
- conexo por arcos *se, dados $x, y \in X$, existe $\varphi : [a, b] \to X$ função contínua tal que $\varphi(a) = x$ e $\varphi(b) = y$.*
- compacto *se toda a sucessão em X admitir uma sub-sucessão convergente.*

É simples ver que se um subconjunto de um espaço métrico, visto ele próprio como espaço métrico, é compacto então é fechado e limitado. É também conhecido, que em $\mathbb{R}^n$, com a métrica d_2 ou d_∞ (que são equivalentes) as noções de compacto e de fechado limitado coincidem.

1.2 Espaços vectoriais normados

Como caso particular de espaços métricos temos os espaços vectoriais normados. A ideia é que estes espaços vectoriais admitem uma distância que é compatível com a a estrutura de espaço vectorial.

Definição 1.6. *Seja X um espaço vectorial sobre $\mathbb{R}$. Uma função $\|\cdot\| : X \longrightarrow \mathbb{R}$ diz-se uma* **norma** *se para todo $x, y \in X$ e para todo $a \in \mathbb{R}$:*

a) $\|x\| \geq 0$;

b) $\|x\| = 0 \Leftrightarrow x = 0$;

c) $\|ax\| = |a|\,\|x\|$;

d) $\|x + y\| \leq \|x\| + \|y\|$.

Vejamos agora como, a partir de uma norma sobre um espaço vectorial, se define uma métrica.

Teorema 1.7. *Seja X um espaço vectorial real e $\cdot : X \times X \longrightarrow \mathbb{R}$ é uma norma. Então a função*

$$\begin{array}{ccc} X \times X & \longrightarrow & \mathbb{R} \\ (x, y) & \mapsto & \|x - y\| \end{array}$$

é uma distância, que se diz "associada à norma".

Demonstração. Vejamos o único passo que não é trivial, a desigualdade triangular. Se $x, y, z \in X$ então, usando a alínea d) da definição de norma, $\|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\|$. □

Como exemplos temos, novamente, $\mathbb{R}^n$ com as normas definidas por

$$\|(x_1, \ldots, x_n)\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}, \quad \|(x_1, \ldots, x_n)\|_\infty = \max\{|x_1|, \ldots, |x_n|\}.$$

Note-se que as distâncias associadas a estas normas são d_2 e d_∞.

1.3 Produto interno

De seguida vamos definir o que se entende por produto interno sobre um espaço vectorial e ver que a esse produto interno podemos associado uma norma.

Definição 1.8. *Se X é um espaço vectorial sobre $\mathbb{R}$, uma função* $\begin{array}{ccc} X \times X & \longrightarrow & X \\ (x, y) & \mapsto & x \cdot y \end{array}$ *diz-se um* **produto interno** *se para todo $x, y, z \in X$ e para todo $a \in \mathbb{R}$:*

a) $x \cdot x \geq 0$;

b) $x \cdot x = 0 \Leftrightarrow x = 0$;

c) $x \cdot y = y \cdot x$;

d) $(ax) \cdot y = a(x \cdot y)$;

e) $(x + y) \cdot z = (x \cdot z) + (x \cdot z)$.

Muitos autores preferem usar a notação $x|y$ ou $< x, y >$ em vez de $x \cdot y$, para designar o "produto interno entre x e y". Como exemplo de um produto interno temos a função definida em $\mathbb{R}^n$ por

$$(x_1, \dots, x_n) \cdot (y_1, \dots, y_n) = x_1 y_1 + \cdots + x_n y_n.$$

Antes de definir a norma associada a um produto interno vamos dar uma ideia da demonstração da chamada desigualdade de Cauchy-Schwarz, que será generalizada mais tarde (desigualdade de Hölder).

Proposição 1.9 (Desigualdade de Cauchy-Schwarz)**.** *Se X é um espaço vectorial real e $\cdot : X \times X \longrightarrow \mathbb{R}$ é um produto interno então:*

$$\forall x, y \in X \quad |x \cdot y| \leq \sqrt{(x \cdot x)\,(y \cdot y)}.$$

Além disso, há igualdade se e só se x e y forem linearmente dependentes.

Demonstração. Podemos supor que $x, y \neq \vec{0}$, pois caso contrário temos trivialmente a igualdade. Atendendo à alínea a) da definição de norma e usando a alínea d) dessa mesma definição temos

$$0 \leq \big((y \cdot y)\, x - (x \cdot y)\, y\big) \cdot \big((y \cdot y)\, x - (x \cdot y)\, y\big) = \cdots = (y \cdot y)\Big((x \cdot x)(y \cdot y) - (x \cdot y)^2\Big)$$

E a desigualdade segue imediatamente e a segunda parte segue "facilmente". □

Teorema 1.10. *Se X é um espaço vectorial real e $\cdot : X \times X \longrightarrow \mathbb{R}$ é um produto interno então a função*

$$\begin{array}{rccc} \| \ \| : & X & \longrightarrow & \mathbb{R} \\ & x & \mapsto & \sqrt{x \cdot x} \end{array}$$

é uma norma, que se diz "associada ao produto interno".

Demonstração. A "única dificuldade" é a demonstração de que a função $\| \ \|$ satisfaz a condição d) da definição de norma.

Sejam $x, y \in X$. Então,

$$\begin{array}{rcl} \|x + y\| \leq \|x\| + \|y\| & \Leftrightarrow & (\|x + y\|)^2 \leq (\|x\| + \|y\|)^2 \\ & \Leftrightarrow & (x + y) \cdot (x + y) \leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2 \\ & \Leftrightarrow & x \cdot x + 2\, x \cdot y + y \cdot y \leq x \cdot x + 2\|x\|\,\|y\| + y \cdot y \\ & \Leftrightarrow & x \cdot y \leq \|x\|\,\|y\|. \end{array}$$

Para concluir basta notar que esta última desigualdade é válida pois é uma consequência da Desigualdade de Cauchy-Schwarz. □

Usando a notação dada por este resultado, a desigualdade de Cauchy-Schwarz pode ser escrita como $|x \cdot y| \leq \|x\|\,\|y\|$, se a norma provem de um produto interno. A prova é muito simples: basta desenvolver ambos os lados da igualdade que queremos provar.

Proposição 1.11 (Regra do paralelogramo)**.** *Se X é um espaço vectorial com um produto interno então*

$$\forall x, y \in X \quad \|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Apenas para dar uma ideia intuitiva, o paralelogramo tem como vértices $\vec{0}$, x, y e $x + y$. A regra dos paralelogramo o que diz é: a soma dos quadrados das medidas das diagonais do paralelogramo é igual à soma dos quadrados das medidas dos lados do paralelogramo.

Como exemplo, podemos ver que a norma $\| \ \|_\infty\|$ definida em $\mathbb{R}^n$ não provem de um produto interno. Isto é verdade porque, se x for o vector que tem 1 na primeira coordenada e 0 nas outras, y for o vector que tem 1 na segunda coordenadas e 0 não outras então

$$\|x\|_\infty = \|y\|_\infty = \|x + y\|_\infty = \|x - y\|_\infty = 1$$

e, portanto,

$$\|x + y\|_\infty^2 + \|x - y\|_\infty^2 = 2 \quad \text{e} \quad 2\left(\|x\|_\infty^2 + \|y\|_\infty^2\right) = 4.$$

1.4 Exemplos

Vamos ver alguns exemplos de espaços vectoriais normados de dimensão infinita. Esses exemplos serão de espaços de funções reais de variável real ou de espaços de sucessões reais. Em alguns desses exemplo não é feita a prova (que é simples) de que se trata realmente de uma norma.

1.4.1 O espaço l^∞

l^∞ é o espaço vectorial formado pelas sucessões reais limitadas com a norma (terá de ser provado que estamos de facto na presença de um espaço vectorial normado) definida por

$$\forall (a_n)_n \in l^\infty \quad \|(a_n)_n\| = \sup_n |a_n|.$$

Note-se, se tivermos uma sucessão $(x_n)_n$ em l^∞ cujos elementos são $x_n = (x_{n,k})_k$ então $(x_n)_n$ converge para $x \in l^\infty$ se e só se

$$\forall \varepsilon > 0 \; \exists p \in \mathbb{N} \; \forall n \geq p \; \|(x_n - x)\|_\infty \leq \varepsilon.$$

Se $x = (a_k)_k$ então, isto é o mesmo que

$$\forall \varepsilon > 0 \; \exists p \in \mathbb{N} \; \forall n \geq p \; \sup_k |(x_{n,k} - a_k| \leq \varepsilon.$$

Em particular, se fixar um $s \in \mathbb{N}$ então $|(x_{n,s} - a_s| \leq \sup_k |(x_{n,k} - x_k|$, e, portanto,

$$\forall \varepsilon > 0 \; \exists p \in \mathbb{N} \; \forall n \geq p \; |x_{n,s} - a_s| \leq \varepsilon.$$

Fica assim visto que, se $(x_n)_n$ converge para x em l^∞ então, para todo $s \in \mathbb{N}$, $(x_{n,s})_n$ converge para $a_s \in l^\infty$.

Exemplo 1.12. *Para cada $n \in \mathbb{N}$ seja $x_n = (x_{n,k})_k$ o elemento de l^∞ cujo termo de ordem n é 1 e os outros termos são todos iguais a 0. Na tabela abaixo represento o elemento x_i na linha i, sendo que x está representado na última linha.*

$$\begin{array}{c||cccccc}
x_1 & 1 & 0 & 0 & 0 & \dots \\
x_2 & 0 & 1 & 0 & 0 & \dots \\
x_3 & 0 & 0 & 1 & 0 & \dots \\
x_4 & 0 & 0 & 0 & 1 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x & 0 & 0 & 0 & 0 & \cdots
\end{array}$$

Deste modo, se esta sucessão convergisse para um elemento $x = (a_k)_k$ então, usando o que é dito acima, $a_s = \lim_n x_{n,s} = 0$, para todo s. Deste modo x seria a sucessão nula. Mas, de facto, a sucessão $(x_n)_n$ não converge para x pois $\|x_n - x\| = 1$, para todo $n \in \mathbb{N}$. Poderíamos também ter notado que a sucessão $(x_n)_n$ não é de Cauchy e, portanto, não é convergente.

Ou outro exemplo, bastante parecido é o seguinte

Exemplo 1.13. *Para cada $n \in \mathbb{N}$ seja $x_n = (x_{n,k})_k$ o elemento de l^∞ cujos termos de ordem menor ou igual a n é 1 e os outros termos são todos iguais a 0. É claro que a sucessão não é de Cauchy pois $\|x_n - x_p\| = 1$, para todo n, p inteiros positivos diferentes (como acontece no exemplo anterior). Deste modo, a sucessão não converge.*

Usando o raciocínio acima podemos ver que, se $(x_n)_n$ fosse convergente então convergiria para a sucessão constante e igual a 1.

$$\begin{array}{c||cccccc}
x_1 & 1 & 0 & 0 & 0 & \dots \\
x_2 & 1 & 1 & 0 & 0 & \dots \\
x_3 & 1 & 1 & 1 & 0 & \dots \\
x_4 & 1 & 1 & 1 & 1 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x & 1 & 1 & 1 & 1 & \cdots
\end{array}$$

Um outro exemplo "curioso" é o dado pela figura

$$\begin{array}{c||llll l}
x_1 & a_1 & 0 & 0 & 0 & \dots \\
x_2 & a_1 & a_1+a_2 & 0 & 0 & \dots \\
x_3 & a_1 & a_1+a_2 & a_1+a_2+a_3 & 0 & \dots \\
x_4 & a_1 & a_1+a_2 & a_1+a_2+a_3 & a_1+a_2+a_3+a_4 & \dots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}$$

em que $(a_k)_k$ é uma sucessão que gera uma série convergente. É simples ver que a sucessão $(x_n)_n$ não é de Cauchy. Do mesmo modo o "candidato" a limite da sucessão (usando o que é dito antes dos exemplos) é uma sucessão que não pertence a l^∞.

A norma $\| \ \|_\infty$ pode também ser definida sobre o conjunto das sucessões convergentes, denotado por c, ou sobre o conjunto das sucessões convergentes para 0, denotado por c_0, que são sub-espaços de l^∞. Todos estes espaços são espaços completos.

1.4.2 O espaço l^p, com $p \geq 1$

. Para $p \geq 1$ consideremos l^p o conjunto de todas as sucessões reais $(a_n)_n$ tais que $\sum_{n=1}^{\infty} |a_n|^p < \infty$. Para ver que l^p é de facto um espaço vectorial é preciso verificar que é fechado para a soma (as outras propriedades são simples de mostrar). Para o demonstrar vamos usar a desigualdade de números reais

$$\forall p \geq 1 \ \forall a, b \in \mathbb{R} \quad |a+b|^p \leq 2^p \Big(|a|^p + |b|^p \Big), \tag{1.1}$$

cuja demonstração, no caso em que $|a| \leq |b|$ (o outro é semelhante) pode ser feita notando que $|a+b|^p \leq (|a|+|b|)^p \leq (2|b|)^p = 2^p |b|^p$.

Deste modo, se $(a_n)_n, (b_n)_n \in l^p$ então $(a_n+b_n)_n \in l^p$ porque

$$\sum_{n=1}^{\infty} |a_n+b_n|^p \leq \sum_{n=1}^{\infty} 2^p \Big(|a_n|^p + |b_n|^p \Big) \leq 2^p \left(\sum_{n=1}^{\infty} |a_n|^p + \sum_{n=1}^{\infty} |a_n|^p \right) < \infty.$$

Nota 1.14. *De facto a desigualdade que usamos pode ser melhorada. Usando a convexidade da função $f : \mathbb{R}_0^+ \to \mathbb{R}$ definida por $f(x) = x^p$ podemos mostrar que*

$$\forall p \geq 1 \ \forall a, b \in \mathbb{R} \quad |a+b|^p \leq 2^{p-1} \Big(|a|^p + |b|^p \Big).$$

Sobre este espaço vamos definir uma função, que veremos ser uma norma,

$$\begin{array}{rccc}
\| \ \|_p : & l^p & \to & \mathbb{R}. \\
& (a_n)_n & \mapsto & \left(\displaystyle\sum_{n=1}^{\infty} |a_n|^p \right)^{\frac{1}{p}}
\end{array}$$

A prova de que esta função é uma norma é a chamada desigualdade de Minkowski que vamos provar usando as desigualdades de Young e a desigualdade de Hölder.

Note-se que, se $p = 1$ o resultado é simples (basta usar a desigualdade triangular em $\mathbb{R}$). No caso em que $p = 2$ também é fácil de ver que estamos na presença de uma norma pois é fácil de ver que a função

$$\begin{array}{ccc}
l^2 \times l^2 & \to & \mathbb{R}. \\
\big((a_n)_n, (b_n)_n \big) & \mapsto & \displaystyle\sum_{n=1}^{\infty} a_n b_n
\end{array}$$

é um produto interno sobre l^2 cuja norma associada é a função $\| \ \|_2$.

Definição 1.15. *Se $p \in [1,\infty]$ chamamos conjugado de p ao único elemento p' de $[1,\infty]$ tal que*

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Na definição acima estamos a considerar $\frac{1}{\infty} = 0$. Deste modo p' é igual a: ∞, se $p = 1$; 1, se $p = \infty$; $p' = \frac{p-1}{p}$, caso contrário. Note-se que $\frac{1}{2}$ é conjugado de si próprio e que o conjugado do conjugado de p é p.

Lema 1.16 (Desigualdade de Young)**.** *Se $p, q > 1$ são tais que $\frac{1}{p} + \frac{1}{q} = 1$ (p e q conjugados um do outro) então*

$$\forall a, b \in \mathbb{R}_0^+ \quad a\,b \leq \tfrac{1}{p}\,a^p + \tfrac{1}{q}\,b^q.$$

Demonstração. Podemos supor que $a, b > 0$ pois caso contrário a desigualdade é trivial.

Usando o facto de a função logaritmo ser convexa, porque a sua segunda derivada é estritamente negativa, e a igualdade $\frac{1}{p} + \frac{1}{q} = 1$, temos

$$\log(\tfrac{1}{p}\,a^p + \tfrac{1}{q}\,b^q) \geq \tfrac{1}{p}\log(\,a^p) + \tfrac{1}{q}\log(\,b^q) = \log(a) + \log(b)$$

de onde obtemos a desigualdade pretendida aplicando a exponencial a ambos os membros. $\square$

Lema 1.17 (Desigualdade de Hölder)**.** *Se $p, q > 1$ são tais que $\frac{1}{p} + \frac{1}{q} = 1$ (p e q conjugados um do outro), $(x_n)_n \in l^p$ e $(y_n)_m \in l^q$ então $(x_n y_n)_n \in l^1$ e*

$$\|(x_n y_n)_n\|_1 \leq \|(x_n)_n\|_p\,\|(y_n)_n\|_q.$$

Demonstração. Se $\|(x_n)_n\|_p = 0$ ou $\|(y_n)_n\|_q = 0$ então $(x_n)_n$ ou $(y_n)_n$ é a sucessão nula e a desigualdade é trivialmente satisfeita. Suponhamos então $\|(x_n)_n\|_p \neq 0$ e $\|(y_n)_n\|_q \neq 0$, Aplicando a desigualdade de Young temos, se $i \in \mathbb{N}$

$$\frac{|x_i|}{\|(x_n)_n\|_p} \cdot \frac{|y_i|}{\|(y_n)_n\|_q} \leq \frac{1}{p}\frac{|x_i|^p}{\|(x_n)_n\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|(y_n)_n\|_q^q}.$$

Somando em i obtemos

$$\frac{1}{\|(x_n)_n\|_p \cdot \|(y_n)_n\|_q}\sum_{i=1}^{\infty}|x_i y_i| \leq \frac{1}{p} + \frac{1}{q} = 1$$

e a conclusão segue imediatamente. $\square$

Estamos agora em condições de mostrar a desigualdade de Minkowski, ou seja, o que faltava para mostrar que as funções $\|\ \ \|_p$ são, de facto, normas.

Lema 1.18 (Desigualdade de Minkowski)**.** *Seja $p > 1$. Se $(x_n)_n, (y_n)_n \in l^p$ então*

$$\|(x_n + y_n)_n\|_p \leq \|(x_n)_n\|_p + \|(y_n)_n\|_p.$$

Demonstração. Recorde-se que já sabemos que $(x_n + y_n)_n \in l^p$. No que segue vamos usar a desigualdade de Hölder na segunda desigualdade e denotar o conjugado de p por q. Temos

$$\begin{aligned}
\sum_{i=1}^{\infty}|x_i + y_i|^p &= \sum_{i=1}^{\infty}|x_i + y_i|\,|x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^{\infty}|x_i|\,|x_i + y_i|^{p-1} + \sum_{i=1}^{\infty}|y_i|\,|x_i + y_i|^{p-1} \\
&\leq \left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{\infty}|x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}} + \left(\sum_{i=1}^{\infty}|y_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{\infty}|x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}} \\
&= \left[\left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty}|y_i|^p\right)^{\frac{1}{p}}\right]\left(\sum_{i=1}^{\infty}|x_i + y_i|^{(p-1)q}\right)^{\frac{1}{q}}.
\end{aligned}$$

Deste modo, como $(p-1)q = p$,

$$\|(x_n + y_n)_n\|_p^p \leq \Big[\|(x_n)_n\|_p + \|(y_n)_n\|_p\Big] \|(x_n + y_n)_n\|_p^{\frac{p}{q}}$$

e a conclusão segue porque $p - \frac{p}{q} = p(1 - \frac{1}{q}) = 1$. □

Nota 1.19. *Consideremos uma sucessão* $(x_n)_n$ *em* l^p *e denotemos, para* $n \in \mathbb{N}$, $x_n = (x_{n,s})_s$. *Com argumentos semelhantes aos que foram usados em* l^∞ *é simples de ver que, se* $(x_n)_n$ *converge para* $x = (a_s)_s$ *em* l^p, *então, para todo* $s \in \mathbb{N}$, $(x_{n,s})_n$ *converge para* $a_s \in l^\infty$.

O seguinte resultado relaciona os diversos espaços l^p e c_0.

Proposição 1.20. *Se* $1 \leq p \leq q$ *então* $l^p \subsetneqq l^q \subsetneqq c_0 \subsetneqq l^\infty$. *Além disso, se* $1 \leq p \leq q \leq \infty$ *então* $\|x\|_q \leq \|x\|_p$, *para todo* $x \in l^p$.

Demonstração. A primeira inclusão pode ser vista como uma consequência da segunda parte da proposição. Em alternativa podemos seguir o seguinte raciocínio. Se $x = (x_n)_n \in l^p$ então a sucessão $(x_n)n$ converge para 0 (pelo que sabemos sobre séries) e, portanto, existe uma ordem a partir da qual $x_n \leq 1$. Daqui concluímos que (a partir dessa ordem) $|x_n|^p \leq |x_n|^q$, o que é suficiente para ver que $x \in l^q$.

As outras inclusões são imediatas.

Para mostrar que não há as igualdades entre os diversos espaços considerados basta considerar; a sucessão $\left(\frac{1}{n^{\frac{1}{p}}}\right)_n$ que pertence a l^q e não a l^p; a sucessão $\left(\frac{1}{n^{\frac{1}{q}}}\right)_n$ que pertence a c_0 e não a l^q; a sucessão constante e igual a 1 que pertence a l^∞ e não a c_0.

Para a segunda parte da proposição consideremos $x \in l^p$ e suponhamos primeiro que $q =\neq \infty$. Se x é a sucessão nula então a desigualdade é trivialmente verdadeira. Se x não é a sucessão nula então $\|x\|_p \neq 0$ e $\frac{|x_n|}{\|x\|_p} \leq 1$, para todo $n \in \mathbb{N}$. Deste modo, como $q > p$ temos $\left(\frac{|x_n|}{\|x\|_p}\right)^q \leq \left(\frac{|x_n|}{\|x\|_p}\right)^p$, ou seja, $\|x\|_p^p\,|x_n|^q \leq \|x\|_p^q\,|x_n|^p$. Deste modo, somando em n,

$$\|x\|_p^p \sum_{n=1}^{\infty} |x_n|^q = \|x\|_p^q \sum_{n=1}^{\infty} |x_n|^p \quad \text{ou seja} \quad \|x\|_p^p \|x\|_q^q = \|x\|_p^q \|x\|_p^p$$

e a conclusão segue imediatamente.

Se $q = \infty$ então $|x_n| \leq \|x\|_p$ para todo o n e, portanto, $\|x\|_\infty = \sup_n |x_n| \leq \|x\|_p$. □

É claro que em $\mathbb{R}^N$ a desigualdade $\|x\|_p \leq \|x\|_q$ também é válida, se $x = (x_1, \ldots, x_N)$ e $1 \leq p \leq q \leq \infty$ (basta usar a proposição anterior aplicada à sucessão obtida de $(x_1, \ldots, x_n, 0, \ldots)$ que pertencem a l^1. É claro que as desigualdades de Hölder e de Minkowski também se aplicam nestas condições.

No caso particular de $\mathbb{R}^N$ temos uma desigualdade entre as normas em sentido contrário.

Proposição 1.21. *Se* $1 \leq p < q \leq \infty$ *e* $x \in \mathbb{R}^N$ *então*

$$\|x\|_p \leq N^{\frac{1}{p} - \frac{1}{q}} \|x\|_q \qquad (\textit{entendendo-se que } \tfrac{1}{\infty} = 0).$$

Demonstração. Se $q = \infty$ então, para $n = 1, \ldots, N$, $|x_n| \leq \|x\|_\infty$ e, portanto, $\|x\|_p \leq \left(\sum_{n=1}^N \|x\|_\infty^p\right)^{\frac{1}{p}} = \left(\|x\|_\infty^p \sum_{n=1}^N 1\right)^{\frac{1}{p}} = N^{\frac{1}{p}} \|x\|_\infty$.

Se $q < \infty$ então, aplicando a desigualdade de Hölder, notando que o conjugado de $\frac{q}{p}$ é $\frac{q}{q-p}$,

$$\|x\|_p^p = \sum_{n=1}^N |x_n|^p \cdot 1 \leq \left(\sum_{n=1}^N |x_n|^q\right)^{\frac{p}{q}} \left(\sum_{n=1}^N 1\right)^{\frac{q-p}{q}} = \|x\|_q^p N^{\frac{q-p}{q}}$$

e a conclusão segue de imediato. □

1.4.3 Alguns espaços de funções reais de variável real

Seja A um conjunto e denotemos por $B(A)$ o espaço vectorial real cujos elementos são as funções reais limitadas definidas em A. Sobre $B(A)$ definimos a função

$$\begin{array}{cccc} \| \ \|_\infty : & B(A) & \longrightarrow & \mathbb{R}. \\ & f & \mapsto & \sup_{x\in A}|f(x)| \end{array}$$

Tal como no caso de l^∞, é simples de ver que estamos na presença de uma norma. Esta norma é também chamada a norma da convergência uniforme.

Note-se que se uma sucessão de funções $(f_n)_n$ em $B(A)$ convergir para uma função f então ela converge pontualmente, isto é, para todo $x \in A$, $(f_n(x))_n$ converge para $f(x)$. Sendo assim, se quisermos ver se uma dada sucessão em $B(A)$ é convergente podemos começar por estudar a convergência pontual dessa sucessão de modo a obtermos um candidato para o limite.

É claro que esta norma pode ser usada em sub-espaços vectoriais de $B(A)$. Por exemplo, se tivermos um intervalo I então podemos considerar o conjunto formado pelas funções reais definidas em I que são contínuas e limitadas. No caso de I ser um intervalo fechado então, é claro, estamos a considerar $C(I,\mathbb{R})$, ou simplesmente, $C(I)$, o conjunto formado por todas as funções contínuas de I em $\mathbb{R}$.

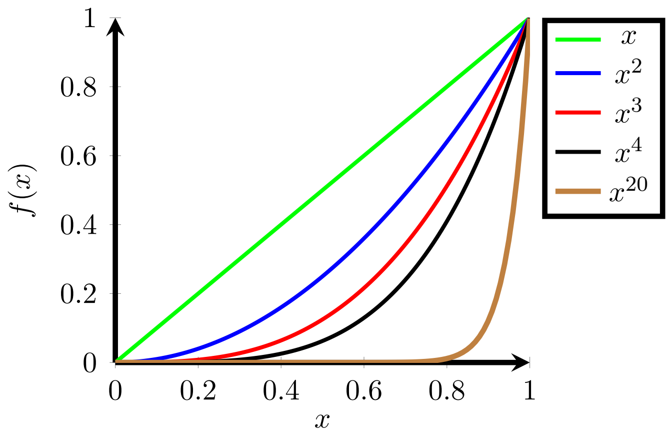
Neste contexto, é bem conhecido que o limite uniforme de funções contínuas é contínuo.

Exemplos 1.22. *Vejamos dois exemplos.*

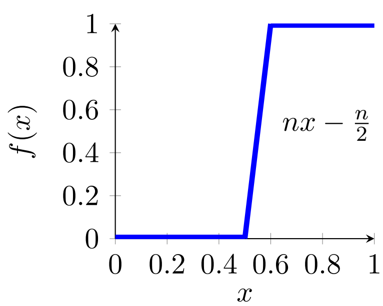
1. *Consideremos $(f_n)_n$ em $C([0,1])$ definidas por $f_n(x) = x^n$.*

   Deste modo, se $x \in [0,1]$, $(f_n(x))_n$ converge para 0, se $x \neq 1$, e converge para 1, se $x = 1$. Deste modo a sucessão converge pontualmente para uma função descontínua, mostrando que a sucessão original $(f_n)_n$ não é convergente em $C([0,1])$.

   É claro que poderíamos, usando conhecimentos de Cálculo em $\mathbb{R}$, mostrar que a sucessão $(f_n)_n$ não é de Cauchy.



2. *Seja $f_n : [0,1] \to \mathbb{R}$ definida por*



$$f_n(x) = \begin{cases} 0 & se\ x \leq \frac{1}{2} \\ 1 & se\ x \geq \frac{1}{2} + \frac{1}{n} \\ nx - \frac{n}{2} & se\ \frac{1}{2} < x < \frac{1}{2} + \frac{1}{n} \end{cases}$$

   É claro que, se $x \in [0, \frac{1}{2}]$ então $(f_n(x)_n$ é a sucessão constante e igual a 0 e, se $x > \frac{1}{2}$ então a partir de certa ordem $f_n(x) = 1$. Deste modo $(f_n)_n$ converge pontualmente para uma função descontínua e, portanto, não converge em $C([0,1])$ com a norma $\| \ \|_\infty$. É claro que é também simples de ver que a sucessão $(f_n)_n$ não é de Cauchy uma vez que, para $m \geq n$, $\|f_m - f_n\|_\infty \geq |f_m(\frac{1}{2} + \frac{1}{m}) - f_n(\frac{1}{2} + \frac{1}{m})| = 1 - \frac{n}{m}$. Em particular $\|f_{2n} - f_n\|_\infty \geq \frac{1}{2}$, mostrando que a sucessão não é de Cauchy.

Podemos agora considerar, se I é um intervalo limitado de $\mathbb{R}$, o espaço vectorial formada pelas funções reais limitadas definidas em I, munida da norma (verifique que assim é!) $\| \ \|_1$ definida por

$$\|f\|_1 = \int_I |f(t)|\,dt.$$

Exemplos 1.23. *Voltemos aos exemplos acima, mas agora considerando a norma* $\| \ \|_1$.

A sucessão $(x^n)_n$ *é convergente para a função nula uma vez que*

$$\int_0^1 |x^n - 0| \, dx = \int_0^1 x^n \, dx = \frac{1}{n+1} \underset{n}{\longrightarrow} 0.$$

No segundo exemplo, se a sucessão convergisse para uma função f *contínua limitada então como*

$$0 \underset{n}{\longleftarrow} \int_0^1 |f_n(x) - f| \, dx \geq \int_0^{\frac{1}{2}} |f_n(x) - f| \, dx = \int_0^{\frac{1}{2}} |f| \, dx \geq 0$$

concluímos que $\int_0^{\frac{1}{2}} |f| \, dx = 0$ *e, portanto, como* f *é contínua,* f *é nula em* $[0, \frac{1}{2}]$. *Do mesmo modo, se* $a > \frac{1}{2}$ *e* k *é tal que* $\frac{1}{2} + \frac{1}{k} < a$ *então, para* $n \geq k$

$$\int_0^1 |f_n(x) - f| \, dx \geq \int_a^1 |f_n(x) - f| \, dx = \int_a^1 |1 - f| \, dx$$

e, de modo análogo que acima, concluímos que f *é constante e igual a* 1 *em* $[a, 1]$. *Como* a *é qualquer em* $]\frac{1}{2}, 1]$ *concluímos que* $f(x) = 1$, *se* $x > \frac{1}{2}$ *e, portanto,* f *é descontínua em* $\frac{1}{2}$, *contrariando a hipótese.*

Por sua vez a sucessão $(f_n)_n$ *é de Cauchy. Para ver isso basta notar que, se* $n, m \in \mathbb{N}$ *então*

$$\|f_n - f_m\|_1 = \int_{\frac{1}{2}}^{\frac{1}{2} + \frac{1}{\min\{n,m\}}} |f_m(x) - f_n(x)| \, dx \leq \int_{\frac{1}{2}}^{\frac{1}{2} + \frac{1}{\min\{n,m\}}} 1 \, dx = \frac{1}{\min\{n, m\}}.$$

Deste modo, dado ε, *tomando* $p \in \mathbb{N}$ *tal que* $\frac{1}{p} \leq \varepsilon$, *para todo* $n, m \geq p$, $\|f_n - f_m\|_1 \frac{1}{\min\{n,m\}} \leq \frac{1}{p} \leq \varepsilon$.

Note-se que poderíamos ter calculado explicitamente $\|f_n - f_m\|_1$

Com o mesmo tipo de argumento que o usado no caso do espaço l^p, cm $p \geq 1$ podemos também definir sobre o espaço das funções reais contínuas e limitadas, definidas num intervalo I (por enquanto) a norma $\| \ \|_p$, do seguinte modo:

$$\|f\|_p = \left(\int_I |f(x)|^p \, dx \right)^{\frac{1}{p}}.$$

Nada muda se considerarmos as sucessões de funções dos exemplos acima. Apenas como curiosidade, no segundo exemplo $\|f_n.f_m\|_p = \frac{(n-m)^p}{(p+1)mn^p}$, se $n \geq m$.

Concluímos assim que $C([0, 1])$ com a norma $\| \ \|_p$ não é um espaço completo, qualquer que seja $p > 1$.

Capítulo 2

Aplicações lineares contínuas

Qualquer aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}^m$ é contínua (considerando a norma usual nestes espaços). De facto, veremos que o resultado é válido independentemente de quais forem as normas aqui definidas. O mesmo não se pode dizer se considerarmos aplicações lineares definidas em espaços vectoriais normados gerais. Vejamos alguns exemplos.

Exemplos 2.1.

1. *A função de $\Phi : C([0,1]) \to \mathbb{R}$ definida por $\Phi(f) = f(1)$ é descontínua, se considerarmos a norma $\|\ \|_p$ em $C([0,1])$, para $p \geq 1$. Para ver isto basta notar que, se $f_n(x) = x^n$ e f é a função nula, então $(f_n)_n$ converge para f e $\big(\Phi(f_n)\big)_n$ não converge para $\Phi(f)$. É claro que se considerarmos a norma $\|\ \|_\infty$ em $C([0,1])$, Φ é contínua.*

2. *A função identidade de $C([0,1])$, que vamos denotar por I, em que consideramos a norma $\|\ \|_p$, com $p \geq 1$, no domínio e a norma $\|\ \|_\infty$ no conjunto de chegada é descontínua. Para ver isso basta considerar a sucessão $(f_n)_n$ do exemplo anterior. Ou então considerar $\Phi \circ I$, em que Φ foi definido no exemplo anterior.*

3. *Consideremos $C([a,b]])$, com a norma $\|\ \|_\infty$, e $\Phi : C([a,b]) \to \mathbb{R}$ definida por $\Phi(f) = \int_a^b f(t)\,dt$. É claro que Φ é linear e é contínua uma vez que*

$$\forall f \in C([a,b]) \quad |\Phi(f)| \leq \int_a^b |f(t)|\,dt \leq \int_a^b \|f\|_\infty dt = (b-a)\|f\|_\infty.$$

4. *Sejam $p \in [1,\infty]$ e q o seu conjugado. Usando a desigualdade de Hölder, que também é válida se $p=1$ ou $p=\infty$, concluímos que, se $y = (y_n)_n \in l^q$ então a função*

$$\begin{array}{rccc} \Phi_y : & l^p & \longrightarrow & \mathbb{R} \\ & (x_n)_n & \mapsto & \sum_{n=1}^\infty x_i y_i \end{array}$$

é contínua, uma vez que, se $x = (x_n)_n \in l^p$ então $|\Phi(x)| \leq \|y\|_q \|x\|_p$.

Note-se ainda que, se $y = (y_n)_n$ e $z = (z_n)_n$ pertencem a l^q e $\Phi_y = \Phi_z$ então $y = z$. Para ver que isto acontece basta calcular Φ_y e Φ_z aplicado nos elementos $(1,0,\ldots),(0,1,0,\ldots),\ldots$.

5. *Consideremos $C^1([0,1])$ e $C([0,1])$ com a norma $\|\ \|_\infty$ e $\Phi : C^1([0,1]) \to C([0,1])$ definida por $\Phi(f) = f'$. É natural pensar que esta função não é contínua pois "se duas funções estão perto uma da outra nada nos garante que as suas derivadas também estejam perto uma da outra". Pensemos numa sucessão de funções que se aproximem (uniformemente) da função nula e que oscilem cada vez mais, ao variar o n. Para um exemplo concreto consideremos $\left(\frac{1}{n}\operatorname{sen}(nx)\right)_n$.*

6. *Veremos mais tarde (Exemplos 2.14) que, se X e Y forem espaços vectoriais normados e X tiver dimensão infinita, então existe uma aplicação linear descontínua de X em Y.*

Definição 2.2. *Dados espaços vectoriais reais X e Y denotamos por $\mathcal{L}(X, Y)$ o conjunto das aplicações lineares contínuas de X em Y.*

No caso de $X = \mathbb{R}^n$ e $Y = \mathbb{R}^m$ com as normas usuais, a continuidade é uma consequência da linearidade e este espaço é identificado com o conjunto das matrizes reais $m \times n$. Esta identificação não é "natural", uma vez que depende da base escolhida sobre os espaços.

É claro que, nas condições referidas, $\mathcal{L}(X, Y)$ é um espaço vectorial sobre $\mathbb{R}$. Veremos que podemos definir uma norma "natural" sobre $\mathcal{L}(X, Y)$, mas antes vamos ver condições necessárias e suficientes para que uma função linear seja contínua.

Teorema 2.3. *Se X e Y são espaços vectoriais normados, $T : X \to Y$ é uma aplicação linear e $a \in X$ então as seguintes condições são equivalentes:*

a) T é uniformemente contínua;

b) T é contínua;

c) T é contínua em a;

d) T é contínua em $\vec{0}$:

e) T transforma conjuntos limitados em conjuntos limitados;

f) existe $M > 0$ tal que $\|T(x)\|_Y \leq M\|x\|_X$.

Demonstração. Vamos denotar por $\| \ \ \|$ as normas em X e em Y, esperando que isso não leve a confusões. Vamos também denotar por $\vec{0}$ o vector nulo, tanto de X como de Y.

É claro que $a) \Rightarrow b) \Rightarrow c)$ e isto, sem usar a linearidade de T. A equivalência entre $c)$ e $d)$ é simples consequência da linearidade (basta escrever a definição da continuidade em a) e em b).

Vejamos que $d) \Rightarrow e)$. Seja U um conjunto limitado de X e consideremos $r > 0$ tal que $U \subseteq B(\vec{0}, r)$. Como T é contínua em $\vec{0}$ e $T(\vec{0}) = \vec{0}$ então existe $\delta > 0$ tal que $T(B(\vec{0}, \delta)) \subseteq B(\vec{0}, 1)$.

Vejamos que $T(U) \subseteq B(\vec{0}, \frac{r}{\delta})$. Se $x \in U$ então $\frac{\delta}{r}x \in B(\vec{0}, \delta)$ porque $U \subseteq B(\vec{0}, r)$ e, portanto, $T(\frac{\delta}{r}x) \in B(\vec{0}, 1)$ ou seja $\|T(\frac{\delta}{r}x)\| < 1$ que é equivalente, usando a linearidade e a definição de norma, a $\|T(x)\| < \frac{r}{\delta}$.

Para mostrar que $e) \Rightarrow f)$ comecemos por considerar, usando $e)$, $M > 0$ tal que $T(D(0, 1)) \subseteq D(0, M)$. Vejamos que, para todo $x \in X$, $\|T(x)\| \leq M\|x\|$. Se $x = \vec{0}$ a desigualdade é trivial. Se $x \neq \vec{0}$ então $\frac{x}{\|x\|} \in D(0, 1)$ e, portanto $T\big(\frac{x}{\|x\|}\big) \in D(0, M)$, ou seja, $\big\|T\big(\frac{x}{\|x\|}\big)\big\| \leq M$ que é equivalente ao que pretendemos.

Finalmente se T satisfaz f) então, para todo $x, y \in X$, $\|T(x - y)\| \leq M\|x - y\|$, ou seja, $\|T(x) - T()y)\| \leq M\|x - y\|$, o que implica a continuidade uniforme (neste passo a hipótese de T ser linear já não é relevante). $\square$

Vamos de seguida ver uma caracterização de normas equivalentes sobre um dado espaço vectorial. Diz-se que duas normas (ou distâncias) sobre um espaço vectorial (ou conjunto) X são equivalentes se as topologias associadas forem iguais. Dito de outra forma, se a função identidade de X em X é contínua com inversa contínua se considerar uma das normas (ou métricas) no domínio e a outra no conjunto de chegada. Atendendo a que a função identidade é linear obtemos o seguinte resultado, usando $f)$ do teorema anterior.

Proposição 2.4. *Duas normas $\| \ \ \|_1$ e $\| \ \ \|_2$ sobre um espaço vectorial normado X são equivalentes se e só se*

$$\exists M, N \in \mathbb{R} \quad \forall x \in X \quad \|x\|_1 \leq M\|x\|_2, \qquad \|x\|_2 \leq N\|x\|_1.$$

Note-se que M e N são necessariamente números estritamente positivos.

Podemos ter dois espaços métricos homeomorfos e um deles ser completo e o outro não. De modo equivalente, podemos ter duas métricas equivalentes num espaço e esse espaço ser completo com uma métrica e não ser com a outra. Costuma-se dizer que a completude não é uma propriedade topológica dos espaços métricos. Já nos casos dos espaços normados a situação é diferente conforme mostra o seguinte resultado que é consequência da proposição anterior.

Corolário 2.5. *Se $(X, \| \ \|_X)$ e $(Y, \| \ \|_Y)$ são espaços vectoriais normados admitindo uma aplicação linear bijectiva contínua com inversa contínua entre eles então, se um deles for completo o outro também é.*

Demonstração. Basta essencialmente notar que qualquer função linear contínua entre espaços vectoriais normados envia sucessões de Cauchy em sucessões de Cauchy atendendo à condição f) do Teorema 2.3. □

Proposição 2.6. *Se X e Y são espaços vectoriais normados então a função*

$$\begin{array}{rccc} \| \ \| : & \mathcal{L}(X,Y) & \longrightarrow & \mathbb{R} \\ & T & \mapsto & \inf \Big\{ M \in \mathbb{R} : \forall x \in X \ \|T(x)\| \leq M\|x\| \Big\} \end{array}$$

é uma norma.

Demonstração. Note-se que o ínfimo referido acima é de facto mínimo (porquê?).

Vejamos apenas a desigualdade triangular. Sejam $T_1, T_2 \in \mathcal{L}(X,Y)$. Para $x \in X$ temos

$$\|(T_1 + T_2)(x)\| = \|T_1(x) + T_2(x)\| \leq \|T_1(x)\| + \|T_2(x)\| \leq \|T_1\| \, \|x\| + \|T_2\| \, \|x\| = (\|T_1\| + \|T_2\|) \, \|x\|.$$

Por definição de $\|T_1 + T_2\|$, concluímos que $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$. □

Com argumentos semelhantes aos que foram usados acima podemos mostrar que, se $T : X \to Y$ e $S : Y \to Z$ são aplicações lineares entre espaços vectoriais normados então $\|S \circ T\| \leq \|S\| \, \|T\|$.

A partir de agora sempre que nos referirmos a $\mathcal{L}(X,Y)$ estaremos a pensar neste espaço vectorial com a norma definida acima.

Nota 2.7. *Quando tivermos uma função $T : X \to Y$, em que X e Y são espaços vectoriais normados e houver dúvidas sobre qual a norma que estamos a considerar em cada espaço escreveremos: $T : (X, \| \ \|_1) \to Y$, para dizer que estamos a considerar em X a norma $\| \ \|_1$; $T : X \to (Y, \| \ \|_2)$, para dizer que estamos a considerar em Y a norma $\| \ \|_2$; $T : (X, \| \ \|_1) \to (Y, \| \ \|_2)$, para dizer que estamos a considerar em X a norma $\| \ \|_1$ e em Y a norma $\| \ \|_2$.*

Proposição 2.8. *Se $T : X \to Y$ é uma aplicação linear entre espaços vectoriais normados então*

$$\|T\| = \sup\{\|T(x)\| : \|x\| \leq 1\} = \sup\{\|T(x)\| : \|x\| = 1\}.$$

Demonstração. Basta usar a equivalência, para $x \neq \vec{0}$, entre $\|T(x)\| \leq M\|x\|$ e $\|T\big(\frac{x}{\|x\|}\big)\| \leq M$ e a definição de supremo. □

Para o cálculo da norma de uma aplicação linear, muitas vezes usa-se o seguinte raciocínio:

- começa-se por mostrar que $\|T(x)\| \leq M\|x\|$ para todo o x, tentando ao máximo minimizar M (isto garante que $\|T\| \leq M$);
- de seguida tentamos (analisando os passos da demonstração do ponto anterior) encontrar um elementos x_0 tal que $\|T(x_0)\| = M\|x_0\|$ (isto garante que $\|T\| \geq M$).

Em geral não se consegue encontrar x_0 nas condições acima, mas há sempre a possibilidade de "encontrar uma aproximação". Isto é, conseguimos encontrar $(x_n)_n$ tais que $\lim_n \frac{\|T(x_n)\|}{\|x_n\|} = M$. Deste modo, da desigualdade (sempre verdadeira) $\|T(x_n)\| \leq \|T\| \, \|x_n\|$, ou equivalentemente, $\frac{\|T(x_n)\|}{\|x_n\|} \leq \|T\|$, obtemos, aplicando limite em n, $M \leq \|T\|$.

Exemplos 2.9. *Vejamos alguns exemplos.*

1. *Se X for um espaço vectorial com um produto interno e $y \in X$ então a função $T : X \to \mathbb{R}$ definida por $T(x) = x \cdot y$ é uma aplicação linear contínua cuja norma é $\|y\|$. Para ver isto basta usar a desigualdade de Cauchy-Schwarz e calcular $T(y)$.*

2. *A função $T : C([0,1]) \to \mathbb{R}$ definida por $T(f) = f(\frac{1}{2})$ é contínua, se considerarmos a norma $\|\ \|$ e tem norma igual a 1, uma vez que $|T(f)| \leq \|f\|_\infty$ e $|T(f_0)| \leq \|f_0\|_\infty$, se f_0 for constante e igual a $\frac{1}{2}$ (ou se f_0 não for a função nula e $\frac{1}{2}$ for um ponto de máximo global de f_0.*

3. *Sejam $a, b, c \in \mathbb{R}$ e consideremos a aplicação linear $T : \mathbb{R}^3 \to \mathbb{R}$ definida por $T(x, y, z) = ax + by + cz$. Consideremos sobre $\mathbb{R}^3$ a norma $\|\ \|_p$. Se $p = 2$ então este é um caso particular do exemplo 2. acima. Se $p = \infty$ então $\|T\| = |a| + |b| + |c|$. O que acontece para os outros valores de p? Usando a desigualdade de Hölder obtemos $\|T(x, y, z)\| \leq (|a|^q + |b|^q + |c|^q)^{\frac{1}{q}} \|(x, y, z)\|_p$, sendo q o conjugado de p. O que mostra que $\|T\| \leq (|a|^q + |b|^q + |c|^q)^{\frac{1}{q}}$.*

4. *A função Φ definida em 3 de Exemplos 2.1 tem norma $b - a$, atendendo ao que está escrito e ao facto de, se f_0 for constante e igual a 1 (por exemplo) $|\Phi(f_0)| = (b - a)\|f_0\|$.*

5. *Sejam $N \in \mathbb{N}$ e $I : \mathbb{R}^N \to \mathbb{R}^N$ a função identidade. Pelo que vimos acima na Proposição 1.20 e na Proposição 1.21, se $p, q \in [1, \infty]$, as normas $\|\ \|_p$ e $\|\ \|_q$ são equivalentes. Além disso, se considerarmos a norma $\|\ \|_p$ no domínio e a $\|\ \|_q$ no conjunto de chegada, então $\|I\| = 1$, se $p \leq q$ (basta usar a Proposição 1.20 e considerar o elemento $(1, 0, \ldots, 0)$, e $\|I\| = N^{\frac{1}{p} - \frac{1}{q}}$, se $p \geq q$ (basta usar a Proposição 1.21 e considerar o elemento $(1, 1, \ldots, 1)$.*

Definição 2.10. *Seja X um espaço vectorial real de dimensão finita e $\mathcal{B} = \langle \vec{e}_1, \ldots, \vec{e}_n \rangle$ uma sua base. Para cada $p \in [1, \infty]$ definimos a norma $\|\ \|_p$ (de facto ela depende também de $\mathcal{B}$) por $\|x\| = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$, se $x = x_1\vec{e}_1 + \cdots x_n\vec{e}_n$.*

Nas condições da definição acima é imediato ver que a função $T : (\mathbb{R}^n, \|\ \|_p) \to (X, \|\ \|_p)$ definida por $T(x_1, \ldots, x_n) = x_1\vec{e}_1 + \cdots x_n\vec{e}_n$ é uma isometria. Em particular os compactos de $(X, \|\ \|_p)$ são os seus fechados limitados.

O seguinte resultado generaliza o que é referido no primeiro exemplo acima.

Proposição 2.11. *Se X é um espaço vectorial de dimensão finita então duas quaisquer normas sobre X são equivalentes.*

Demonstração. Se fixarmos uma base $\mathcal{B} = \langle \vec{e}_1, \ldots, \vec{e}_n \rangle$ de X e consideremos a norma $\|\ \|_1$ associada a esta base. Seja $\|\ \|$ uma qualquer norma em X e vejamos que essa norma é equivalente à norma $\|\ \|_1$ definida acima.

Se $x = a_1\vec{e}_1 + \cdots a_n\vec{e}_n$ e $M = \max\{\|\vec{e}_1\|, \ldots, \|\vec{e}_1\|\}$ então

$$\|x\| = \|a_1\vec{e}_1 + \cdots a_n\vec{e}_n\| \leq |a_1|\,\|\vec{e}_1\| + \cdots |a_n|\,\|\vec{e}_n\| \leq M\,(|a_1| + \cdots + \cdots |a_n|)$$

o que mostra que a função $I : (X, \|\ \|_1) \to (X, \|\ \|)$ é contínua. Deste modo a imagem por I de $S^1 = \{x \in X : \|x\|_1\}$, que vimos ser um compacto, é um compacto porque I é contínua. Como a função $\|\ \| : (X, \|\ \|) \to \mathbb{R}$ é contínua e $I(S^1)$ é um compacto então a imagem por esta função do conjunto $I(S^1)$ tem um mínimo c que não pode ser 0 pois $\vec{0} \notin I(S^1)$. Concluímos assim que, para todo $x \in X$ tal que $\|x\|_1 = 1$, $\|x\| \geq c$.

Sendo assim, se $x \in X$ e $x \neq \vec{0}$ então $\left\|\frac{x}{\|x\|_1}\right\| \geq c$, ou seja $\|x\| \geq c\|x\|_1$. Como esta última desigualdade é também válida para $x = \vec{0}$ concluímos que $\|x\|_1 \leq \frac{1}{c}\|x\|$, concluindo a demonstração. □

Contrariamente ao que acontece para espaços métricos, se tivermos um espaço vectorial normado X com duas normas e uma sucessão em X, se ela for de Cauchy considerando uma das normas então também o é considerando a outra. Este resultado é uma consequência da Proposição 2.4. Os seguintes resultados são agora imediatos.

Corolário 2.12. *Se Y é um sub-espaço de dimensão finita de um espaço vectorial normado X então Y é completo.*

Demonstração. Y é um espaço vectorial normado considerando sobre ele a restrição da norma definida em X e também considerando a norma $\| \ \|_2$ relativamente a uma sua qualquer base. Como Y tem dimensão finita, estas duas normas são equivalentes e o resultado segue do que é dito acima, antes do enunciado deste corolário. □

Proposição 2.13. *Se $T : X \to Y$ é uma aplicação linear entre dois espaços vectoriais normados e X tem dimensão finita então T é contínua.*

Demonstração. De acordo com a Definição 2.10 consideremos uma norma $\| \ \|_1$ relativa a uma base $\langle \vec{e}_1, \ldots, \vec{e}_n \rangle$ de $\mathcal{B}$. Deste modo, se $x = x_1\vec{e}_1 + \cdots, x_n\vec{e}_n$ e $M = \max\{\|T(\vec{e}_1), \ldots, T(\vec{e}_n)\}$ então

$$\begin{aligned}\|T(x)\| &= \|T(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n)\| \leq \|x_1T(\vec{e}_1) + \cdots + x_nT(\vec{e}_n)\| \\ &\leq |x_1|\,\|T(\vec{e}_1)\| + \cdots + |x_n|\,\|T(\vec{e}_n)\| \leq M(|x_1| + \cdots + |x_n|) = M\|x\|_1.\end{aligned}$$

A conclusão segue agora da proposição anterior que diz que a norma previamente definida em X é equivalente à norma $\| \ \|_1$. □

As propriedades referidas nestas duas últimas proposições relativa a espaços vectoriais de dimensão finita são falsas se estivermos com um qualquer espaço vectorial de dimensão infinita.

Exemplos 2.14. *Seja X um qualquer espaço vectorial de dimensão infinita e $\mathcal{B}$ uma sua base (garantida pelo Lema de Zorn, que é equivalente ao Axioma da Escolha).*

1. *Vejamos que X admite duas normas não equivalentes.*

   Uma vez que cada elemento $x \in X$ se escreve de maneira única como uma soma $\boxed{\boxed{\textit{finita}}}$ *de elementos de $\mathcal{B}$, tem sentido definir a norma $\| \ \|_1$ e a norma $\| \ \|_\infty$ em X (verifique!). Deste modo é simples verificar que $\|x\|_\infty \leq \|x\|_1$, qualquer que seja $x \in X$. Por outro lado, não existe M tal que $\|x\|_1 \leq M\|x\|_\infty$, para todo $x \in X$. Para ver isto basta considerar, $n \in \mathbb{N}$ com $n > M$ e x um elemento de X que se escreve como uma soma de n elementos de $\mathcal{B}$. Neste caso $\|x\|_1 > M\|x\|_M$, pois $\|x\|_1 = n$ e $\|x\|_\infty = 1$.*

2. *Vejamos agora que, se $\| \ \|$ for uma norma sobre X e Y um espaço vectorial normado então existe $T : X \to Y$ linear e descontínua.*

   Recorde-se que definir T equivale a definir a imagem dos elementos de uma base de X. Consideremos um subconjunto numerável $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n, \ldots, \}$ de elementos de $\mathcal{B}$ e y um elemento qualquer não nula de Y. Deste modo a função

$$\begin{array}{ccc}\mathcal{B} & \longrightarrow & Y \\ \vec{e} & \mapsto & \begin{cases} n\|\vec{e}\|\ y & \textit{se existe } n \in \mathbb{N} \textit{ tal } \vec{e} = \vec{e}_n \\ \vec{0} & \textit{caso contrário}\end{cases}\end{array}$$

   pode ser prolongada a uma função linear $T : X \to Y$.

   T é descontínua pois não existe $M > 0$ tal que $|T(x)| \leq M\|x\|$, para todo $x \in X$. Se isto acontecesse então $|T(\vec{e}_n)| \leq M\|\vec{e}_n\|$, para todo $n \in \mathbb{N}$ ou equivalentemente $n\|\vec{e}_n\|\,\|y\| \leq M\|\vec{e}_n\|$, o que é absurdo.

Com a intenção de mostrar que espaço normado de dimensão infinita os discos nunca são compactos, vamos ver um resultado preliminar.

Lema 2.15 (Riesz)**.** *Sejam X um espaço normado, Y, Z sub-espaços de X com $Y \subseteq Z$, $Y \neq Z$ e Y fechado. Então*

$$\forall \theta \in]0,1[\ \exists z \in Z : \ \|z\| = 1, \quad \inf_{y \in Y} \|z - y\| \geq \theta.$$

Demonstração. Começamos por considerar um qualquer $v \in Z \setminus Y$ e seja $a = \inf_{y\in Y} \|v - y\|$. Note-se que, como Y é fechado e $v \notin Y$ então $a > 0$.

Por definição de ínfimo, como $\frac{a}{\theta} > a$ existe $y_0 \in Y$ tal que $a \leq \|v - y_0\| \leq \frac{a}{\theta}$. Resta agora verificar que $z = \frac{v-y_0}{\|v-y_0\|}$ satisfaz as condições pretendidas. É claro que $\|Z\| = 1$. Por outro lado, se $y \in Y$ então

$$\|z - y\| = \left\| \frac{v - y_0}{\|v - y_0\|} - y \right\| = \frac{1}{\|v - y_0\|} \big\|v - y_0 - \|v - y_0\|y\big\| = \frac{1}{\|v - y_0\|} \big\|v - (y_0 + \|v - y_0\|y)\big\|.$$

Uma vez que $y_0 + \|v - y_0\|y \in Y$ (porque Y é um sub-espaço vectorial de X), $\big\|y - (y_0 + \|v - y_0\|y)\big\| \geq a$ e, portanto $\|z - y\| \geq \frac{a}{\|v-y_0\|} \geq \theta$. □

Não é relevante para o assunto, mas este resultado para o espaço vectorial $\mathbb{R}^n$, com a norma euclidiana, é uma consequência de resultados bem conhecidos de álgebra linear e geometria analítica. De facto, neste caso, o resultado vale para $\theta = 1$.

Corolário 2.16. *Se X é um espaço vectorial normado de dimensão infinita , $a \in X$ e $r > 0$ então $S(a, r)$ e qualquer subconjunto de X que o contenha não é compacto.*

Demonstração. Como a função $T : X \to X$ definida por $T(x) = \frac{1}{r}(x - a)$ é contínua com inversa contínua e envia $S(a, r)$ em $S(\vec{0}, 1)$ basta-nos considerar o vaso em que $a = \vec{0}$ e $r = 1$.

Vamos definir uma sucessão $(x_n)_n$ em $S(\vec{0}, 1)$ que mostraremos não admitir sub-sucessão convergente. Para cada $n \in \mathbb{N}$ chamamos Y_n ao sub-espaço de X gerado por $\{x_1, \ldots, x_n\}$. Estes espaços sendo de dimensão finita são necessariamente fechados e são diferentes de X, que tem dimensão infinita. A definição da sucessão $(x_n)_n$ é feita por recorrência do seguinte modo:

- x_1 é um qualquer elemento de X de norma 1;
- estando definidos $x_1, \ldots, x_n$ definimos x_{n+1} como um qualquer elemento de X, dado pelo lema anterior, tal que $\|x_{n+1}\| = 1$ e $\inf_{y\in Y_n} \|x_{n+1} - y\| \geq \frac{1}{2}$ (o que implica que $\|x_{n+1} - x_k\| \geq \frac{1}{2}$, se $k \leq n$).

Note-se que a sucessão $(x_n)_n$ pertence a $S(\vec{0}, 1)$. Do modo como foi definida a sucessão $(x_n)_n$, qualquer dos seus elementos está à distância pelo menos $\frac{1}{2}$ de todos os outros. Em particular a sucessão $(x_n)_n$ não admite nenhuma sub-sucessão convergente, não apenas em $S(\vec{0}, 1)$ mas em X. □

Nota 2.17. *Em qualquer espaço topológico um subconjunto fechado de um espaço compacto é compacto. Em particular, se um disco num espaço métrico for compacto então a esfera correspondente também é. Sem usarmos o corolário acima é simples verificar que, se uma esfera num qualquer espaço vectorial normado é compacta então o disco correspondente também é. De facto, se $a \in X$ e $r > 0$ e $S(a, r)$ é compacto a função*

$$\begin{array}{rccc} T: & [0,1] \times S(a,r) & \longrightarrow & D(a,r) \\ & (t,x) & \mapsto & t(x-a)+a \end{array}$$

é contínua e sobrejectiva. Como $[0, 1] \times S(a, r)$ é um compacto, a sua imagem por T, que é $D(a, r)$ também é compacto.

Nota 2.18. *Para vermos a "raridade" dos conjuntos compactos em espaços vectoriais normados note-se que qualquer subconjunto de um espaço vectorial normado com interior não vazio nunca é compacto, uma vez que ele conterá sempre um disco.*

Capítulo 3

Espaços duais

3.1 Introdução

Consideremos $\mathbb{R}^n$ com uma qualquer norma. Sabemos que, se $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ então $T_a : \mathbb{R}^n \to \mathbb{R}$ definida por $T_a(x_1, \ldots, x_n) = a_1x_1 + \cdots + a_nx_n$, se $x = (x_1, \ldots, x_n)$, é linear e que toda a aplicação linear de $\mathbb{R}^n$ em $\mathbb{R}$ é definida desta forma. É também claro que, se $a \neq b$ então $T_a \neq T_b$ (basta calcular ambas as funções nos vectores da base canónica). Sabemos também estas funções são todas contínuas, independentemente da norma que se considera em $\mathbb{R}^n$, uma vez que $\mathbb{R}^n$ é um espaço vectorial normado de dimensão finita.

Fica assim definida uma função bijectiva linear contínua e com inversa contínua (verifique!)

$$\begin{array}{cccc} T: & \mathbb{R}^n & \longrightarrow & \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \\ & a & \mapsto & T_a \end{array}$$

Recorde-se que, se a norma em $\mathbb{R}^n$ for a norma euclidiana então T é uma isometria (ver Exemplos 2.9).

O mesmo se pode fazer com qualquer espaço vectorial X de dimensão finita. Fixamos uma base $\mathcal{B}$, escrevemos os elementos de X nesta base e tudo funciona como se estivéssemos em $\mathbb{R}^n$. Assim X é identificado linear e topologicamente com $\mathcal{L}(X, \mathbb{R})$. Note-se que esta identificação varia com a escolha de uma base em X. Repetindo o processo, os espaços X, $\mathcal{L}(X, \mathbb{R})$, $\mathcal{L}(\mathcal{L}(X, \mathbb{R}), \mathbb{R})$ estão identificado.

3.2 Dual de um espaço vectorial normado

Se X é um espaço vectorial normado chama-se **espaço dual de** X e denota-se por X', ao espaço $\mathcal{L}(X, \mathbb{R})$.

Recorda-se que em X' a considerar a norma definida na Proposição 2.6. A notação usual para a norma de um elemento $f \in X$ é $\|f\|_{X'}$, mas será usado $\|f\|$ se esta notação não levantar dúvidas.

Contrariamente ao que acontece, no caso em que X tem dimensão finita, no caso geral, não existe um isomorfismo linear contínuo com inversa contínua entre X e X'. De facto, isso acontece sempre que X não é completo.

Teorema 3.1. *Se X e Y são espaços vectoriais normados e Y é completo então $\mathcal{L}(X, Y)$ é completo. Em particular X' é um espaço completo.*

Demonstração. Consideremos uma sucessão de Cauchy $(T_n)_n$ em $\mathcal{L}(X, Y)$ e mostremos que a sucessão é convergente.

Fixemos $x \in X$ e vejamos que a sucessão $(T_n(x))_n$ é de Cauchy em Y. Uma vez que, por definição de norma de uma aplicação linear,

$$\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \, \|x\|$$

concluímos que a sucessão $(T_n(x))_n$ é uma sucessão de Cauchy em Y.

Como Y é completo, então a sucessão $(T_n(x))_n$ é convergente para um elemento a que vamos chamar $T(x)$. Fica assim definida uma função $T : X \to Y$ por $T(x) = \lim_n T_n(x)$, que, naturalmente é candidata para limite da sucessão $(T_n)_n$. Atendendo à propriedade do limite de sucessões em Y é claro que T é linear.

Como a sucessão $(T_n)_n$ é de Cauchy então

$$\forall \varepsilon \; \exists p \in \mathbb{N} \; \forall n, m \geq p \; \|T_n - T_m\| \leq \varepsilon.$$

Fixando $\varepsilon > 0$ e usando a definição de norma de uma aplicação linear, existe $p \in \mathbb{N}$ tal que

$$\forall n, m \geq p \; \forall x \in X \; \|(T_n - T_m)(x)\| \leq \varepsilon \|x\|.$$

Como a função norma é contínua, então para todo $x \in X$, n, m como acima,

$$\|(T - T_m)(x)\| = \lim_n \|(T_n - T_m)(x)\| \leq \varepsilon \|x\|,$$

o que mostra que $T - T_m$ é contínua e, portanto, $T = (T - T_m) + T_m$ é contínua.

Voltando acima mostramos que,

$$\forall \varepsilon \; \exists p \in \mathbb{N} \; \forall m \geq p \; \|T - T_m\| \leq \varepsilon$$

ou seja, a sucessão $(T_m)_m$ converge para T em $\mathcal{L}(X, Y)$. □

3.2.1 O dual de l^p

Vamos estudar um caso particular importante. Seja $p \in [1, \infty]$ e consideremos a função já definida em 4) de Exemplos 2.1, $\Phi : l^{p'} \longrightarrow (l^p)'$ tal que, se $y = (y_n)_n \in l^{p'}$, $\Phi(y) : \begin{array}{ccc} l^p & \longrightarrow & \mathbb{R}. \\ (x_n)_n & \mapsto & \sum_{n=1}^{\infty} x_i y_i \end{array}$

Vimos que Φ é linear injectiva e que $|\Phi(y)(x)| \leq \|y\|_{p'} \|x\|_p$, para todo $y \in l^{p'}$ e $x \in l^p$. Isto mostra que $\|\Phi(y)\| \leq \|y\|_{p'}$ e, portanto, Φ é contínua e $\|\Phi\| \leq 1$.

Vamos agora ver que Φ é uma isometria. Começamos por introduzir a notação de $\operatorname{sgn}(a)$ (sinal de a) como sendo: 1, se $a > 0$; -1, se $a < 0$ e 0 se $a = 0$. Os casos em que $p = 1$ e $p = \infty$ serão vistos em separado. Consideremos $y = (y_n)_n$ um elemento não nulo de $l^{p'}$ e mostremos que $\|\Phi(y)\| \geq \|y\|_{p'}$.

- Suponhamos que $p \in]1, \infty[$. Seja $x = (x_n)_n$ tal que $x_n = \operatorname{sgn}(y_n) |y_n|^{\frac{p'}{p}}$. Note-se que $x \in l^p$ porque

  $$\sum_{n=1}^{\infty} |x_n|^p = \sum_{n=1}^{\infty} |y_n|^{p'}.$$

  Daqui obtemos $\|\Phi(y)\| \geq \|y\|_{p'}$ porque $\|x\|_p = (\|y\|_p)^{\frac{p'}{p}}$ e

  $$|\Phi_y(x)| = \sum_{n=1}^{\infty} |y_n|^{\frac{p'}{p}} |y_n| = \sum_{n=1}^{\infty} |y_n|^{1 + \frac{p'}{p}} = \sum_{n=1}^{\infty} |y_n|^{p'} = \|y\|_{p'}^{p'} = \|y\|_{p'} \|x\|_p.$$

- Suponhamos que $p = \infty$. Seja $x = (x_n)_n$ tal que $x_n = \operatorname{sgn}(y_n)$. Note-se que $x \in l^\infty$, $\|x\|_\infty = 1$ e

  $$|\Phi_y(x)| = \sum_{n=1}^{\infty} |y_n| = \|y\|_1 \|x\|_\infty$$

  Fica assim mostrado que $\|\Phi(y)\| \geq \|y\|_1$.

- Suponhamos que $p = 1$. Para $k \in \mathbb{N}$ seja $\vec{e}_k$ a sucessão cujos termos são todos iguais a 0 excepto o de ordem k que é igual a 1. Note-se que $\vec{e}_k \in l^1$, $\|\vec{e}_k\| = 1$ e $|\Phi_y(\vec{e}_k)| = |y_k| = |y_k| \, \|\vec{e}_k\|_1$. Fica assim mostrado que $\|\Phi(y)\| \geq |y_k|$, qualquer que seja $k \in \mathbb{N}$. Deste modo $\|\Phi(y)\| \geq \|y\|_\infty$.

Conclusão: Para todo $p \in [1, \infty]$, Φ é uma isometria.

Juntando toda a informação que temos sobre Φ, podemos desde já dizer que $l^{p'}$ é isometricamente isomorfo a um sub-espaço vectorial normado de l^p.

Nota 3.2. *No que foi feito acima, no caso em que $p = \infty$ poderíamos ter considerado que o conjunto de chegada de Φ era c_0'. Recorde-se que c_0 é o sub-espaço de l^∞ formado pelas sucessões convergentes para 0 (com a norma do supremo). Na prova teria de haver uma pequena alteração porque a sucessão $x = (sgn(y_n))_n$ não pertencem em geral a c_0. De qualquer maneira podemos considerar, para $N \in \mathbb{N}$, $x^N = (x_n^N)_n$ em que x_n^N é igual a $sgn(y_n)$, se $n \leq N$ e igual a 0 caso contrário. Daqui concluiríamos que $\|\Phi_y\| \geq \sum_{n=1}^N |y_n|$ para todo N e, portanto $\|\Phi(y)\| \geq \|y\|_1$.*

Teorema 3.3. *Se $p \in [1, \infty[$ então a função $\Phi : l^{p'} \longrightarrow (l^p)'$ definida acima é uma isometria bijectiva e, portanto, $(l^p)' = l^{p'}$ [1]. Por outro lado l^1 é identificado com c_0', o dual do espaço das sucessões reais convergentes para 0.*

Demonstração. Suponhamos que $p \in [1, \infty[$. Resta-nos mostrar que Φ é uma função sobrejectiva. Seja $f \in (l^p)'$ e consideremos $y = (y_n)_n$ onde $y_n = f(\vec{e}_n)$. Para $N \in \mathbb{N}$ sejam

$$y^N = \sum_{n=1}^{N} y_n \vec{e}_n, \quad x^N = \sum_{n=1}^{N} \text{sgn}(y_n)|y_n|^{p'-1}\vec{e}_n.$$

Note-se que

$$f(x^N) = \sum_{n=1}^{N} |y_n|^{p'} \quad f(x^N) \leq \|f\| \, \|x^N\|_p = \|f\| \left(\sum_{n=1}^{N} |y_n|^{p'} \right)^{\frac{1}{p}}.$$

Daqui obtemos,

$$\left(\sum_{n=1}^{N} |y_n|^{p'} \right)^{\frac{1}{p'}} \leq \|f\|.$$

Deste modo, $\left(\sum_{n=1}^{\infty} |y_n|^{p'}\right)^{\frac{1}{p'}} \leq \|f\|$ e, portanto, $y = (y_n)_n \in l^{p'}$.

Para concluir resta mostrar que, para todo $x = (x_n)_n \in l^p$, $f(x) = \sum_{n=1} x_n y_n$. Como esta igualdade é válida para qualquer sucessão quase nula, este conjunto de sucessões é denso em l^p e f é contínua então a igualdade vale em l^p.

Suponhamos agora que $p = \infty$. Note-se que o raciocínio acima não funciona neste caso porque o conjunto das sucessões quase nulas não é denso em l^∞. Mas, de facto, este conjunto é denso em c_0. Deste modo tudo funciona bem, se o conjunto de chegada da função Φ, quando $p = \infty$ fosse c_0' (ver Nota 3.2). ☐

3.3 Teorema de Hahn -Banach

É sempre fácil prolongar uma aplicação linear definida num sub-espaço F de um espaço vectorial E. Para isso basta considerar uma qualquer base de E que contenha uma base de F e definir a imagem dos elementos de $E \setminus F$ como entendermos.

A situação muda de figura se considerarmos estivermos a trabalhar com espaços normados de dimensão infinita, pois nesse caso teremos de garantir a continuidade das funções.

No que segue vamos exigir sobre as funções a linearidade e uma espécie de generalização de continuidade.

Definição 3.4. *Seja X um espaço vectorial e $p : X \to \mathbb{R}$. Diz-se que p é um funcional sublinear se:*

a) $p(x + y) \leq p(x) + p(y)$ para todo $x, y \in X$ (sub-aditividade);

b) $p(ax) = ap(x)$ para todo $x \in X$ e $a \in \mathbb{R}_0^+$ (homogenidade positiva).

[1] Estamos a abusar da notação. De facto identificamos os dois espaços.

É claro que uma norma ou um múltiplo de uma norma é um funcional sublinear. No que se segue vamos considerar funções que estão "abaixo" de funções sublineares. Por exemplo, se $f \in X'$ então existe $M > 0$ tal que $f(x) \leq M\|x\|$, para todo $x \in X$.

A prova do teorema seguinte usa o Lema de Zorn, que é equivalente ao Axioma da Escola. Este Lema já foi usado implicitamente quando foi dito que todos os espaços vectoriais admitem uma base.

Lema de Zorn. *Se $(\mathcal{P}, \leq)$ é um conjunto parcialmente ordenado cujas cadeias têm majorante então $(\mathcal{P}, \leq)$ admite elementos maximais.*

Teorema 3.5 (Hahn-Banach). *Seja X um espaço vectorial normado, $p : X \to \mathbb{R}$ uma função sublinear. Se Y é um sub-espaço de X e $f_0 : Y \to \mathbb{R}$ é linear e tal que $f_0(x) \leq p(x)$ para todo $x \in Y$, então existe $f : X \to \mathbb{R}$ que é um prolongamento de f_0 tal que $f(x) \leq p(x)$ para todo $x \in X$.*

Demonstração. Com a intenção de usar o Lema de Zorn consideramos $\mathcal{P}$ o conjunto formado por todos os prolongamentos g de f_0 satisfazendo a condição $g(x) \leq p(x)$, para todo x do domínio de g. Definimos um ordem parcial sobre $\mathcal{P}$ do seguinte modo: se $g, h \in \mathcal{P}$ dizemos que $g \leq h$ se h é um prolongamento de g.

Consideremos uma cadeia $(g_i)_{i\in I}$ de elemento de $\mathcal{P}$ e, para cada $i \in I$ seja Y_i o domínio de g_i. Então a função $h : \bigcup_{i\in I} Y_i \to \mathbb{R}$ por $h(x) = g_j(x)$ se $x \in Y_j$, está bem definida (porque $(g_i)_{i\in I}$ é uma cadeia) e pertence a $\mathcal{P}$.

Aplicando o Lema de Zorn, sabemos que existe um elemento maximal em $\mathcal{P}$. Seja $g : Y \to \mathbb{R}$ esse elemento maximal e vejamos que $Y = X$.

Suponhamos que $Y \neq X$ e consideremos $x_0 \in X \setminus Y$ e Z o sub-espaço de X gerado por $Y \cup \{x_0\}$. Note-se que Z é formado pelos elementos da forma $x = y + ax_0$, com $y \in Y$ e $a \in \mathbb{R}$. Além disso esta escrita é única. Deste modo, para $c \in \mathbb{R}$ a função $f_c : Z \longrightarrow \mathbb{R}$ tal que, nas condições acima, $f_c(x) = g(x) + ca$ está bem definida, prolonga g e é linear (verifique!). Vejamos que existe $c \in \mathbb{R}$ tal que $f_c \in \mathcal{P}$, ou seja, tal que $f_c(x) \leq p(x)$, para todo $x \in Z$. É claro que só precisamos de mostrar a desigualdade se $a \neq 0$.

Temos

$$f_c(x) \leq p(x) \Leftrightarrow g(y) + ca \leq p(y + ax_0) \Leftrightarrow \begin{cases} g(\frac{y}{a}) + c \leq p(x_0 + \frac{y}{a}) & \text{se } a > 0 \\ g(-\frac{y}{a}) - c \leq p(-x_0 - \frac{y}{a}) & \text{se } a < 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} c \leq p(x_0 + \frac{y}{a}) - g(\frac{y}{a}) & \text{se } a > 0 \\ c \geq -p(-x_0 - \frac{y}{a}) + g(-\frac{y}{a}) & \text{se } a < 0 \end{cases}$$

Sejam

$$d = \sup_{y_1 \in Y} -p(-x_0 + y_1) + g(y_1), \qquad e = \inf_{y_2 \in Y} p(x_0 + y_2) - g(y_2)$$

Vamos agora mostrar que $d \leq e$ ou equivalentemente que

$$\forall y_1, y_2 \in Y \quad -p(-x_0 + y_1) + g(y_1) \leq p(x_0 + y_2) - g(y_2).$$

Esta desigualdade é verdadeira porque, usando a linearidade de g, para todo $y_1, y_2 \in Y$

$$g(y_1) + g(y_2) = g(y_1 + y_2) \leq p(y_1 + y_2) = p(y_1 - x_0 + y_2 + x_0) \leq p(-x_0 + y_1) + p(x_0 + y_2).$$

Deste modo qualquer $c \in [d, e]$ satisfaz as condições pretendidas, isto é $f_c \in \mathcal{P}$ e prolonga g, o que é absurdo pois g é um elemento maximal de $\mathcal{P}$. $\square$

Como consequência deste teorema podemos provar a existência de elementos do dual de um espaço vectorial normado com certas condições. Em particular veremos que todo o um espaço vectorial normado está isometricamente mergulhado no seu bidual (dual do dual).

Nota 3.6. *Seja X um espaço vectorial normado. No que segue iremos utilizar funções sublineares definidas por $p(x) = M\|x\|$, para todo $x \in X$, em que M é um número real fixo. Note-se que neste caso, se $f : X \to \mathbb{R}$ é linear então dizer que $f(x) \leq p(x)$, para todo $x \in X$ é o mesmo que dizer que $|f(x)| \leq p(x)$, para todo $x \in X$. Isto acontece porque, se $x \in X$, $f(-x) = -f(x)$ e $p(-x) = p(x)$.*

Se tivermos um espaço vectorial normado X, um seu sub-espaço Y e $f \in X'$ então é claro que a restrição de f a Y pertence a Y' e tem norma menor ou igual à norma de f. No que segue vamos prolongar funções sem alterar a "sua" norma.

Corolário 3.7. *Seja X um espaço vectorial normado e Y um seu sub-espaço. Se $f \in Y'$ então f pode ser prolongada a um elemento $g \in X'$ tal que $\|g\|_{X'} = \|f\|_{Y'}$.*

Demonstração. Seja $f \in Y'$. Como a função $p : X \to \mathbb{R}$ definida por $p(x) = \|f\|_{Y'}\|x\|$, se $x \in X$, é sublinear e $|f(x)| \leq p(x)$, para todo $x \in Y$ então o teorema de Hahn-Banach garante que existe um prolongamento de f a uma função linear $g : X \to \mathbb{R}$ tal que $|g(x)| \leq \|f\|_{Y'}\|x\|$ para todo $x \in X$. Isto mostra que $g \in X'$ e que $\|g\|_{X'} \leq \|f\|_{Y'}$. □

Proposição 3.8. *Se X é um espaço vectorial normado, Y um sub-espaço próprio e fechado de X, $z \in X \setminus Y$ e δ a distância de z a Y, então existe $f \in X'$ com norma igual a 1 e tal que $f(z) = \delta$ e $f(y) = 0$, para todo $y \in Y$.*

Demonstração. Consideramos Z o espaço gerado por $Y \cup \{z\}$, que é igual a $\{y + \lambda z : \lambda \in \mathbb{R}\}$ e definimos $g : Z \to \mathbb{R}$ por que $g(y + \lambda z) = \lambda\delta$. É claro que g é linear, g restrita a Y é a função nula e que $g(z) = \delta$. Vejamos que g é contínua. Se $\lambda \neq 0$,

$$\|y + \lambda z\| = \| - \lambda(-\tfrac{y}{\lambda} - z)\| = |\lambda|\| - \tfrac{y}{\lambda} - z| \geq |\lambda|\delta = |g(y + \lambda z)|,$$

o que mostra que g é contínua e que $\|g\| \leq 1$.

Por outro lado, se $(y_n)_n$ for uma sucessão de elementos de Y tais que $\lim_n \|z - y_n\| = \delta$ e $z_n = z - y_n$ então $|g(z_n)| = \delta = \frac{\delta}{\|z_n\|}\|z_n\|$, o que mostra que $\|g\| = 1$.

Resta agora prolongar g usando o Corolário 3.7. □

É claro que, nesta proposição, tudo funcionaria mesmo que Y não fosse fechado, desde que z não pertencesse à aderência de Y.

Corolário 3.9. *Se X é um espaço vectorial normado e $z \in X \setminus \{0\}$ então existe $f \in X'$ com norma igual a 1 e tal que $f(z) = \|z\|$.*

Demonstração. Basta considerar na proposição anterior $Y = \{0\}$. □

Nota 3.10. *Como consequência do corolário anterior, se $x, z \in X$ existe sempre $f \in X'$ que os separa, isto é, tal que $f(x) \neq f(z)$. Para ver isto, basta aplicar o corolário anterior a $y = x - z$.*

3.4 Espaços reflexivos

Definição 3.11. *Seja X um espaço vectorial normado. Chamamos bidual de X e denotamos por X'' ao dual do dual de X.*

Veremos que, em certas condições, X e X'' são identificados de um modo "natural".

Definição 3.12. *Seja X um espaço vectorial normado. Chamamos homomorfismo canónico entre X e X'' à função*

$$J_X : X \longrightarrow X'' \tag{3.1}$$

em que, se $x \in X$, $J(x) : X' \to \mathbb{R}$ e $J_X(x)(f) = f(x)$, para todo $f \in X'$.

Não havendo dúvidas sobre o espaço que se está a considerar escreveremos apenas J em vez de J_X para indicar o homomorfismo canónico de X e X''.

Proposição 3.13. *Nas condições da definição acima, J é uma isometria linear.*

Demonstração. Vejamos apenas a prova de que J preserva a norma. Se x é um elemento não nulo de X então, como $|J(x)(f)| = |f(x)| \leq \|x\| \, \|f\|_{X'}$, para todo $f \in X'$ e, portanto, $\|J(x)\|_{X''} \leq \|x\|$. Para mostrar a outra desigualdade, consideramos, usando o Corolário 3.9, $f \in X'$ tal que $\|f\|_{X'} = 1$ e $f(x) = \|x\|$. Assim, $|J(x)(f)| = |f(x)| = \|x\| \, \|f\|_{X'}$, concluindo a demonstração. □

Como J é uma isometria então é injectiva e, portanto, J identifica X com um sub-espaço de X''. Se X tiver dimensão finita então X'' tem a mesma dimensão de X e, portanto, J identifica X com X''.

Definição 3.14. *Um espaço vectorial normado X diz-se reflexivo se a função $J : X \to X''$, definida acima, for bijectiva.*

Há aqui uma pequena observação a fazer: há exemplos (muito complicados) de espaços vectoriais normados X não reflexivos embora exista uma isometria entre X e X''.

Recorde-se que, se considerarmos o espaço l^p, com $p \in]1, \infty[$ então

$$(l^p)'' = \left((l^p)'\right)' = \left(l^{p'}\right)' = l^{p''} = l^p. \tag{3.2}$$

Analisando os passos acima é simples verificar que a identificação referida é a que é dada por J. Ou seja, $J : l^p \to (l^p)''$ é bijectiva, se $p \in]1, \infty[$.

Vejamos alguns consequências simples das definições:

- se E é reflexivo então E é completo (ver Teorema 3.1);
- como $(c_0)'' = (l^1)' = l^\infty \neq c_0$, então c_0 não é um espaço reflexivo, embora seja completo e, se l^1 fosse reflexivo então o dual de l^∞ seria l^1 e l^∞ seria reflexivo (verifique!).

Veremos mais tarde que qualquer espaço vectorial X com um produto interno e completo é reflexivo. Nestes casos acontece o mesmo que em $\mathbb{R}^n$ com o produto interno usual: toda a aplicação linear $f : X \to \mathbb{R}$ é definida por $f(x) = z \cdot x$, para algum $z \in X$. Esta propriedade está na base da prova de que estes espaços são reflexivos.

Teorema 3.15. *Seja X um espaço vectorial normado completo. Então X é reflexivo se e só se X' é reflexivo.*

Demonstração. Vamos denotar por J e $\tilde{J}$ os homomorfismo canónicos de X em X' e de X' em X''', respectivamente. Deste modo, se $f \in X$, $\tilde{J}(f) : X'' \to \mathbb{R}$ é definido por $\tilde{J}(f)(\phi) = \phi(f)$, se $\phi \in X''$.

Suponhamos que X é reflexivo e vejamos X', ou seja, que $\tilde{J}$ é sobrejectivo. Seja $L \in X'''$. Neste caso, se $f \in X'$ então

$$\begin{aligned}
\tilde{J}(f) = L &\Leftrightarrow \forall \phi \in X'' \quad \tilde{J}(f)(\phi) = L(\phi) \\
&\Leftrightarrow \forall \phi \in X'' \quad \phi(f) = L(\phi) \\
&\Leftrightarrow \forall x \in X \quad J(x)(f) = L(J(x)) \quad \text{porque, como } J \text{ é sobrejectiva, } X'' = \{J(x) : x \in X\} \\
&\Leftrightarrow \forall x \in X \quad f(x) = (L \circ J)(x).
\end{aligned}$$

Concluímos assim que, para todo $L \in X'''$ existe $f = L \circ J$ tal que $\tilde{J}(f) = L$, o que mostra que $\tilde{J}$ é sobrejectivo, ou seja, que X' é reflexivo.

Suponhamos agora que X' é reflexivo (ou seja, $\tilde{J}$ é sobrejectivo) e vejamos que X é reflexivo (ou seja, J é sobrejectivo). Se J não fosse sobrejectivo então $J(X)$ era um sub-espaço próprio de X'' que é fechado porque X é completo.

Pela Proposição 3.8 existe L um elemento X''' com norma 1 que é nulo em $J(X)$. Como X' é reflexivo existe $f \in X'$ tal que $\tilde{J}(f) = L$. Como L é nulo em $J(X)$ então

$$\forall x \in X \quad 0 = L(J(x)) = \tilde{J}(f)(J(x)) = J(x)(f) = f(x),$$

o que é absurdo pois, se f for a função nula (elemento nulo de X') então $L = \tilde{J}(f)$ é o elemento nulo de X''' e, portanto, tem norma igual a 0. □

3.5 Topologia fraca

Seja X um espaço vectorial normado. Vamos definir uma topologia sobre X que é mais grossa (menos fina) que a topologia associada à norma em X. A principal propriedade que esta topologia tem (em termos de aplicações) é que os discos de um espaço reflexivo são compactos para esta topologia. De facto no lugar do disco podemos colocar qualquer subconjunto de X que seja fechado (para a norma), limitado e convexo. Em particular qualquer sucessão limitada em X admite uma sub-sucessão convergente nesta nova topologia.

Recorde-se a noção de topologia inicial para um conjunto de funções definidas num mesmo conjunto.

Definição 3.16. *Sejam X um conjunto não vazio, $(Y_i)_{\in I}$ uma família de espaços topológicos e $(f_i)_{i\in I}$ uma família de funções tais que, para cada $i \in I$, $f_i : X \to Y_i$.*

Chama-se topologia inicial em X relativamente a $(f_i)_i$ à menor topologia sobre X tal que f_i é contínua para todo $i \in I$.

O seguinte resultado é apenas consequência da definição de topologia.

Proposição 3.17. *Nas condições da definição acima, se $\mathcal{T}$ é a topologia fraca referida então*

$$\mathcal{B} = \left\{ \bigcap_{i\in J} f_j^{-1}(W_j) : J \subseteq I, J \text{ é finito, } W_j \text{ é um aberto de } Y_j, \text{ para } j \in J \right\},$$

é uma base de $\mathcal{T}$, isto é, $\mathcal{T}$ é formado pelas uniões (quaisquer) de elementos de $\mathcal{B}$.

Demonstração. É claro que os elementos de $\mathcal{B}$ são necessários para garantir a continuidade da família de funções $(f_i)_i$. Para a segunda parte basta notar que a intersecção finita de uniões de elementos de $\mathcal{B}$ é ainda uma união de elementos de $\mathcal{B}$. □

A caracterização das sucessões fracamente convergentes pode ser feita de modo simples.

Proposição 3.18. *Nas condições acima, se $x \in X$ e $(x_n)_n$ for uma sucessão em X então a sucessão converge para x, relativamente à topologia fraca, se e só se, para todo $i \in I$, a sucessão $(f_i(x_n))_n$ converge para $f_i(x)$.*

Demonstração. A prova da condição necessária usa apenas o facto de uma função contínua transformar sucessões convergentes em sucessões convergentes.

Suponhamos agora que, para todo $i \in I$, $(f_i(x_n))_n$ converge para $f_i(x)$ e vejamos que $(x_n)_n$ converge para x. Para isso basta mostrar que (com as notações acima) que, para qualquer elemento $V \in \mathcal{B}$, com $x \in V$,

$$\exists p \in \mathbb{N}\ \forall n \geq p \quad x_n \in V.$$

Sejam $J \subseteq I$ finito e, para $j \in J$, W_j aberto de Y_j tal que $V = \bigcap_{j\in J} f_j^{-1}(W_j)$. Como, para $j \in J$, $f_j(x) \in W_j$ e $(f_j(x_n))_n$ converge para $f_j(x)$, então

$$\exists p_i \in \mathbb{N}\ \forall n \geq p_i \quad x_n \in f_j^{-1}(W_j).$$

Deste modo, se $p = \max_j p_j$ e $n \geq p$, $x_n \in \bigcap_j f_j^{-1}(W_j) = V$. □

A partir daqui, vamos sempre considerar que X é um espaço vectorial normado e considerar a topologia fraca sobre X relativa a todos os elementos de X'. Esta topologia costuma-se denotar por $\sigma(X, X')$ para evidenciar qual o espaço e quais as funções que estamos a considerar.

À topologia da norma chamamos topologia forte sobre X. É claro, das definições, que a topologia forte é mais fina do que a topologia fraca. Por outro lado as topologias são, em geral, diferentes (ver Exemplo 3.21). De facto, pode-se mostrar, que a topologia fraca em espaços vectoriais normados de dimensão infinita não é metrizável e, portanto, é diferente da topologia da forte. Já no caso dos espaços vectoriais de dimensão finita a situação é diferente.

Proposição 3.19. *Se X é um espaço vectorial normado de dimensão finita então as topologias, sobre X, forte e fraca coincidem.*

Demonstração. Atendendo à Proposição 2.10 podemos considerar que a norma em X é a norma $\| \ \|_\infty$ (ver Definição 2.11) relativamente a uma dada base $\mathcal{B}$ de X. Suponhamos que a dimensão de X é igual a n.

Basta-nos provar que qualquer bola contem um aberto da topologia fraca. Seja $y = (y_1, \dots, y_n)_\mathcal{B} \in X$ e $r > 0$. Consideremos, para $i = 1, \dots, n$ a função $f_i : X \to \mathbb{R}$ que a cada $x = (x_1, \dots, x_n)_\mathcal{B}$ faz corresponder x_i. Note-se que, se $x \in X$, $\|x\|_\infty = \max_{i=1,\dots,n} |f_i(x)|$. Deste modo

$$
\begin{aligned}
B(y, r) = \{x \in X : \|x - y\|_\infty < r\} &= \{x \in X : \max_{i=1,\dots,n} |x_i - y_i| < r\} = \bigcap_{i=1}^n \{x \in X : |x_i - y_i| < r\} \\
&= \bigcap_{i=1}^n \left\{x \in X : x_i \in \left]y_i - r, y_i + r\right[\right\} = \bigcap_{i=1}^n \left\{x \in X : f_i(x) \in \left]y_i - r, y_i + r\right[\right\} \\
&= \bigcap_{i=1}^n f_i^{-1}\left(\left]y_i - r, y_i + r\right[\right).
\end{aligned}
$$

Deste modo $B(y, r)$ é um aberto para a topologia $\sigma(X, X')$. □

Proposição 3.20. *Se X é um espaço vectorial normado então a topologia $\sigma(X, X')$ é Hausdorff.*

Demonstração. Sejam y e z dois elementos diferentes de X e consideremos $f \in X'$ tal que $f(y) \neq f(z)$ (ver Nota 3.10). Seja α um número real entre $f(y)$ e $f(z)$. Então $f^{-1}(]-\infty, \alpha[)$ e $f^{-1}(]\alpha, \infty[)$ são dois abertos de X para a topologia fraca que separam y de z. □

Dada uma sucessão $(x_n)_n$ em X e $x \in X$ escreveremos:

- $x_n \underset{n}{\longrightarrow} x$ ou $x_n \longrightarrow x$ para significar que a sucessão converge forte para x;
- $x_n \underset{n}{\rightharpoonup} x$ ou $x_n \rightharpoonup x$ para significar que a sucessão converge fraco para x.

É claro das definições que a convergência forte (de sucessões) implica a sua convergência fraca. Recorde que a topologia forte, sendo metrizável, é definida pelas suas sucessões convergentes. O mesmo pode não ser válido relativamente à topologia fraca. De facto, em l^1 as sucessões fortemente convergentes e as fracamente convergentes são as mesmas. Vejamos um exemplo em l^p, em que as duas noções não são equivalentes.

Exemplo 3.21. *Seja $p \in]1, \infty[$ e consideremos a sucessão $(\vec{e}_n)_n$ de elementos de l^p (recorde-se que $\vec{e}_n$ é a sucessão cujos termos são todos nulos excepto o de ordem n que é igual a 1).*

Note-se que a sucessão $(\vec{e}_n)_n$ não é (fortemente) convergente pois $\|\vec{e}_n - \vec{e}_m\|_p = 1$, se $n \neq m$. Por outro lado se $f \in (l^p)'$ então, atendendo ao Teorema 3.3, existe $y = (y_1, y_2, \dots) \in l^{p'}$ tal que $f(x) = \sum_i x_i y_i$ para todo $x = (x_1, x_2, \dots) \in l^p$. Em particular $f(\vec{e}_n) = y_n$, o que mostra que a sucessão $\big(f(\vec{e}_n)\big)_n$ converge para 0, pois a série $\sum_n |y_n|^{p'}$ é convergente. Concluímos assim que a sucessão $(\vec{e}_n)_n$ converge fracamente em X.

Proposição 3.22. *Seja X um espaço vectorial normado completo, $(x_n)_n$ uma sucessão em X, $(f_n)_n$ uma sucessão em X', $x \in X$ e $f \in X'$. Então:*

a) se $x_n \underset{n}{\rightharpoonup} x$ então $(\|x_n\|)_n$ é uma sucessão limitada e $\|x\| \leq \liminf_n \|x_n\|$;

b) se $x_n \underset{n}{\rightharpoonup} x$ e $f_n \underset{n}{\longrightarrow} f$ então $f_n(x_n) \underset{n}{\longrightarrow} f(x)$.

Demonstração. A primeira parte alínea a) não será demonstrada. Para a segunda parte se aplicarmos $\liminf_n$ à desigualdade $|f(x_n)| \leq \|f\| \, \|x_n\|$ obtemos $|f(x)| \leq \|f\| \liminf_n \|x_n\|$ ou seja, $|J(x)(f)| \leq \liminf_n \|x_n\| \, \|f\|$, que diz que $\|J(x)\|_{X''} \leq \liminf_n \|x_n\|$ e a conclusão segue porque J é uma isometria.

A segunda alínea é uma consequência das desigualdades

$$|f_n(x_n) - f(x)| \leq |(f_n - f)(x_n)| + |f(x_n - x)| \leq \|f_n - f\| \|x_n\| + |f(x_n - x)|,$$

da definição de convergência fraca e da limitação de $(\|x_n\|)_n$. □

Capítulo 4

Espaços de Hilbert

Na Secção 1.3 foi dada a definição de produto interno sobre um espaço vectorial e algumas das suas propriedades: desigualdade de Cauchy-Schwarz, e a igualdade (regra) do paralelogramo. Como exemplo típico temos $\mathbb{R}^n$, com o produto interno que induz a norma usual. Do mesmo modo, a norma em l^2 provem de um produto interno. Podemos também ver que, se X for um espaço vectorial de qualquer dimensão maior do que 1 e $\mathcal{B}$ for uma sua base então a norma $\| \quad \|_p$ associada a esta base não é induzida por um produto interno se $p \neq 2$, uma vez que não satisfaz a regra do paralelogramo. Para mostrar isto basta considerar x e y dois quaisquer elementos diferentes de $\mathcal{B}$ e notar que

$$\|x+y\|^2 = \|x-y\|^2 = 2^{\frac{2}{p}}, \qquad \|x\|^2 = \|y\|^2 = 1.$$

A regra do paralelogramo caracteriza de facto o produto interno.

Proposição 4.1. *Se X for um espaço vectorial normado com uma norma $\| \quad \|$ que satisfaz a regra do paralelogramo então a função definida em $X \times X$ por*

$$\forall x, y \in X \quad x \cdot y = \tfrac{1}{4}\left(\|x+y\|^2 - \|x-y\|^2\right)$$

define um produto interno, cuja norma associada é a norma original do espaço.

Demonstração. As 3 primeiras condições referidas na Definição 1.8 são fáceis de verificar. Resta-nos ver que, fixado $z \in X$ a função $F : X \to \mathbb{R}$ definida por $F(x) = x \cdot z$ $\left(= \frac{1}{4}\left(\|x+z\|^2 - \|x-z\|^2\right)\right)$ é linear.

Sejam $x, y \in X$ e mostremos que $F(x+y) = F(x) + F(y)$. Pela regra do paralelogramo temos, se $w \in X$,

$$\|x+y+w\|^2 = 2\|x+w\|^2 + 2\|y\|^2 - \|x-y+w\|^2, \quad \|x+y+w\|^2 = 2\|y+w\|^2 + 2\|x\|^2 - \|y-x+w\|^2.$$

Somando as duas igualdades (e dividindo por 2) obtemos

$$\|x+y+w\|^2 = \|x\|^2 + \|y\|^2 + \|x+w\|^2 + \|y+w\|^2 - \tfrac{1}{2}\|x-y+w\|^2 - \tfrac{1}{2}\|y-x+w\|^2.$$

Substituindo w por z e depois por $-z$ e subtraindo, ficamos com

$$\|x+y+z\|^2 - \|x+y+z\|^2 = \|x+z\|^2 - \|x-z\|^2 + \|y+z\|^2 - \|y-z\|^2,$$

que é equivalente a $F(x+y) = F(x) + F(y)$.

Vejamos agora que $F(\lambda x) = \lambda F(x)$ se $\lambda \in \mathbb{R}$ e $x \in X$. Fazendo $x = y = 0$, $y = -x$ e $y = x$ na igualdade que acabamos de mostrar, obtemos $F(0) = 0$, $F(-x) = -F(x)$ e $F(2x) = 2\,F(x)$ e, por indução, $F(nx) = n\,F(x)$ e $F(-nx) = -n\,F(x)$, para $n \in \mathbb{N}$. Se $\lambda = \frac{p}{q}$, com $p \in \mathbb{Z}$ e $q \in \mathbb{N}$ então, usando o que foi feito,

$$F(\tfrac{p}{q}\,x) = \tfrac{p}{q}\,F(x) \iff q\,F(\tfrac{p}{q}\,x) = q\,\tfrac{p}{q}\,F(x) \iff F(q\,\tfrac{p}{q}\,x) = p\,F(x) \iff F(p\,x) = p\,F(x)$$

e esta última afirmação é verdadeira.

Se $\lambda \in \mathbb{R}$ consideremos uma sucessão $(\lambda_n)_n$ em $\mathbb{Q}$ convergindo para λ. Deste modo

$$F(\lambda\,x) = F(\lim_n \lambda_n\,x) = \lim_n F(\lambda_n x) = \lim_n \lambda_n F(x) = \lambda\,F(x).$$

Na segunda igualdade usamos a continuidade de F (verifique!) e na terceira, usamos o ponto anterior. $\square$

Definição 4.2. *A um espaço vectorial com um produto interno chamamos espaço pré-hilbertiano. Se o espaço for completo dizemos que ele é um espaço de Hilbert.*

Proposição 4.3. *Se H é um espaço pré-hilbertiano então a função produto interno é uma função contínua.*

Demonstração. Como, para $x, y, a, b \in H$ temos, usando a desigualdade triangular e a desigualdade de Cauchy-Schwarz,

$$|(x \cdot y) - (a \cdot b)| = |(x - a) \cdot y + a \cdot (y - b)| \leq |(x - a) \cdot y| + |a \cdot (y - b)| \leq \|x - a\| \, \|y\| + \|a\| \, \|y - b\|.$$

de onde resulta a continuidade da função no ponto (a, b). □

4.1 Teorema da projecção

Seja X um espaço vectorial, A um subconjunto não vazio de X e $x \in X$. Será que existe um ponto de A que está à distância mínima de x? E, caso exista, esse ponto será único?

É claro que, se A não for fechado então a resposta à primeira pergunta é não. Basta considerar $x \in \overline{A} \setminus A$. Vamos então assumir que A é fechado. É muito simples encontrar exemplos que mostrem que a distância de x a A pode ser atingida em muitos pontos, mesmo numa infinidade deles, ou até em todos: se X é qualquer, $x \in X$ e $A = S_d(x, r)$ então todos os pontos de A estão à mesma distância de x.

Antes de dar um exemplo mais geral, vamos recordar que um conjunto A contido num espaço vectorial X é convexo se qualquer segmento de recta unindo dois pontos de A está contido em A, ou seja,

$$\forall x, y \in A \quad \forall \alpha \in [0, 1] \quad \alpha x + (1 - \alpha) y \in A.$$

Note-se que todo o sub-espaço de um espaço vectorial é convexo.

Nota 4.4. *Consideremos um subconjunto A de $\mathbb{R}^2$ que seja fechado e um ponto $x \in X \setminus A$. Parece ser óbvio que existe a circunferência centrada em x de menor raio que intersecta A (a prova vem a seguir). Esses pontos são os pontos de A que estão à distância mínima de x. Se o conjunto A não for convexo, parece ser possível escolher x tal que a circunferência referida intersecta A em mais do que um ponto.*

Vamos enunciar um resultado que usa apenas conhecimentos de topologia em $\mathbb{R}^n$.

Lema 4.5. *Consideremos sobre $\mathbb{R}^n$ a métrica usual e sejam A um fechado de $\mathbb{R}^n$ e $x_0 \in \mathbb{R}^n$. Então existe $a \in A$ tal que $d(x_0, A) = d(x_0, a)$.*

Demonstração. Consideremos $r > 0$ tal que $D(x_0, r) \cap A \neq \emptyset$. Então é claro que $d(x, A) = d(x, D(x, r) \cap A)$. O que mudou agora foi que o conjunto $D(x, r) \cap A$ é compacto e, portanto, a função que a cada $x \in \mathbb{R}^n$ associa $d(x, x_0)$, sendo contínua, envia $D(x_0, r) \cap A$ num compacto de $\mathbb{R}$. Esse compacto tem necessariamente um mínimo. □

O que está em causa neste resultado não é A ser fechado mas sim o facto de a intersecção de A com um disco centrado em x_0 ser compacta.

Sendo assim não e de admirar que este resultado não seja válido em espaços vectoriais de dimensão infinita, onde os discos de raio positivo nunca são compactos (ver Corolário 2.16). Basta considerar (por exemplo) o espaço l^p, com $1 \leq p \leq \infty$, $x = \vec{0}$ e $A = \{\frac{n+1}{n} \vec{e}_n : n \in \mathbb{N}\}$. Podemos ver que A é fechado (porque os elementos de A estão a uma distância maior do que 1 uns dos outros), $d(x, A) = 1$ e $d(x, a) > 1$, para todo $a \in A$.

Estamos agora em condições de enunciar e demonstrar o teorema da projecção para espaços de Hilbert.

Teorema 4.6 (da projecção de Hilbert). *Sejam H um espaço de Hilbert, K um subconjunto não vazio de H que é convexo e fechado. Nestas condições, para qualquer $x \in H$ existe um e um só $a \in K$ tal que*

$$\|x - a\| = \min_{z \in K} \|x - z\| \quad (= d(x, K)).$$

Se K for um sub-espaço de H, então a é o único elemento de K tal que $x - a$ é perpendicular a K.

Demonstração. Sejam $\delta = d(x, K)$ e, usando a definição de $d(x, K)$, $(a_n)_n$ uma sucessão de elementos de K tais que $\|x - a_n\| \leq \delta + \frac{1}{n}$, para $n \in \mathbb{N}$. Vejamos que a sucessão $(a_n)_n$ é de Cauchy.

Para cada $n, m \in \mathbb{N}$, aplicando a regra do paralelogramo, temos

$$\|(a_n - x) - (a_m - x)\|^2 + \|(a_n - x) + (a_m - x)\|^2 = 2\|a_n - x\|^2 + 2\|a_m - x\|^2$$

ou seja

$$\|a_n - a_m\|^2 + 4\|\tfrac{1}{2}(a_n + a_m) - x\|^2 = 2\|a_n - x\|^2 + 2\|a_m - x\|^2.$$

Como K é convexo, $\frac{1}{2}(a_n + a_m) \in K$ e portanto $\|\frac{1}{2}(a_n + a_m) - x\| \geq \delta$. Substituindo acima, obtemos

$$\|a_n - a_m\|^2 \leq -4\delta^2 + 2\|a_n - x\|^2 + 2\|a_m - x\|^2 \leq -4\delta^2 + 2(\delta + \tfrac{1}{n})^2 + 2(\delta + \tfrac{1}{m})^2 = \tfrac{4\delta}{n} + \tfrac{4\delta}{m} + \tfrac{2}{n^2} + \tfrac{2}{m^2}$$

o que mostra que a sucessão $(a_n)_n$ é de Cauchy. Como H é completo e K é fechado, a sucessão converge para um elemento $a \in K$. Como a função norma é contínua concluímos que $\|x - a\| = \|x - \lim_n a_n\| = \lim_n \|x - a_n\| = \delta$.

Suponhamos agora que $a, b \in K$ são tais que $\|x - a\| = \|x - b\| = d(x, K)$. Usando a regra do paralelogramo, como acima, mas agora com a no lugar de a_n e b no lugar de a_m temos, usando o mesmo tipo de raciocínio que acima,

$$\|a - b\|^2 \leq -4\delta^2 + 2\|a - x\|^2 + 2\|b - x\|^2 = -4\delta^2 + 2\delta^2 + 2\delta^2 = 0,$$

o que mostra que $a = b$. $\square$

Nas condições do teorema acima, vamos denotar por $p_K(x)$, projecção de x em K, como o único elemento de K tal que $d(x, K) = \|x - a\|$.

A projecção de um elemento sobre um convexo fechado tem uma outra caracterização. A ideia é que se $x \in K$ então qualquer triângulo de vértices x, $p_K(x)$ e z, com $z \in K$ é obtuso em $p_K(x)$. Vamos usar uma forma generalizada do teorema de Pitágoras.

Proposição 4.7. *Se H é um espaço de Hilbert, K é um convexo fechado não vazio de H e $x \in H$. Nestas condições, se $y \in K$ as seguintes condições são equivalentes:*

a) $y = p_K(x)$;

b) $(x - y) \cdot (z - y) \leq 0$, *para todo* $z \in K$.

Demonstração. Dado $z \in K$ vamos "pensar" no triângulo de vértices x, y, z que queremos que seja obtuso em y. Temos

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 + 2(x - y) \cdot (y - z).$$

Na hipótese b), $2(x - y) \cdot (y - z) \geq 0$ e, portanto, $\|x - z\|^2 \geq \|x - y\|^2$, de onde concluímos a).

Suponhamos agora a hipótese a) e usemos a igualdade acima para $w = tz + (1 - t)y$, com $t \in]0, 1]$ que pertence a K, pois K é convexo. Obtemos

$$\|x - y\|^2 \leq \|x - w\|^2 = \|x - y\|^2 + \|y - w\|^2 + 2(x - y) \cdot (y - w),$$

ou seja, porque $y - w = t(y - z)$,

$$\|x - y\|^2 \leq \|x - y\|^2 + t^2\|y - z\|^2 + 2t(x - y) \cdot (y - z).$$

Deste modo $t^2\|y - z\|^2 + 2t(x - y) \cdot (y - z) \geq 0$, ou seja, $t\|y - z\|^2 + 2(x - y) \cdot (y - z) \geq 0$. Aplicando limite, quando $t \to 0$, obtemos o pretendido. $\square$

No caso em que o convexo é um sub-espaço vectorial de H podemos ir um pouco mais além.

Corolário 4.8. *Se H é um espaço de Hilbert, M é um sub-espaço vectorial fechado de H e $x \in H$. Nestas condições, se $y \in M$, as seguintes condições são equivalentes:*

a) $y = p_M(x)$;

b) $(x - y) \cdot z = 0$, *para todo* $z \in M$

Além disso, a função $p_M : H \to M$ é linear e com norma 1.

Demonstração. Basta mostrar que as condições b) na proposição anterior e neste corolário são equivalentes, se o convexo for um sub-espaço vectorial de H. Assumindo a condição b) deste corolário então, como $z - y \in M$, obtemos $(x - y) \cdot (z - y) = 0$. Inversamente, se assumirmos a condição b) da proposição então, dado $z \in M$, $z+y$ e $-z+y$ pertencem a M e, portanto, $(x-y)\cdot((z+y)-y) \leq 0$ e $(x-y)\cdot((-z+y)-y) \leq 0$, ou seja $(x - y) \cdot z = 0$.

Para provar que a função p_M é linear basta mostrar, usando o que acabamos de provar, que, se $x, y \in H$ e $\alpha, \beta \in \mathbb{R}$ então $\big((\alpha x + \beta y) - (\alpha p_M(x) + \beta p_M(y))\big) \cdot z = 0$ para todo $z \in M$. Nestas condições, usando a caracterização de $p_M(x)$ e de $p_M(y)$ temos

$$\big((\alpha x + \beta y) - (\alpha p_M(x) + \beta p_M(y))\big) \cdot z = \alpha(x - p_M(x)) \cdot z + \beta(y - p_M(y)) \cdot z = 0,$$

concluindo a demonstração. □

4.2 Complementar ortogonal

Vamos introduzir alguma notação. Se H é um espaço pré-hilbertiano, $x, y \in H$ e $A, B \subseteq H$ diremos que:

- $x, y \in H$ são (mutuamente) ortogonais (perpendiculares) se $x \cdot y = 0$ (notação $x \perp y$);
- x é ortogonal a A se $x \perp a$, para todo $a \in A$ (notação $x \perp A$);
- A e B são ortogonais se $a \perp b$, para todo $a \in A$ e $b \in B$ (notação $A \perp B$).

Escreveremos ainda $A^\perp$, o ortogonal de A, para designar o conjunto $\{x \in H : x \perp A\}$. Se $A = \{a\}$, escreveremos $a^\perp$ em vez de $\{a\}^\perp$.

Lema 4.9. *Se H é um espaço pré-hilbertiano e $A \subseteq H$ então $A^\perp$ é um sub-espaço vectorial fechado de H. Além disso, se $\langle A \rangle$ for o sub-espaço gerado por A então $\langle A \rangle^\perp = A^\perp$.*

Demonstração. $A^\perp$ é um sub-espaço de H, pois, se $x, y \in A^\perp$ e $\alpha, \beta \in \mathbb{R}$ então, para todo $a \in A$, $(\alpha\, x + \beta\, y) \cdot a = \alpha(x \cdot a) + \beta(y \cdot a) = 0$. Para ver que $A^\perp$ é um fechado, note-se que $A^\perp = \bigcap_{a \in A} a^\perp$ e que $a^\perp$ é um fechado pois é igual a $T_a^{-1}(0)$, sendo $T_a : H \to \mathbb{R}$ definida por $T_a(x) = a \cdot x$, para $x \in H$ (Ver Exemplo 2.9).

É claro que, $\langle A \rangle^\perp \subseteq A^\perp$. Por outro lado, se $x \in A^\perp$ então x é ortogonal a qualquer combinação linear de elementos de A, ou seja, $x \in \langle A \rangle^\perp$. Fica assim mostrado que $\langle A \rangle^\perp = A^\perp$. □

Atendendo à igualdade $\langle A \rangle^\perp = A^\perp$, basta-nos considerar o caso em que A é um sub-espaço. Note-se que, se A for um sub-espaço de H então $A \subseteq \left(A^\perp\right)^\perp$ mas a igualdade falha, pelo menos, sempre A não for fechado.

É claro que, se considerarmos $\mathbb{R}^n$ com o produto interno usual, então, usando apenas álgebra linear,podemos mostrar que, se A é um sub-espaço de $\mathbb{R}^n$ então

$$A = \left(A^\perp\right)^\perp, \qquad \mathbb{R}^n = A \oplus A^\perp.$$

Vamos ver que estas igualdades também valem em espaços de Hilbert, desde que os sub-espaços considerados sejam fechados.

Teorema 4.10 (complementar ortogonal)**.** *Se H é um espaço de Hilbert e M um sub-espaço fechado de H então H é a soma directa de M com $M^\perp$.*

Demonstração. Se $x \in M \cap M^\perp$ então, em particular, $x \cdot x = 0$, ou seja $x = 0$. Por outro lado se $x \in H$ então $x = (x - p_M(x)) + p_M(x)$ e $p_M(x) \in M$ e $x - p_M(x) \in M^\perp$ porque, se $z \in M$ então $(x - p_M(x)) \cdot z = 0$, pelo Corolário 4.8. □

Nota 4.11. *É claro que, nas condições do teorema anterior, $p_M^2 = p_M$ e p_M restrita a $M^\perp$ é constante e igual a 0. Para além disso, se $x = y + z$ com $y \in M$ e $z \in M^\perp$ então $\|x\| = \sqrt{\|y\|^2 + \|z\|^2}$. Isto acontece porque $\|y+z\|^2 = \|y\|^2 + \|z\|^2 + 2y \cdot z$.*

Corolário 4.12. *Se H é um espaço de Hilbert e M é um sub-espaço vectorial então:*

a) *M é denso se e só se $M^\perp = \{0\}$;*

b) *$M^{\perp\perp} = \overline{M}$.*

Demonstração. Recorde-se que $M^\perp = \overline{M}^\perp$. Para provar a alínea a) basta usar a igualdade $H = \overline{M} \oplus M^\perp$.

Para a alínea b), consideremos $N = \overline{M}$ e mostremos que $N^{\perp\perp} = N$. De facto resta mostrar que $N^{\perp\perp} \subseteq N$. Como $N^{\perp\perp}$ é um sub-espaço vectorial fechado, então é um espaço de Hilbert. Desse modo, pelo Teorema 4.10, $N^{\perp\perp} = N \oplus P$, em que $P = \{x \in N^{\perp\perp} : x \in N^\perp\} = N^{\perp\perp} \cap N^\perp = \{0\}$. □

4.3 Dualidade em espaços de Hilbert

Se H é um espaço de Hilbert definimos $R_H : H \to H'$ tal que, se $z \in H$, $R_H(z) : H \to \mathbb{R}$ é tal que $R(z)(x) = x \cdot z$, se $x \in H$.

Atendendo à desigualdade de Cauchy-Schwarz $|R_H(z)(x)| \leq \|z\| \, \|x\|$ e ao facto de $R_H(z)(z) = \|z\| \, \|z\|$, $\|R_H(z)\|_{H'} = \|z\|$. Deste modo, R_H é uma aplicação linear que é uma isometria (e, portanto, $\|R_H\| = 1$).

Teorema 4.13 (Representação de Riesz)**.** *Com a notação acima, R_H é um isomorfismo isométrico.*

Demonstração. Resta ver que R_H é sobrejectivo. Seja $f \in H'$ não nulo e N o núcleo de f. N é um sub-espaço fechado não nulo de H. Pelo Teorema 4.10, $H = N \oplus N^\perp$. Seja $w \in N^\perp \setminus \{0\}$ e $z = \frac{w}{\|w\|}$.

Se $x \in H$ então é simples verificar que $-f(x)z + f(z)x \in N$ e, portanto, $(-f(x)z + f(z)x) \cdot z = 0$. Daqui obtemos $f(x) = x \cdot (f(z)z$. Concluímos assim que f é igual a $R_H(f(z)z)$. □

Note-se que, se $f \in H'$ e $x \in H$ então $R^{-1}(f) \cdot x = f(x)$.

Corolário 4.14. *Se H é um espaço de Hilbert então a H' é um espaço de Hilbert.*

Demonstração. Atendendo ao teorema anterior basta transportar o produto interno definido em H para H'. Mais especificamente, a função

$$\begin{array}{rcl} \langle\ ,\ \rangle : H' \times H' & \longrightarrow & \mathbb{R} \\ (f,g) & \mapsto & R_H^{-1}(f) \cdot R_H^{-1}(g) \end{array}$$

é um produto interno que induz a norma definida em H'. □

Corolário 4.15 (Reflexividade espaços de Hilbert)**.** *Todo o espaço de Hilbert é reflexivo.*

Demonstração. Resta-nos mostrar que a função J_H (ver Definição 3.12) é sobrejectiva. Seja $\phi \in H''$. Como H' é um espaço de Hilbert então $\phi = R_{H'}(g)$ para algum $g \in H'$. Como R_H é sobrejectiva, existe $x \in H$ tal que $g = R_H(x)$. Deste modo, para $f \in H'$, temos

$$\phi(f) = R_{H'}(g)(f) = \langle g, f\rangle = R_H^{-1}(g) \cdot R_H^{-1}(f) = x \cdot R_H^{-1}(f) = f(x)$$

e, portanto, $\phi = J_H(x)$. □

Capítulo 5

Espaços de medida

5.1 Preliminares

Pensemos no conjunto $\mathbb{R}$ e na noção que temos de *medida de um intervalo* I, que é a sua amplitude e que vamos denotar por $l(I)$. O que se pretende é prolongar o máximo possível esta noção a outros subconjuntos de $\mathbb{R}$ satisfazendo certas propriedades. Por exemplo, se o conjunto for a união disjunta de intervalos queremos que a medida do conjunto seja a soma das medidas desses intervalos. Deste modo a medida de um conjunto finito deverá ser 0 e, por exemplo, a medida do conjunto $]-3,2] \cup [5,9] \cup \{12\}$ será 9. Mas há subconjuntos de $\mathbb{R}$ bem mais complicados do que estes: $\mathbb{Q} \cap [0,1]$, $[0,1] \cap \mathbb{Q}$, o conjunto de Cantor, etc..

O ideal seria definir definir $\mathcal{A} \subseteq \mathcal{P}(\mathbb{R})$, o maior possível, e uma função $\mu : \mathcal{A} \to [0,\infty]$, tal que

1. $\mu(\emptyset) = 0$;
2. $\mu([a,b]) = b - a$, se $a, b \in \mathbb{R}$ e $a \leq b$;
3. μ é invariante por translações;
4. μ é numeravelmente aditiva, isto é, se $(A_n)_{n\in\mathbb{N}}$ é uma sucessão de subconjuntos de $\mathbb{R}$ disjuntos (dois a dois) então $\mu(\cup_n A_n) = \sum_n \mu(A_n)$.

É claro que $\mathcal{A}$ teria de conter todos os intervalos, o conjunto vazio e ser fechado para a união disjunta numerável e para a translação. Uma tal função, a existir, teria as seguintes propriedades (entre outras).

- $\mu(\{a\}) = 0$, se $a \in \mathbb{R}$ (porque $\{a\} = [a,a]$).
- μ é finitamente aditiva, isto é, se $N \in \mathbb{N}$ e $A_1, \ldots, A_N$ são subconjuntos de $\mathcal{A}$ disjuntos (dois a dois) então $\mu(\cup_{n=1}^N A_n) = \sum_{n=1}^N \mu(A_n)$ (basta considerar $A_n = \emptyset$, se $n > N$).
- Se A é um conjunto contável, isto é, finito ou numerável, então $\mu(A) = 0$ (porque $A = \cup_{a\in A}\{a\}$).
- $\mu([a,b[) = \mu(]a,b[) = \mu(]a,b])$, se $a, b \in \mathbb{R}$ e $a < b$ (porque, $[a,b]$ é a união disjunta de cada um desses intervalos e de um conjunto com 1 ou 2 pontos).
- Se A é um intervalo ilimitado então $\mu(A) = \infty$ (porque A é uma união disjunta de intervalos de amplitude 1, por exemplo).
- Se $A, B, A \setminus B, A \cap B \in \mathcal{A}$ então $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$ (porque μ é aditiva).
- Se $A, B, A \setminus B, A \cap B \in \mathcal{A}$ e $\mu(A \cap B) < \infty$ então então $\mu(A \setminus B) = \mu(A) - \mu(A \cap B)$.
- Se $A, B, A \setminus B \in \mathcal{A}$ e $B \subseteq A$ então $\mu(B) \leq \mu(A)$ (A é a união disjunta de B e $A \setminus B$).
- Se $A, B, A \setminus B, B \setminus A, A \cap B \in \mathcal{A}$ e $\mu(A \Delta B) = 0$ então $\mu(A) = \mu(B)$ (porque daqui se obtém que $\mu(A \setminus B) = \mu(B \setminus A) = 0$.

De agora em diante quando escrevermos $\dot\cup$ para significar união disjunta.

Vejamos que não existe nenhuma função $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ nas condições acima.

Exemplo 5.1. *Consideremos a relação de equivalência sobre $[0, 1]$, definida por $a \sim b$ se e só se $a - b \in \mathbb{Q}$. É claro que, para cada $a \in [0, 1]$, a classe de equivalência de a é $a + \mathbb{Q} = \{a + q : q \in \mathbb{Q}\}$ e, portanto $a \sim b$ se e só se $a + \mathbb{Q} = b + \mathbb{Q}$.*

Esta relação de equivalência induz uma partição em $\mathbb{R}$. Consideremos (usando o axioma da escolha) um elemento em cada uma das classes e consideremos V o conjunto formado por esses elementos. Este conjunto diz-se um conjunto de Vitali.

Suponhamos que existe uma função μ nas condições referidas acima e consideremos $(q_n)_n$ uma sucessão formada por <u>todos</u> os elementos de $[-1, 1] \cap \mathbb{Q}$. Nestas condições os conjuntos $V_n = q_n + V$ são disjuntos dois a dois e

$$[0, 1] \subseteq \dot{\bigcup_n} V_n \subseteq [-1, 2].$$

Para ver a primeira inclusão, considere-se $x \in [0, 1]$ e $v \in V$ tal que $x \sim v$. Deste modo $x = (x - v) + v$ e $x - v \in \mathbb{Q}$ porque $x \sim v$ e $x - y \in [-1, 1]$, porque $x, y \in [0, 1]$.

Como μ é invariante por translação e numeravelmente aditiva

$$1 \leq \sum_n \mu(V_n) \leq 3 \quad \text{ou seja} \quad 1 \leq \sum_n \mu(V) \leq 3,$$

o que é impossível pois a última soma é 0 ou ∞.

O exemplo mostra que se queremos uma função μ nas condições acima teremos de restringir o domínio. De qualquer maneira esse domínio terá de ser fechado para algumas operações, como já vimos.

Notação. *Se $A \subseteq X$ vamos designar, para simplificar a escrita, por A^c, o complementar de A em X, isto é $A^c = X \setminus A$. É claro que só usaremos esta notação se não houver dúvidas sobre qual é o conjunto X.*

Chegamos assim à definição de álgebra e de σ-álgebra (de conjuntos).

Definição 5.2. *Seja X um conjunto não vazio. Uma família Σ de subconjuntos de X diz-se uma álgebra (de conjuntos) se:*

a) $\emptyset, X \in \Sigma$;

b) *é fechado para a passagem ao complementar;*

c) *é fechado para as operações de união e intersecção.*

Se, além disso, Σ for fechada para a união e intersecção numeráveis, Σ diz-se uma σ-álgebra.

De facto, se Σ satisfizer a) e b) então Σ é fechada para a união (finita ou numerável) se e só se for fechada para a intersecção (finita ou numerável). É também claro que, se $A, B \in \Sigma$ então $A \setminus B \in \Sigma$ porque $A \setminus B = A \cap B^c$. Note-se que $\mathcal{P}(X)$ e $\{\emptyset, X\}$ são σ-álgebras e que a intersecção de álgebras (ou σ-álgebras) é uma álgebra (ou σ-álgebra).

Deste modo, dado $\mathcal{A} \subseteq \mathcal{P}(X)$ existe sempre a menor álgebra (ou σ-álgebra) que contém $\mathcal{A}$.

Por exemplo, seja X um qualquer conjunto não vazio e consideremos a menor álgebra $\mathcal{A}$ e a menor σ-álgebra $\mathcal{B}$ que contem todos os conjuntos singulares de X. Se X for finito então $\mathcal{A} = \mathcal{B}$. Se X não for finito então $\mathcal{A}$ está contido propriamente em $\mathcal{B}$. De facto

- $\mathcal{A} = \{A \in \mathcal{P}(X) : A \text{ ou } A^c \text{ é finito}\}$;
- $\mathcal{B} = \{A \in \mathcal{P}(X) : A \text{ ou } A^c \text{ é finito ou numerável}\}$.

Um exemplo que será usado continuamente será a σ-álgebra gerada pelos abertos de um espaço topológico.

Definição 5.3. *Seja $(X, \mathcal{T})$ um espaço topológico. A mais pequena σ-álgebra de X que contém todos os abertos de X chama-se σ-álgebra de X de Borel e denota-se por $\mathcal{B}(X)$. Os elementos de $\mathcal{B}(X)$ dizem-se borelianos de X.*

É claro que na definição acima podemos substituir "aberto" por "fechado" que nada se altera.
Suponhamos agora que o espaço topológico é $\mathbb{R}$ com a topologia usual. Note-se que, todo o:

- aberto de $\mathbb{R}$ é uma união numerável de intervalos abertos;
- intervalo aberto é uma união numerável de intervalos fechados (por exemplo ($]a, b[= \cup_n [a + \frac{1}{n}, b - \frac{1}{n}]$);
- intervalo fechado é uma intersecção numerável de intervalos fechados à esquerda e abertos à direita e vice-versa (por exemplo ($[a, b] = \cap_n [a, b + \frac{1}{n}[$);
- intervalo fechado à esquerda e aberto à direita e vice-versa é uma intersecção numerável de intervalos abertos (por exemplo ($[a, b[= \cap_n]a + \frac{1}{n}, b[$).

Deste modo, $\mathcal{B}(\mathbb{R})$ é gerado pela família formada pelos intervalos de um destes 4 tipos. Com argumentos semelhantes podemos ver que $\mathcal{B}(\mathbb{R})$ é gerado pelos intervalos ilimitados à direita ou à esquerda e que são fechados (ou abertos). Podemos ainda supor que todos os extremos dos intervalos considerados são números racionais, obtendo assim um conjunto gerador de $\mathcal{B}(\mathbb{R})$ que é numerável. Mais geralmente o conjunto $\{B(x, \frac{1}{k}) : x \in \mathbb{Q}^n,\ k \in \mathbb{N}\}$ é numerável e gera $\mathcal{B}(\mathbb{R}^n)$.

5.2 Medida

Vamos começar por definir os espaços onde as medidas vão ser definidas.

Definição 5.4. *Chamamos espaço mensurável a um par (X, Σ) onde X é um conjunto não vazio e Σ uma σ-álgebra sobre X. Os elementos de Σ dizem-se conjuntos mensuráveis.*

É sobre estes espaços que vamos definir medidas.

Definição 5.5. *Uma medida sobre um espaço mensurável (X, Σ) é uma função $\mu : \Sigma \to [0, \infty]$ tal que:*

- *$\mu(\emptyset) = 0$;*
- *μ é numeravelmente aditiva, isto é, se $(A_n)_{n \in \mathbb{N}}$ é uma sucessão de elementos de Σ disjuntos (dois a dois) então $\mu(\cup_n A_n) = \sum_n \mu(A_n)$.*

μ diz-se completa se qualquer subconjunto de um elemento de Σ com medida 0 pertence a Σ (tendo, por consequência, medida 0). Se $\mu(X) < \infty$ a medida diz finita e uma medida de probabilidade se $\mu(X) = 1$. μ diz-se σ-finita, se X for a união de uma família numerável de conjuntos de medida finita.

Definição 5.6. *Um espaço de medida é um triplo (X, Σ, μ) onde (X, Σ) é um espaço mensurável e μ é uma medida sobre (X, Σ). Se $\mu(X) = 1$, dizemos que (X, Σ, μ) é um espaço de probabilidade.*

Vejamos quatro exemplos de espaços de medida $(X, \mathcal{P}(X), \mu)$ muito simples.

- A **medida trivial** - μ é constante e igual a 0.
- A **medida da contagem** - $\mu(A)$ é o cardinal de A, se A for finito, e igual a ∞, caso contrário. Esta medida é σ-finita se e só se X for contável.
- A **medida uniforme**, se X tem N elementos - μ é igual a $\frac{1}{N}$ vezes a medida da contagem sobre X. Esta medida de probabilidade pode ser generalizada dando pesos $p_1, \ldots, p_N$ aos elementos de X.
- Se $a \in X$, a **medida de Dirac** centrada em a, denotada por δ_a é definida por: $\delta_a(A)$ é igual a 1, se $a \in A$, e igual a 0, caso contrário.

Nota 5.7. *A aditividade numerável não pode ser substituída pela aditividade. Consideremos* $\mu : \mathcal{P}(\mathbb{N}) \to [0,\infty]$ *tal que:* $\mu(A) = 0$, *se* A *é finito;* $\mu(A) = \infty$, *caso contrário.* μ *é aditiva mas não é numeravelmente aditiva uma vez que* $\mu(\mathbb{N}) = \infty$, $\mathbb{N} = \cup_n \{n\}$ *e* $\sum_n \mu(\{n\}) = 0$.

Vejamos algumas propriedades sobre espaços de medida. Recorde-se que o limite superior/limite inferior de uma família numerável de subconjuntos de um conjunto X $(A_n)_n$ é

$$\limsup_{n\to\infty} A_n = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k, \qquad \liminf_{n\to\infty} A_n = \cup_{n=1}^{\infty} \cap_{k=n}^{\infty} A_k.$$

Teorema 5.8. *Se* (X, Σ, μ) *é um espaço de medida então:*

a) μ *é finitamente aditiva;*

b) $\mu(A) \leq \mu(B)$ *se* $A, B \in \Sigma$ *e* $A \subseteq B$*;*

c) $\mu(A) = \mu(A \setminus B) + \mu(A \cap B)$ *se* $A, B \in \Sigma$*;*

d) $\mu(A \setminus B) = \mu(A) - \mu(A \cap B)$ *se* $A, B \in \Sigma$ *e* $\boxed{\mu(A) < \infty}$*;*

e) $\mu(\cup_n A_n) \leq \sum_n \mu(A_n)$, *se* $(A_n)_n$ *for uma sucessão em* Σ*;*

f) $\mu(\cup_n A_n) = \lim_n \mu(A_n)$, *se* $(A_n)_n$ *for uma sucessão crescente em* Σ*;*

g) $\mu(\cap_n A_n) = \lim_n \mu(A_n)$, *se* $(A_n)_n$ *for uma sucessão decrescente em* Σ *e* $\boxed{\mu(A_1) < \infty}$*;*

h) $\mu(\limsup_n A_n) = 0$, *se* $(A_n)_n$ *for uma sucessão em* Σ *tal que* $\boxed{\sum_n \mu(A_n) < \infty}$.

Demonstração. As três primeiras alíneas já foram "essencialmente" provadas e a quarta é uma consequência da terceira.

e) Consideremos a sucessão $(B_n)_n$ definida por $B_0 = A_1$ e $B_n = A_n \setminus \cup_{i=1}^{n-1} A_i$. Deste modo $(B_n)_n$ é uma família de conjuntos disjuntos (dois a dois) e cuja união é $\cup_n A_n$. Sendo assim, usando a definição de medida e a alínea b), obtemos

$$\mu(\cup_n A_n) = \mu(\cup_n B_n) = \sum_n \mu(B_n) \leq \sum_n \mu(A_n).$$

f) Consideremos $(B_n)_n$ tal que $B_1 = A_1$ e $B_n = A_n \setminus A_{n-1}$ (esta família é a que foi considerada na alínea anterior, aplicada a este caso). Note-se que, para cada $k \in \mathbb{N}$, $A_k = \cup_{n=1}^{k} B_n$. Deste modo, usando a aditividade numerável na segunda igualdade e a aditividade finita na penúltima,

$$\mu(\cup_n A_n) = \mu(\cup_n B_n) = \sum_n \mu(B_n) = \lim_{k\to\infty} \sum_{n=1}^{k} \mu(B_n) = \lim_{k\to\infty} \mu(\cup_{n=1}^{k} B_n) = \lim_{k\to\infty} \mu(A_k).$$

g) Usemos a alínea anterior considerando a família $(A_1 \setminus A_n)_n$ que satisfaz as condições exigidas na alínea anterior, sendo que a sua união é $A_1 \setminus \cap_n A_n$. Deste modo

$$\mu(A_1 \setminus \cap_n A_n) = \lim_{n\to\infty} \mu(A_1 \setminus A_n),$$

e a conclusão segue da alínea d), porque $\mu(A_1)$ é finito.

h) Notando que a família $\left(\cup_{k=n}^{\infty} A_k\right)_n$ é decrescente, podemos aplicar as alíneas g) e e), uma vez que $\mu\left(\cup_{k=1}^{\infty} A_k\right) \leq \sum_k \mu(A_k) < \infty$, temos

$$\mu(\limsup_n A_n) = \mu(\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} A_k) = \lim_n \mu\left(\cup_{k=n}^{\infty} A_k\right) \leq \lim_n \sum_{k=n}^{\infty} \mu(A_k) = 0,$$

por definição de série convergente. □

Note-se que, se tivermos um espaço de medida (X, Σ, μ) pode existir uma família de conjuntos $(A_n)_n$ decrescente com medida infinita cuja intersecção tem medida finita. Deste modo $\lim_n \mu(A_n) \neq \mu(\cap_n A_n)$. Um exemplo muito simples é o dado pela medida da contagem sobre $\mathbb{R}$ (ou $\mathbb{N}$), considerando $A_n = [n, \infty[$.

5.3 Medidas exteriores

Dado um conjunto não vazio X, vamos agora definir o que se entende por uma medida exterior $\mu^* : \mathcal{P}(X) \to [0, \infty]$. Definiremos de seguida conjunto mensurável relativamente a μ^* e mostraremos que a restrição de μ^* ao conjunto formado pelos conjuntos mensuráveis é de facto uma medida completa.

Um caso particular desta construção levar-nos-á a uma medida definida numa σ-álgebra que contem os borelianos de $\mathbb{R}$ satisfazendo as 4 condições referidas na página 30, para além de ser σ-finita e completa.

Definição 5.9. *Seja X um conjunto não vazio. Uma função $\mu^* : \mathcal{P}(X) \to [0, \infty]$ diz-se uma medida exterior sobre X se:*

a) $\mu^*(\emptyset) = 0$;

b) é monótona, isto é, $\mu^*(A) \leq \mu^*(B)$, *se* $A \subseteq B$;

c) é numeravelmente sub-aditiva, isto é, se $(A_n)_n$ for uma sucessão em $\mathcal{P}(X)$ então $\mu^*(\cup_n A_n) \leq \sum_n \mu^*(A_n)$.

É claro que qualquer medida é uma medida exterior. Dado um conjunto não vazio X as seguintes funções são medidas exteriores:

$$\mu_1^* = \begin{cases} 1 & \text{se } A \neq \emptyset \\ 0 & \text{se } A = \emptyset, \end{cases} \qquad \mu_2^* = \begin{cases} 0 & \text{se } A \text{ é contável} \\ 1 & \text{caso contrário} \end{cases}$$

Além disso, μ_1^* é uma medida se e só se A for singular, e μ_2^* é uma medida se e só se X for contável (ou seja μ_2^* é a função nula).

Vamos agora definir o que se entende por um conjunto E ser μ^*-mensurável, se μ^* for uma medida exterior.

Definição 5.10. *Seja X um conjunto não vazio e μ^* uma medida exterior. Um subconjunto E de X diz-se μ^*-mensurável se*

$$\forall A \in \mathcal{P}(X) \quad \mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c).$$

Note-se que a exigência da igualdade acima é natural, uma vez que os conjuntos $A \cap E$ e $A \cap E^c$ são disjuntos e a sua união é A. Note-se que, se E e F forem μ^*-mensuráveis e disjuntos então, tomando $A = E \cup F$ na igualdade acima, obtemos $\mu^*(E \cup F) = \mu^*(E) + \mu^*(F)$, que é uma igualdade necessária.

Atendendo a que μ^* é sub-aditiva, a desigualdade $\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap E^c)$ é sempre verdadeira.

É claro que os conjuntos $\emptyset$ e X são sempre μ^*-mensurável e que, E é μ^*-mensurável se e só se E^c é μ^*-mensurável.

Notação. *Se X for um conjunto não vazio e μ^* for uma medida exterior sobre X, vamos denotar por $\mathcal{M}_{\mu^*}$ ou simplesmente por $\mathcal{M}$, se não houver dúvidas sobre a que medida exterior nos estivermos a referir, o conjunto formado por todos os subconjuntos de X que são μ^*-mensuráveis.*

Lema 5.11. *Nas condições da definição acima, seja $E \in \mathcal{P}(X)$ tal que $\mu^*(E) = 0$. Então $E, E^c \in \mathcal{M}_{\mu^*}$ e, se $F \in \mathcal{P}(X)$ então $\mu^*(E \cup F) = \mu^*(F)$.*

Demonstração. Se $A \in \mathcal{P}(X)$ então $\mu^*(A \cap E) + \mu^*(A \cap E^c) = \mu^*(A \cap E^c) \leq \mu^*(A)$ (usando a igualdade $\mu^*(\emptyset) = 0$ e (duas vezes) a monotonia de μ^*). Isto mostra que $E \in \mathcal{M}_{\mu^*}$ e, pelo que vimos, $E^c \in \mathcal{M}_{\mu^*}$. Por outro lado, $\mu^*(F) \leq \mu^*(E \cup F) \leq \mu^*(E) + \mu^*(F) = \mu^*(F)$. $\square$

Vejamos quais os conjuntos mensuráveis relativamente a μ_1^* e μ_2^* definidos acima.

No primeiro caso, apenas $\emptyset$ e X são mensuráveis. De facto se $E \subseteq X$ é μ_1^*-mensurável então $\mu_1^*(X) = \mu_1^*(X \cap E) + \mu_1^*(X \cap E^c) = \mu_1^*(E) + \mu_1^*(E^c)$, o que acontece se e só se $E = \emptyset$ ou $E^c = \emptyset$.

Relativamente a μ_2^*, seja E um conjunto contável. Como $\mu_2^*(E) = 0$ então, pela nota anterior, E é μ_2^*-mensurável, bem como os seu complementar. Por outro lado, se E for μ_2^*-mensurável então $\mu_2^*(X) = \mu_2^*(X \cap E) + \mu_2^*(X \cap E^c) = \mu_2^*(E) + \mu_2^*(E^c)$, o que implica que $\mu_2^*(E) = 0$ ou $\mu_2^*(E^c) = 0$. Isto prova que os conjuntos μ_2^*-mensuráveis são os conjuntos contáveis e os seus complementares.

O seguinte resultado corresponde à definição de conjunto μ^*-mensurável no caso em que $n = 2$ e $E_2 = E_1^c$.

Proposição 5.12. *Seja X um conjunto não vazio e μ^* uma medida exterior. Se $E_1, \dots, E_n$ é uma família finita disjunta de elementos de $\mathcal{M}_{\mu^*}$ então*

$$\forall A \in \mathcal{P}(X) \quad \mu^*\big(A \cap \bigcup_{i=1}^{n} E_i\big) = \sum_{i=1}^{n} \mu^*(A \cap E_i).$$

Demonstração. Vejamos o caso em que $n = 2$. O caso geral segue por indução.

Como $E_1 \in \mathcal{M}_{\mu^*}$ temos, dado $A \in \mathcal{P}(X)$,

$$\mu^*(A \cap (E_1 \cup E_2)) = \mu^*(A \cap (E_1 \cup E_2) \cap E_1) + \mu^*(A \cap (E_1 \cup E_2) \cap E_1^c).$$

Para concluir basta notar que $A \cap (E_1 \cup E_2) \cap E_1 = A \cap E_1$ e $A \cap (E_1 \cup E_2) \cap E_1^c = A \cap E_2$. □

Teorema 5.13. *Se X é um conjunto não vazio e μ^* é uma medida exterior então o conjunto $\mathcal{M}_{\mu^*}$ é uma σ-álgebra.*

Demonstração. Denotemos $\mathcal{M}_{\mu^*}$ por $\mathcal{M}$. Comecemos por mostrar que $\mathcal{M}$ é fechado para a união. Sejam $E, F \in \mathcal{M}$ e consideremos $A \in \mathcal{P}(X)$. Então, como $A \cap (E \cup F) = (A \cap E) \cup ((A \cap E^c) \cap F)$ e $A \cap (E \cup F)^c = (A \cap E^c) \cap F^c$ temos

$$\begin{aligned}
\mu^*(A \cap (E \cup F)) + \mu^*(A \cap (E \cup F)^c) &= \mu^*((A \cap E) \cup ((A \cap E^c) \cap F)) + \mu^*((A \cap E^c) \cap F^c) \\
&\leq \mu^*(A \cap E) + \mu^*((A \cap E^c) \cap F)) + \mu^*((A \cap E^c) \cap F^c) \\
&\quad \text{pela sub-aditividade de } \mu^* \\
&= \mu^*(A \cap E) + \mu^*(A \cap E^c) \quad \text{porque } F \text{ é } \mu^*\text{-mensurável} \\
&= \mu^*(A) \quad \text{porque } E \text{ é } \mu^*\text{-mensurável.}
\end{aligned}$$

Atendendo ao que é dito depois da definição de álgebra, podemos concluir que $\mathcal{M}$ é uma álgebra, uma vez que já vimos que $\emptyset, X \in \mathcal{M}$ e que $\mathcal{M}$ é fechado para a passagem ao complementar.

Vejamos que a união disjunta numerável de elementos de $\mathcal{M}$ pertence a $\mathcal{M}$. Seja $(E_n)_n$ uma família disjunta de elementos de $\mathcal{M}$ e $E = \cup_n E_n$. Para cada $A \in \mathcal{P}(X)$ e $k \in \mathbb{N}$, como $\cup_{n=1}^k E_n$ pertence a $\mathcal{M}$,

$$\begin{aligned}
\mu^*(A) &= \mu^*(A \cap \cup_{n=1}^k E_n) + \mu^*(A \cap \big(\cup_{n=1}^k E_n\big)^c) \\
&\geq \mu^*(A \cap \cup_{n=1}^k E_n) + \mu^*(A \cap E^c), \quad \text{porque } \mu^* \text{ é monótona} \\
&= \sum_{n=1}^{k} \mu^*(A \cap E_n) + \mu^*(A \cap E^c) \quad \text{pela Proposição 5.12.}
\end{aligned}$$

Deste modo, aplicando limite quando k tende para ∞, e recordando que μ^* é numeravelmente sub-aditiva,

$$\mu^*(A) \geq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) + \mu^*(A \cap E^c) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c) \geq \mu^*(A), \tag{5.1}$$

o que mostra que $E \in \mathcal{M}$.

Para ver que $\mathcal{M}$ é fechado para a união numerável consideremos uma família $(E_n)_n$ de elementos de $\mathcal{M}$ e notemos que, considerando $E_0 = \emptyset$

$$\cup_n E_n = \dot{\bigcup}_{n=1}^{\infty} (E_n \setminus \cup_{i=0}^{n-1} E_i) = \dot{\bigcup}_{n=1}^{\infty} \Big(E_n \cap \big(\cup_{i=0}^{n-1} E_i\big)^c\Big),$$

o que mostra que $\cup_n E_n \in \mathcal{M}$ pelo que foi provado acima. □

Estamos agora em condições de definir a medida associada a uma medida exterior.

Teorema 5.14 (Caratheodory)**.** *Se X é um conjunto não vazio e μ^* é uma medida exterior então μ^* restrita ao conjunto $\mathcal{M}_{\mu^*}$ é uma medida completa.*

Demonstração. Vejamos que μ^* restrita a $\mathcal{M}_{\mu^*}$ é numeravelmente aditiva. Seja $(E_n)_n$ uma família disjunta de elementos de $\mathcal{M}_{\mu^*}$. No teorema anterior mostramos em (5.1) que, se $A \in \mathcal{P}(X)$ então $\mu^*(A) = \sum_{n=1}^{\infty} \mu^*(A \cap E_n) + \mu^*(A \cap E^c)$. Fazendo $A = \cup_n E_n$ obtemos $\mu^*(\cup_n E_n) = \sum_{n=1}^{\infty} \mu^*(E_n)$ porque $\mu^*(\emptyset) = 0$.

O Lema 5.11 mostra que a medida é completa. □

Notação. *Nas condições acima vamos denotar por $\mathcal{N}_\mu$ o conjunto formado pelos subconjunto de X com medida (μ) nula, isto é, $\mathcal{N}_\mu = \{E \in \mathcal{M}_{\mu^*} : \mu^*(E) = 0\}$.*

Nota 5.15. *Nas condições do teorema anterior, se $\mathcal{M}_{\mu^*} \neq \mathcal{P}(X)$ então μ^* não é uma medida. Isto acontece porque se $E \in \mathcal{P}(X) \backslash \mathcal{M}_{\mu^*}$ então existe $A \in \mathcal{P}(X)$ tal que $\mu^*(A) < \mu^*(A \cap E) + \mu^*(A \cap E^c)$ de onde concluímos que μ^* não é aditiva.*

5.4 Medida de Lebesgue em $\mathbb{R}$

Vamos agora definir uma medida exterior λ^* sobre $\mathbb{R}$ tal que, se $a < b$, $\lambda^*[a,b] = l([a,b])$, a amplitude de I e λ^* é invariante por translação. Veremos que a restrição de λ^* a $\mathcal{L}(\mathbb{R})$ é uma medida que satisfaz as condições referidas no início da capítulo (página 1).

Teorema 5.16. *A função $\lambda^* : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ definida por*

$$\forall A \in \mathcal{P}(X) \quad \lambda^*(A) = \inf \Big\{ \sum_{n=1}^{\infty} l(I_n) : (I_n)_n \text{ é uma família de intervalos abertos de } \mathbb{R} \text{ tal que } A \subseteq \cup_{n=1}^{\infty} I_n \Big\}$$

é uma medida exterior sobre $\mathbb{R}$.

Demonstração. Como $\emptyset \subseteq]0, \frac{1}{n}[$, para todo $n \in \mathbb{N}$, então $\lambda^*(\emptyset) \leq \frac{1}{n}$ e portanto, $\lambda^*(\emptyset) = 0$. Por outro lado λ^* é monótona porque uma cobertura de um conjunto é também uma cobertura de qualquer seu subconjunto.

Para mostrar que λ^* é numeravelmente sub-aditiva, consideremos uma família $(A_n)_n$ de subconjunto de $\mathbb{R}$. É claro que, se $\lambda^*(A_n) = \infty$ para algum n, então $\mu(\cup_n A_n) = \infty$ pela monotonia. Suponhamos agora que $\mu(A_n)$ é finito para todo $n \in \mathbb{N}$. Fixemos $\varepsilon > 0$ e, usando a definição de ínfimo, consideremos, para cada $n \in \mathbb{N}$, $(I_k^n)_k$ uma família de intervalos abertos tais que $A_n \subseteq \cup_{k=1}^{\infty} I_k^n$ e $\sum_{k=1}^{\infty} l(I_k^n) \leq \lambda^*(A_n) + \frac{\varepsilon}{2^n}$. Daqui resulta que $\cup_{n=1}^{\infty} A_n \subseteq \cup_{n=1}^{\infty} \cup_{k=1}^{\infty} I_k^n$ e portanto, usando a definição de λ^* na primeira desigualdade,

$$\lambda^*\left(\cup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} l\big(I_k^n\big) \leq \sum_{n=1}^{\infty} \left(\lambda^*(A_n) + \tfrac{\varepsilon}{2^n}\right) = \sum_{n=1}^{\infty} \lambda^*(A_n) + \varepsilon.$$

de onde se conclui, fazendo $\varepsilon \to 0$, que $\lambda^*\left(\cup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \lambda^*(A_n)$. □

λ^* diz-se medida exterior de Lebesgue e, como referido antes, vamos denotar por $\mathcal{L}(\mathbb{R})$ o conjunto $\mathcal{M}_\lambda^*(\mathbb{R})$.

Definição 5.17. *Chama-se* **medida de Lebesgue**, *e denota-se por λ, à restrição de λ^* a $\mathcal{L}(\mathbb{R})$.*

Veremos no Corolário 5.23 que a definição de conjunto mensurável dada para a medida λ^* é maximal. Ou seja, se a restrição de λ^* a uma σ-álgebra $\mathcal{D}$ de $\mathcal{P}(\mathbb{R})$ contendo $\mathcal{B}(\mathbb{R})$ for uma medida então $\mathcal{D} \subseteq$ está contida em $\mathcal{L}(\mathbb{R})$.

Vamos de seguida tentar perceber que conjuntos pertencem a $\mathcal{L}(\mathbb{R})$. Comecemos com algumas propriedades sobre λ^*. Note-se que a notação $\mathcal{N}_\lambda$ para o conjunto $\{E \in \mathcal{L}(\mathbb{R}) : \lambda(E) = 0\}$ que é igual a $\{E \in \mathcal{P}(\mathbb{R}) : \lambda^*(E) = 0\}$, pelo Lema 5.11.

Proposição 5.18. *λ^* satisfaz as seguintes propriedades:*

a) todo o conjunto contável tem medida nula (pertence a $\mathcal{N}_\lambda$);

b) λ^* é invariante por translações;

c) $\lambda^*(I) = l(I)$, *se* I *for um intervalo de* $\mathbb{R}$.

Demonstração. Comecemos por notar que, se $a \in \mathbb{R}$ então $\{a\} \subseteq]a - \frac{1}{n}, a + \frac{1}{n}[$ e, portanto, $\lambda^*(\{a\}) \leq \frac{2}{n}$ para todo $n \in \mathbb{N}$. Daqui obtemos $\lambda^*(\{a\}) = 0$. Como λ^* é numeravelmente aditiva, $\lambda^*(A) = 0$ se A é contável.

Vejamos que λ^* é invariável por translações. Seja $A \subseteq \mathbb{R}$, $a \in \mathbb{R}$ e $(I_n)_n$ uma família subconjuntos de $\mathbb{R}$. Então $(I_n)_n$ é uma cobertura de A **se e só se** $\{a\} + I_n$ for uma cobertura de $\{a\} + A$. Para concluir basta notar que: I_n é um intervalo aberto se e só se $\{a\} + I_n$ é um intervalo aberto; uma translação de um intervalo tem a mesma amplitude que o intervalo original.

Vejamos agora que $\lambda^*([a, b]) = b - a$, se $a < b$. Como $[a, b] \subseteq]a - \varepsilon, b + \varepsilon[$, para todo $\varepsilon > 0$ então $\lambda^*([a, b]) \leq b - a - 2\varepsilon$ e, portanto, $\lambda^*([a, b]) \leq b - a$.

Para mostrar a outra desigualdade, consideremos $(I_n)_n$, uma família de intervalos abertos de $\mathbb{R}$ tal que $[a, b] \subseteq \cup_{n=1}^{\infty} I_n$, e mostremos que $b - a \leq \sum_{n=1}^{\infty} l(I_n)$. Como $[a, b]$ é compacto, existe J subconjunto finito de $\mathbb{N}$ tal que $[a, b] \subseteq \cup_{n \in J} I_n$. Deste modo, nesta família finita de intervalos, vai existir um da forma $]a_1, b_1[$ que contém a. Se $b \notin]a_1, b_1[$ então consideramos outro intervalo $]a_2, b_2[$ que contem b_1. Se $b \notin]a_2, b_2[$ então consideramos outro intervalo $]a_3, b_3[$ que contem b_2. Como temos uma família finita de abertos que contêm $[a, b]$ vai existir $k \in \mathbb{N}$ tal que $b \in]a_k, b_k[$. Deste modo, como $b_{i-1} \in]a_i, b_i[$, para $i = 1, \ldots, k$,

$$\sum_{i=1}^{\infty} l(I_n) \geq \sum_{i=1}^{k} l(]a_i, b_i[) = \sum_{i=0}^{k} (b_i - a_i) = b_k - a_1 + \sum_{i=2}^{k} (b_{i-1} - a_i) \geq b_k - a_1 \geq b - a.$$

Se I for um outro (não fechado) intervalo limitado de extremos a, b então, para todo $\varepsilon > 0$, $[a + \varepsilon, b - \varepsilon] \subseteq I \subseteq [a - \varepsilon, b + \varepsilon]$ e, portanto, $b - a - 2\varepsilon \leq \lambda^*(I) \leq b - a + 2\varepsilon$, para todo ε. Daqui concluímos que $\lambda^*(I) = b - a$.

Finalmente, se I for um intervalo ilimitado então ele contem intervalos limitados de amplitude tão grande quanto se queira e, portanto, $\lambda^*(I) = \infty$. □

Nota 5.19. *Das alíneas a) e c) da proposição anterior podemos concluir que qualquer intervalo de extremos* a *e* b, *com* $a < b$, *não é um conjunto numerável.*

Proposição 5.20. *Se* A *é um subconjunto de* $\mathbb{R}$ *então, para todo* $\varepsilon > 0$ *existe um aberto* V_ε *contendo* A *e tal que* $\lambda^*(A) \leq \lambda^*(V_\varepsilon) \leq \lambda^*(A) + \varepsilon$.

Demonstração. Basta considerar, usando a definição de $\lambda^*(A)$, V_ε a união de uma família $(I_n)_n$ de intervalos abertos tal que $\sum_{n=1}^{\infty} \lambda^*(I_n) \leq \lambda^*(A) + \varepsilon$. Note-se que $\lambda^*(V_\varepsilon) \leq \sum_{n=1}^{\infty} \lambda^*(I_n)$. □

Nota 5.21. *Apesar de* λ^* *não ser aditiva é simples verificar, usando simplesmente a definição que, se* $A, B \in \mathcal{P}(X)$ *e existir* $a \in \mathbb{R}$ *tal que* $A \subseteq [a, \infty[$ *e* $B \subseteq] - \infty, a]$ *então* $\lambda^*(A \cup B) = \lambda^*(A) + \lambda^*(B)$.

Estamos agora em condições de terminar o que foi referido na página 1.

Teorema 5.22. $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{L}(\mathbb{R})$. *Em particular a função* $\lambda : \mathcal{L}(\mathbb{R}) \to [0, \infty]$ *é uma medida* σ*-finita, invariante por translações e tal que, se* I *for um intervalo de* $\mathbb{R}$, $\lambda(I)$ *é a amplitude de* I.

Demonstração. Atendendo ao que foi dito no último parágrafo da Secção 5.1, basta provar que, para todo $a \in \mathbb{R}$ o intervalo $]a, \infty[\in \mathcal{L}(\mathbb{R})$. Seja então $a \in \mathbb{R}$, $A \in \mathcal{P}(X)$ e mostremos que, se $A_1 = A \cap]a, \infty[$ e $A_2 =] - \infty, a]$, então $\lambda^*(A) \geq \lambda^*(A_1) + \lambda^*(A_2)$. É claro que podemos supor que $\lambda^*(A)$ é finita.

Dado $\varepsilon > 0$, consideremos uma família $(I_n)_n$ de intervalos abertos tais que $A \subseteq \cup_n I_n$ e $\sum_n \lambda^*(I_n) \leq \lambda^*(A) + \varepsilon$. Deste modo

$$A = A_1 \dot{\cup} A_2, \quad A_1 \subseteq \bigcup_n (I_n \cap]a, \infty[), \quad A_2 \subseteq \bigcup_n (I_n \cap] - \infty, a])$$

e portanto, pela monotonia e a sub-aditividade de λ^*, e a Nota 5.21,

$$\lambda^*(A_1) + \lambda^*(A_2) \leq \sum_n \lambda^*(I_n \cap]a, \infty[) + \sum_n \lambda^*(I_n \cap] - \infty, a]) = \sum_n \lambda^*(I_n) \leq \lambda^*(A) + \varepsilon.$$

Fazendo agora ε tender para 0, obtemos a desigualdade pretendida.

λ é σ-finita porque (por exemplo) $\mathbb{R} = \cup_n [-n, n]$. □

Estamos agora em condições de mostrar que a definição de conjunto mensurável para λ^* foi a melhor possível.

Corolário 5.23. *Seja $\mathcal{D}$ uma σ-álgebra de $\mathcal{P}(\mathbb{R})$ contendo propriamente $\mathcal{L}(\mathbb{R})$. Nestas condições, a restrição de λ^* a $\mathcal{D}$ não é uma medida.*

Demonstração. Suponhamos que existe $E \in \mathcal{D} \setminus \mathcal{L}(\mathbb{R})$. Então, por definição de $\mathcal{L}(\mathbb{R})$ existe $A \in \mathcal{P}(X)$ e $\varepsilon > 0$ tal que

$$\lambda^*(A) = \lambda^*(A \cap E) + \lambda^*(A \cap E^c) - \varepsilon.$$

Usando a Proposição 5.20, seja U um aberto contendo A tal que $\lambda^*(U) \leq \lambda^*(A) + \frac{\varepsilon}{2}$. Deste modo

$$\lambda^*(U) \leq \lambda^*(A) + \tfrac{\varepsilon}{2} = \lambda^*(A \cap E) + \lambda^*(A \cap E^c) - \varepsilon + \tfrac{\varepsilon}{2} \leq \lambda^*(U \cap E) + \lambda^*(U \cap E^c) - \tfrac{\varepsilon}{2} = \lambda^*(U) - \tfrac{\varepsilon}{2}.$$

A última igualdade é verdadeira porque $U, E, U \cap E, U \cap E^c$ pertencem a $\mathcal{D}$ e λ^*, restricta a $\mathcal{D}$ é aditiva. $\square$

As σ-álgebras $\mathcal{B}(\mathbb{R})$ e $\mathcal{L}(\mathbb{R})$ são diferentes. De facto, muito diferentes, uma vez que $\mathcal{B}(\mathbb{R})$ tem a mesma cardinalidade do que $\mathbb{R}$ (não será demonstrado) e $\mathcal{L}(\mathbb{R})$ tem a cardinalidade de $\mathcal{P}(\mathbb{R})$.

Exemplo 5.24. *O conjunto de Cantor $\mathcal{C}$ é definido como a intersecção de uma família decrescente de fechados $(C_n)_n$.*

$C_0 = [0, 1]$ e obtemos C_{n+1} à custa de C_n, começando por dividir em 3 partes iguais cada um dos intervalos C_n^k e depois retirar o interior do intervalo do meio. Por exemplo:

$$\begin{aligned}
C_1 &= [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, 1],\\
C_2 &= [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{3}{9}] \cup [\tfrac{6}{9}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1],\\
C_3 &= [0, \tfrac{1}{27}] \cup [\tfrac{2}{27}, \tfrac{3}{27}] \cup [\tfrac{6}{27}, \tfrac{7}{27}] \cup [\tfrac{8}{27}, \tfrac{9}{27}] \cup [\tfrac{18}{27}, \tfrac{19}{27}] \cup [\tfrac{20}{27}, \tfrac{21}{27}] \cup [\tfrac{24}{27}, \tfrac{25}{27}] \cup [\tfrac{26}{27}, 1].
\end{aligned}$$

Deste modo, para $n \in \mathbb{N}$, C_n é uma união disjunta de 2^n intervalos fechados C_n^k, de amplitude $\frac{1}{3^n}$, com $k = 1, \ldots, 2^n$.

Uma outra maneira de ver o conjunto de Cantor é notar que ele formado por todos os números em $[0, 1]$ que admitem uma representação na base 3 que não usa o dígito 1.

Como, para cada $n \in \mathbb{N}$, C_n é uma união finita de elementos de $\mathcal{B}(\mathbb{R})$ e C é uma intersecção numerável de elementos de $\mathcal{B}(\mathbb{R})$, então $\mathcal{C} \in \mathcal{B}(\mathbb{R})$.

Por outro lado $\lambda(\mathcal{C}) = 0$, uma vez que, para todo $n \in \mathbb{N}$, $\mathcal{C} \subseteq C_n$ e $\lambda(C_n) = 2^n \cdot \frac{1}{3^n} = \left(\frac{2}{3}\right)^n$. Deste modo, qualquer subconjunto de $\mathcal{C}$ é mensurável e tem medida nula, porque λ é uma medida completa. Como $\mathcal{C}$ e $\mathbb{R}$ têm o mesmo cardinal então o conjunto $\mathcal{L}(\mathbb{R})$ tem cardinalidade igual à de $\mathcal{P}(\mathbb{R})$ (verifique os detalhes!).

Para finalizar esta secção vamos mostrar um resultado que diz, em particular, que $\mathcal{L}(\mathbb{R})$ é gerada por $\mathcal{B}(\mathbb{R}) \cup \{E \in \mathcal{L}(\mathbb{R}) : \lambda(E) = 0\}$.

Começamos com um resultado na linha da Proposição 5.20, mas aplicada apenas a conjuntos mensuráveis.

Teorema 5.25. *Se $E \in \mathcal{P}(\mathbb{R})$ então as seguintes condições são equivalentes:*

a) $E \in \mathcal{L}(\mathbb{R})$;

b) *para todo $\varepsilon > 0$ existe um aberto U tal que $E \subseteq U$ e $\lambda^*(U \setminus E) \leq \varepsilon$;*

c) *existe W, que é uma intersecção numerável de abertos de $\mathbb{R}$, tal que $E \subseteq W$ e $\lambda^*(W \setminus E) = 0$;*

b)* *para todo $\varepsilon > 0$ existe um fechado F tal que $F \subseteq E$ e $\lambda^*(E \setminus F) \leq \varepsilon$;*

c)* *existe G, que é uma união numerável de fechados de $\mathbb{R}$, tal que $G \subseteq E$ e $\lambda^*(E \setminus G) = 0$.*

Em particular, $\mathcal{L}(\mathbb{R})$ é a σ-álgebra gerada por $\mathcal{B}(\mathbb{R}) \cup \mathcal{N}_\lambda$. Mais concretamente, todo o elemento de $\mathcal{L}(\mathbb{R})$ é a união disjunta de um elemento de $\mathcal{B}(\mathbb{R})$ com um de $\mathcal{N}_\lambda$.

Demonstração. Vamos começar por provar que as 3 primeiras condições são equivalentes.

a)$\Rightarrow$b) Seja $\varepsilon > 0$. Suponhamos primeiro que $\lambda(E) < \infty$. Pela Proposição 5.20 existe um aberto U contendo E e tal que $\lambda^*(U) \leq \lambda^*(E) + \varepsilon$. Como U é aberto então $U \in \mathcal{L}(\mathbb{R})$. Deste modo $\lambda(U) \leq \lambda(E) + \varepsilon$ ou seja, porque λ é aditiva e $\lambda(E) < \infty$, $\lambda(U \setminus E) = \lambda(U) - \lambda(E) \leq \varepsilon$.

Se $\lambda(E) = \infty$, consideremos uma família disjunta $(I_n)_n$ de intervalos limitados de $\mathbb{R}$ cuja união é $\mathbb{R}$ e $(E_n)_n = (E \cap I_n)_n$. Usando o que acabou de ser provado, consideremos $(U_n)_n$ uma família de abertos tais que $\mu(U_n \setminus E_n) \leq \frac{\varepsilon}{e^n}$, para $n \in \mathbb{N}$. Daqui e da sub-aditividade de λ^* resulta que $U = \cup_n U_n$ está nas condições requeridas.

b)$\Rightarrow$c) Para cada $n \in \mathbb{N}$ seja U_n aberto contendo E tal que $\lambda^*(U_n \setminus E) \leq \frac{1}{n}$. O conjunto $W = \cap_n U_n$ está nas condições pretendidas porque $W \setminus E \subseteq U_n \setminus E$, para todo $n \in \mathbb{N}$.

c)$\Rightarrow$a) Como $E = W \cap (W \setminus E)$ e $W \in \mathcal{B}(\mathbb{R}) \subseteq \mathcal{L}(\mathbb{R})$ e $W \setminus E \in \mathcal{L}(\mathbb{R})$, porque $\lambda^*(W \setminus E) = 0$, concluímos que $E \in \mathcal{L}(\mathbb{R})$.

Como $E \in \mathcal{L}(\mathbb{R})$ se e só se $E^c \in \mathcal{L}(\mathbb{R})$ então a) é equivalente a b) e c) aplicadas a E^c, que é o mesmo que c*) e b*) aplicadas a E.

Vejamos a conclusão do teorema. Se $E \in \mathcal{L}(\mathbb{R})$ consideremos G definido pela alínea c*). Deste modo, como $E, G \in \mathcal{L}(\mathbb{R})$ então $E \setminus G, \in \mathcal{L}(\mathbb{R})$ e $E = G \cup \setminus (E \setminus G)$. $\square$

5.5 Integral de Lebesgue

A ideia a reter é que o integral de Lebesgue vai ser definido numa classe de funções maior do que a classe onde está definido o integral de Riemann. De qualquer maneira, quando o integral de Riemann de uma função estiver definido então os dois integrais são iguais.

Recorde-se a definição de função característica de um conjunto A, que é definida por: $\chi_A(x) = 1$, se $x \in A$; $\chi_A(x) = 0$, se $x \notin A$.

No integral de Riemann de uma função positiva definida num intervalo I, começamos por considerar uma partição finita de I, e, se $(I_k)_{1 \leq k \leq n}$ for a família de intervalos definidos por essa partição, calculamos

$$\sum_{k=1}^{n} \inf_{I_k} f \times l(I_k), \quad \sum_{k=1}^{n} \sup_{I_k} f \times l(I_k) \tag{5.2}$$

e depois estudamos estas somas e as suas diferenças conforme vamos refinando a partição.

No integral de Lebesgue, consideramos funções mensuráveis ϕ com imagem finita $\{a_1, \ldots, a_n\}$ que estejam por baixo de f e depois calculamos

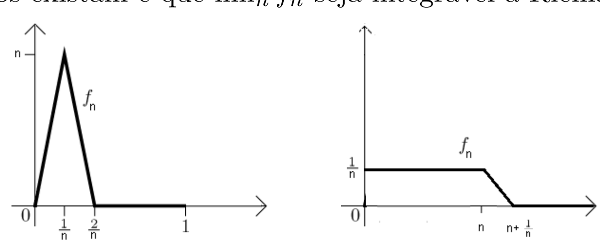
$$\sum_{k=1}^{n} a_k \times \lambda(\phi^{-1}(a_k)) \tag{5.3}$$

e depois estudamos estas somas quando ϕ "se aproxima" de f.

Os seguintes exemplos, um com funções definidas num intervalo fechado limitado e o outro com funções definidas num intervalo fechado ilimitado, mostram que, mesmo para funções contínuas, e independentemente da definição de integral que generalize o integral de Riemann, não é verdade que

$$\int \lim_n f_n = \lim_n \int f_n, \tag{5.4}$$

mesmo que os limites referidos existam e que $\lim_n f_n$ seja integrável à Riemann.



O integral de Riemann das funções das duas sucessões existe e é igual a 1 enquanto que e integral de Riemann da função limite (a função nula) é igual a 0. Note-se também que no segundo caso a convergência é uniforme.

Veremos que o integral de Lebesgue satisfaz a igualdade referida em (5.4) em algumas situações: por exemplo, se a sucessão de funções for crescente. Isto não acontece com o integral de Riemann, como podemos ver no seguinte exemplo.

Exemplo 5.26. *Por exemplo, se $(q_n)_n$ for uma sucessão injectiva que percorre todos os racionais em $[0,1]$ então a sucessão $\left(\chi_{\{q_1,\dots,q_n\}}\right)_n$ é uma sucessão de funções definidas em $[0,1]$ que são integráveis à Riemann, pois têm apenas um número finito de pontos de descontinuidade, mas o seu limite, que é $\chi_{[0,1]\cap\mathbb{Q}}$, não é integrável porque as somas em (5.2) relativas a esta função valem 0 e 1, independentemente da partição. Note-se conjunto de pontos de descontinuidade da função é igual a $[0,1]$.*

Esta situação não acontece com o integral de Lebesgue. De facto, como $\chi_{[0,1]\cap\mathbb{Q}}$ é uma função simples a soma, para esta função, referida em (5.3) é igual a 1 uma vez que $\lambda([0,1]\cap\mathbb{Q}) = 0$ e $\lambda([0,1]\setminus\mathbb{Q}) = 1$. Para o integral de Lebesgue as funções $\chi_{[0,1]\cap\mathbb{Q}}$ e a função constante e igual a 1 "são iguais" uma vez que diferem de um conjunto com medida nula. Note-se a função $\chi_{[0,1]\cap\mathbb{Q}}$ é descontínua em $[0,1]$ e a função igual a 1 não tem pontos de descontinuidade. Isto é decisivo para o integral de Riemann.

Para definir o integral de Lebesgue começamos por considerar um espaço de medida (X,Σ,μ), depois definimos o que se entende por função mensurável. O integral de Lebesgue será definido para funções mensuráveis.

5.5.1 Funções mensuráveis

No que segue consideremos um espaço mensurável (X,Σ) e $(\mathbb{R},\mathcal{B}(\mathbb{R}))$. Muito do que é dito de seguida continua válido se considerarmos um espaço de medida (Y,Σ_Y) no lugar de $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.

Definição 5.27. *Seja (X,Σ) um espaço mensurável. Uma função $f : X \to \mathbb{R}$ diz-se mensurável se a imagem recíproca de qualquer elemento de $\mathcal{B}(\mathbb{R})$ for um elemento de Σ.*

De facto, para ver que uma função é mensurável basta-nos analisar a imagem recíproca dos elementos de uma família geradora de $\mathcal{B}(\mathbb{R})$ e não todos os elementos de $\mathcal{B}(\mathbb{R})$. Isto acontece por causa do seguinte resultado.

Lema 5.28. *Se $f : X \to \mathbb{R}$ então $\{A \in \mathcal{P}(\mathbb{R}) : f^{-1}(A) \in \Sigma\}$ é uma σ-álgebra.*

Demonstração. Basta usar as propriedade satisfeitas por qualquer função: a imagem recíproca da intersecção/união de uma família de conjuntos é igual à intersecção/união das imagens recíprocas dos elementos dessa família; a imagem recíproca do complementar de um conjunto é igual ao complementar da imagem recíproca desse conjunto. □

Atendendo a este lema e à definição de função mensurável temos o seguinte resultado.

Proposição 5.29. *Se $f : X \to \mathbb{R}$ e $\mathcal{A}$ é uma família de subconjuntos de $\mathbb{R}$ tal que a σ-álgebra gerada por $\mathcal{A}$ é igual a $\mathcal{B}(\mathbb{R})$ então f é mensurável se e só se $f^{-1}(A) \in \Sigma$ para todo $A \in \mathcal{A}$.* □

Como exemplos de conjuntos $\mathcal{A}$ nas condições da proposição anterior temos o conjunto formado por todos os intervalos abertos de $\mathbb{R}$, o conjunto formado por todos os intervalos da forma $]a,\infty[$, etc. (ver mais exemplos dados depois da Definição 5.3).

Exemplos 5.30. *Vejamos alguns exemplos simples de funções mensuráveis.*

1. *Qualquer função constante é mensurável.*

2. *Se X for um espaço topológico e considerarmos a σ-álgebra $\mathcal{B}(X)$ então toda a função contínua de X em $\mathbb{R}$ é mensurável (basta usar a proposição anterior).*

3. *Se (X,Σ) é um espaço mensurável e $A \in \mathcal{P}(X)$ então χ_A é mensurável se e só se $A \in \Sigma$. Isto acontece porque, se $B \in \mathcal{B}(\mathbb{R})$ então $\chi^{-1}(B)$ é igual ao conjunto vazio, a X, a A ou a $X \setminus A$, consoante 0 e 1 pertencem ou não a B.*

Se tivermos um espaço mensurável (X,Σ) podemos ter funções definidas, não em X, mas num seu subconjunto E. Só nos vai interessar, se $E \in \Sigma$. Se $f : E \to \mathbb{R}$ podemos considerar:

- (E, Σ_E), em que $\Sigma_E = \{A \cap E : A \in \Sigma\}$;
- a extensão por 0 de f a X.

O seguinte resultado diz que qualquer dos caminhos "vai levar ao mesmo resultado".

Lema 5.31. *Se (X,Σ) é um espaço mensurável e $E \in \Sigma$. Se considerarmos o espaço mensurável $(E, \{A \cap E : A \in \Sigma\})$ temos:*

a) se $f : X \to \mathbb{R}$ é mensurável então a restrição de f a E é mensurável;

b) se $f : E \to \mathbb{R}$ é mensurável então o prolongamento por 0 de f a X é mensurável.

Demonstração. Para a alínea a), basta notar que, se $A \in \mathcal{P}(\mathbb{R})$ então $f_{|E}^{-1}(A) = f^{-1}(A) \cap E$. Para a alínea b), se g for o prolongamento de f referido e $A \in \mathcal{P}(X)$ então $g^{-1}(A)$ é igual a $f^{-1}(A)$, se $0 \notin A$ e é igual a $f^{-1}(A) \cup (X \setminus E)$, caso contrário. □

Atendendo a este resultado podemos essencialmente pensar que todas as funções estão definidas em X. É o que vamos fazer de agora em diante.

Vamos agora ver que o conjunto formado pelas funções mensuráveis definidas num dado espaço mensurável é fechado para várias operações.

Proposição 5.32. *Se (X,Σ) é um espaço mensurável, $f : X \to \mathbb{R}$ mensurável e $\phi : \mathbb{R} \to \mathbb{R}$ é mensurável (considerando a σ-álgebra $\mathcal{B}(\mathbb{R})$ no domínio e no conjunto de chegada) então $\Phi \circ f$ é mensurável.*

Demonstração. Basta notar que, se $A \in \mathcal{P}(\mathbb{R})$ então $(\phi \circ f)^{-1}(A) = f^{-1}\big(\phi^{-1}(A)\big)$ □

Recorde-se que qualquer função contínua de $\mathbb{R}$ em $\mathbb{R}$ é mensurável. Deste modo, se f é mensurável então também são mensuráveis as funções $|f|$ (considerando $\phi(x) = |x|$), cf, com $c \in \mathbb{R}$ (considerando $\phi(x) = cx$), f^2 (considerando $\phi(x) = x^2$), f^+ (considerando $\phi(x) = \max\{x, 0\}$, f^- (considerando $\phi(x) = -\min\{x, 0\}$). Em alguns destes casos poderíamos ter usado alguma das igualdades $|f| = f^+ + f^-$, $f = f^+ - f^-$, de onde resulta $f^+ = \frac{1}{2}(|f| + f)$ e $f^- = \frac{1}{2}(|f| - f)$.

Proposição 5.33. *Se (X,Σ) é um espaço mensurável, $f, g : X \to \mathbb{R}$ são funções mensuráveis então $f + g$, $f \cdot g$, $f \vee g$ e $f \wedge g$ são mensuráveis.*

Demonstração. Como, para qualquer $a \in \mathbb{R}$,

$$(f+g)^{-1}(]-\infty, a[) = \bigcup_{r \in \mathbb{Q}} f^{-1}(]-\infty, r[) \cap g^{-1}(]-\infty, a-r[)$$

concluímos que $f + g$ é mensurável. Resta agora notar que $f + g = \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right]$, $f \vee g = \frac{1}{2}\left[f + g + |f - g|\right]$ e $f \wedge g = \frac{1}{2}\left[f + g - |f - g|\right]$. □

Nas condições da proposição anterior podemos usar o mesmo tipo de argumento para mostrar que as funções cf (para um dado $c \in \mathbb{R}$), f^2, $|f|$, f^+, f^- são mensuráveis. De facto

$$\begin{aligned}
(cf)^{-1}(]-\infty, a[) &=]-\infty, \tfrac{a}{c}[\quad \text{se } c > 0\\
(cf)^{-1}(]-\infty, a[) &=]\tfrac{a}{c}, \infty[\quad \text{se } c < 0\\
(f^2)^{-1}(]a, \infty[) &= X \quad \text{se } a < 0\\
(f^2)^{-1}(]a, \infty[) &= f^{-1}(]\sqrt{a}, \infty[) \cup f^{-1}(]-\infty, -\sqrt{a}[) \quad \text{se } a \geq 0\\
|f|^{-1}(]a, \infty[) &= X \quad \text{se } a < 0\\
|f|^{-1}(]a, \infty[) &= f^{-1}(]a, \infty[) \cup f^{-1}(]-\infty, -a[) \quad \text{se } a \geq 0.
\end{aligned}$$

Proposição 5.34. *Se (X,Σ) é um espaço mensurável e $(f_n)_n$ é uma sucessão de funções mensuráveis então as funções $\sup_n f_n$, $\inf_n f_n$, $\limsup_n f_n$ e $\liminf_n f_n$ são mensuráveis.*

Em particular se $(f_n)_n$ converge (pontualmente) então $\lim_n f_n$ é mensurável.

Demonstração. Dado $a \in \mathbb{R}$ temos $(\sup_n f_n)^{-1}(]a,\infty[) = \bigcup_{n=1}^{\infty} f_n^{-1}(]a,\infty[)$. Resta agora notar que $\inf_n f_n = -\sup(-f_n)$, $\limsup_n f_n = \inf_n \left(\sup_{k\geq n} f_k\right)$ e $\liminf_n f_n = \sup_n \left(\inf_{k\geq n} f_k\right)$. □

O seguinte resultado, relativo a espaços mensuráveis com uma medida completa será usada várias vezes mais à frente. Recorde-se que todas as medidas que forem definidas à custa de uma medida exterior (ver Teorema 5.14), como é o caso da medida de Lebesgue, são completas.

Proposição 5.35. *Se (X,Σ,μ) for um espaço de medida, com μ completa, e $f,g : X \to \mathbb{R}$ diferirem apenas um conjunto I com $\mu(I) = 0$ então f é mensurável se e e só se g for mensurável.*

Demonstração. Suponhamos que f é mensurável (o outro caso, é análogo). Seja $A \in \mathcal{B}(\mathbb{R})$ e vejamos que $g^{-1}(A) \in \Sigma$. Temos

$$g^{-1}(A) = \left(g^{-1}(A) \setminus I\right) \ \dot\cup\ \left(g^{-1}(A) \cap I\right) = \left(f^{-1}(A) \setminus I\right) \ \dot\cup\ \left(g^{-1}(A) \cap I\right).$$

Como μ é uma medida completa então $g^{-1}(A) \cap I \in \Sigma$. Deste modo, como f é mensurável, $g^{-1}(A) \in \Sigma$ pois é a união de dois elementos de Σ. □

O exemplo seguinte mostra que, na proposição anterior, a condição de a medida ser completa é necessária.

Exemplo 5.36. *Sejam (X,Σ,μ) um espaço de medida, com μ não completa, e $f : X \to \mathbb{R}$ mensurável. Consideremos $A, B \in \mathcal{P}(X)$ tais que $A \in \Sigma$, $\mu(A) = 0$, $B \subseteq A$ e $B \notin \Sigma$, e $g = f + \chi_A + \chi_B$. Atendendo a Exemplos 5.30, χ_A é mensurável e χ_B não é mensurável. Como $\chi_B = g - f - \chi_A$, as proposições anteriores garantem-nos que g não é mensurável. Para terminar resta notar que $A = \{x \in X : f(x) \neq g(x)\}$.*

5.5.2 Integral de uma função simples

O integral vai ser definido para funções mensuráveis, começando pelas que são funções simples positivas.

Definição 5.37. *Seja (X,Σ) um espaço mensurável. Uma função $f : X \to \mathbb{R}$ diz-se simples se for mensurável e a sua imagem for um conjunto finito.*

Nota 5.38. *De facto a definição de função simples é mais geral e não exige a mensurabilidade da função. De qualquer modo optou-se por esta definição para simplificar a escrita e porque essa noção só se vai aplicar a funções mensuráveis.*

As funções simples podem ser caracterizadas da seguinte forma.

Lema 5.39. *Se (X,Σ) é um espaço mensurável então uma função $f : X \to \mathbb{R}$ é simples se e só se existirem $n \in \mathbb{N}$, $a_1,\ldots,a_n \in \mathbb{R}$, $E_1,\ldots,E_n \in \Sigma$ tais que*

$$f = \sum_{i=1}^{n} a_i \chi_{E_i}.$$

Demonstração. Se f estiver nas condições referidas então a f toma apenas os valores que são somas finitas dos elementos $a_1,\ldots,a_n$. Por outro lado f é mensurável porque é uma combinação linear de funções mensuráveis, uma vez que $E_1,\ldots,E_n \in \Sigma$.

Inversamente se f for mensurável se a imagem de f tiver n elementos e for igual a $\{a_1,\ldots,an\}$ então, $f = \sum_{i=1}^{n} a_i \chi_{f^{-1}(a_i)}$ e, como f é mensurável então, para todo $i = 1,\ldots,n$, $f^{-1}(a_i) \in \Sigma$. □

Note-se que uma função simples tem muitas representações da forma referida. Por exemplo

$$3\chi_{[0,3]} = 2\chi_{[0,1]} + 2\chi_{[2,3]} + \chi_{[1,3]} + \chi_{[0,2]} + \chi_{[1,2]}.$$

De entre todas as representações nesta forma de uma função simples há uma (referida na demonstração do lema anterior) a que vamos chamar de representação canónica de f.

Definição 5.40. *Se (X, Σ) é um espaço mensurável e $f : X \to \mathbb{R}$ é uma função simples cuja imagem tem n elementos, $a_1, \ldots, a_n$ então*

$$f = \sum_{i=1}^{n} a_i \chi_{f^{-1}(a_i)},$$

e esta representação de f diz-se canónica.

Note-se que, na representação canónica dada acima, a família $\left(f^{-1}(a_i)\right)_{i=1,\ldots,n}$ é uma partição de X e os elementos $a_1 \ldots, a_n$ são todos diferentes. É claro que, como f é mensurável, os conjuntos $f^{-1}(a_i)$, com $i = 1, \ldots, n$, são mensuráveis.

Exemplo 5.41. *Uma representação $f = \sum_{i=1}^{n} a_i \chi_{E_i}$, só é uma representação canónica se $(E_i)_{i=1}^{n}$ for uma partição de X. Se os conjuntos forem disjuntos mas a sua união não for X então a representação canónica de f é $\sum_{i=1}^{n+1} a_i \chi_{E_i}$, em que $a_{i+1} = 0$ e $E_{n+1} = X \setminus (\cup_{i=1}^{n} E_n)^c$.*

No que segue vamos adoptar a convenção de que $0 \times \infty = 0$.

Proposição 5.42. *Se (X, Σ, μ) é um espaço de medida, $f : X \to \mathbb{R}$ é uma função simples não negativa e $f = \sum_{i=1}^{n} a_i \chi_{E_i}$ e $f = \sum_{i=1}^{m} b_i \chi_{F_i}$, em que os conjuntos E_i e F_j são mensuráveis e $a_1, \ldots, a_n, b_1, \ldots, b_m \in \mathbb{R}$, então*

$$\sum_{i=1}^{n} a_i \mu(E_i) = \sum_{j=1}^{m} b_j \mu(F_j).$$

Demonstração. Pelo que vimos acima, podemos considerar que $\cup_{i=1}^{n} E_i$ e $\cup_{i=1}^{m} F_i$ são iguais a X.

Vamos fazer a demonstração em dois passos.

$\boxed{\text{Primeiro passo}}$

Seja $I = \{i \in \{1, 2, \ldots, n\} : E_1 \cap E_i \neq \emptyset\}\}$ e consideremos $\mathcal{A} = \{J \subseteq I : 1 \in J\}$. Deste modo

$$E_1 = \dot{\bigcup}_{J \in \mathcal{A}} E_J, \quad \text{em que} \quad E_J = \cap_{j \in J} E_j \setminus \cup_{j \notin J} E_j, \quad \mu(E_1) = \sum_{J \in \mathcal{A}} E_J.$$

Obtemos assim uma cobertura de X formada pelos conjuntos E_J, com $J \in \mathcal{A}$, e $E_i \setminus E_1$, para $i = 2, \ldots, n$. Por exemplo:

- se $n = 2$, temos $E_1 = \underbrace{(E_1 \setminus E_2)}_{=E_{\{1\}}} \cup \underbrace{(E_1 \cap E_2)}_{=E_{\{1,2\}}}$;
- se $n = 3$, temos $E_1 = \underbrace{(E_1 \setminus (E_2 \cup E_3))}_{=E_{\{1\}}} \cup \underbrace{((E_1 \cap E_2) \setminus E_3)}_{=E_{\{1,2\}}} \cup \underbrace{((E_1 \cap E_3) \setminus E_2)}_{=E_{\{1,3\}}} \cup \underbrace{(E_1 \cap E_2 \cap E_3)}_{=E_{\{1,2,3\}}}$;
- se $n = 4$, temos

$$E_1 = \underbrace{(E_1 \setminus (E_2 \cup E_3 \cup E_4))}_{=E_{\{1\}}} \cup \underbrace{((E_1 \cap E_2) \setminus E_3 \cup E_4)}_{=E_{\{1,2\}}} \cup \cup \underbrace{((E_1 \cap E_3) \setminus E_2 \cup E_4)}_{=E_{\{1,3\}}} \cup \underbrace{((E_1 \cap E_4) \setminus E_2 \cup E_3)}_{=E_{\{1,4\}}} \cup$$
$$\cup \underbrace{((E_1 \cap E_2 \cap E_3) \setminus E_4)}_{=E_{\{1,2,3\}}} \cup \underbrace{((E_1 \cap E_2 \cap E_4) \setminus E_3)}_{=E_{\{1,2,4\}}} \cup \underbrace{((E_1 \cap E_3 \cap E_4) \setminus E_3)}_{=E_{\{1,3,4\}}} \cup \underbrace{(E_1 \cap E_2 \cap E_3 \cap E_4)}_{=E_{\{1,2,3,4\}}}.$$

Voltando ao caso geral, vemos que,

$$\sum_{i\in I} a_i\chi_{E_i} = \sum_{J\subseteq I}\left(\sum_{j\in J} a_j\right)\chi_{E_J} + \sum_{i=2}^{n} a_i\chi_{E_i\setminus E_1}, \qquad \text{(note-se que, se } i\notin I \text{ então } E_i\setminus E_1 = E_1).$$

Nesta nova partição de X apenas os últimos $n-1$ conjuntos referidos se podem intersectar. Sendo assim, repetindo o processo, obtemos uma partição de X, $(G_i)_{1\le i\le r}$, e números, $c_1,\ldots,c_r\ge 0$, tais que

$$\sum_{i=1}^{n} a_i\mu(E_i) = \sum_{j=1}^{r} c_j\mu(G_j).$$

Segundo passo

Suponhamos que $(E_i)_{1\le i\le n}$ e $(F_i)_{1\le j\le m}$ s ão partições de X.
Deste modo, para cada $i=1,\ldots,n$ e $j=1,\ldots,m$

$$E_i = E_i\cap X = E_i\cap \dot\cup_{j=1}^{m} F_j = \dot\cup_{j=1}^{m}(E_i\cap F_j), \qquad F_j = F_j\cap X = E_i\cap \dot\cup_{i=1}^{n} E_i = \dot\cup_{i=1}^{n}(E_i\cap F_j).$$

Note-se que, se $E_i\cap F_j\neq\emptyset$ então $a_i=b_j$, porque ambos são iguais a $f(x)$, se $x\in E_i\cap F_j$. Deste modo

$$\sum_{i=1}^{n} a_i\mu(E_i) = \sum_{i=1}^{n} a_i\mu\left(\dot\cup_{j=1}^{m}(E_i\cap F_j)\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i\mu(E_i\cap F_j) = \sum_{j=1}^{m}\sum_{i=1}^{n} b_j\mu(E_i\cap F_j) = \cdots = \sum_{j=1}^{m} b_j\mu(E_j),$$

usando a aditividade numerável de μ (duas vezes). □

Atendendo a este resultado a definição seguinte tem sentido.

Definição 5.43. *Seja (X,Σ,μ) um espaço de medida e $f: X\to\mathbb{R}$ uma função simples não negativa. Se $f=\sum_{i=1}^{n} a_i\chi_{E_i}$, com $E_1,\ldots,E_n\in\Sigma$, define-se integral de f relativamente a μ e denota-se por $\int_X f\,d\mu$ como sendo*

$$\int_X f\,d\mu = \sum_{i=1}^{n} a_i\mu(E_i).$$

Note-se que, no espaço de medida $(\mathbb{R},\mathcal{L}(\mathbb{R}),\lambda)$,

$$\int_{\mathbb{R}} \chi_{[0,1]\cap\mathbb{Q}}\,d\lambda = 0, \qquad \int_{\mathbb{R}} \chi_{[0,1]\setminus\mathbb{Q}}\,d\lambda = \int_{\mathbb{R}} \chi_{[0,1]}\,d\lambda = 1.$$

Mais geralmente, num espaço de medida (X,Σ,μ), se $f=\sum_{i=1}^{n} a_i\chi_{E_i}$, com $E_1,\ldots,E_n\in\Sigma$ e A for um conjunto com medida nula então

$$\begin{aligned}\int_X f\,d\mu &= \sum_{i=1}^{n} a_i\mu(E_i) = \sum_{i=1}^{n} a_i\left[\mu(E_i\cap A)+\mu(E_i\cap A^c)\right] = \sum_{i=1}^{n} a_i\mu(E_i\cap A) + \sum_{i=1}^{n} a_i\mu(E_i\cap A^c)\\ &= \sum_{i=1}^{n} a_i\mu(E_i\cap A^c).\end{aligned}$$

Isto implica que, se alterar f em A, a função obtida é mensurável e o integral não se altera. Mais concretamente, se duas funções simples forem iguais excepto num conjunto de medida nula então têm o mesmo integral.

5.5.3 Integral de uma função mensurável não negativa

Vamos agora definir o que se entende por integral de uma função mensurável não negativa à custa de se saber que estas funções são limite pontual de funções simples. Comecemos por provar este resultado.

Teorema 5.44. *Seja (X, Σ, μ) um espaço de medida. Se $f : X \to \mathbb{R}$ é uma função mensurável não negativa então existe uma sucessão monótona crescente de funções simples a convergir pontualmente para f. Se f for majorada, então a convergência é uniforme.*

Demonstração. Para cada $n \in \mathbb{N}$ vamos dividir o intervalo $[0, n[$ em $n2^n$ intervalos abertos à direita e fechados à esquerda e com a mesma amplitude. Temos assim os intervalos $I_{n,k} = [\frac{k-1}{2^n}, \frac{k}{2^n}[$, com $k = 1, \dots, n2^n$. Note-se que, nas condições acima, se $E_{n,k} = f^{-1}(I_{n,k})$,

$$I_{n,k} = I_{n+1,2k-1} \cup I_{n+1,2k} \quad \text{e portanto} \quad E_{n,k} = E_{n+1,2k-1} \cup E_{n+1,2k}.$$

Consideremos ainda $F_n = f^{-1}([n, \infty[)$. Vejamos que a sucessão de funções $(\phi_n)_n$ definida por

$$\phi_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} \chi_{E_{n,k}} + n\chi_{F_n},$$

está nas condições pretendidas. Note-se que, se $n \in \mathbb{N}$ então $E_{n,k}$ e F_n são mensuráveis pois são a imagem recíproca de intervalos de f por uma função mensurável.

Por outro lado

$$0 \leq f(x) - \phi_n(x) \leq \begin{cases} \frac{1}{2^n} & \text{se } x \in F_n^c \\ f(x) - n & \text{se } x \in F_n \end{cases} \quad \text{porque } \phi_n(x) = \frac{k-1}{2^n} \text{ e } f(x) < \frac{k}{2^n}, \text{ se } x \in E_{n,k}. \tag{5.5}$$

Deste modo, se $x \in X$ e $n_0 \in \mathbb{N}$ for tal que $n_0 > f(x)$ então, para todo $n \geq n_0$, $0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n}$ de onde se conclui que sucessão $(\phi_n(x))_n$ converge para $f(x)$. Se f for majorado por uma constante $M > 0$ então para $n > M$, temos $0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n}$, para todo $x \in X$, e portanto $(\phi_n)_n$ converge uniformemente para f.

Vejamos agora que a sucessão $(\phi_n)_n$ é crescente. Para todo $n \in \mathbb{N}$ e todo $x \in X$,

$$\phi_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{se } x \in E_{n,k} \\ n & \text{se } x \in F_n \end{cases} \leq \begin{cases} \frac{2k-2}{2^{n+1}} & \text{se } x \in E_{n+1,2k-1} \\ \frac{2k-1}{2^{n+1}} & \text{se } x \in E_{n+1,2k} \\ \phi_{n+1}(x) & \text{se } x \in f^{-1}([n, n+1[) \\ n+1 & \text{se } x \in F_{n+1} \end{cases}$$

e portanto $\phi_n(x) \leq \phi_{n+1}(x)$, uma vez que se $f(x) \in [n, n+1[$ então $x \in E_{n+1,k}$, para algum $k \geq n2^{n+1} + 1$, o que implica que $\phi_{n+1}(x) \geq n$. $\square$

Definição 5.45. *Seja (X, Σ, μ) um espaço de medida e $f : X \to \mathbb{R}$ uma função mensurável não negativa. O integral de Lebesgue de f relativamente a μ é definido como*

$$\int_X f \, d\mu = \sup \left\{ \int_X \phi \, d\mu : 0 \leq \phi \leq f, \ \phi \text{ é uma função simples} \right\}.$$

Se $A \in \Sigma$, define-se

$$\int_A f \, d\mu = \int_X \chi_A f \, d\mu.$$

É claro que, se f for uma função simples então esta e a anterior definição coincidem.

Se não houver dúvidas quando ao domínio de f podemos escrever $\int f \, d\mu$ em vez de $\int_X f \, d\mu$.

Note-se ainda que, com as notações da definição, se f for simples e igual a $\sum_{i=1}^n a_i \chi_{E_i}$ então $\chi_A f$ também é simples e $\int_A f \, d\mu = \sum_{i=1}^n a_i \mu(A \cap E_i)$ e, em particular,

$$\int_A d\mu = \int_X \chi_A \, d\mu = \mu(A).$$

Vamos ver algumas propriedades do integral.

Nota 5.46. *Se $(X\sigma)$ for um espaço mensurável e $f, g : X \to \mathbb{R}$ forem duas funções simples não negativas então existe uma partição de X, $(E_i)_{1\leq i\leq n}$ $(n \in \mathbb{N})$, formada por elementos de Σ, e $a_1, \dots, a_n, b_1, \dots, b_n \geq 0$ tais que*

$$f = \sum_{i=1}^{n} a_i \chi_{E_i}, \quad g = \sum_{i=1}^{n} b_i \chi_{E_i}.$$

Isto acontece porque, se $f = \sum_{i=1}^{r} c_i \chi_{F_i}$ e $g = \sum_{j=1}^{m} d_j \chi_{G_j}$ forem as representações canónicas de duas funções simples, então, $F_i = \dot\cup_{j=1}^{m} F_i \cap G_j$ e $G_j = \dot\cup_{i=1}^{r} F_i \cap G_j$

$$f = \sum_{i=1}^{r} \sum_{j=1}^{m} a_i \chi_{F_i \cap G_j} \quad g = \sum_{i=1}^{n} \sum_{j=1}^{m} b_j \chi_{F_i \cap G_j}.$$

Notação. *Sejam (X, Σ, μ) e $f, g : X \to \mathbb{R}$ funções mensuráveis. Diz-se que f e g são iguais quase sempre e escreve-se $f = g$ q.s., se $\mu(\{x \in X : f(x) \neq g(x)\}) = 0$. Notação análoga será usada se uma sucessão de funções convergir excepto num conjunto de medida nula.*

Proposição 5.47. *Sejam (X, Σ, μ) um espaço de medida, $c \in \mathbb{R}_0^+$, $A, B \in \Sigma$, $(A_k)_{k\in\mathbb{N}}$ uma família disjunta de elementos de Σ e $f, g : X \to \mathbb{R}$ funções mensuráveis não negativas. Nestas condições:*

a) $\int cf \, d\mu = c \int f \, d\mu$;

b) *se* $f \leq g$ *então* $\int f \, d\mu \leq \int g \, d\mu$;

c) *se* $A \subseteq B$ *então* $\int_A f d\mu \leq \int_B f \, d\mu$;

d) $\int_{\cup_k A_k} f d\mu = \sum_k \int_{A_k} f \, d\mu$;

e) *se* $\mu(A) = 0$ *então* $\int_A f = 0$;

f) *(desigualdade de Chebyshev)* $c\,\mu(\{x \in X : f(x) \geq c\}) \leq \int f \, d\mu$;

g) $\int f \, d\mu = 0$ *se e só se* $f = 0$ *quase sempre;*

h) *se* $f = g$ *quase sempre então* $\int f \, d\mu = \int g \, d\mu$.

Demonstração. Suponhamos primeiro que f e g são funções simples. Usando a Nota 5.46 f e g admitem representações da forma $f = \sum_{i=1}^{n} a_i \chi_{E_i}$ e $g = \sum_{i=1}^{n} b_i \chi_{E_i}$, sendo $(E_i)_{1\leq i\leq n}$ uma partição de X formada por elementos de Σ. Deste modo, se $D \in \Sigma$,

$$\int f \, d\mu = \sum_{i=1}^{n} c a_i \mu(E_i), \ \int g \, d\mu = \sum_{i=1}^{n} b_i \mu(D \cap E_i) \ \int cf \, d\mu = \sum_{i=1}^{n} c a_i \mu(E_i), \ \int_D f \, d\mu = \sum_{i=1}^{n} a_i \mu(D \cap E_i).$$

Daqui resultam imediatamente a), c) e e). Para mostrar b), basta notar que, se $x \in E_i$ então $f(x) = a_i$ e $g(x) = b_i$.

Para alínea d), como a família $(A_k)_k$ é disjunta, temos $\sum_k \chi_{A_k} f = \chi_{\cup_k A_k} f$ e portanto

$$\begin{aligned}
\int_{\cup_k A_k} f d\mu &= \sum_{i=1}^{n} a_i \mu((\cup_k A_k) \cap E_i) = \sum_{i=1}^{n} a_i \mu(\dot\cup_k (A_k \cap E_i)) \\
&= \sum_{i=1}^{n} a_i \sum_{k=1}^{\infty} \mu(A_k \cap E_i) \\
&= \sum_{k=1}^{\infty} \sum_{i=1}^{n} a_i \mu(A_k \cap E_i) = \sum_{k=1}^{\infty} \int_{A_k} f \, d\mu.
\end{aligned}$$

Vejamos agora o caso geral.

a) Podemos supor que $c \neq 0$, pois se $c = 0$ a igualdade é trivial. Consideremos ϕ, uma função simples tal que $0 \leq \phi \leq f$. Como $0 \leq c\phi \leq cf$ concluímos, por definição de integral, que $\int cf\, d\mu \geq \int c\psi d\mu = c \int \psi\, d\mu$ (porque ϕ é simples). Deste modo $\frac{1}{c} \int cf\, d\mu \geq \int \psi\, d\mu$. Novamente por definição de integral, $\frac{1}{c} \int cf\, d\mu \geq \int f\, d\mu$ ou seja $\int cf\, d\mu \geq c \int f\, d\mu$

Repetindo o argumento para a função cf, no lugar de f, e para a constante $\frac{1}{c}$, no lugar de c, obtemos $\int f\, d\mu \geq \frac{1}{c} \int cf\, d\mu$ ou seja $c \int f\, d\mu \geq \int cf\, d\mu$.

b) Basta notar que toda a função simples que esteja abaixo de f também está abaixo de g.

c) Basta usar a alínea anterior aplicada às funções $\chi_A f$ e $\chi_B f$.

d) Seja ϕ simples tal que $0 \leq \phi \leq \chi_{\cup_k A_k} f$. Então, se $A = \cup_k A_k$

$$\chi_{A_k} \phi \leq \chi_{A_k} f, \qquad \phi = \chi_A \phi \leq \chi_A f$$

e portanto, pelo que já foi provado para funções simples,

$$\int \phi\, d\mu = \int_A \phi\, d\mu = \sum_k \int_{A_k} \phi = \sum_k \int \chi_{A_k} \phi \leq \sum_k \int \chi_{A_k} f = \sum_k \int_{A_k} f.$$

Deste modo, pela definição de integral, $\int_A f d\mu \leq \sum_k \int_{A_k} f\, d\mu$.

Para mostrar o inverso basta ver que, para todo $N \in \mathbb{N}$, $\sum_{k=1}^N \int_{A_k} f\, d\mu \leq \int_A f d\mu$, ou seja, atendendo à definição de integral, se $\phi_1, \ldots, \phi_N$ são funções simples tais que $0 \leq \phi_k \leq \chi_{A_k} f$ para $k = 1, \ldots, N$, então $\sum_{k=1}^N \int \phi_k\, d\mu \leq \int_A f$.

Sejam então $\phi_1, \ldots, \phi_N$ nas condições referidas e consideremos a função simples $\phi = \sum_{k=1}^N \phi_k$. Pelo que já vimos para funções simples, e notando que ϕ e ϕ_k são iguais em A_k,

$$\sum_{k=1}^N \int \phi_k\, d\mu = \sum_{k=1}^N \int_{A_k} \phi_k\, d\mu = \sum_{k=1}^N \int_{A_k} \phi\, d\mu = \int_{\cup_{k=1}^N A_k} \phi d\mu \leq \int_{\cup_{k=1}^N A_k} f d\mu \leq \int_A f d\mu.$$

e) Basta usar a definição de integral, pois já vimos que $\int_A \phi\, d\mu = 0$, se ϕ for uma função simples.

f) Basta notar que $c\chi_A \leq f$, sendo $A = \{x \in X : f(x) \geq c\}$, e usar alíneas anteriores.

g) Sejam $A = \{x \in X : f(x) = 0\}$ e $B = X \setminus A$. Como $X = A \dot{\cup} B$ então, por d), $\int f\, d\mu = \int_A f\, d\mu + \int_B f\, d\mu = \int_B f\, d\mu$.

Se $\mu(B) = 0$ então, por e), $\int_B f\, d\mu = 0$.

Inversamente, se $\int f\, d\mu = 0$, consideremos, para $n \in \mathbb{N}$, $B_n = \{x \in X : f(x) \geq \frac{1}{n}\}$. Pela alínea anterior $\mu(B_n) = 0$. Como $(B_n)_n$ é uma sucessão crescente cuja união é B concluímos que $\mu(B) = \lim_n \mu(E_n) = 0$.

h) Sejam $A = \{x \in X : f(x) = g(x)\}$ e $B = X \setminus A$. Usando as alíneas d) e e), temos

$$\int f\, d\mu = \int_A f\, d\mu + \int_B f\, d\,\mu = \int_A f\, d\mu = \int_A g\, d\mu = \int_A g\, d\mu + \int_B g\, d\,\mu = \int g\, d\mu$$

uma vez que $\mu(B) = 0$ e f é igual a g em A.

□

Como corolário das alíneas c) e d) desta proposição, vejamos como podemos associar a cada função mensurável não negativa uma medida.

Proposição 5.48. *Seja (X, Σ, μ) um espaço de medida e $f : X \to \mathbb{R}$ uma função mensurável não negativa. Então a função*

$$\begin{array}{rccc} \mu_f: & \Sigma & \longrightarrow & [0, \infty] \\ & A & \mapsto & \int_A f\, d\mu \end{array}$$

é uma medida. □

Vejamos agora a primeira das condições suficientes para que a igualdade referida em (5.4) seja válida.

Teorema 5.49 (da convergência monótona)**.** *Sejam (X, Σ, μ) um espaço de medida e $(f_n)_n$ uma sucessão crescente de funções mensuráveis reais não negativas definidas em X convergindo quase sempre para uma função f. Nestas condições*

$$\lim_{n\to\infty} \int f_n d\mu = \int f\, d\mu.$$

Demonstração. Seja $N \subseteq X$ tal que $(f_n)_n$ converge para f em N^c e consideremos $(g_n)_n$ e g definidos por $g_n = \chi_{N^c}$ e $g = \chi_{N^c} f$, com $n \in \mathbb{N}$. Atendendo à alínea h) da Proposição 5.47, o que se pretende é equivalente à igualdade $\lim_{n\to\infty} \int f_n d\mu = \int f\, d\mu$.

Pela Proposição 5.34, g é mensurável e, como $(g_n)_n$ é crescente e majorada por g, então $\left(\int g_n d\mu\right)_n$ é uma sucessão crescente (em $\mathbb{R}_0^+ \cup \{+\infty\}$) majorada por $\int g d\mu$. Deste modo $\lim_n \int g_n d\mu$ existe (em $\mathbb{R}_0^+ \cup \{+\infty\}$) e $\lim_n \int g_n d\mu \leq \int g d\mu$.

Vejamos a desigualdade contrária.

Seja ϕ uma função simples tal que $0 \leq \phi \leq g$. Para $\varepsilon > 0$ e $n \in \mathbb{N}$ consideremos

$$A_{\varepsilon,n} = \{x \in X : g_n(x) \geq (1-\varepsilon)\phi(x)\}.$$

É então claro que, para todo $n \in \mathbb{N}$, $A_n \in \Sigma$ e $A_n \subseteq A_{n+1}$. Para além disso, como $\phi \leq g$ e $(g_n)_n$ converge para g, $X = \cup_n A_{n,\varepsilon}$.

Consideremos $\mu_\phi : \Sigma \to [0, \infty]$ definida por $\mu_\phi(A) = \int_A \phi\, d\mu$, para $A \in \Sigma$. Pela Proposição 5.48, esta função é uma medida. Deste modo, usando a alínea f) do Teorema 5.8,

$$\mu_\phi(X) = \mu_\phi\left(\cup_n A_{\varepsilon,n}\right) = \lim_n \mu_\phi(A_{\varepsilon,n}) \quad \text{ou seja} \quad \int \phi\, d\mu = \lim_n \int_{A_{\varepsilon,n}} \phi d\mu,$$

e, pela definição de $A_{\varepsilon,n}$,

$$\lim_n \int g_n\, d\mu \geq \lim_n \int_{A_{\varepsilon,n}} g_n\, d\mu \geq (1-\varepsilon) \lim_n \int_{A_{\varepsilon,n}} \phi\, d\mu = (1-\varepsilon) \int \phi\, d\mu.$$

Fazendo ε tender para 0, obtemos $\lim_n \int g_n\, d\mu \geq \int \phi\, d\mu$, para toda a função simples não negativa tal que $\phi \leq g$, ou seja $\lim_n \int g_n\, d\mu \geq \int g\, d\mu$. □

Estamos agora em condições de demonstrar a aditividade do integral. Se as funções em questão forem simples, esta demonstração não necessita do teorema da convergência dominada.

Corolário 5.50. *Se (X, Σ, μ) é um espaço de medida e $f, g : X \to \mathbb{R}$ são funções mensuráveis não negativas então*

$$\int_X (f+g) d\mu = \int_X f\, d\mu + \int_X g\, d\mu.$$

Demonstração. Suponhamos primeiro que f e g são simples. Pela Nota 5.46, existe uma partição de X, $(E_i)_{1\leq i\leq n}$ ($n \in \mathbb{N}$), com $E_1, \ldots, E_n \in \Sigma$, e $a_1, \ldots, a_n, b_1, \ldots, b_n \geq 0$ tais que $f = \sum_{i=1}^n a_i \chi_{E_i}$, $g = \sum_{i=1}^n b_i \chi_{E_i}$, e, portanto, $f + g = \sum_{i=1}^n (a_i + b_i)\chi_{E_i}$. O resultado pretendido segue da definição de integral de funções simples.

Vejamos agora o caso geral.

Usando o Teorema 5.44 sejam $(\phi_n)_n$ e $\psi_n)_n$ sucessões de funções simples tais que

$$0 \leq \phi_n \leq f, \quad \lim_n \phi_n = f, \quad 0 \leq \psi_n \leq g, \quad \lim_n \psi_n = g.$$

Usando o teorema anterior e o que já foi provado para funções simples, como $(\phi_n + \psi_n)_n$ converge para $f + g$

$$\int (f+g)\, d\mu = \int \lim_{n\to\infty} (\phi_n + \psi_n)\, d\mu = \lim_{n\to\infty} \int (\phi_n + \psi_n)\, d\mu = \lim_{n\to\infty} \left[\int \phi_n\, d\mu + \int \psi_n\, d\mu \right]$$

e a conclusão segue por aplicação do teorema anterior às sucessões $(\phi_n)_n$ e $(\psi_n)_n$. □

Corolário 5.51. *Se (X, Σ, μ) é um espaço de medida, $(f_n)_n$ é uma sucessão de funções mensuráveis não negativas e definidas em X e $f : X \to \mathbb{R}$ tal que $\sum_n f_n = f$ então*

$$\int_X f\, d\mu = \sum_{n=1}^{\infty} \int_X f_n\, d\mu.$$

Demonstração. Basta aplicar o teorema da convergência monótona à sucessão $(g_n)_n$ em que $g_n = \sum_{k=1}^{n} f_k$ e depois o corolário anterior. □

Como consequência deste corolário temos um resultado sobre as medidas da forma μ_f, definidas na Proposição 5.48.

Proposição 5.52. *Seja (X, Σ, μ) um espaço de medida e $f : X \to \mathbb{R}$ uma função mensurável não negativa e μ_f a medida associada a f (ver Proposição 5.48). Se $g : X \to \mathbb{R}$ é não negativa e mensurável então*

$$\int_X g\, d\mu_f = \int_X gf\, d\mu.$$

Demonstração. Vamos começar por mostrar que, se ϕ for uma função simples não negativa então $\int_X \phi\, d\mu_f = \int_X \phi f\, d\mu$. Consideremos a representação canónica de ϕ, $\phi = \sum_{i=1}^{n} \chi_{E_i}$. Deste modo

$$\int_X \phi f\, d\mu = \int_X \left(\sum_{i=1}^{n} a_i \chi_{E_i} \right) f\, d\mu = \sum_{i=1}^{n} a_i \int_X \chi_{E_i} f\, d\mu = \sum_{i=1}^{n} a_i \int_{E_i} f\, d\mu = \sum_{i=1}^{n} a_i \mu_f(E_i)$$

$$\int_X \phi\, d\mu_f = \int_X \left(\sum_{i=1}^{n} a_i \chi_{E_i} \right) d\mu_f = \sum_{i=1}^{n} a_i \int_X \chi_{E_i} d\mu_f = \sum_{i=1}^{n} a_i \int_{E_i} d\mu_f = \sum_{i=1}^{n} a_i \mu_f(E_i).$$

No caso geral escrevemos $g = \lim_n \phi_n$, em que $(\phi_n)_n$ é uma sucessão crescente e aplicamos o que acabamos de mostrar e o teorema da convergência monótona. □

O seguinte é um resultado que mostra, em particular, que uma das desigualdades em (5.4) é sempre verdadeira (se os limites referidos existirem)

Teorema 5.53 (Lema de Fatou)**.** *Sejam (X, Σ, μ) um espaço de medida e $(f_n)_n$ uma sucessão de funções mensuráveis reais não negativas definidas em X. Nestas condições*

$$\int \liminf_{n\to\infty} f_n\, d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

Demonstração. Para cada $n \in \mathbb{N}$, seja $g_n = \inf_{k\geq n} f_k$, que é uma função mensuráveis pela Proposição 5.34. Deste modo, como $(g_n)_n$ é crescente e $\lim_n g_n = \liminf_n f_n$, aplicando o teorema da convergência monótona, obtemos

$$\int \liminf_{n\to\infty} f_n\, d\mu = \lim_{n\to\infty} \int g_n d\mu.$$

Como $g_n \leq f_k$, para todo $k \geq n$, temos $\int g_n d\mu \leq \int f_k d\mu$ e, portanto, $\int g_n d\mu \leq \inf_{k\geq n} \int f_k d\mu$. Daqui e da igualdade anterior obtemos $\int \liminf_n f_n\, d\mu \leq \lim_n \inf_{k\geq n} \int f_k d\mu = \liminf_n \int f_n d\mu$. □

5.5.4 Funções integráveis

Recorde-se que, se (X, Σ) for um espaço mensurável e $f : X \to \mathbb{R}$ então $f = f^+ - f^-$ e $|f| = f^+ + f^-$. Deste modo, f é mensurável se e só se f^+ e f^- forem mensuráveis e que, se f for mensurável então $|f|$ também é. O contrário não é, em geral, verdadeiro. Por exemplo, se A não for mensurável, a função $f = \chi_A - \chi_{A^c}$ não é mensurável mas $|f|$ é mensurável.

Note-se ainda que, se $f = g - h$, com g e h não negativas, então $f^+ \leq g$ e $f^- \leq h$.

Definição 5.54. *Seja (X, Σ, μ) um espaço de medida e $f : X \to \mathbb{R}$. Se $\int f^+ d\mu$ $\boxed{\text{ou}}$ $\int f^- d\mu$ forem finitos define-se integral de f, denotado por $\int_X f d\mu$ ou simplesmente $\int f d\mu$, como o número, em $\mathbb{R} \cup \{-\infty, +\infty\}$,*

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu.$$

f diz-se integrável se $\int f^+ d\mu$ $\boxed{\text{e}}$ $\int f^- d\mu$ forem finitos.

Note-se que, nas condições da definição anterior, f é integrável se e só se $\int |f| d\mu$ é finito.

Notação. *Seja (X, Σ, μ) um espaço de medida. Denota-se por $\mathcal{L}^1(X, \mu)$, ou simplesmente $\mathcal{L}^1(X)$, se não houver dúvidas quanto à medida em questão, ao conjunto formado pelas funções integráveis de X em $\mathbb{R}$.*

O seguinte resultado é consequência de resultados análogos para funções não negativas. As duas primeiras alíneas dizem que $\mathcal{L}^1(X, \mu)$ é um espaço vectorial sobre $\mathbb{R}$ (relativamente à soma e multiplicação por um escalar usuais).

Teorema 5.55. *Sejam (X, Σ, μ) um espaço de medida, $f, g \in \mathcal{L}^1(X, \mu)$ e $c \in \mathbb{R}$. Então:*

a) $cf \in \mathcal{L}^1(X, \mu)$ e $\int (cf) d\mu = c \int f d\mu$;

b) $f + g \in \mathcal{L}^1(X, \mu)$ e $\int (f + g) d\mu = \int f d\mu + \int g d\mu$;

c) se $f \leq g$ quase sempre então $\int f d\mu \leq \int g d\mu$.

Demonstração. Comecemos por notar que

$$(cf)^+ = \begin{cases} cf^+ & \text{se } c \geq 0 \\ -cf^- & \text{se } c \leq 0, \end{cases} \qquad (cf)^- = \begin{cases} cf^- & \text{se } c \geq 0 \\ -cf^+ & \text{se } c \leq 0 \end{cases}$$
$$(f+g)^+ \leq f^+ + g^+, \qquad (f+g)^- \leq f^- + g^-.$$

Daqui decorre imediatamente que $cf, f+g \in \mathcal{L}^1(X, \mu)$ e a igualdade na alínea a), usando os resultados já conhecidos para funções não negativas. Para a alínea b) basta usar os resultados já conhecidos para funções não negativas, a definição de integral e a igualdade $(f+g)^+ - (f+g)^- = (f^+ - f^-) + (g^+ - g^-)$ (que apenas diz que $f + g = f + g$).

Para alínea c) basta notar que $g - f \geq 0$ quase sempre e, portanto $\int (g - f) d\mu \geq$ e a conclusão segue das alíneas anteriores. □

O seguinte resultado, muito simples, vai ser útil na demonstração do teorema que se segue.

Lema 5.56. *Seja (X, Σ, μ) um espaço de medida. Se $g \in \mathcal{L}^1(X, \mu)$ e $f : X \to \mathbb{R}$ é mensurável e $|f| \leq g$ então $f \in \mathcal{L}^1(X, \mu)$.*

Demonstração. Da desigualdade $|f| \leq g$ obtemos $f^+, f^- \leq g$. Deste modo o integral de f^+ e o de f^- são positivos e menores que o integral de g que é finito. Concluímos assim que $f^+, f^- \in \mathcal{L}^1(X, \mu)$ e, pela proposição anterior, $f \in \mathcal{L}^1(X, \mu)$. □

Estamos agora em condições de mostrar uma outra condição suficiente para que a igualdade (5.4) seja satisfeita.

Teorema 5.57 (Convergência dominada de Lebesgue)**.** *Sejam (X, Σ, μ) um espaço de medida e $(f_n)_n$ uma sucessão de funções reais definidas em X convergindo q.s. para uma função f. Se existe $g \in \mathcal{L}^1(X, \mu)$ tal que $|f_n| \leq g$ q.s., então $f \in \mathcal{L}^1(X, \mu)$ e*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Demonstração. Como $(f_n)_n$ converge q.s. para f então $|f| \leq g$ e, pelo lema anterior, $f \in \mathcal{L}^1(X, \mu)$. Aplicando o Lema de Fatou às sucessões $(g + f_n)_n$ e $(g - f_n)_n$ (que são não negativas) obtemos

$$\int g d\mu + \int f d\mu = \int (g + f) d\mu \leq \liminf_n \int (g + f_n) d\mu = \int g d\mu + \liminf_n \int f_n$$
$$\int g d\mu - \int f d\mu = \int (g - f) d\mu \leq \liminf_n \int (g - f_n) d\mu = \int g d\mu - \limsup_n \int f_n.$$

Como $\int g d\mu \in \mathbb{R}$ obtemos as desigualdades

$$\limsup_n \int f_n d\mu \leq \int f d\mu \leq \liminf_n \int f_n d\mu,$$

que são equivalentes à igualdade pretendida. □

Note-se que as sucessões de funções referidas nos exemplos após (5.4) não admitem uma função de $\mathcal{L}^1(X, \mu)$ que as "domine".

Estamos agora em condições de demonstrar que o integral de Lebesgue generaliza o integral de Riemann.

Teorema 5.58. *Seja I um intervalo fechado limitado de $\mathbb{R}$ e $f : I \to \mathbb{R}$ uma função limitada. Se f é integrável à Riemann então também é integrável à Lebesgue.*

Demonstração. Para cada partição $\mathcal{P} = (x_0, x_1, \ldots, x_n)$ de I denote-se por $S^f_{\mathcal{P}}$ e por $s^f_{\mathcal{P}}$ as funções simples

$$S^f_{\mathcal{P}} = \sum_{i=1}^{n} \big(\sup_{[x_{i-1},x_i]} f\big) \chi_{[x_{i-1},x_i]}, \quad s^f_{\mathcal{P}} = \sum_{i=1}^{n} \big(\inf_{[x_{i-1},x_i]} f\big) \chi_{[x_{i-1},x_i]},$$

e por $S(f, \mathcal{P})$ e $s(f, \mathcal{P})$ a soma superior e a soma inferior de f relativamente a $\mathcal{P}$, respectivamente.

Note-se que $s^f_{\mathcal{P}} \leq f \leq S^f_{\mathcal{P}}$ e, por definição de integral de Lebesgue de funções simples,

$$\int S^f_{\mathcal{P}} d\lambda = S(f, \mathcal{P}), \quad \int s^f_{\mathcal{P}} d\lambda = s(f, \mathcal{P}).$$

Como f é integrável á Riemann, existe uma sucessão $(\mathcal{P}_n)_n$ de partições de I tal que, se $n \in \mathbb{N}$, $\mathcal{P}_{n+1}$ é um refinamento de $\mathcal{P}_n$ e

$$\int_I f(x) dx = \lim_n S(f, \mathcal{P}_n) = \lim_n s(f, \mathcal{P}_n). \tag{5.6}$$

É claro que as sucessões $(s^f_{\mathcal{P}_n})_n$ e $(S^f_{\mathcal{P}_n})_n$ são convergentes pois a primeira é crescente e majorada por f e a segunda é decrescente e minorada por f. Deste modo os limites g, da primeira sucessão e h, da segunda, são funções mensuráveis. Como $s^f_{\mathcal{P}_n}$ e $S^f_{\mathcal{P}_n}$ são limitadas em módulo pela mesma constante que limita $|f|$ podemos usar o teorema da convergência dominada de Lebesgue para obter

$$\int g d\mu = \lim_n \int s^f_{\mathcal{P}_n} d\mu = \lim_n s(f, \mathcal{P}_n) = \int_I f(x) dx$$
$$\int h d\mu = \lim_n \int S^f_{\mathcal{P}_n} d\mu = \lim_n S(f, \mathcal{P}_n) = \int_I f(x) dx.$$

Daqui obtemos $\int (h - g) d\mu = 0$, o que implica, porque $h \geq g$, que $h = g$ quase sempre (relativamente a μ). Como $g \leq f \leq h$ concluímos que $g = f = h$ quase sempre. Deste modo f é integrável à Lebesgue e a conclusão final segue pois $f = g$ q. s. e $\int g d\mu = \int_I f(x) dx$. □

5.6 Espaços $L^p(X,\mu)$

Os espaços $L^p(X,\mu)$, com $1 \leq p \leq \infty$ vão ser definidos como um quociente de um espaço $\mathcal{L}^p(X,\mu)$ por uma relação de equivalência que identifica duas funções que sejam iguais quase sempre.

Comecemos pelo caso em que $p = \infty$.

Definição 5.59. *Seja (X,Σ,μ) um espaço de medida. Uma função mensurável $f : X \to \mathbb{R}$ diz-se essencialmente limitada se existe $M > 0$ tal que $|f| \leq M$ quase sempre. Nestas condições, denota-se por $\|f\|_\infty$ o supremo essencial de $|f|$, isto é,*

$$\|f\|_\infty = \inf\{M > 0 : |f| \leq M \ a.e. \ \}.$$

Definição 5.60. *Sejam (X,Σ,μ) um espaço de medida. Denota-se por $\mathcal{L}^\infty(X,\mu)$ o conjunto*

$$\mathcal{L}^\infty(X,\mu) = \{f : X \to \mathbb{R} : \ f \text{ é mensurável e essencialmente limitada}\}.$$

Definição 5.61. *Sejam (X,Σ,μ) um espaço de medida e $p \in [1,\infty[$. Denota-se por $\mathcal{L}^p(X,\mu)$ o conjunto*

$$\mathcal{L}^p(X,\mu) = \left\{f : X \to \mathbb{R} : \ f \text{ é mensurável e } \int |f|^p d\mu < \infty\right\},$$

e por $\|f\|_p$, o valor $\left(\int |f|^p d\mu\right)^{\frac{1}{p}}$.

As demonstrações dos resultados que vamos enunciar têm demonstração aos análogos para os espaços l^p. De facto os espaços l^p pode ser visto como o espaço de medida $(\mathbb{N},\mathcal{P}(\mathbb{N}),\mu)$ em que μ é a medida da contagem. Isto acontece porque uma sucessão real não é mais do que uma função $f : \mathbb{N} \to \mathbb{R}$ e que todas estas funções são mensuráveis no espaço de medida $(\mathbb{N},\mathcal{P}(\mathbb{N}),\mu)$. Resta assim ver que, se $f : \mathbb{N} \to \mathbb{R}$ então

$$\int_{\mathbb{N}} f d\mu = \sum_{n=1}^{\infty} f(n).$$

Deste modo, se $1 \leq p \leq \infty$ e $f : \mathbb{N} \to \mathbb{R}$ então $f \in \mathcal{L}(\mathbb{N},\mu)$ se e só se $\big(f(n)\big)_{n\in\mathbb{N}} \in l^p$.

Proposição 5.62. *Se (X,Σ,μ) é um espaço de medida e $p \in [1,\infty]$ então $\mathcal{L}^p(X,\mu)$ é um espaço vectorial sobre $\mathbb{R}$.*

Demonstração. Sejam $f, g \in \mathcal{L}^p(X,\mu)$ e $c \in \mathbb{R}$. Suponhamos primeiro que $p < \infty$. Então $cf, f+g \in \mathcal{L}^p(X,\mu)$ porque

$$\int |cf|^p d\mu = \int |c|^p| \int |f|^p d\mu < \infty$$
$$\int |f+g|^p d\mu \leq 2^{p-1}\left(\int |f|^p d\mu + \int |g|^p d\mu\right) < \infty$$

usando a desigualdade 1.1.

Se $p = \infty$, basta usar a igualdade $|cf| = |c||f|$, a desigualdade $|f+g| \leq |f| + |g|$ e o facto de a união de dois conjuntos com medida nula ter medida nula. □

Teorema 5.63 (Desigualdade de Hölder)**.** *Sejam (X,Σ,μ) um espaço de medida e $1 \leq p \leq \infty$. Se $f \in \mathcal{L}^p(X,\mu)$ e $g \in \mathcal{L}^q(X,\mu)$, sendo q o conjugado de p então $fg \in \mathcal{L}^1(X,\mu)$ e*

$$\|fg\|_1 \leq \|f\|_p\|g\|_q.$$

Demonstração. Suponhamos primeiro que $1 < p < \infty$. Se $\|f\|_p = 0$ (ou $\|g\|_p = 0$) então, pela alínea h) da Proposição 5.47, $|f|^p = 0$ (ou $\|g\|_p = 0=$ quase sempre e, portanto $|f| = 0$ (ou $|g| = 0$) quase sempre e a desigualdade pretendida é verdadeira.

Se $\|f\|_p \neq 0 \neq \|g\|_p = 0$, aplicamos a desigualdade de Young (Lema 1.16) a consideramos $\frac{f}{\|f\|_p}$ e $\frac{g}{\|g\|_p}$ e depois integramos para obter

$$\int \frac{|fg|}{\|f\|_p \, \|g\|_p} d\mu \leq \frac{1}{p} \int \frac{|f|^p}{\|f\|_p^p} + \frac{1}{q} \int \frac{|g|^q}{\|g\|_p^q} d\mu = \frac{1}{p} \frac{1}{\|f\|_p^p} \int |f|^p + \frac{1}{q} \frac{1}{\|g\|_p^q} \int |g|^q d\mu = \frac{1}{p} + \frac{1}{q} = 1$$

e, portanto $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Se $p = 1$ e $q = \infty$ então, como $|g| \leq \|g\|_\infty$ quase sempre, $\|fg\|_1 = \int |fg| d\mu \leq \int \|g\|_\infty |f| d\mu = \|g\|_\infty \int |f| d\mu = \|g\|_\infty \|f\|_1$. □

Em geral, dado um espaço de medida (X, Σ, μ), não é de esperar qualquer relação de inclusão entre $\mathcal{L}^p(X, \mu)$ e $\mathcal{L}^q(X, \mu)$ para $p \neq q$, embora isso possa acontecer em alguns casos particulares. Por exemplo, vimos na Proposição 1.20 que $l^p \subseteq l^q$, se $p \leq q$. Veremos de seguida que, se a medida for finita temos a desigualdade inversa.

Se estivermos em $\mathbb{R}$ com a medida de Lebesgue as funções constantes e as da forma $\frac{1}{x^a}\chi_{]0,1]}$ e $\frac{1}{x^a}\chi_{[1,\infty[}$ são suficientes para mostrar que, se $1 \leq q < p < \infty$, os espaços $\mathcal{L}^p(\mathbb{R}, \lambda)$ e $\mathcal{L}^q(\mathbb{R}, \lambda)$ não estão relacionados pela inclusão.

Proposição 5.64. *Se (X, Σ, μ) é um espaço de medida, μ é finita e $1 \leq q \leq p \leq \infty$ então $\mathcal{L}^p(X, \mu) \subseteq \mathcal{L}^q(X, \mu)$. Mais concretamente, se $f \in \mathcal{L}^p(X, \mu)$ então $f \in \mathcal{L}^q(X, \mu)$ e, interpretando $\frac{1}{\infty}$ como ∞,*

$$\|f\|_q \leq \mu(X)^{\frac{1}{q} - \frac{1}{p}} \, \|f\|_p.$$

Demonstração. Seja $f \in \mathcal{L}^p(X, \mu)$. Usando a desigualdade de Hölder no caso em que $p < \infty$, notando que o conjugado de $\frac{p}{q}$ é $\frac{p}{p-q}$, obtemos

$$\int |f|^q d\mu \begin{cases} = \int |f|^q \cdot 1 d\mu \leq \left(\int |f|^{q\frac{p}{q}} d\mu \right)^{\frac{q}{p}} \left(\int 1^{\frac{p}{p-q}} d\mu \right)^{\frac{p-q}{p}} & \text{se } p < \infty \\ \leq \int \|f\|_\infty^q d\mu = \|f\|_\infty^q \int 1 d\mu & \text{se } p = \infty, \end{cases}$$

e a desigualdade pedida segue pois $\int 1 d\mu = \mu(X)$. □

O seguinte resultado tem uma demonstração semelhante à demonstração do seu caso particular (ver Lema 1.18).

Teorema 5.65 (Desigualdade de Minkowski)**.** *Sejam (X, Σ, μ) um espaço de medida e $1 \leq p \leq \infty$. Se $f, g \in \mathcal{L}^p(X, \mu)$ então $f + g \in \mathcal{L}^p(X, \mu)$ e*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_q.$$

Demonstração. Basta na demonstração do Lema 1.18 substituir: x_i por g; y_i por g e os somatórios por integral relativamente a $d\mu$). □

Num espaço de medida (X, Σ, μ), se $1 \leq p \leq \infty$ e $f \in \mathcal{L}^p(X, \mu)$ então $\|f\|_p = 0$ se e só se $f = 0$ quase sempre (no caso $p = \infty$ é simplesmente a definição e no caso em que $p < \infty$ basta usar a alínea h) da Proposição 5.47). Deste modo, $\| \ \|_p$ só define uma norma se a igualdade quase sempre for equivalente à igualdade, ou equivalentemente, se o único subconjunto de X com medida nula for o conjunto vazio.

Recorde-se que espaço l^p pode ser visto como $\mathcal{L}^p(\mathbb{N}, \mu)$, em que μ é a medida da contagem. Por esta razão, se $A \subseteq \mathbb{N}$ então $\mu(A) = 0$ se e só se $A = \emptyset$. Daqui concluímos que $\mathcal{L}^p(\mathbb{N}, \mu)$ é um espaço vectorial normado.

Uma maneira de construir um espaço vectorial normado a partir de $\mathcal{L}^p(X, \mu)$ é definir uma relação de equivalência que identifique duas funções que são iguais quase sempre. Vamos chamar $\sim$ a esta relação que, facilmente se mostra ser uma relação de equivalência.

Definição 5.66. *Sejam* (X, Σ, μ) *um espaço de medida,* $1 \leq p \leq \infty$ *e* $\sim$ *a relação sobre* $\mathcal{L}^p(X, \mu)$ *definida por*

$$f \sim g \;\Leftrightarrow\; f = g \text{ quase sempre.}$$

Define-se $L^p(X, \mu)$ *como sendo* $\mathcal{L}^p(X, \mu)/\sim$.

Atendendo ao que foi feito aqui aqui temos o seguinte resultado.

Corolário 5.67. *Se* (X, Σ, μ) *e* $1 \leq p \leq \infty$ *então a função* $\begin{array}{rccc} \|\ \|_p : & L^p(X,\mu) & \to & \mathbb{R} \\ & f & \mapsto & \|f\|_p \end{array}$ *é uma norma.*

Teorema 5.68. *Se* (X, Σ, μ) *é um espaço de medida e* $1 \leq p \leq \infty$ *então* $L^p(X, \mu)$*, com a norma* $\|\ \|_p$ *é um espaço normado completo.*

Demonstração. Vejamos o caso em que $p < \infty$. O caso em que $p = \infty$ pode ser tratado como na demonstração de que o limite uniforme de funções contínuas é contínua.

Seja $(f_n)_n$ uma sucessão de Cauchy em $L^p(X, \mu)$. Para mostrar que a sucessão converge é suficiente mostrar que existe uma sua subsucessão convergente. Consideremos, para cada $k \in \mathbb{N}$, $\phi(k) \in \mathbb{N}$ tal que $\|f_n - f_{\phi(k)}\|_p \leq \frac{1}{2^k}$ para todo $n \geq \phi(k)$.

Para cada $n \in \mathbb{N}$ consideremos g_n definida por $g_n(x) = \sum_{k=1}^n |f_{\varphi(k+1)} - f_{\varphi(k)}|$. Como $(g_n)_n$ é uma sucessão crescente de funções não negativas e $g_n(x)| \leq \sum_{k=1}^n \frac{1}{2^k} = 1$ então, usando o teorema da convergência monótona $g = \lim_n g_n = \sum_{k=1} \infty |f_{\varphi(k+1)} - f_{\varphi(k)}|$ é mensurável e $\|g\|_p < \infty$. Como consequência a série $\sum_{k=1}^\infty (f_{\varphi(k+1)} - f_{\varphi(k)})$ é mensurável e pertence a $L^p(X, \mu)$.

Consideremos f definida por $f(x) = f_{\phi(1)} + \sum_{k=1}^\infty (f_{\varphi(k+1)} - f_{\varphi(k)})$. Como f é a soma de duas funções de $L^p(X, \mu)$ então $f \in L^p(X, \mu)$. Por outro lado $(f_{\phi(n)}(x))_n$ converge para $f(x)$ porque

$$(f_{\phi(n)}(x))_n = \left(f_{\phi(1)}(x) + \sum_{k=1}^n (f_{\varphi(k+1)}(x) - f_{\varphi(k)}(x)) \right)_n \to_n f(x).$$

Como $|f_{\phi(n)} - f| \leq |f_{\phi(n)}| + |f| \leq |f_{\phi(n)} - f_{\phi(1)}| + |f_{\phi(1)}| + |g| \leq 1 + |f_{\phi(1)}| + |g|$, podemos usar o teorema da convergência dominada para concluir que a sucessão $(f_{\phi(n)})_n$ converge para f. $\square$