Versão geométrica do teorema de Hahn-Banach

Eduardo Marques de Sá
Centro de Matemática da Universidade de Coimbra
Outubro 2013

Convexos e Funcionais de Minkowski

V é um espaço vetorial sobre $\mathbb{R}$. Dado um convexo $K \subseteq V$, a *bitola* ("*gauge*", em inglês) ou *funcional de Minkowski* de K é a função

$$\mu_K : V \to [0, +\infty], \quad \mu_K(x) := \inf\{\lambda > 0 : \ x \in \lambda K\}.$$

(Convenção: $\inf \varnothing = +\infty$.) Interessa apenas o caso em que, para cada x, o conjunto $\{\lambda > 0 : \ x \in \lambda K\}$ é não vazio; em tal caso, que suporemos sempre, o contradomínio de μ_K é $\mathbb{R}_+ = [0, +\infty[$. A funcional de Minkowski é *sub-linear*, o que significa ser

positivamente homogénea:	$\mu_K(\lambda x) = \lambda \mu_K(x)$, para $\lambda > 0$
e *sub-aditiva*:	$\mu_K(x + y) \leqslant \mu_K(x) + \mu_K(y)$.

Toda a funcional sub-linear, $\mu : V \to \mathbb{R}_+$, determina um conjunto $B_\mu := \{x : \mu(x) \leqslant 1\}$, chamado *bola unitária* de μ.[1]

Coisas giras de digestão fácil. Diz-se que a é *ponto interno* de K se a interseção de K com qualquer reta que passa por a é um intervalo do qual a é ponto não extremo;[2] o conjunto desses pontos denota-se $K^{\triangleright}$. Chama-se *fronteira*-1 de K ao conjunto das extremidades das interseções de K com retas de V. Diz-se que K é 1-*fechado* se contém toda a sua fronteira-1; e diz-se *encorpado*[3] se é 1-fechado e tem um ponto interno.

Exercícios

(a) Se K tem um ponto interno, são internos todos os seus pontos exceto os da fronteira-1.

(b) B_μ é um convexo encorpado e tem μ por funcional de Minkowski.

(c) $K \rightsquigarrow \mu_K$ produz uma bijecção do conjunto dos convexos encorpados sobre o conjunto das funcionais sub-lineares.

[1] Se, para certo $v \neq 0$, vale $\mu(v) \leqslant 0$, a semi-recta $\mathbb{R}_+ v$ está contida em B_μ.

[2] Em dimensão finita, *interno* é sinónimo de *interior* para a topologia canónica.

[3] Recorde-se que, em dimensão finita, chama-se *corpo convexo* a todo o convexo, compacto, de interior não vazio.

Conjuntos lineares, funcionais lineares e separação

Conjunto linear de V é todo o transladado dum subespaço de V, *i.e.*, um conjunto da forma $L = u+S$, onde $u \in V$ e S é subespaço de V. A *dimensão* de L é, por definição, a dimensão de S.

Vejamos o caso em que V tem dimensão finita n. *Hiperplano* é um conjunto linear de dimensão $n-1$. Seja $H = u + S$, onde $\dim S = n-1$. Fixemos $w \in V \setminus H$; então $V = \mathbb{R}w \oplus S$; isto significa que todo o $x \in V$ se escreve como $x = rw + s$, onde $r \in \mathbb{R}$ e $s \in S$ são unicamente determinados. Fica assim definida uma funcional $f : V \to \mathbb{R}$ tal que

$$x = f(x)\, w + s\,, \quad \text{onde } s \in S.$$

Claro que f é linear e S é o núcleo de f. Pondo $\alpha := f(u)$, tem-se

$$H = \{x : \ f(x) = \alpha\}. \tag{1}$$

Como há uma infinidade de escolhas de $w \in V \setminus H$, para todo o hiperplano existe uma infinidade de pares (f, α) para os quais vale (1). Usamos $H_{f,\alpha}$ para denotar $\{x : f(x) = \alpha\}$. Um hiperplano $H = H_{f,\alpha}$ origina quatro conjuntos, ditos *semi-espaços*, dados por:

$$S_{f\Box\alpha} = \{x : f(x)\Box\alpha\}$$

onde as duas ocorrências de $\Box$ representam um dos símbolos $\leqslant, \geqslant, <, >$. *Exercícios*: (*a*) $S_{f\leqslant\alpha}$ é complementar de $S_{f>\alpha}$; (*b*) $V = H \dot\cup S_{f<\alpha} \dot\cup S_{f>\alpha}$.

A representação (1) para os hiperplanos é a que costuma adotar-se como definição de hiperplano quando V tem dimensão arbitrária, finita ou não. Isso facilita o tratamento analítico. Por exemplo, diz-se que $H_{f,\alpha}$ *separa em sentido lato* dois conjuntos $X, Y \subset V$, se um deles está contido no semi-espaço $S_{f\leqslant\alpha}$ e o outro está contido em $S_{f\geqslant\alpha}$. Se neste enunciado utilizarmos os semi-espaços $S_{f<\alpha}$ e $S_{f>\alpha}$, obtemos o conceito de *separação em sentido estrito*. Estas definições têm outras *nuances* conforme a utilização que se faça dos símbolos $\leqslant, \geqslant, <, >$.

Outra definição muito popular: dizemos que $H_{f,\alpha}$ *separa fortemente* dois conjuntos $X, Y \subset V$, se existem reais r, s tais que: $r < \alpha < s$, um dos conjuntos está contido em $S_{f\leqslant r}$ e o outro está contido em $S_{f\geqslant s}$. Em tal caso, todos os hiperplanos $H_{f,t}$, com $r < t < s$, separam fortemente X e Y.

Axioma da Escolha

Eis três formas muito populares do axioma:

> Princípio da Escolha. *Todo o conjunto $\mathcal{C}$ de conjuntos não vazios tem uma função de escolha, i.e., existe uma função ε de domínio $\mathcal{C}$, tal que $\varepsilon(X) \in X$ para todo o $X \in \mathcal{C}$.*
>
> Princípio da Boa Ordenação. *Todo o conjunto pode ser ordenado de modo a que todo o subconjunto não vazio possua um elemento mínimo.*
>
> Princípio Maximal de Hausdorff. *Todo o conjunto X parcialmente ordenado tem uma cadeia maximal.*
>
> Lema de Zorn. *Todo o conjunto X parcialmente ordenado, no qual cada cadeia é superiormente limitada, tem um elemento maximal.*

Estes enunciados são equivalentes entre si, em face dos restantes axiomas habituais da teoria dos conjuntos. Usar um ou outro faz-se de acordo com a conveniência do utilizador. As implicações PE $\Rightarrow$ PBO, PE $\Rightarrow$ PMH e PE $\Rightarrow$ LZ devem-se, respetivamente, a E. Zermelo (1904), a F. Hausdorff (1914) e a K. Kuratowski (1922)-M. Zorn (1935). As recíprocas são óbvias.

Teoremas de Hahn-Banach

Sejam μ e φ funcionais sub-lineares em V. Dizemos que μ *domina* φ *em* S ($S \subseteq V$) se $\varphi(x) \leqslant \mu(x)$ para todo o $x \in S$. É importante notar que, no caso $S = V$, a dominância equivale a $B_\mu \subseteq B_\varphi$.

Teorema de extensão de Hahn-Banach. *Sejam μ uma funcional sublinear em V, S um subespaço de V e $\varphi : S \to \mathbb{R}$ uma funcional linear dominada por μ em S. Existe uma extensão linear de φ a V que é dominada por μ em V.*

Para demonstrar o teorema, começamos por um lema de extensão elementar.

Lema do Passo Indutivo. *Se $w \in V \setminus S$, então podemos estender φ a uma funcional linear $\psi : S + \mathbb{R}w \to \mathbb{R}$ dominada por μ em $S + \mathbb{R}w$.*

Demonstração. É fácil provar que $\varphi(x) - \mu(x - w) \leqslant \varphi(y) + \mu(y - w)$, para $x, y \in S$. Escolhe-se qualquer real a tal que

$$\sup_{x \in S} \big[\varphi(x) - \mu(x - w)\big] \leqslant a \leqslant \inf_{y \in S} \big[\varphi(y) + \mu(y - w)\big].$$

Define-se a funcional $\psi(s + \lambda w) = \varphi(s) + \lambda a$, para $s \in S$ e $\lambda \in \mathbb{R}$, a qual satisfaz todas as condições requeridas. $\square$

Demonstração do Teorema de Extensão. Podemos iterar o passo de indução obtendo sucessivas extensões de φ a espaços cada vez maiores. Mas o terminus deste processo não está ao alcance da intuição... precisa de um rigoroso passe de magia!

> Na teoria dos conjuntos, uma função $f : A \to B$ é o sub-conjunto de $A \times B$ constituído pelos (x, y) onde $x \in A$ e y é a imagem de x por f; e y costuma denotar-se por $f(x)$.[4] Diz-se que uma função g *estende* f se $f \subseteq g$; isto significa que $\mathrm{Dom}(f) \subseteq \mathrm{Dom}(g)$ e $f(x) = g(x)$ em $\mathrm{Dom}(f)$.

Seja $\mathfrak{F}$ o conjunto das funcionais lineares, com domínios $\subseteq V$, dominadas por μ e que estendem φ. Este conjunto está ordenado pela inclusão (= extensão).

Usando o Lema de Zorn. Tome-se uma cadeia $\mathscr{C}$ de $\mathfrak{F}$. A união de $\mathscr{C}$, $\bigcup_{f \in \mathscr{C}} f$, é funcional linear que estende φ e é dominada por μ. Portanto $\mathfrak{F}$ é indutivo. Pelo Lema de Zorn $\mathfrak{F}$ tem pelo menos um elemento maximal, m. Se $\mathrm{Dom}(m) \neq V$, o Passo Indutivo permitiria estender propriamente m a um elemento de $\mathfrak{F}$, contradizendo a maximalidade. Portanto $\mathrm{Dom}(m) = V$. $\square$

Usando o Princípio Maximal de Hausdorff. Seja $\mathscr{M}$ uma cadeia maximal de $\mathfrak{F}$. É fácil ver que a união de $\mathscr{M}$ é elemento maximal de $\mathfrak{F}$. A prova pode terminar-se, pois, como anteriormente. $\square$

A utilização direta do Princípio da Escolha produziria uma prova de grande complexidade (o óbvio consiste em provar previamente o Lema de Zorn...).

Teorema de Hahn-Banach, em versão geométrica. *Sejam $K \subseteq V$ um convexo com um ponto interno e L um conjunto linear que não interseta $K^{\triangleright}$. Existe um hiperplano extensão de L que não interseta $K^{\triangleright}$.*

Demonstração da versão geométrica. Sem perda de generalidade podemos supor que 0 é ponto interno de K. Sejam S subespaço de V e f uma funcional linear em S tais que $L = \{s \in S : \ f(s) = 1\}$. A funcional de Minkowski μ_K domina f em S. Por H-B existe uma extensão linear de f, digamos $\bar{f} : V \to \mathbb{R}$ dominada por μ em V. Então $\bar{L} := \{x \in V : \ \bar{f}(x) = 1\}$ contém L e a dominância significa que $\bar{L}$ não interseta $K^{\triangleright}$. $\square$

[4] Na análise costuma chamar-se *gráfico* de f a este conjunto de pares. Mas na teoria dos conjuntos a função e o seu gráfico são a mesma coisa.

Veja-se a consequência óbvia, mas subtil: todo o ponto u não interno a K produz um conjunto linear, $L = \{u\}$, que não interseta $K^{\triangleright}$. O teorema garante a existência dum hiperplano H que passa por u mas não toca $K^{\triangleright}$; portanto H separa u de K em sentido lato. Em particular, todo o ponto u da fronteira-1 de K pode separar-se latamente de K por um hiperplano H; este H é uma espécie de hiperplano "tangente"a K no ponto u.

Normas, Continuidade e dominância

Chama-se *norma* em V a uma funcional μ que é

$$\begin{array}{ll} \textit{homogénea}: & \mu(\lambda x) = |\lambda| \mu(x) \\ \textit{subaditiva}: & \mu(x+y) \leqslant \mu(x) + \mu(y). \\ \textit{não degenerada}: & \mu(x) = 0 \Rightarrow x = 0. \end{array}$$

Isto equivale a acrescentar à sublinearidade de μ o facto de B_μ ser simétrica ($B_\mu = -B_\mu$) e não conter semi-retas. Uma norma induz uma distância entre vetores e, *a fortiori*, uma topologia em V.[5]

Seja φ uma funcional em V. A continuidade de φ em $0 \in V$ escreve-se:

$$\forall_{\epsilon>0}\ \exists_{\delta>0}\ \forall_x \quad \mu(x) \leqslant \delta \Rightarrow |\varphi(x)| \leqslant \epsilon.$$

Se φ é positivamente homogénea, isto equivale à existência dum $C \geqslant 0$ tal que $|\varphi|$ é dominada por $C\mu$. Vamos admitir que φ é linear. Então as coisas simplificam-se muito pois φ é contínua na origem sse é globalmente contínua; sse $|\varphi|$ é limitada na bola unitária B_μ. Define-se, então,

$$\|\varphi\|_\mu := \sup_{x \in B_\mu} |\varphi(x)|\,.$$

A funcional $\|\cdot\|_\mu$ é uma norma em V^*, o espaço das funcionais lineares *contínuas* em V. Note-se a dominância $|\varphi| \leqslant \|\varphi\|_\mu \mu$. De facto, $\|\varphi\|_\mu$ é a menor das constantes C tais que $|\varphi|$ é dominada por $C\mu$.

> *Problema.* Provar que uma funcional linear f em V é contínua sse B_f tem um ponto interior, sse $H_{f,1}$ é fechado.

[5] Distância$(u,v) := \mu(u-v)$. Por definição, diz-se *aberto* todo o conjunto $A \subseteq V$ tal que $\forall_{a\in A}\ \exists_{\delta>0}\ a + \delta B_\mu \subset A$.

Teorema de separação de Hahn-Banach. *Sejam A, B convexos disjuntos do espaço normado V, com norma $\|\cdot\|$. Se A é aberto, existe uma funcional linear contínua, $f : V \to \mathbb{R}$, e um real r, tais que $A \subseteq S_{f<r}$ e $B \subseteq S_{f\geqslant r}$.*

Demonstração. Fixemos $a_0 \in A$, $b_0 \in B$, e seja $x_0 = b_0 - a_0$. A origem pertence ao conjunto $K := x_0 + A - B$, que é convexo e aberto; portanto, a origem é ponto interno de K, pelo que podemos aplicar o teorema de extensão de H-B com $\mu = \mu_K$, a funcional de Minkowski de K.

Seja S o subespaço $\mathbb{R}x_0$ e defina-se $\varphi : S \to \mathbb{R}$ por $\varphi(\lambda x_0) = \lambda$. Como $x_0 \notin K$, $\mu(x_0) \geqslant 1$; portanto μ domina φ em S. Por H-B, existe uma funcional linear f em V que estende φ e é dominada por μ.

O semi-espaço B_f contém B_μ; como 0 é ponto interior de B_μ, existe $\delta > 0$ tal que B_f contém $\delta B_{\|\cdot\|}$; portanto f é contínua.

Sejam $a \in A, b \in B$. Como $a-b+x_0 \in K$, vale $f(a-b+x_0) \leqslant \mu(a-b+x_0) \leqslant 1$; como $f(x_0) = 1$, temos $f(a) \leqslant f(b)$. Assim, sendo

$$\alpha = \sup_{a\in A} f(a) \quad \text{e} \quad \beta = \inf_{b\in B} f(b),$$

verifica-se $\alpha \leqslant \beta$. É fácil provar que o $\sup_{a\in A} f(a)$ não é atingido em A, pelo que vale $f(a) < \alpha \leqslant \beta \leqslant f(b)$. O teorema vale para qualquer $r \in [\alpha, \beta]$. □

Um caso simples. Seja $V = \mathbb{R}^2$, A o semiplano dos (x, y) tais que $y < 0$ e B o complementar de A. Seguindo a notação da prova anterior, escolhamos $a_0 = (0, -1)$ e $b_0 = (0, 0)$. Temos $x_0 = (0, 1)$ e K é o semiplano definido por $y < 1$. Temos $\mu(x, y) = \max\{0, y\}$, e φ está definida no eixo OY por $\varphi(0, y) = y$. As extensões lineares de φ são da forma $f(x, y) = cx + y$; a dominação por μ obriga $c = 0$. Portanto há apenas uma extensão de φ dominada por μ (que corresponde a uma só reta separadora de K e x_0). Feitas as contas, $\alpha = \beta = 0$, donde a reta de equação $f(x, y) = 0$ — o eixo OX— separa oa dois semiplanos, estando contida num deles!