Il teorema di Hahn-Banach

Stefano Pascolutti

4 aprile 2005

1 Il teorema di Hahn - Banach – Forma analitica

Sia X uno spazio vettoriale su $\mathbb{R}$. Il teorema di Hahn - Banach riguarda il prolungamento di forme lineare definite su un sottospazio vettoriale di X. Con questo teorema, sotto opportune condizioni, potremo estendere la forma definita sul sottospazio a una forma lineare definita su tutto X.

Teorema 1.1 (di Hahn - Banach, forma analitica). *Sia* $p\colon X \to \mathbb{R}$ *una applicazione che verifichi*

$$p(\lambda x) = \lambda p(x) \quad \forall x \in X \quad e \quad \forall \lambda > 0 \tag{1}$$

$$p(x+y) \leq p(x) + p(y) \quad \forall x, y \in X \tag{2}$$

Sia inoltre $X_0 \subset X$ *un sottospazio vettoriale e sia* $f_0\colon X_0 \to \mathbb{R}$ *una applicazione lineare tale che*

$$f_0(x) \leq p(x) \quad \forall x \in X_0. \tag{3}$$

Allora esiste una forma lineare $\tilde{f}$ *definita su* X *che prolunga* f_0*, cioè*

$$\tilde{f}(x) = f_0(x) \quad \forall x \in X_0$$

e tale che

$$\tilde{f}(x) \leq p(x) \quad \forall x \in X. \tag{4}$$

Dimostrazione. Diciamo che $g\colon W \to \mathbb{R}$ è una estensione ammissibile per f_0 se $X_0 \subseteq W \subseteq X$ e $g|_{X_0} = f_0$ e $g(x) \leq p(x)$ per ogni $x \in X$. Consideriamo $\mathcal{E} = \{g\colon g \text{ è una estensione ammissibile per } f_0\}$. $\mathcal{E} \neq \emptyset$, visto che $f_0 \in \mathcal{E}$. Inoltre $\mathcal{E}$ è dotato dell'ordine parziale $\prec$: $g_1 \prec g_2$ se e solo se $\operatorname{dom} g_1 \subseteq \operatorname{dom} g_2$ e $g_2|_{\operatorname{dom} g_1} = g_1$. $(\mathcal{E}, \prec)$ è un insieme induttivo perché data una catena $g_1 \prec g_2 \prec \ldots \prec g_n \ldots$, l'elemento maggiorante sarà $\mathcal{G}$ dove $\operatorname{dom} \mathcal{G} = \bigcup_n \operatorname{dom} g_n$

e $\mathcal{G}(x) = y$ se e solo se $x \in \operatorname{dom} g_i$ e $g_i(x) = y$ per qualche i. Per il lemma di Zorn $(\mathcal{E}, \prec)$ ammette elemento massimale $\tilde{f}$. Basta dimostrare che ogni estensione ammissibile a un sottospazio proprio di X può essere ulteriormente prolungata in modo ammissibile.
Sia, infatti, $X_0 \subseteq W \subset X$. Sia $z \in X \backslash W$ e prendiamo $W \subset \widehat{W} = \operatorname{span}(W, z)$. Quindi, ogni $y \in \widehat{W}$ si scrive come $y = w + tz$, con $w \in W$ e $t \in \mathbb{R}$. Sia $h \colon \widehat{W} \to \mathbb{R}$ tale che $h(y) = h(w + tz) = h(w) + th(z)$. Vogliamo che h sia una estensione ammissibile per g, con $g \in E$ e $\operatorname{dom} g = W$. In particolare, $h(w) = g(w)$. Dobbiamo assegnare un valore a $h(z)$, in modo che tutto torni, cioè $h(y) \le p(y)$ per ogni $y \in \widehat{W}$.
Se $t = 0$ non c'è nessun problema.
Sia $h(z) = c \in \mathbb{R}$: vogliamo che

$$h(w + tz) \le p(w + tz).$$

Sia inizialmente $t > 0$. Allora $h(w + tz) = th(\frac{w}{t} + z) = t(h(\frac{w}{t}) + h(z)) = t(g(\frac{w}{t}) + c)$. D'altra parte, $p(w + tz) = tp(\frac{w}{t} + z)$. Quindi, ponendo $w_1 = \frac{w}{t}$ e dividendo per t le due condizioni appena trovate:

$$c \le p(w_1 + z) - g(w_1) \qquad \forall w_1 \in W \tag{5}$$

D'altra parte, sia $t < 0$. Allora sia $s = -t > 0$. Quindi $h(w - sz) = sh(\frac{w}{s} - z) = s(h(\frac{w}{s}) - h(z)) = s(g(\frac{w}{s}) - c)$. Ma $p(w - sz) = sp(\frac{w}{s} - z)$. Quindi, ponendo $w_2 = \frac{w}{s}$ e dividendo per s le due condizioni appena trovate:

$$c \ge g(w_2) - p(w_2 - z) \qquad \forall w_2 \in W \tag{6}$$

Mettendo assieme (5) e (6), otteniamo che:

$$g(w_2) - p(w_2 - z) \le c \le p(w_1 + z) - g(w_1) \qquad \forall w_1, w_2 \in W. \tag{7}$$

Basta vedere che esiste un c che soddisfa (7). E' sufficiente che $g(w_2) - p(w_2 - z) \le p(w_1 + z) - g(w_1)$, cioè $g(w_1) + g(w_2) \le p(w_1 - z) + p(w_2 + z)$. Ma questo è vero perché p e g sono lineari nella somma e quindi: $g(w_1 + w_2) \le p(w_1 + z + w_2 - z) = p(w_1 + w_2)$, che è vero per ipotesi. $\square$

Esiste una estensione del teorema di Hahn-Banach che oltre a estendere il funzionale, conserva una invarianza rispetto all'azione di un semigruppo abeliano di operatori:

Teorema 1.2 (di Hahn-Banach - funzionali invarianti per un semigruppo abeliano di operatori). *Siano X, X_0, p, f_0 come nel teorema precedente. Sia G un semigruppo abeliano di operatori lineari tali che valgano le seguenti proprietà:*

1. $A(X_0) \subseteq X_0 \qquad \forall A \in G$

2. $f_0(A(x)) = f_0(x) \qquad \forall A \in G$

3. $p(A(x)) \leq p(x) \qquad \forall x \in X, \forall A \in G$

Allora esiste $\tilde{f}$ *come nel teorema (1.1) che in più soddisfa* $\tilde{f}(A(x)) = \tilde{f}(x)$ *per ogni* $x \in X$*.*

Non dimostrerò il teorema ma indicherò una applicazione:

1.1 Il limite di Banach

Consideriamo l'insieme l_∞ delle successioni limitate a valori in $\mathbb{R}$. Quindi $l_\infty = \{(x_n)_n \colon |x_n| < +\infty \quad \forall n \in \mathbb{N}\}$. Non possiamo certo assegnare un limite a ogni successione: la successione $(-1)^n \in l_\infty$ è limitata ma non ammette limite (ha una sottosuccessione convergente a $+1$ e una sottosuccessione convergente a -1). Inoltre l_∞ è uno spazio vettoriale di dimensione infinita su $\mathbb{R}$. Consideriamo il sottospazio delle successioni che hanno limite: $\bar{l}$. $\bar{l}$ è effettivamente un sottospazio di l_∞ e ammette, come semigruppo abeliano di operatori G quello delle traslazioni: $(x_n)_n \xrightarrow{L_k} (x_{n+k})_n$. È un semplice esercizio verificare che se $(x_n)_n \in \bar{l}$ allora anche $L_k((x_n)_n) \in \bar{l}$. Sia quindi $l_\infty = X$, $\bar{l} = X_0$. Rimangono da introdurre f_0 e p, come nel teorema (1.2):

1. $f_0 = \lim\limits_{n\to\infty}$. Infatti abbiamo che $\lim\limits_{n\to\infty} : \bar{l} \to \mathbb{R}$. Inoltre, per le proprietà del limite, questo è lineare. Inoltre si ha che $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} x_{n+k}$. Pertanto f_0 è invariante per l'azione di G su $\bar{l}$

2. $p : \limsup\limits_{n\to\infty}$. Infatti $\limsup\limits_{n\to\infty} : l_\infty \to \mathbb{R}$; per definizione di $\limsup\limits_{n\to\infty}$ e per le proprietà del sup, è un operatore lineare e maggiora il $\lim\limits_{n\to\infty}$ nel caso di successioni con limite. Vale anche la proprietà 3) e la verifica è banale.

Pertanto si può estendere il funzionale $\lim\limits_{n\to\infty}$ a tutto l_∞. Sia quindi $\tilde{f} = \mathrm{Lim}$. Per il teorema di Banach, questo funzionale ha le proprietà desiderate ed estende il limite solito. Tanto per fare un esempio, calcoliamo il $\mathrm{Lim}(-1)^n$: per traslazione, $\mathrm{Lim}(-1)^n = \mathrm{Lim}(-1)^{n+1}$, e quindi $\mathrm{Lim}(-1)^n - \mathrm{Lim}(-1)^{n+1} = 0$. Per linearità, $\mathrm{Lim}(-1)^n + \mathrm{Lim}(-1)^n = 2\,\mathrm{Lim}(-1)^n = 0$. Da cui $\mathrm{Lim}(-1)^n = 0$. Purtroppo, l'operatore Lim non ammette l'unicità del limite. Ma questo è un altro discorso.

2 Il teorema di Hahn-Banach – Forma geometrica

Vediamo ora due versioni geometriche del teorema di Hahn-Banach, che ci danno informazioni sulla topologia dello spazio che stiamo considerando. In particolare, vedremo che due convessi aperti non vuoti disgiunti sono sempre separati da un iperpiano. Cosa significa questo? Vediamo le definizioni.

Definizione 2.1. *Sia X uno spazio vettoriale su $\mathbb{R}$ e sia $f \in X^\star$, cioè un funzionale lineare da X in $\mathbb{R}$ (non necessariamente continuo). Un iperpiano di X è un insieme $H = \{x \in X : f(x) = \alpha\}$, per qualche $\alpha \in \mathbb{R}$. Diremo che H è l'iperpiano di equazione $[f = \alpha]$.*

Lemma 2.1. *L'iperpiano di equazione $[f = \alpha]$ è chiuso sse f è continua.*

La dimostrazione di un verso è banale, l'altra non è utile ai fini della dimostrazione del teorema di Hahn-Banach. Nel seguito, tutti gli iperpiani che considereremo saranno chiusi. Per praticità li chiamerò semplicemente iperpiani.

Definizione 2.2. *Siano $A, B \subseteq X$. Diciamo che l'iperpiano H di equazione $[f = \alpha]$ separa A e B in senso* largo *se*

$$f(x) \leq \alpha \quad \forall x \in A \qquad e \qquad f(x) \geq \alpha \quad \forall x \in B. \tag{8}$$

Diciamo che H separa A e B in senso stretto *se esiste un $\varepsilon > 0$ tale che*

$$f(x) \leq \alpha - \varepsilon \quad \forall x \in A \qquad e \qquad f(x) \geq \alpha + \varepsilon \quad \forall x \in B. \tag{9}$$

Prima di procedere con il teorema vero e proprio, vediamo altri due lemmi che risulteranno utili per la dimostrazione del teorema.

Lemma 2.2 (jauge di un convesso). *Sia $C \in X$ un convesso aperto e $0 \in C$. Per ogni $x \in X$ definiamo*

$$p(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in C\}.$$

Allora verifichiamo le seguenti proprietà per $p : X \to \mathbb{R}_0$:

1. $p(\lambda x) = \lambda p(x) \quad \forall x \in X$
2. $p(x + y) \leq p(x) + p(y) \quad \forall x, y \in X$
3. $\exists M > 0$ *tale che* $0 \leq p(x) \leq M \|x\|$

4. $C = \{x \in X : p(x) < 1\}$.

Una tale funzione si dice la jauge *del convesso C.*

Dimostrazione. La 1) è banale, date le proprietà del sup. Vediamo di dimostrare la 3). $0 \in C$, per ipotesi. Dal fatto che C è aperto, segue che esiste un $r > 0$ tale che $B(0, r) \subseteq C$. Quindi, posto $M = \frac{1}{r}$, si ha che $p(x) \leq M \|x\|$ per ogni $x \in X$.
Proviamo la 4). Sia $x \in C$. Allora esiste un $\varepsilon > 0$ tale che $x(1 + \varepsilon) \in C$. In particolare, $p(x) \leq \frac{1}{1+\varepsilon} < 1$. Se invece $p(x) < 1$, allora, in particolare, $1 \cdot x = x \in C$.
Vediamo infine la 2), che richiede un po' di lavoro in più. Per ogni $\varepsilon > 0$ si ha che $\frac{x}{p(x)+\varepsilon} \in C$ e $\frac{y}{p(y)+\varepsilon} \in C$. Essenso C convesso, anche una combinazione lineare di $\frac{x}{p(x)+\varepsilon}$ e $\frac{y}{p(y)+\varepsilon}$ sta in C. Quindi, per $t \in [0, 1]$, si ha che $t\frac{x}{p(x)+\varepsilon} + (1-t)\frac{y}{p(y)+\varepsilon} \in C$. In particolare, se $t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$, si ha che $\frac{x+y}{p(x)+p(y)+2\varepsilon} \in C$. E quindi $p(x + y) \leq p(x) + p(y) + 2\varepsilon$, da cui la tesi. □

Lemma 2.3. *Sia $C \subset X$ e $x_0 \in C^c$. Allora esiste $f \in X^\star$ tale che $f(x) < f(x_0)$ per ogni $x \in C$. Cioè l'iperpiano di equazione $[f = f(x_0)]$ separa $\{x_0\}$ e C in senso largo.*

Dimostrazione. Per traslazione, possiamo supporre che $0 \in C$. A questo punto, introduciamo la jauge di C e consideriamo $X_0 = \mathbb{R}x_0 = \{tx_0 : t \in \mathbb{R}\}$. Prendiamo la funzione $g : X_0 \to \mathbb{R}$, definita nel modo seguente:

$$g(tx_o) = tp(x_0).$$

Una tale funzione è lineare e ovviamente $g(x) \leq p(x)$ per ogni $x \in X_0$. Applichiamo il teorema (1.1) alla funzione f_0. Esiste un funzionale $f \in X^\star$ come nella tesi del teorema. Vediamo che caratteristiche ha questo funzionale. Sicuramente $f(x) \leq p(x)$ per ogni $x \in X$. Pertanto $f(x) < 1$ per ogni $x \in C$ per la proprietà 4) del lemma (2.2). D'altra parte, $f(x_0) = p(x_0) > 1$. Inoltre, grazie al punto 3) del lemma (2.2), possiamo concludere che la f è anche continua (è limitata in norma). Quindi f determina un iperpiano e questo separa $\{x_0\}$ e C. □

Siamo pronti, ora, a enunciare e dimostrare il teorema di Hahn-Banach in forma geometrica.

Teorema 2.1 (di Hahn-Banach – Prima forma geometrica). *Siano $A, B \subset X$ convessi, non vuoti disgiunti. Sia A aperto. Allora esiste un iperpiano che separa A e B in senso largo.*

Dimostrazione. Consideriamo la differenza algebrica tra i due insiemi: $C = A - B$. Siccome A e B sono disgiunti, allora $0 \notin C$. Inoltre C è convesso. Non solo: $C = \bigcup_{b \in B}(A - b)$ è una unione qualsiasi di aperti e quindi è aperto. Per il lemma (2.3), esiste una funzione $f \in X^\star$ tale che $[f = f(0)]$ separa $\{0\}$ e C. In particolare, $f(c) < 0$ per ogni $c \in C$. Per linearità, $f(a) < f(b)$ per ogni $a \in A$ e $b \in B$. Pertanto esiste α tale che:

$$\limsup_{a \in A} f(a) \leq \alpha \leq \liminf_{b \in B} f(b)$$

Allora l'iperpiano di equazione $[f = \alpha]$ separa A e B in senso largo. $\square$

Esiste un'altra versione geometrica del teorema di Hahn-Banach

Teorema 2.2 (di Hahn-Banach – Seconda forma geometrica). *Siano $A, B \subset X$ due insiemi convessi, non vuoti, disgiunti. Sia A chiuso e B compatto. Allora esiste un iperpiano che separa A e B in senso stretto.*

Dimostrazione. Consideriamo $A_\varepsilon = A + B(0, \varepsilon)$ e $B_\varepsilon = B + B(0, \varepsilon)$, con $\varepsilon > 0$. Siccome A è chiuso e B è compatto, allora esiste un ε per cui $A_\varepsilon \cap B_\varepsilon = \emptyset$. Infatti, se così non fosse si potrebbero prendere tre successioni ε_n, a_n, b_n, con $\varepsilon_n > 0$ e decrescenti, $a_n \in A$ e $b_n \in B$ in modo che $|a_n - b_n| \leq 2\varepsilon_n$. Ponendo $c_n = a_n - b_n$, si ha che $a_n = b_n + c_n$. Ora, $c_n \leq 2\varepsilon_n$ e quindi tende a 0. b_n è una successione in un compatto e quindi ha una sottosuccessione convergente a $\tilde{b} \in B$. Quindi anche a_n converge a $\tilde{b}$. Ma A è chiuso e quindi $\tilde{b}$ sta in A. Quindi $\tilde{b} \in A \cap B$. Assurdo. Quindi, per un certo ε, siamo nelle condizioni del teorema (2.1) e pertanto esiste un iperpiano che separa A_ε e B_ε. Cioè esistono f, α tali che l'iperpiano $[f = \alpha]$ separa in senso largo A_ε e B_ε. $f(a + \varepsilon c) \leq \alpha \leq f(b + \varepsilon c$ per ogni $a \in A, b \in B, c \in B(0, 1)$. Ma allora $f(a) + \varepsilon \|f\| \leq \alpha \leq f(b) - \varepsilon \|f\|$. Da cui la tesi. $\square$