P. Maurer

ENS Rennes

Recasages : 181, 208, 253.

Référence : Rouvière, Petit guide du calcul différentiel (pour le lemme surtout).

Inspiré du travail de Théo Pierron.

Théorèmes de Hahn-Banach en dimension finie

Le développement est constitué d'un lemme sur la jauge d'un ouvert connexe, du théorème de Hahn-Banach analytique et de sa version géométrique.

Dans la leçon 181, on insistera sur la preuve du lemme et de la version géométrique en allant vite (voire en admettant) la version analytique. A l'inverse, dans les leçons 208 et 253, on pourra admettre le lemme et se concentrer sur les deux théorèmes de Hahn-Banach.

On se donne $(E, \|\cdot\|)$ un espace vectoriel normé sur $\mathbb{R}$, de dimension finie $n \in \mathbb{N}^*$.

Définition 1.

Soit $\mathcal{C}$ un ouvert convexe de E contenant zéro. On définit la jauge de $\mathcal{C}$ comme l'application

$$j_C\colon \begin{cases} E & \to & \mathbb{R}_+ \\ x & \mapsto & \inf\{\lambda > 0 \,:\, x \in \lambda\mathcal{C}\} \end{cases}.$$

Lemme 2. *La jauge de $\mathcal{C}$ est bien définie sur E. Elle vérifie les propriétés suivantes, pour x et y des vecteurs de E :*

1. $\mathcal{C} = \{x \in E \,:\, j_C(x) < 1\}$,
2. $\forall \mu > 0 \quad j_C(\mu x) = \mu j_C(x)$, *($j_C$ est positivement homogène)*
3. $\exists M \geq 0,\ \forall x \in E \quad 0 \leq j_C(x) \leq M\,\|x\|$,
4. $j_C(x+y) \leq j_C(x) + j_C(y)$. *($j_C$ est sous-additive)*

Démonstration.

Comme $\mathcal{C}$ est un ouvert contenant zéro, il existe $r > 0$ tel que $B(0, r) \subset \mathcal{C}$. Aussi, pour $x \in E$ et $\lambda > \frac{\|x\|}{r}$, on a $\left\|\frac{x}{\lambda}\right\| < r$ donc $x \in \lambda\mathcal{C}$. On en déduit que $I(x) := \{\lambda > 0 \,:\, x \in \lambda\mathcal{C}\}$ est non vide, donc il admet une borne inférieure.

1. Comme $\mathcal{C}$ est ouvert, $I(x)$ est un ouvert de $]0, +\infty[$. En effet, c'est l'image réciproque de $\mathcal{C}$ par l'application continue $\lambda \mapsto \frac{1}{\lambda}x$ de $]0, +\infty[$ dans E.

   Par ailleurs, $I(x)$ est une demi-droite. En effet, pour $\lambda \in I(x)$ et $\mu \geq \lambda$, on a

$$\begin{aligned} \frac{1}{\mu}x &= \left(\frac{\lambda}{\mu}\right) \cdot \frac{1}{\lambda}x \\ &= \left(\frac{\lambda}{\mu}\right) \cdot \frac{1}{\lambda}x + \left(1 - \frac{\lambda}{\mu}\right) \cdot 0 \\ &\in \mathcal{C}. \end{aligned}$$

On en déduit que $I(x)$ est de la forme $]j_C(x), +\infty[$. Par définition de $I(x)$, x appartient à $\mathcal{C}$ si et seulement si $1 \in I(x)$, c'est-à-dire si et seulement si $1 > j_C(x)$. On a donc bien

$$\mathcal{C} = \{x \in E \;:\; j_C(x) < 1\}.$$

2. Soit $\mu > 0$. Montrons que $j_C(\mu x) = \mu j_C(x)$.

    - $\boxed{j_C(\mu x) \geq \mu j_C(x).}$

      Soit $\lambda \in I(\mu x)$. Alors $\dfrac{\mu x}{\lambda} \in \mathcal{C}$, donc $\dfrac{\lambda}{\mu} \in I(x)$, donc $\dfrac{\lambda}{\mu} \geq j_C(x)$. Par passage à la borne inférieure, on en déduit que $j_C(\mu x) \geq \mu j_C(x)$.

    - $\boxed{j_C(\mu x) \leq \mu j_C(x).}$

      Soit $\lambda \in I(x)$. Alors $\dfrac{x}{\lambda} \in \mathcal{C}$ donc $\dfrac{\mu x}{\mu \lambda} \in \mathcal{C}$ donc $\mu \lambda \in I(\mu x)$, donc $\mu \lambda \geq j_C(\mu x)$. Par passage à la borne inférieure, on en déduit que $\mu j_C(x) \geq j_C(\mu x)$.

3. Montrons qu'il existe $M \geq 0$ tel que $\forall x \in E \quad 0 \leq j_C(x) \leq M \, \|x\|$. Ce résultat est trivial pour $x = 0$, on suppose donc x non nul.

   On se donne $r > 0$ tel que $B(0, r) \subset \mathcal{C}$. Pour $x > 0$ et $\varepsilon \in]0, r[$, on a

   $$\frac{r - \varepsilon}{\|x\|} \cdot x \;\in\; \mathcal{C}.$$

   Donc par définition de la jauge, $j_C(x) \leq \dfrac{\|x\|}{r - \varepsilon}$. Par passage à la limite quand $\varepsilon \to 0$, on en déduit que $j_C(x) \leq \dfrac{1}{r} \|x\|$.

4. Soit $x, y \in E$, et $\varepsilon > 0$. Par caractérisation de la borne inférieure, il existe $\lambda_x, \lambda_y > 0$ tels que $\dfrac{x}{\lambda_x} \in \mathcal{C}$, $\dfrac{y}{\lambda_y} \in \mathcal{C}$ et $\lambda_x \leq j_C(x) + \varepsilon$, $\lambda_y \leq j_C(y) + \varepsilon$.

   On pose alors :

   $$\begin{aligned} z &= \frac{\lambda_x}{\lambda_x + \lambda_y} \cdot \frac{x}{\lambda_x} + \frac{\lambda_x}{\lambda_x + \lambda_y} \cdot \frac{x}{\lambda_y} \in \mathcal{C} \\ &= \frac{x + y}{\lambda_x + \lambda_y}. \end{aligned}$$

   Alors comme $z \in \mathcal{C}$, on a $j_C(z) < 1$ d'après 1, donc par homogénéité, on a

   $$\frac{j_c(x + y)}{\lambda_x + \lambda_y} < 1.$$

   Il s'en suit que $j_c(x + y) < \lambda_x + \lambda_y \leq j_c(x) + j_c(y) + 2\varepsilon$. Ceci étant vrai pour tout $\varepsilon > 0$, il vient :

   $$j_C(x + y) \;\leq\; j_C(x) + j_C(y).$$

$\square$

Théorème 3. *(Hahn-Banach, forme analytique)*

Soit $p: E \to \mathbb{R}$ positivement homogène et sous-additive. Soit G un sous-espace vectoriel de E et g une forme linéaire sur G telle que $g \le p$.

Alors il existe une forme linéaire f sur E telle que $f_{|G} = g$ et $f \le p$ sur E.

Démonstration.

Etape 1 : On se donne $F \subset E$ un sous-espace vectoriel de dimension maximum pour lequel il existe une forme linéaire $f: F \to \mathbb{R}$ vérifiant $f_{|G} = g$ et $f \le p$ sur F. On suppose par l'absurde que $F \subsetneq E$.

Donnons nous un vecteur $x_0 \in E \setminus F$ et posons alors, pour $\alpha = \alpha(x_0) \in \mathbb{R}$,

$$f_\alpha: \begin{cases} F \oplus \mathbb{R}x_0 & \to & \mathbb{R} \\ x + t x_0 & \mapsto & f(x) + t\alpha \end{cases}.$$

Il est clair que $f_{\alpha|G} = g$. L'objectif est donc de trouver un paramètre $\alpha \in \mathbb{R}$ tel que $f_\alpha \le p$, c'est-à-dire tel que

$$\forall x \in F \quad \forall t \in \mathbb{R} \quad f(x) + t\alpha \le p(x + t x_0). \quad (\star)$$

Etape 2 : Montrons qu'il suffit de prendre α vérifiant

1. $\forall x \in F \quad \forall t \in \mathbb{R} \quad f(x) + \alpha \le p(x + x_0)$,
2. $\forall x \in F \quad \forall t \in \mathbb{R} \quad f(x) - \alpha \le p(x - x_0)$.

Supposons ces deux conditions réunies. On a alors

$$\begin{aligned} \forall t > 0 \quad f(x) + t\alpha &= t\left(f\left(\frac{x}{t}\right) + \alpha\right) \\ &\le t p\left(\frac{x}{t} + x_0\right) \\ &= p(x + t x_0), \end{aligned}$$

et de même,

$$\begin{aligned} \forall t < 0 \quad f(x) + t\alpha &= -t\left(f\left(\frac{x}{-t}\right) - \alpha\right) \\ &\le -t\left(p\left(\frac{x}{-t} - x_0\right)\right) \\ &= p(x + t x_0), \end{aligned}$$

où l'on a utilisé dans ces deux calculs la positive homogénéité de p, étant donné que dans le premier cas, $t > 0$ et dans le second, $-t > 0$. Par ailleurs, lorsque $t = 0$, l'inégalité $(\star)$ s'écrit simplement $f(x) \le p(x)$, ce qui est toujours vrai pour $x \in F$ par hypothèse sur F.

Remarquons alors que α vérifie les conditions 1 et 2 si et seulement si on a

$$s := \sup_{y \in F} (f(y) - p(y - x_0)) \le \alpha \le \inf_{y \in F} (p(y + x_0) - f(y)) := i.$$

Pour qu'un tel α existe, il suffit que le segment $[s, i]$ soit non vide, donc que l'on ait $s \le i$.

Etape 3 : Montrons que $s \le i$.

On va utiliser la sous-additivité de p. Pour $x, y \in F$, on a en effet

$$\begin{aligned} f(x) + f(y) &= f(x+y) \\ &\leq p(x+y) \\ &\leq p(x+x_0) + p(y-x_0). \end{aligned}$$

Aussi, il vient

$$f(x) + f(y) \leq p(x+y) \leq p(x+x_0) + p(y-x_0),$$

donc pour tout $x, y \in F$,

$$f(y) - p(y-x_0) \leq p(x+y) - p(y-x_0) - f(x) \leq p(x+x_0) - f(x).$$

Par passage à la borne supérieure, on en déduit que

$$\forall x \in F \quad \sup_{y \in F} \left(f(y) - p(y-x_0) \right) \ \leq \ p(x+x_0) - f(x),$$

puis par passage à la borne inférieure,

$$\sup_{y \in F} f(y) - p(y-x_0) \ \leq \ \inf_{x \in F} p(x+x_0) - f(x),$$

ce qui conclut que $s \leq i$.

Etape 4 : Conclusion.

Ainsi, on peut choisir $\alpha \in [s, i]$, et f_α est une forme linéaire sur $F \oplus \mathbb{R}x_0$ qui vérifie $f_{\alpha|G} = g$ et $f_\alpha \leq p$ sur $F \oplus \mathbb{R}x_0$. Comme $x_0 \notin F$, on a $\dim(F \oplus \mathbb{R}x_0) > \dim(F)$, et ceci contredit la maximalité de F.

□

Définition 4. *Soit φ une forme linéaire sur E. On appelle hyperplan affine de E tout ensemble H de la forme $H = \mathrm{Ker}(\varphi - c)$, où c est un réel quelconque.*

H est alors un espace affine dirigé par l'hyperplan vectoriel $\mathrm{Ker}(\varphi)$.

Définition 5.

Soit H un hyperplan affine de E. On se donne $\varphi \in E^*$ et $c \in \mathbb{R}$ tels que $H = \mathrm{Ker}(\varphi - c)$.

- *On appelle les demi-espaces limités par H les deux ensembles*

$$E_1 = \{x \in E : \ \varphi(x) \leq c\} \quad \textit{et} \quad E_2 = \{x \in E : \ \varphi(x) \geq c\}.$$

- *Etant donné A et B deux parties de E, on dit que H sépare A et B si $A \subset E_1$ et $B \subset E_2$ ou $A \subset E_2$ et $B \subset E_1$.*

Lemme 6. *Soit C un convexe ouvert de E non vide et $x_0 \in E \setminus C$. Alors il existe un hyperplan affine H de E séparant $\{x_0\}$ et C.*

Démonstration.

Quitte à considérer $C' = C - x$ et $\{x_0 - x\}$ où $x \in C$, on peut considérer que $0 \in C$. On pose

$$g : \left\{ \begin{array}{lll} \mathbb{R}x_0 & \to & \mathbb{R} \\ t x_0 & \mapsto & t \end{array} \right. .$$

Comme $x_0 \notin C$, la jauge j_C de C vérifie $j_C(x_0) \geq 1 = g(x_0)$, donc pour tout $t > 0$, on a

$$g(tx_0) \leq j_C(tx_0).$$

Par ailleurs, pour $t \leq 0$, on a $g(tx_0) = t \leq 0 \leq j_C(tx_0)$, donc on a bien $g \leq j_C$. D'après le théorème de Hahn-Banach analytique, il existe une forme linéaire f sur E telle que $f_{|\mathbb{R}x_0} = g$ et $f \leq p$.

On a alors $f(x_0) = g(x_0) = 1$ et pour tout $x \in C$, $f(x) \leq j_C(x) = 1$, donc en notant $E_1 = \{x \in E : f(x) \leq 1\}$ et $E_2 = \{x \in E : f(x) \geq 1\}$ on a $\{x_0\} \subset E_2$ et $C \subset E_1$. Ceci montre que l'hyperplan affine $H = \mathrm{Ker}(f-1)$ sépare $\{x\}$ et C.

□

Théorème 7. *(Hahn-Banach, forme géométrique)*

Soit A et B deux convexes de E disjoints et non vides. Si A est ouvert, il existe un hyperplan affine H qui sépare A et B.

Démonstration. On pose $C = A - B = \{x - y : x \in A, \ y \in B\}$.

Alors C est un convexe de E, et comme $A \cap B = \varnothing$, on a $0 \notin C$. Par ailleurs, C peut s'écrire sous la forme

$$C = \bigcup_{y \in B} A - y,$$

où les ensembles $A - y$ sont ouverts puisque A est ouvert. Aussi, C est un ouvert en tant que réunion d'ouverts. D'après le lemme 6, il existe un hyperplan affine $H = \mathrm{Ker}(f - c)$ qui sépare $\{0\}$ et C.

Posons $E_1 = \{x \in E : f(x) \leq c\}$ et $E_2 = \{x \in E : f(x) \geq c\}$. Par symétrie, on peut supposer que $C \subset E_1$ et $\{0\} \subset E_2$. On a alors $f(0) \geq c$, donc $0 \geq c$, et pour tout $x \in C$, $f(x) \leq c$.

Aussi, il vient que $f(x) \leq 0$ pour tout $x \in C$, donc pour $x \in A$ et $y \in B$, on a, par définition de C :

$$f(x - y) \leq 0 \quad \text{donc} \quad f(x) \leq f(y).$$

En passant à la borne supérieure puis à la borne inférieure, on en déduit que

$$s := \sup_{x \in A} f(x) \leq \inf_{y \in B} f(y) := i.$$

Donc il existe $\alpha \in [s, i]$. L'hyperplan affine $H' = \mathrm{Ker}(f - \alpha)$ sépare donc A et B. □