J.-P. CROISILLE
18/09/2014

Convexité et hyperplans : l'alternative de Farkas

Soit $A \in \mathbb{M}_{m,n}(\mathbb{R})$, $b \in \mathbb{R}^m$.
On s'intéresse aux solutions positives du système linéaire

$$(1) \qquad Ax = b$$

La question est de savoir s'il existe des $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ solutions de (1). Ce problème intervient en programmation linéaire.

Théorème 1 [Alternative de Farkas 1904]
On a l'alternative suivante :
(i) ou bien $Ax = b$ possède (au moins) une solution $x \geq 0$
(c'est-à-dire $x_i \geq 0$ $i = 1 \ldots n$)

(ii) ou bien le problème

$$(2) \qquad y^T A \geq 0, \quad y^T b < 0$$

possède une solution $y \in \mathbb{R}^m$.

démonstration : Une "alternative" est un "ou exclusif".
Montrons d'abord que (i) et (ii) ne peuvent être vrais simultanément. Supposons que (i) et (ii) soient vrais. Alors, en multipliant à gauche l'égalité $Ax = b$ par y^T on obtient

$$y^T (Ax) = y^T b,$$

c'est-à-dire

$$\underbrace{(y^T A)}_{\geq 0} \underbrace{x}_{\geq 0} = y^T b.$$

$y^T A \in \mathbb{R}^n$ a toutes ses composantes ≥ 0 et x a toutes ses composantes ≥ 0, donc

$(y^T A)\, x \in \mathbb{R}$ est positif. Mais $y^T b < 0$ par (ii). Contradiction. D'où le résultat.

Il faut ensuite montrer que (i) OU (ii) est VRAI. Supposons donc que (i) est faux. Il faut construire un $y \in \mathbb{R}^m$ solution de (2).

Nous utilisons le théorème de séparation d'un convexe fermé et d'un point que l'on démontrera ensuite. Il s'énonce ainsi :

Théorème 2 [Séparation d'un convexe fermé et d'un point]

Soit $C \subset \mathbb{R}^n$ un convexe fermé et soit $x_0 \in \mathbb{R}^n$, $x_0 \notin C$. Alors il existe un hyperplan $H = \{y \in \mathbb{R}^n \text{ t.q. } u^T y = c,\ u \in \mathbb{R}^n, u \neq 0, c \in \mathbb{R}\}$ qui sépare C et x_0. C'est-à-dire, en notant H^+ et H^- les demi-espaces

$$H^+ = \{y \in \mathbb{R}^n \text{ t.q. } u^T y > c\}$$

$$H^- = \{y \in \mathbb{R}^n \text{ t.q. } u^T y < c\},$$

on a $C \subset H^+$ et $x_0 \in H^-$.

démonstration du Thm 1 :

Supposons que (i) soit faux : cela signifie que b n'est pas dans l'ensemble $C(A)$ défini par

$$C(A) = \{ y \in \mathbb{R}^m , \underset{y_i \geq 0 \quad i=1\dots n}{y = y_1 A^1 + \dots + y_n A^n} \}$$

où $A^1 \dots A^n$ = les colonnes de A. En effet Ax s'écrit :

$$\begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \\ a_{m1} & & a_{mn} \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = y_1 \underbrace{\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}}_{A^1} + \dots + y_n \underbrace{\begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}}_{A^n}$$

$C(A)$ est l'ensemble des combinaisons linéaires à coefficients ≥ 0 des colonnes de A.

1) $C(A)$ est convexe : car si $y, z \in C(A)$

$$y = y_1 A^1 + \dots + y_n A^n \quad , y_i \geq 0 \; ,$$
$$z = z_1 A^1 + \dots + z_n A^n \quad ; z_i \geq 0 \; ,$$

on a pour tout $\lambda \in [0,1]$:

$$\lambda y + (1-\lambda) z = \lambda (y_1 A^1 + \dots + y_n A^n) + (1-\lambda)(z_1 A^1 + \dots + z_n A^n)$$
$$= \underbrace{(\lambda y_1 + (1-\lambda) z_1)}_{\geq 0} A^1 + \dots + \underbrace{(\lambda y_n + (1-\lambda) z_n)}_{\geq 0} A^n$$

donc $\lambda y + (1-\lambda) z \in C(A)$

2) De plus $C(A)$ est fermé. Pour le voir, montrons que $C(A)^c$ est ouvert.

. Si $A = 0_{M_{m,n}(\mathbb{R})}$, on a $C(A) = \{0\}$ qui est fermé.

. Supposons à présent $A \neq 0_{M_{m,n}(\mathbb{R})}$. Soit $y \in C(A)^c$ (c.a.d $y \notin C(A)$). Alors il existe une colonne $A^{i_0} \neq 0_{\mathbb{R}^m}$ t.q. dans la décomposition $y = y_1 A^1 + \dots + y_n A^n$, $y_{i_0} < 0$.

Alors tout $z \in B(y, |y_{i_0}|)$, boule de centre y et de rayon $|y_{i_0}|$ pour la norme $\|y\|_\infty = \max_{1 \leq i \leq n} |y_i|$ satisfait $z_{i_0} < y_{i_0} + |y_{i_0}| = 0$, donc $z_{i_0} < 0$ et donc $z \in C(A)^c$. Ceci prouve que $C(A)^c$ est ouvert et donc que $C(A)$ est fermé.

Revenons à la preuve de l'alternative. On suppose donc que (i) est faux : c'est-à-dire que $b \notin C(A)$. $C(A)$ est convexe et fermé d'après ce qui précède. Donc on applique le théorème 2 : Il existe un hyperplan H défini par $u \neq 0$ et $c \in \mathbb{R}$, d'équation

$$u^T x = c$$

qui sépare strictement $C(A)$ et b.

C'est-à-dire :

$$u^T z > c \qquad \forall z \in C(A) \quad (3)$$

et $\quad u^T z < c \qquad$ pour $z = b \qquad (4)$

On prend alors $z = A(\lambda x)$ pour $x \in \mathbb{R}^n$, $x \geqslant 0$ et λ paramètre > 0. On a bien $z \in C(A)$! Donc

$$u^T z = u^T A(\lambda x) = \lambda\, u^T A x > c$$

Donc en divisant par $\lambda > 0$:

$u^T A x > \dfrac{c}{\lambda}$. On fait ensuite $\lambda \to +\infty$.

Cela donne :

$$(3) \quad u^T A x \geqslant 0 \qquad \text{pour tout } x \in \mathbb{R}^n, x \geqslant 0.$$

Puisque (3) est vraie pour tout $x \geqslant 0$, on a que le vecteur $(u^T A)^T$ a toutes ses composantes $\geqslant 0$.

Donc $\quad u^T A \geqslant 0$. De plus (4) donne $u^T b < c$. Mais $0_{\mathbb{R}^n} \in C(A)$, donc $0 > c$ par (3) ce qui donne

$$u^T b < 0 \quad \text{dans (4)}.$$

D'où le résultat.

Nous prouvons ensuite le théorème de séparation (Thm 2). On a besoin pour cela du résultat suivant, le théorème de "projection sur un convexe fermé" dont on donne également la preuve.

<u>Lemme</u> [Projection sur un convexe fermé]

Soit $C \subset \mathbb{R}^n$ ensemble convexe fermé et $x_0 \notin C$. Alors il existe $\overline{x} \in C$ qui réalise le minimum de la distance $\{\|x_0 - x\|,\ x \in C\}$ c'est-à-dire

$$\|x_0 - \overline{x}\| = \min_{x \in C} \|x_0 - x\|$$

<u>dém</u>: On peut supposer que $x_0 = 0_{\mathbb{R}^n}$. Il suffit d'opérer un changement de coordonnés $x \mapsto x' = x - x_0$ pour se ramener à ce cas.

Donc on montre qu'il existe $\overline{x} \in C$ tel que

$$\|\overline{x}\| = \min_{x \in C} \|x\|$$ — On note $\delta = \inf_{x \in C} \|x\|$

Soit $x_1, x_2 \ldots x_k$ une suite de points de C tel que

$$\|x_k\| \xrightarrow[k \to +\infty]{} \delta \qquad (5)$$

Par définition de "inf", une telle suite existe. On veut montrer que x_k est une suite de Cauchy. Cela donnera que x_k converge vers une certaine limite $\overline{x}$ et puisque C est fermé, on aura $\overline{x} \in C$. Et en passant à la limite dans (I), on aura

$$\|\bar{x}\| = \delta .$$

Soit k_1 et k_2 deux indices. On a

$$\|x_{k_1} - x_{k_2}\|^2 + \|x_{k_1} + x_{k_2}\|^2 = 2\|x_{k_1}\|^2 + 2\|x_{k_2}\|^2$$

par la relation du parallélogramme (c'est une relation connue...)

On a $x_{k_1}, x_{k_2} \in C$, donc $\frac{1}{2}(x_{k_1} + x_{k_2}) \in C$
par convexité de C. Donc

$$\left\|\frac{1}{2}(x_{k_1} + x_{k_2})\right\| \geq \delta$$ Donc

$$\|x_{k_1} - x_{k_2}\|^2 + (2\delta)^2 \leq 2\|x_{k_1}\|^2 + 2\|x_{k_2}\|^2$$

Mais $\|x_k\| \to \delta$. Donc pour $k_1, k_2 \geq k_0(\varepsilon)$

On a $\|x_{k_1}\| < \delta + \varepsilon$, $\|x_{k_2}\| < \delta + \varepsilon$

Donc pour $k_1, k_2 \geq k_0(\varepsilon)$

$$\|x_{k_1} - x_{k_2}\|^2 + 4\delta^2 \leq 2(\delta+\varepsilon)^2 + 2(\delta+\varepsilon)^2$$

ce qui donne : $k_1, k_2 \geq k_0(\varepsilon) \Longrightarrow$

$$\|x_{k_1} - x_{k_2}\|^2 \leq 8\delta\varepsilon + 4\varepsilon^2$$

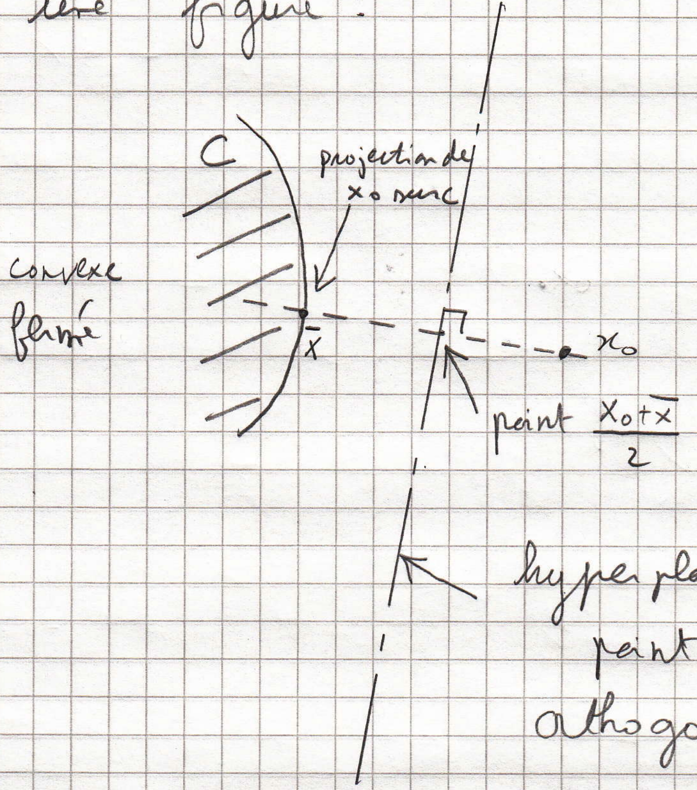
Ceci permet de conclure que : pour tout $\varepsilon' > 0$
il existe $k_0(\varepsilon')$ t.q. $k_1, k_2 \geq k_0(\varepsilon')$

$$\|x_{k_1} - x_{k_2}\| < \varepsilon'$$

La suite est donc de Cauchy, d'où le résultat.
(car $\mathbb{R}^n$ est complet...)

démonstration du théorème de séparation d'un convexe fermé et d'un point. (Thm 2)

La construction de l'hyperplan se comprend sur une figure :



hyperplan: passe par le point $\frac{x_0+\bar{x}}{2}$ et orthogonal à $\overrightarrow{x_0\bar{x}} = \bar{x} - x_0$

Soit $x_0 \notin C$. On considère $\bar{x}$ = le point qui réalise le $\inf_{x \in C} \|x_0 - x\| = \delta$

Soit $u = \frac{\bar{x} - x_0}{\delta}$ et $m = \frac{1}{2}(x_0 + \bar{x})$ = point milieu de x_0 et $\bar{x}$.

On définit l'hyperplan H par

$$H = \{ y \in \mathbb{R}^n \text{ tel que } u^T(y-m) = 0 \}$$

On vérifie ensuite que:

1) $C \subset H^+ = \{ y \in \mathbb{R}^n \text{ t.q. } u^T(y-m) > 0 \}$

2) $x_0 \in H^- = \{ y \in \mathbb{R}^n \text{ t.q. } u^T(y-m) < 0 \}$

• On a: $u^T(x_0 - m) = \frac{1}{\delta}(\bar{x} - x_0)\frac{x_0 - \bar{x}}{2} = -\frac{1}{2\delta}\|x_0 - \bar{x}\|^2 < 0$

donc 2) est VRAI

Montrons ensuite 1). Soit $x \in C$. Alors pour $\lambda \in [0,1]$, $\lambda \bar{x} + (1-\lambda)x \in C$ par convexité de C. On a donc

$$\|\bar{x} - x_0\|^2 \leqslant \|\lambda \bar{x} + (1-\lambda)x - x_0\|^2$$

c'est-à-dire :

$$\|\bar{x} - x_0\|^2 \leq \|\bar{x} - x_0 + (1-\lambda)(x-\bar{x})\|^2$$
$$\leqslant \|\bar{x} - x_0\|^2 + 2(1-\lambda)(\bar{x}-x_0)^T(x-\bar{x}) + (1-\lambda)^2\|x-\bar{x}\|^2$$

On en déduit :

$$2(1-\lambda)(\bar{x}-x_0)^T(x-\bar{x}) + (1-\lambda)^2\|x-\bar{x}\|^2 \geqslant 0$$

En divisant par $2(1-\lambda) > 0$

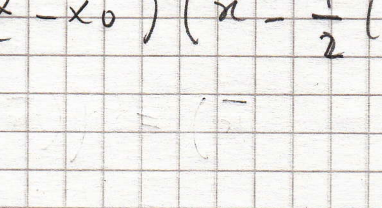
$$(\bar{x}-x_0)^T(x-\bar{x}) + \frac{1}{2}(1-\lambda)\|x-\bar{x}\|^2 \geqslant 0$$

et avec $\lambda \to 1^-$, on obtient

$$(\bar{x}-x_0)^T(x-\bar{x}) \geqslant 0$$

C'est-à-dire $-(x_0-\bar{x})(x-\bar{x}) \leqslant 0$

l'angle $(x_0-\bar{x}, x-\bar{x})$ est obtus.



Ensuite on évalue $u^T(x-m)$:

$$u^T(x-m) = \frac{1}{\delta}(\bar{x}-x_0)\left(x - \frac{1}{2}(x_0+\bar{x})\right)$$. On a :

On a :

$$(\bar{x} - x_0)^T \Big(x - \overbrace{\tfrac{1}{2}\big(x_0 + \bar{x}\big)}^{m}\Big) =$$

$$(\bar{x} - x_0)^T \Big(x - \bar{x} + \tfrac{1}{2}\big(\bar{x} - x_0\big)\Big) =$$

$$\underbrace{(\bar{x} - x_0)^T (x - \bar{x})}_{\geqslant 0} + \frac{1}{2} \underbrace{\|\bar{x} - x_0\|^2}_{> 0}$$

donc $(\bar{x} - x_0)^T (x - m) > 0$, c'est-à-dire

$C \subset H^+$, ce qui conclut.
