Une preuve élémentaire du lemme de Farkas

PAR PATRICK TELLER

Résumé

Le lemme de Farkas est bien connu et joue un rôle très important tant en analyse convexe qu'en optimisation linéaire;.

Habituellement sa démonstration a recours à des méthodes de l'analyse convexe (projection sur un convexe compact [1], séparation des convexes [2]).

Il existe aussi des démonstrations dans le cadre de l'optimisation linéaire (par exemple [3] et [4]); celles-ci sont constructives mais elles exigent l'emploi de toute la machinerie de la méthode du simplexe, y compris la règle de Bland.

La démonstration proposée ici ne fait appel qu'à des méthodes linéaires, elle est donc élémentaire et s'applique par exemple à $\mathbb{Q}$; ce qui n'est pas le cas des méthodes analytiques.

Dans ce qui suit K représentera au choix $\mathbb{R}$ ou $\mathbb{Q}$.

Proposition 1. *Quelques éléments géométrico-algébriques*

Soit $A \in \mathcal{M}_{\mathrm{mn}}(K)$ *une matrice de rang m,* $b \in \mathcal{M}_{n1}(K)$ *et le système* $\left\{\begin{array}{l}\mathrm{AX}=b\\ X \geqslant 0\end{array}\right.$.

i) L'ensemble des solutions du système AX=b est un sous-espace affine dirigé par Ker(A).

ii) L'ensemble Σ *des solutions du système* $\left\{\begin{array}{l}\mathrm{AX}=b\\ X \geqslant 0\end{array}\right.$ *est appelé polytope; s'il n'est pas borné il contient des demi-droites.*

iii) Si S est une matrice extraite de A de rang m et si on désigne T la matrice complément de S, c'est à dire formée par les autres colonnes de A, $AX=b \Longleftrightarrow \exists (x_{t_1}, \ldots.x_{t_{n-m}}) \in K^{(n-m)}$, $\begin{pmatrix} x_{s_1}\\ \ldots.\\ \ldots\\ x_{s_m}\end{pmatrix} = S^{-1}b - S^{-1}T\begin{pmatrix} x_{t_1}\\ \ldots.\\ \ldots\\ x_{t_{n-m}}\end{pmatrix}$ *(*) et* Σ *est paramétré par* $(x_{t_1}, \ldots.x_{t_{n-m}}) \in K^{+(n-m)} \longmapsto \begin{pmatrix} x_{s_1}\\ \ldots.\\ \ldots\\ x_{s_m}\end{pmatrix} = S^{-1}b - S^{-1}T\begin{pmatrix} x_{t_1}\\ \ldots.\\ \ldots\\ x_{t_{n-m}}\end{pmatrix}$, *sous la contrainte* $\begin{pmatrix} x_{s_1}\\ \ldots.\\ \ldots\\ x_{s_m}\end{pmatrix} \geqslant 0$.

Quitte à renuméroter les variables, (*) s'écrira sous la forme

$$\left\{\begin{array}{l}\begin{pmatrix} 1 & 0 & \ldots & \ldots. & 0 & * & \ldots & \ldots. & \ldots & *\\ 0 & 1 & 0 & \ldots. & \ldots & * & \ldots & \ldots. & \ldots & *\\ & 0 & \ldots & \ldots & \ldots & \ldots. & \ldots & \ldots & \ldots. & *\\ & & 0 & \ldots\ldots\ldots & \ldots & * & \ldots & \ldots & \ldots & *\\ & & & 0 & 1 & * & \ldots & \ldots & \ldots & * \end{pmatrix} X' = S^{-1}b\\ X' \geqslant 0\end{array}\right.$$

iv) Les points définis par $(x_{t_1}=0, \ldots.x_{t_{n-m}}=0)$ *et* $\begin{pmatrix} x_{s_1}\\ \ldots.\\ \ldots\\ x_{s_m}\end{pmatrix} = S^{-1}b$ *sont appelés les sommets du polytope et (*) est appelée la paramétrisation du polytope à partir du sommet(0,* $S^{-1}b$*);* $(x_{t_1}, \ldots.x_{t_{n-m}})$ *sont dites variables de base, les autres sont dites hors-base.*

v) Si f est une forme linéaire dont la restriction à une demi-droite est majorée elle admet un maximum à l'origine de la demi-droite.

vi) Si f est une forme linéaire dont la restriction à Σ *est majorée elle admet un maximum, atteint entre autres, en un sommet de* Σ.

Démonstration.

Les seuls points à préciser sont

v) Il suffit de remarquer que la restriction d'une forme linéaire à une demi-droite est une application affine, majorée donc décroissante.

vi) Le seul cas à détailler est celui où Σ ne serait pas borné: la restriction de f à chacune des génératrices étant décroissante on peut se ramener à un compact, d'où le résultat. $\square$

Remarque 2.

Dans le cas K=$\mathbb{Q}$ on remarquera que, malgré l'usage du mot « compact » les valeurs prises par ces fonctions à coefficients dans $\mathbb{Q}$ en des points à coordonnées rationnelles sont des rationnels.

Définition 3. *Matrice régulièrement échelonnée et réduite*

Une matrice rectangulaire M=(mij) est dite régulièrement échelonnée lorsque

i) si elle possède des lignes nulles ce sont les dernières

ii) dans toute ligne non nulle le premier terme non nul est un 1, appelé pivot

iii) il existe un entier q tel que les pivots ont pour coordonnées (1,1),....,(q,q)

elle est dite réduite lorsque, de plus, dans les colonnes des pivots il n'y a qu'un terme non nul.

Il est bien connu que

Théorème 4.

Quelle que soit $A\in\mathcal{M}_{\mathrm{mn}}(K)$, il existe $P\in\boldsymbol{GL}_m(K)$ et une matrice de permutation Q, telles que PAQ est régulièrement échelonnée et réduite.

Remarque 5.

La permutation Q (renumérotation des variables) n'est destinée qu'au confort de l'écriture de la matrice après échelonnement et réduction.

Ces rappels élémentaires fournissent une démonstration du lemme de Farkas

Théorème 6. *(lemme de Farkas)*

Soient $A\in\mathcal{M}_{\mathrm{mn}}(K)$ et $b\in\mathcal{M}_{m1}(K)$, où K=Q,R ou tout corps commutatif ordonné.

Une et une seule de ces deux propositions est vraie

i) $\exists X\in\mathcal{M}_{m1}(K), X\geqslant 0, \mathrm{AX}=b$

ii) $\exists Y\in\mathcal{M}_{m1}(K), {}^t\mathrm{YA}\leqslant 0$ et ^{t}Yb>0.

Démonstration.

* Montrons i)=>non ii).

Si i) est vrai alors pour tout Y ^{t}YAX=^{t}Yb, d'où ^{t}YAX et ^{t}Yb sont nécessairement de même signe.

La seconde implication sera démontrée dans l'esprit de la méthode du simplexe.

** Montrons que non i)=>ii)

Supposons le i) faux.

Nous supposerons effectuées les opérations élémentaires sur les lignes et la permutation des variables de telle sorte que le système AX=b est de la forme $\begin{cases} (\, I_m \quad N\,)X = b \\ X \geqslant 0 \end{cases}$

En désignant les équations du système AX=b par ($L_1X = b_1$, ..., $L_mX = b_m$) soit p le minimum des i, tels que le système $\begin{cases} L_1X = b_1 \\ .... \\ .... \\ L_iX = b_i \\ X \geqslant 0 \end{cases}$ ne possède pas de solutions; bien entendu $p \leqslant m$.

Si p=1 l'équation $L_1X = b_1$ est de la forme $x_1 + \sum_{j>1} a_{1j}x_j = b_1$, et nécessairement $b_1 < 0$ et les $a_{1j} \geqslant 0$ ce qui établit le ii) avec $Y = \begin{pmatrix} 1 \\ 0 \\ ... \\ .... \\ 0 \end{pmatrix}$.

Si $p \geqslant 2$ on désignera l'ensemble des solutions du système $\begin{cases} L_1X = b_1 \\ .... \\ .... \\ L_{p-1}X = b_{p-1} \\ X \geqslant 0 \end{cases}$ par D.

Soit la p ième équation du système : $x_p + \sum_{j>p} a_{\mathrm{pj}}x_j = b_p$ comme elle n'a pas de solution dans D l'application $f = b_p - \sum_{j>p} a_{\mathrm{pj}}x_j$ y reste strictement négative et y possède donc un maximum $\gamma < 0$, nécessairement atteint en un sommet $(0, S^{-1}b)$ de D.

D est décrit « à partir du sommet $(0, S^{-1}b)$» et relativement à une renumérotation des variables par le système $\begin{matrix} \begin{pmatrix} 1 & 0 & ... & .... & 0 & 0 & * & ... & .... & ... & * \\ 0 & 1 & 0 & .... & ... & 0 & * & ... & .... & ... & * \\ & 0 & ... & ... & ... & ... & .... & ... & ... & .... & * \\ & & 0 & .... & ... & ... & * & ... & ... & ... & * \\ & & & 0 & 1 & 0 & * & ... & ... & ... & * \end{pmatrix} X' = b' \\ X' \geqslant 0 \end{matrix}$ que l'on désignera par $\begin{cases} L'_1X' = b'_1 \\ .... \\ .... \\ L'_{p-1}X' = b'_{p-1} \\ X' \geqslant 0 \end{cases}$.

(on remarquera la colonne de zéros qui exprime le fait que x_p n'intervient pas dans ce système, et donc $x'_p = x_p$).

Relativement à la nouvelle base $x'_p = x_p$ est défini par l'équation $x'_p + \sum_{j \neq p} a'_{\mathrm{pj}}x'_j = b_p$ (que l'on notera $L'_pX' = b_p$) ou toute équation de la forme $(L'_p + \sum_{i<p} \lambda_i L_i)X' = (b_p + \sum_{i<p} \lambda_i b'_i)$; si on pose pour chaque $i < p$, $\lambda_i = -a'_{\mathrm{pi}}$ le premier membre est de la forme $x'_q + \sum_{j>p} \alpha_j x'_j$ et le second est la valeur de x'_p lorsque $x'_{p+1} = ... = x'_n = 0$, c'est à dire en $(0, S^{-1}b)$, c'est à dire γ.

Par définition la restriction à D de f est égale à $x_p (= x'_p)$, qui a comme maximum $\gamma < 0$, donc sur D $\sum_{j>p} \alpha_j x'_j = \gamma - x'_p \geqslant 0$; par suite les α_j sont positifs tandis que $\gamma < 0$.

Les deux étapes de notre construction, pivot de Gauss (sur les lignes), multiplication à gauche par S^{-1} puis ajout à la p-ième équation d'une combinaison linéaire des précédentes, se traduisent par une multiplication de la p-ième ligne de (A,b) par un vecteur ligne ${}^tY$, tel que ${}^tYA \geqslant 0$ et ${}^tYb < 0$; on ne citera pas les permutations des colonnes qui n'ont pas d'influence sur la question qui nous intéresse. □

[1] Bernd Gärtner, Jiri Matousek, Understanding and Using Linear Programming, Springer, 2006.

[2] Jonathan M. Borwein et Adrian S. Lewis, Convex Analysis, coll. « Ouvrages de mathématiques de la Société mathématique du Canada », vol. 3, 2e édition, Springer, 2006.

[3] M.S. Bazaran, J.J. Jarvis, H.D. Sherali, Linear Programming and Network Flows, Wiley, 2010.

Paris, mars 2019.