Cours n°2

Optimisation linéaire. Méthode du simplexe.

Solutions de base d'un système linéaire

Dans ce second cours consacré à l'optimisation linéaire, on formalise l'étude des problèmes d'optimisation linéaire en introduisant des termes de vocabulaire. On aborde ensuite la notion de solutions de base d'un système linéaire. Cette notion apparaît lorsque l'on s'intéresse aux sommets du polyèdre des contraintes d'un problème d'optimisation linéaire, dont on a entraperçu lors du cours précédent le rôle dans la résolution d'un problème d'optimisation linéaire.

1 Forme générale d'un problème d'optimisation linéaire

1.1 Quelques rappels d'algèbre linéaire

Définition 1 (Forme linéaire)

Soit $U = \mathbb{R}^Q$ un espace vectoriel ($Q \in \mathbb{N}^*$) et $F : U \to \mathbb{R}$ une application. On dit que F est une *forme linéaire* sur U si F est une application linéaire, c'est-à-dire si

$$\forall\,(x,y) \in U^2, \forall\,(\lambda,\mu) \in \mathbb{R}^2, \quad F(\lambda\,x + \mu\,y) = \lambda\,F(x) + \mu\,F(y)$$

Autrement dit, une forme linéaire est une application linéaire à valeurs dans l'ensemble des réels $\mathbb{R}$ (et non dans un espace vectoriel quelconque).

Exercice

Forme linéaire. Montrer que l'application définie par

$$(x_1,x_2) \mapsto 2\,x_1 + 4\,x_2$$

est une forme linéaire sur $\mathbb{R}^2$.

Définition 2 (Application affine)

Soit $U = \mathbb{R}^Q$, $V = \mathbb{R}^P$ deux espaces vectoriels avec $(Q,P) \in (\mathbb{N}^*)^2$ et $G : U \to V$ une application. On dit que G est une *application affine* sur U (à valeurs dans V) s'il existe un vecteur $b \in V$ et une application linéaire $H : U \to V$ tels que

$$\forall\,x \in U, \quad G(x) = H(x) + b$$

Autrement dit, G est une application affine sur U s'il existe un vecteur $b \in V$ tel que $G - b$ est une application linéaire sur U. Si $b = 0$, alors G est une application linéaire (les applications linéaires sont donc des applications affines). Le vecteur $b \in V$ est unique, et vaut $b = G(0)$.

Exemple

Les applications définies par

- $(x_1,x_2) \mapsto -x_1 + x_2 - 5$
- $(x_1,x_2) \mapsto -x_1 - x_2 + 8$
- $(x_1,x_2) \mapsto x_1 - 11$

sont des applications affines sur $\mathbb{R}^2$ à valeurs réelles, tandis que l'application définie par

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -x_1 + 3\,x_2 - 5 \\ -x_1 - x_2 + 8 \\ x_1 - 11 \end{pmatrix} = \begin{pmatrix} -x_1 + 3\,x_2 \\ -x_1 - x_2 \\ x_1 \end{pmatrix} - \begin{pmatrix} 5 \\ -8 \\ 11 \end{pmatrix}$$

est une application affine sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}^3$.

1.2 Problème d'optimisation linéaire sous forme générale

Les notions rappelées dans le paragraphe précédent permettent de définir les problèmes d'optimisation linéaire sous forme générale.

Définition 3 (Forme générale d'un problème d'optimisation linéaire)

On appelle problème d'optimisation linéaire *sous forme générale* un problème d'optimisation de la forme

$$\begin{array}{ll} \text{Maximiser} & F(X) \\ \text{sous les contraintes} & X \in \mathbb{R}^Q, G_1(X) \leq 0, \dots, G_P(X) \leq 0 \end{array}$$

où

- $(P,Q) \in (\mathbb{N}^*)^2$
- la fonction objectif $F : \mathbb{R}^Q \to \mathbb{R}$ est une forme linéaire sur $\mathbb{R}^Q$
- les contraintes $G_i : \mathbb{R}^Q \to \mathbb{R}$ pour $i = 1,\dots,P$ sont des applications affines sur $\mathbb{R}^Q$ à valeurs réelles.

On peut évidemment également considérer des problèmes de minimisation.

Exemple

Problème d'optimisation linéaire sous forme générale. Le problème d'optimisation

$$\begin{array}{ll} \text{Maximiser} & z = 2\,x_1 + 4\,x_2 \\ \text{sous les contraintes} & \begin{array}{rrcr} -x_1 + & x_2 & \leq & 5 \\ -x_1 - & x_2 & \leq & -8 \\ x_1 & & \leq & 11 \\ x_1\ , & x_2 & \geq & 0 \end{array} \end{array}$$

est un problème d'optimisation linéaire sous forme générale, avec

- $F : (x_1, x_2) \mapsto 2\,x_1 + 4\,x_2$ la fonction objectif
- $G_1 : (x_1, x_2) \mapsto -x_1 + x_2 - 5$,
- $G_2 : (x_1, x_2) \mapsto -x_1 - x_2 + 8$,
- $G_3 : (x_1, x_2) \mapsto x_1 - 11$,
- $G_4 : (x_1, x_2) \mapsto -x_1$,
- $G_5 : (x_1, x_2) \mapsto -x_2$ les contraintes affines.

On remarquera que dans la définition **??**, la convention adoptée est que les contraintes affines s'écrivent comme des contraintes de *négativité*.

1.3 Contraintes égalité, contraintes inégalité

Dans la définition **??**, les contraintes du problème d'optimisation considéré sont de la forme

$$G_i(X) \leq 0$$

On parle de *contraintes d'inégalités*. Il est évidemment possible de considérer des *contraintes d'égalités* de la forme

$$H_j(X) = 0$$

par exemple dans des problèmes de la forme

$$\begin{array}{ll} \text{Maximiser} & F(X) \\ \text{sous les contraintes} & X \in \mathbb{R}^Q, G_1(X) \leq 0, \ldots, G_P(X) \leq 0 \\ & \qquad H_1(X) = 0, \ldots, H_R(X) = 0 \end{array} \tag{1}$$

On remarque que, pour tout $X \in \mathbb{R}^Q$ et pour tout $j = 1, \ldots, R$,

$$H_j(X) = 0 \qquad \Longleftrightarrow \qquad \begin{cases} H_j(X) \leq 0 \\ -H_j(X) \leq 0 \end{cases}$$

Ainsi, la forme générale de ce problème d'optimisation linéaire est donnée par

$$\begin{array}{lll} \text{Maximiser} & F(X) & \\ \text{sous les contraintes} & X \in \mathbb{R}^Q, & G_1(X) \leq 0, \ldots, \quad G_P(X) \leq 0 \\ & & H_1(X) \leq 0, \ldots, \quad H_R(X) \leq 0 \\ & & -H_1(X) \leq 0, \ldots, -H_R(X) \leq 0 \end{array}$$

soit, en posant $G_{P+j} = H_j$ et $G_{P+R+j} = -H_j$ pour $j = 1, \ldots, R$

$$\begin{array}{ll} \text{Maximiser} & F(X) \\ \text{sous les contraintes} & X \in \mathbb{R}^Q, G_1(X) \leq 0, \ldots, G_P(X) \leq 0 \\ & \qquad G_{P+1}(X) \leq 0, \ldots, G_{P+R}(X) \leq 0 \\ & \qquad G_{P+R+1}(X) \leq 0, \ldots, G_{P+2R}(X) \leq 0 \end{array} \tag{2}$$

On pourra vérifier en exercice que les deux problèmes (1) et (2) sont équivalents.

1.4 Forme étendue d'un problème sous forme générale

On va maintenant écrire de manière plus explicite les problèmes d'optimisation linéaire sous forme générale, à l'aide de coefficients réels.

Considérons d'abord une forme linéaire F sur $\mathbb{R}^Q$. Montrons qu'il existe un vecteur $f = {}^{\mathrm{t}}(f_1, \ldots, f_Q) \in \mathbb{R}^Q$ tels que

$$\forall\, X \in \mathbb{R}^Q, \quad F(X) = \sum_{\ell=1}^{Q} f_\ell X_\ell$$

En effet, si on considère $(e_1, \ldots, e_Q)$ la base canonique de $\mathbb{R}^Q$, alors la linéarité de F permet d'écrire

$$\forall\, X \in \mathbb{R}^Q, \quad F(X) = F\left(\sum_{\ell=1}^{Q} X_\ell e_\ell\right) = \sum_{\ell=1}^{Q} F(e_\ell) X_\ell$$

On obtient alors le résultat annoncé en posant $f_\ell = F(e_\ell)$ pour tout $\ell = 1, \ldots, Q$.

Intéressons-nous maintenant aux contraintes affines. Soit $k = 1, \ldots, Q$ etG_k une application affine sur $\mathbb{R}^Q$ à valeurs réelles. Par définition, il existe un réel g_k tel que $H_k = G_k + g_k$ est une forme linéaire. La constante g_k est déterminée en écrivant

$$0 = H_k(0) = G_k(0) + g_k \qquad \Longleftrightarrow \qquad g_k = -G_k(0)$$

En appliquant ce qui précède à la forme linéaire H_k, on peut montrer l'existence de Q réels $G_{k1}, \ldots, G_{kQ}$ tel que

$$\forall\, X \in \mathbb{R}^Q, \quad G_k(X) = H_k(X) + G_k(0) = \sum_{j=1}^{Q} G_{kj} X_j + G_k(0)$$

où $G_{kj} = H_k(e_j) = G_k(e_j) - g_k = G_k(e_j) - G_k(0)$.

La forme étendue d'un problème sous forme générale est donc donnée par

$$\begin{array}{ll} \text{Maximiser} & \displaystyle\sum_{\ell=1}^{Q} f_\ell X_\ell \\ \text{sous les contraintes} & X \in \mathbb{R}^Q, \displaystyle\sum_{j=1}^{Q} G_{1j} X_j \leq g_1 \\ & \qquad \ldots \\ & \qquad \displaystyle\sum_{j=1}^{Q} G_{Pj} X_j \leq g_P \end{array}$$

1.5 Écriture matricielle

On se propose maintenant d'écrire la forme étendue d'un problème d'optimisation linéaire sous forme générale de manière plus compacte, en utilisant l'écriture matricielle. On pose

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_Q \end{pmatrix}, \; f = \begin{pmatrix} f_1 \\ \vdots \\ f_Q \end{pmatrix}, \; G = \begin{pmatrix} G_{11} & \ldots & G_{1Q} \\ \vdots & \ddots & \vdots \\ G_{P1} & \ldots & G_{PQ} \end{pmatrix}, \; g = \begin{pmatrix} g_1 \\ \vdots \\ g_P \end{pmatrix}$$

Alors $$\sum_{j=1}^{Q} f_j\, X_j = {}^{\mathrm{t}}f\, X \quad \text{et} \quad \sum_{j=1}^{Q} G_{ij}\, X_j = (G\, X)_i \qquad \text{pour } i = 1, \ldots, P$$

Notation 1 (Comparaison de vecteurs)

Si $a = (a_i)_{1 \leq i \leq P}$ et $b = (b_i)_{1 \leq i \leq P}$ sont deux vecteurs de $\mathbb{R}^P$, alors $a \leq b$ signifie que $a_i \leq b_i$ pour tout $i = 1, \ldots, P$.

Ainsi, en utilisant cette notation, on en déduit que tout problème d'optimisation linéaire sous forme générale peut s'écrire sous la forme

$$\begin{array}{ll} \text{Maximiser} & {}^{t}\!f\,X \\ \text{sous les contraintes} & G\,X \leq g \end{array}$$

où $X, f \in \mathbb{R}^Q$, $g \in \mathbb{R}^P$, $G \in \mathcal{M}_{P,Q}(\mathbb{R})$.

Dans les deux sections qui suivent, on va introduire deux notions qui vont nous permettre de mieux appréhender l'ensemble réalisable d'un problème d'optimisation linéaire, ainsi que ces solutions optimales.

2 Polyèdres

2.1 Définition

Définition 4 (Polyèdre)

Soit $\mathcal{C} \subset \mathbb{R}^Q$ avec $Q \in \mathbb{N}^*$. On dit que $\mathcal{C}$ est un *polyèdre* dans $\mathbb{R}^Q$ s'il existe une matrice $G \in \mathcal{M}_{P,Q}(\mathbb{R})$ et un vecteur $g \in \mathbb{R}^P$ tels que

$$\mathcal{C} = \{X \in \mathbb{R}^Q \mid G\,X \leq g\}$$

Remarque : Attention, cette définition n'est pas standard. Certains auteurs parleront plus préférentiellement de *polytopes* pour désigner cet objet.

Une autre manière de définir un polyèdre est de le définir comme une intersection *finie* des demi-espaces fermés de $\mathbb{R}^Q$. En effet, on peut écrire

$$\mathcal{C} = \left\{ X \in \mathbb{R}^Q \;\middle|\; \forall\, k = 1, \ldots, P,\ \sum_{j=1}^{Q} G_{kj} X_j \leq g_k \right\} = \bigcup_{k=1}^{P} \left\{ X \in \mathbb{R}^Q \;\middle|\; \sum_{j=1}^{Q} G_{kj} X_j \leq g_k \right\}$$

(si h est une application affine à valeurs réelles, et H l'hyperplan d'équation $h(X) = 0$, alors l'espace $\mathbb{R}^Q$ est "partagé en deux" par l'hyperplan H, les deux demi-espaces fermés qui en résultent étant les ensembles définis par $h(X) \leq 0$ et $h(X) \geq 0$).

Exemple

Polyèdres bornés. Voici quelques exemples de polyèdres *bornés*.

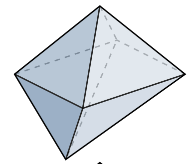
- dans $\mathbb{R}$: tout segment (intervalle fermé borné) $[a, b]$, qui est l'intersection entre les deux demi-droites $[a, +\infty[$ et $]-\infty, b]$;
- dans $\mathbb{R}^2$: tout domaine délimité par un polygone (triangle, carré, rectangle, losange, hexagone etc.) ;
- dans $\mathbb{R}^3$: cube, pyramide, tétraèdre, ...

Il faut noter que des polyèdres non bornés existent également en toute dimension (puisque tout demi-espace est par définition un polyèdre).

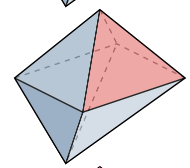
2.2 Description d'un polyèdre

On va décrire de manière informelle les éléments constitutifs d'un polyèdre. On va se placer pour ce faire dans $\mathbb{R}^3$. Leur définition rigoureuse est loin d'être triviale, et sort du cadre de ce cours.

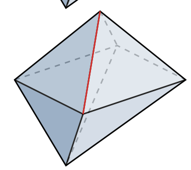
Un polyèdre est constitué de quatre éléments distincts :



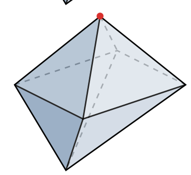
Le *corps* de ce polyèdre est le volume représenté par le schéma ci-contre. Il est délimité par les faces du polyèdre (voir point suivant).



En rouge, on peut apercevoir une *face* de ce polyèdre. Il s'agit d'une surface (ici bornée), plane. Dans cet exemple, le polyèdre possède huit faces.



On a tracé ici en rouge une *arête* de ce polyèdre. Il s'agit ici d'un segment de droite. Dans cet exemple, le polyèdre possède douze arêtes.



Enfin, on a représenté en rouge un *sommet* de ce polyèdre, qui est un point particulier du polyèdre. Dans cet exemple, le polyèdre possède six sommets.

Dans ce cours, ce sont les sommets d'un polyèdre qui nous intéressent, et plus spécifiquement ceux d'une classe de polyèdres particuliers, ceux contenus dans $(\mathbb{R}^+)^Q$.

2.3 Sommets d'un polyèdre $\{MX \leq b, X \geq 0\}$ dans $\mathbb{R}^2$

Dans cette section, on va étudier les sommets d'un polyèdre dans $(\mathbb{R}^+)^2$. Cette étude justifiera plus tard les notions qui seront introduites pour décrire la méthode du simplexe.

On considère le polyèdre suivant :

$$\mathcal{C}_+ = \left\{ X \in \mathbb{R}^2 \;\middle|\; MX \leq b \text{ et } X \geq 0 \right\}, \text{ avec } M = \begin{pmatrix} -1 & -1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ et } b = \begin{pmatrix} -8 \\ 11 \\ 5 \end{pmatrix}$$

Il s'agit d'un polyèdre entièrement contenu dans le quart de plan $(\mathbb{R}^+)^2$ (voir Figure **??**). Il s'écrit sous forme étendue :

$$\mathcal{C}_+ = \left\{ X = (X_1, X_2) \in \mathbb{R}^2 \;\middle|\; \begin{array}{rrcr} -X_1 & -X_2 & \leq & -8 \\ X_1 & & \leq & 11 \\ -X_1 & +X_2 & \leq & 5 \\ X_1 & & \geq & 0 \\ & X_2 & \geq & 0 \end{array} \right\}$$

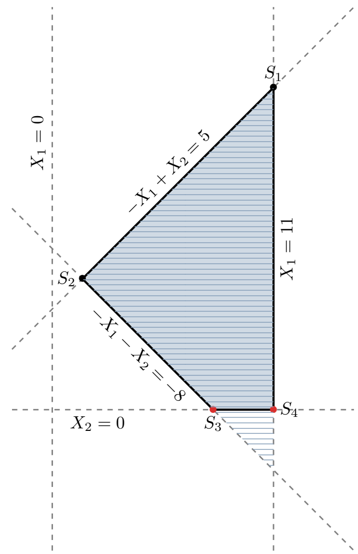


FIGURE 1 – Polyèdre $\mathcal{C}_+$ en bleu pâle. La région hachurée est le polyèdre $\{MX \leq b\}$.

Intéressons-nous à ses sommets. Il en possède quatre, représentés par des points dans la figure **??**. On peut les classer en deux types : ceux, représentés en noir, qui sont également sommets du polyèdre

$$\left\{X \in \mathbb{R}^2 \;\middle|\; MX \leq b\right\}$$

et ceux, en rouge, qui ne sont pas des sommets de ce second polyèdre. Les quatre sommets peuvent être définis comme intersection de deux droites, et sont respectivement donnés par les systèmes d'équations linéaires suivants :

$$S_1 : \left\{\begin{array}{rcr} X_1 & = & 11 \\ -X_1 + X_2 & = & 5 \end{array}\right. \qquad S_2 : \left\{\begin{array}{rcr} -X_1 - X_2 & = & -8 \\ -X_1 + X_2 & = & 5 \end{array}\right.$$

$$S_3 : \left\{\begin{array}{rcr} -X_1 - X_2 & = & -8 \\ X_2 & = & 0 \end{array}\right. \qquad S_4 : \left\{\begin{array}{rcr} X_1 & = & 11 \\ X_2 & = & 0 \end{array}\right.$$

Si on considère par exemple le sommet S_1, la résolution du système linéaire associé montre que $S_1 = (11, 16)$. On peut vérifier que la première contrainte inégalité du polyèdre $\mathcal{C}_+$ est bien satisfaite ($-11 - 16 = -27 \leq -8$), mais de manière stricte, de même que les deux contraintes de positivité $X_1, X_2 \geq 0$, alors que les deuxième et troisième contraintes sont dites *saturées* (satisfaites sous forme d'égalité). Pour saturer la première contrainte, il manque une quantité *positive* (notons-la X_3), qui vaut :

$$-X_1 - X_2 + X_3 = -8 \qquad \Longleftrightarrow \qquad X_3 = -8 - \big(-X_1 - X_2\big) = -8 - (-27) = 19$$

Ainsi, cette quantité (appelée *variable d'écart*) permet d'intégrer la première contrainte dans la définition du sommet S_1 :

$$\left\{\begin{array}{l} -X_1 - X_2 + X_3 = -8 \\ X_1 = 11 \\ -X_1 + X_2 = 5 \end{array}\right.$$

Ce système linéaire admet une unique solution, $(X_1, X_2, X_3) = (11, 16, 19)$. Remarquons qu'extraire les deux premières composantes permet d'obtenir les coordonnées du sommet S_1.

Si on procède de même avec le sommet S_2, c'est-à-dire si on ajoute une variable d'écart X_4 dans la deuxième contrainte (qui est la seule contrainte – en-dehors des contraintes de positivité – à ne pas être saturée), alors on aboutit au système linéaire

$$\left\{\begin{array}{l} -X_1 - X_2 = -8 \\ X_1 + X_4 = 11 \\ -X_1 + X_2 = 5 \end{array}\right.$$

qui admet comme unique solution $(X_1, X_2, X_4) = (3/2, 13/2, 19/2)$. Une fois encore, on peut vérifier que le point $(X_1, X_2) = (3/2, 13/2)$ correspond au sommet S_2.

Qu'en est-il des deux autres sommets ? Contrairement aux sommets S_1 et S_2, ils ne saturent qu'une seule des trois premières contraintes. Le sommet S_3, par exemple, sature la première contrainte. En extrapolant la démarche précédente, on va ajouter *deux* variables d'écart, une par contrainte non saturée. On obtient alors le système linéaire suivant :

$$\left\{\begin{array}{l} -X_1 - X_2 = -8 \\ X_1 + X_4 = 11 \\ -X_1 + X_2 + X_5 = 5 \end{array}\right.$$

Il s'agit cette fois d'un système de trois équations à quatre inconnues. Sa résolution donne l'ensemble de solutions suivant

$$(X_1, X_2, X_4, X_5) \in \left\{(8 - X_2, X_2, 3 + X_2, 13 - 2\,X_2) \;\middle|\; X_2 \in \mathbb{R}\right\}$$

En particulier, l'unique solution vérifiant $X_2 = 0$ est le point $(8, 0, 3, 13)$, dont les deux premières coordonnées sont exactement celles du sommet S_3. Enfin, ajouter deux variables d'écart dans les première et troisième contraintes (contraintes non saturées par le sommet S_4 amène à considérer le système

$$\left\{\begin{array}{l} -X_1 - X_2 + X_3 = -8 \\ X_1 = 11 \\ -X_1 + X_2 + X_5 = 5 \end{array}\right.$$

dont les solutions sont

$$(X_1, X_2, X_3, X_5) \in \left\{(11, X_2, 3 + X_2, 16 - X_2) \;\middle|\; X_2 \in \mathbb{R}\right\}$$

parmi lesquelles l'unique solution satisfaisant $X_2 = 0$ vaut $(11, 0, 3, 16)$ et permet d'obtenir les coordonnées du sommet S_4.

Notons avant de poursuivre que, pour les deux derniers sommets, annuler le paramètre X_2 dans la solution revient en fait à considérer respectivement les systèmes linéaires

$$\left\{\begin{array}{l} -X_1 = -8 \\ X_1 + X_4 = 11 \\ -X_1 + X_5 = 5 \end{array}\right. \qquad \text{et} \qquad \left\{\begin{array}{l} -X_1 + X_3 = -8 \\ X_1 = 11 \\ -X_1 + X_5 = 5 \end{array}\right.$$

Il est intéressant de noter que les solutions obtenues en résolvant les différents systèmes linéaires (pour les sommets S_3 et S_4, celles vérifiant $X_2 = 0$) sont toutes à composantes *positives*. On verra que cette particularité permet de caractériser les sommets du polyèdre.

Dans la section suivante, on va introduire des outils qui permettront d'exploiter les observations de cette section pour formaliser la caractérisation des sommets d'un polyèdre dans $(\mathbb{R}^+)^Q$.

3 Solutions de base d'un système linéaire

3.1 Extraction de sous-vecteurs et de sous-matrices

Définition 5 (Rang d'une matrice)

Soit A une matrice de $\mathcal{M}_{m,n}(\mathbb{R})$ (m lignes, n colonnes). Son *rang*, noté $\mathrm{rang}(A)$, est la dimension du sous-espace vectoriel généré par ses vecteurs colonnes.

Commençons par quelques remarques sur cette notion déjà connue.

- Par définition, on a $\mathrm{rang}(A) \leq n$, puisque la dimension générée par une famille de vecteurs ne peut excéder le cardinal de cette famille.
- On a $\mathrm{rang}({}^{\mathrm{t}}A)$, donc en particulier $\mathrm{rang}(A) \leq m$ (en appliquant la remarque précédente à la matrice ${}^{\mathrm{t}}A$).
- Si A est de la forme (avec $n \geq m$)

$$A = \begin{pmatrix} M_{11} & \dots & M_{1(n-m)} & 1 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \dots & \vdots & \vdots & \ddots & \ddots & 0 \\ M_{m1} & \dots & M_{m(n-m)} & 0 & \dots & 0 & 1 \end{pmatrix} = \begin{pmatrix} M & I_m \end{pmatrix}$$

  alors $\mathrm{rang}(A) = m$.

Rappelons un résultat connu concernant le rang : si $\mathrm{rang}(A) = p$, alors on peut extraire parmi les vecteurs colonnes de A p vecteurs linéairement indépendantes. En particulier, si $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et que $\mathrm{rang}(A) = m$, alors on peut extraire parmi les vecteurs colonnes de A m vecteurs qui forment une base de $\mathbb{R}^m$.

Notation 2 (Extraction de vecteurs colonnes)

Soit A une matrice de $\mathcal{M}_{m,n}(\mathbb{R})$ (m lignes, n colonnes). Pour tout $j = 1, \dots, n$, on note A_j la j-ième colonne de A

$$A_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

On note $\Gamma = \Big\{\gamma : \{1, \dots, m\} \to \{1, \dots, n\} \text{ strictement croissante}\Big\}$

et pour tout $\gamma \in \Gamma$, on note A_γ la matrice carrée de taille m

$$A_\gamma = \Big(A_{\gamma(1)}, \dots, A_{\gamma(m)}\Big)$$

Si $j_1, \dots, j_m \in \{1, \dots, n\}$ sont m indices distincts, alors il existe une unique application $\gamma \in \Gamma$ telle que $\gamma(\{1, \dots, m\}) = \{j_1, \dots, j_m\}$. L'ensemble Γ permet donc de caractériser de manière unique l'extraction d'une sous-matrice carrée de A.

Notation 3 (Indices complémentaires)

Soit $\gamma \in \Gamma$, on note $\hat{\gamma} : \{1, \dots, n-m\} \to \{1, \dots, n\}$ l'unique application strictement croissante telle que

$$\gamma\left(\{1, \dots, m\}\right) \cup \hat{\gamma}\left(\{1, \dots, n-m\}\right) = \{1, \dots, n\}$$

Autrement dit, $\hat{\gamma}$ donne les indices des vecteurs colonnes restant après avoir extrait la sous-matrice A_γ.

Exemple

Soit $(m, n) = (3, 5)$. Considérons $\gamma : \{1, 2, 3\} \to \{1, 2, 3, 4, 5\}$ définie par

$$\gamma(1) = 1, \quad \gamma(2) = 3 \quad \text{et} \quad \gamma(3) = 4$$

qui appartient à Γ. L'application $\hat{\gamma}$ associée est alors donnée par

$$\hat{\gamma}(1) = 2 \quad \text{et} \quad \hat{\gamma}(2) = 5$$

Notation 4

Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ telle que $\mathrm{rang}(A) = m$. On note

$$\mathcal{B} = \Big\{\gamma \in \Gamma \mid \mathrm{rang}(A_\gamma) = m\Big\}$$

L'ensemble $\mathcal{B}$ est non vide car on a rappelé que, de toute matrice de rang m, on pouvait extraire m vecteurs colonnes linéairement indépendantes. Ainsi, on peut également interpréter $\mathcal{B}$ de la manière suivante : $\gamma \in \mathcal{B}$ si $\{A_{\gamma(1)}, \dots, A_{\gamma(m)}\}$ forme une base de $\mathbb{R}^m$, ou encore si A_γ est inversible.

Remarque : Par inclusion, on a

$$\mathrm{card}(\mathcal{B}) \leq \mathrm{card}(\Gamma) \quad \text{avec} \quad \mathrm{card}(\Gamma) = \binom{n}{m} = \frac{n!}{m!(n-m)!}$$

En particulier, l'ensemble $\mathcal{B}$ est *fini*.

Définition 6 (Variables en base / variables hors base)

Soit $\gamma \in \mathcal{B}$ et $\hat{\gamma}$ associé. On appelle *variables en base (pour γ)* les variables $x_{\gamma(1)}$, ..., $x_{\gamma(m)}$ et *variables en hors base (pour γ)* les variables $x_{\hat{\gamma}(1)}$, ..., $x_{\hat{\gamma}(n-m)}$.

Notation 5

Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ de rang m. Pour tout vecteur $x \in \mathbb{R}^n$ et $\gamma \in \mathcal{B}$ avec $\hat{\gamma}$ associé, on note

$$x_B = {}^{\mathrm{t}}(x_{\gamma(1)}, \dots, x_{\gamma(m)}) \quad \text{et} \quad x_N = {}^{\mathrm{t}}(x_{\hat{\gamma}(1)}, \dots, x_{\hat{\gamma}(n-m)})$$

et $$B = A_\gamma \quad \text{et} \quad N = A_{\hat{\gamma}}$$

Remarque : B est inversible par définition de $\mathcal{B}$.

Exemple

Considérons l'application introduite dans l'exemple 5, définie par $\gamma(1) = 1$, $\gamma(2) = 3$, $\gamma(3) = 4$. Dans ce cas, si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{pmatrix}$$

alors on a $$B = \begin{pmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} a_{12} & a_{15} \\ a_{22} & a_{25} \\ a_{32} & a_{35} \end{pmatrix}$$

Soit $x = {}^{\mathrm{t}}(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5$. Alors

$$A\,x = \begin{pmatrix} a_{11}\,x_1 + a_{12}\,x_2 + a_{13}\,x_3 + a_{14}\,x_4 + a_{15}\,x_5 \\ a_{21}\,x_1 + a_{22}\,x_2 + a_{23}\,x_3 + a_{24}\,x_4 + a_{25}\,x_5 \\ a_{31}\,x_1 + a_{32}\,x_2 + a_{33}\,x_3 + a_{34}\,x_4 + a_{35}\,x_5 \end{pmatrix}$$

tandis que

$$B\,x_B = \begin{pmatrix} a_{11}\,x_1 + a_{13}\,x_3 + a_{14}\,x_4 \\ a_{21}\,x_1 + a_{23}\,x_3 + a_{24}\,x_4 \\ a_{31}\,x_1 + a_{33}\,x_3 + a_{34}\,x_4 \end{pmatrix} \quad \text{et} \quad N\,x_N = \begin{pmatrix} a_{12}\,x_2 + a_{15}\,x_5 \\ a_{22}\,x_2 + a_{25}\,x_5 \\ a_{32}\,x_2 + a_{35}\,x_5 \end{pmatrix}$$

Ainsi, $$A\,x = B\,x_B + N\,x_N$$

Ainsi que vu dans l'exemple précédent, on peut montrer que, pour tout $x \in \mathbb{R}^n$,

$$A\,x = B\,x_B + N\,x_N$$

En particulier, si x est solution du système linéaire $A\,x = b$, alors $b = B\,x_B + N\,x_N$, soit

$$x_B = B^{-1}b - B^{-1}N\,x_N$$

De la même manière, pour tout $x \in \mathbb{R}^n$, on a

$${}^{\mathrm{t}}c\,x = {}^{\mathrm{t}}c_B\,x_B + {}^{\mathrm{t}}c_N\,x_N$$

3.2 Retour sur les sommets d'un polyèdre $\{MX \leq b, X \geq 0\}$

Utilisons les notations introduites dans les deux sections précédentes pour réécrire les calculs de la section 2.3. Les quatre systèmes linéaires avec variables d'écart permettant d'obtenir les coordonnées des sommets S_1, S_2, S_3 et S_4 s'écrivent matriciellement :

$$\begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} -8 \\ 11 \\ 5 \end{pmatrix}, \qquad \begin{pmatrix} -1 & -1 & 0 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_4 \end{pmatrix} = \begin{pmatrix} -8 \\ 11 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_4 \\ X_5 \end{pmatrix} = \begin{pmatrix} -8 \\ 11 \\ 5 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_3 \\ X_5 \end{pmatrix} = \begin{pmatrix} -8 \\ 11 \\ 5 \end{pmatrix}$$

En introduisant $$A = \begin{pmatrix} -1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} M & I_3 \end{pmatrix} \qquad (\star)$$

et $X = {}^{\mathrm{t}}(X_1, X_2, X_3, X_4, X_5)$, ces quatre systèmes linéaires s'écrivent de manière plus compacte, pour $i = 1, \dots, 4$,

$$B_i\,X_{B_i} = b \qquad \text{avec } B_i = A_{\gamma_i}$$

avec $$\begin{cases} \gamma_1(1) = 1, & \gamma_1(2) = 2 \quad \text{et} \quad \gamma_1(3) = 3 \\ \gamma_2(1) = 1, & \gamma_2(2) = 2 \quad \text{et} \quad \gamma_2(3) = 4 \\ \gamma_3(1) = 1, & \gamma_3(2) = 4 \quad \text{et} \quad \gamma_3(3) = 5 \\ \gamma_4(1) = 1, & \gamma_4(2) = 3 \quad \text{et} \quad \gamma_4(3) = 5 \end{cases}$$

Puisque ces quatre systèmes linéaires admettent une unique solution, on en déduit que les matrices $B_i = A_{\gamma_i}$ sont inversibles, donc de rang 3. Autrement dit, les γ_i appartiennent à l'ensemble $\mathcal{B}$. Par ailleurs, si on complète le vecteur X par des zéros pour les variables hors base, c'est-à-dire

$$X_{N_i} = 0 \qquad N_i = A_{\hat{\gamma}_i}$$

alors $$A\,X = b$$

De tels vecteurs X associés aux quatre γ_i sont respectivement

$$X = \begin{pmatrix} 11 \\ 16 \\ 19 \\ 0 \\ 0 \end{pmatrix}, \quad X = \begin{pmatrix} 3/2 \\ 13/2 \\ 0 \\ 19/2 \\ 0 \end{pmatrix}, \quad X = \begin{pmatrix} 8 \\ 0 \\ 0 \\ 3 \\ 13 \end{pmatrix} \quad \text{et} \quad X = \begin{pmatrix} 11 \\ 0 \\ 3 \\ 0 \\ 16 \end{pmatrix} \qquad (\star\star)$$

À nouveau, notons que les deux premières coordonnées de ces quatre vecteurs donnent précisément les coordonnées des quatre sommets du polyèdre $\{MX \leq b \text{ et } X \geq 0\}$. De plus, ces quatre vecteurs sont à composantes positives.

3.3 Solutions de base

Définition 7 (Solution de base)

Soit $\gamma \in \mathcal{B}$. On appelle *solution de base* du système $A\,x = b$ associée à γ la solution x^* du système linéaire $A\,x = b$ définie par

$$x_B^* = B^{-1}b \qquad \text{et} \qquad x_N^* = {}^{\mathrm{t}}(0,\dots,0)$$

Remarque : Les variables *hors base* d'une solution de base sont toujours nulles.

Exemple

Soit A donné par $(\star)$. Les quatre vecteurs dans la ligne $(\star\star)$ sont des solutions de base de la matrice A. Les autres solutions de base sont

$$X = \begin{pmatrix} 11 \\ -3 \\ 0 \\ 0 \\ 19 \end{pmatrix}, \quad X = \begin{pmatrix} -5 \\ 0 \\ -13 \\ 16 \\ 0 \end{pmatrix}, \quad X = \begin{pmatrix} 0 \\ 5 \\ -3 \\ 11 \\ 0 \end{pmatrix}, \quad X = \begin{pmatrix} 0 \\ 8 \\ 0 \\ 11 \\ -3 \end{pmatrix} \quad \text{et} \quad X = \begin{pmatrix} 0 \\ 0 \\ -8 \\ 11 \\ 5 \end{pmatrix}$$

On remarquera que les cinq derniers vecteurs possèdent au moins une composante négative.

Ainsi, dans l'exemple considéré dans ce cours, les sommets du polyèdre

$$\left\{X \in \mathbb{R}^2 \;\middle|\; MX \le b \text{ et } X \ge 0\right\}$$

sont caractérisés par les solutions de base *à composantes positives* du système $A\,X = b$ associé, où la matrice A est obtenue en accolant la matrice M et la matrice identité de taille 3. On admettra que ce résultat se généralise à tous les polyèdres dans $(\mathbb{R}^+)^Q$. Une telle caractérisation n'étant pas possible pour des polyèdres quelconques, cela justifie que l'on cherche dans la section suivante à s'intéresser à des problèmes d'optimisation linéaire de contraintes incluant la positivité des variables. En effet, on a entraperçu au cours précédent le rôle prépondérant que les sommets du polyèdre des contraintes pouvaient jouer dans la résolution des problèmes d'optimisation linéaire.

4 Forme canonique d'un problème d'optimisation linéaire

L'objectif de ce dernier paragraphe est de montrer que tout problème d'optimisation linéaire peut s'écrire sous la forme d'un problème d'optimisation linéaire avec contraintes de positivité sur les variables.

4.1 Définition

Définition 8 (Forme canonique d'un problème d'optimisation linéaire)

On appelle problème d'optimisation linéaire *sous forme canonique* un problème d'optimisation de la forme

$$\begin{array}{ll} \text{Maximiser} & {}^{\mathrm{t}}c\,x \\ \text{sous les contraintes} & A\,x \le b \\ & \quad x \ge 0 \end{array}$$

où $x, c \in \mathbb{R}^q$, $b \in \mathbb{R}^p$, $A \in \mathcal{M}_{p,q}(\mathbb{R})$.

Remarque : Les problèmes d'optimisation linéaire sous forme canonique sont des cas particuliers de problèmes d'optimisation linéaire sous forme générale, car les contraintes de positivité $x \ge 0$ sont des contraintes affines.

4.2 Équivalence entre les formes générale et canonique

Proposition 1

Soit $(\mathcal{P}_g)$ un problème d'optimisation linéaire sous forme générale. Alors il existe $(q,p) \in (\mathbb{N}^*)^2$ et $x, c \in \mathbb{R}^q$, $b \in \mathbb{R}^p$, $A \in \mathcal{M}_{p,q}(\mathbb{R})$ tels que $(\mathcal{P}_g)$ est équivalent au problème sous forme canonique

$$\begin{array}{ll} \text{Maximiser} & {}^{\mathrm{t}}c\,x \\ \text{sous les contraintes} & A\,x \le b \\ & \quad x \ge 0 \end{array}$$

Plus précisément, on va montrer que le problème d'optimisation linéaire

$$(\mathcal{P}_g) \qquad \begin{array}{ll} \text{Maximiser} & {}^{\mathrm{t}}f\,X \\ \text{sous les contraintes} & G\,X \le g \end{array}$$

avec $f = (f_j)_{1 \le j \le Q}$, $X = (X_j)_{1 \le j \le Q}$, $g = (g_k)_{1 \le k \le P}$ et $G = (G_{kj})_{\substack{1 \le k \le P \\ 1 \le j \le Q}}$, et le problème

$$(\mathcal{P}_c) \qquad \begin{array}{ll} \text{Maximiser} & {}^{\mathrm{t}}c\,x \\ \text{sous les contraintes} & A\,x \le g \\ & \quad x \ge 0 \end{array}$$

avec $c = (c_\ell)_{1 \le \ell \le 2Q}$, $x = (x_\ell)_{1 \le \ell \le 2Q}$ et $A = (a_{k\ell})_{\substack{1 \le k \le P \\ 1 \le \ell \le 2Q}}$ tels que

$$c = \begin{pmatrix} f_1 \\ \vdots \\ f_Q \\ -f_1 \\ \vdots \\ -f_Q \end{pmatrix}, \quad A = \begin{pmatrix} G_{11} & \dots & G_{1Q} & -G_{11} & \dots & -G_{1Q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ G_{P1} & \dots & G_{PQ} & -G_{P1} & \dots & -G_{PQ} \end{pmatrix}$$

sont équivalents.

DÉMONSTRATION : On va décomposer la preuve en trois étapes.
Étape 0 : Soit $X = {}^{\mathrm{t}}\big(X_1, \dots, X_Q\big) \in \mathbb{R}^Q$ et $x = {}^{\mathrm{t}}\big(x_1, \dots, x_{2Q}\big) \in \mathbb{R}^{2Q}$ tels que

$$\forall\, j = 1, \dots, Q, \qquad X_j = x_j - x_{Q+j}$$

Montrons que ${}^{\mathrm{t}}c\,x = {}^{\mathrm{t}}f\,X$ et que $\displaystyle\sum_{\ell=1}^{2Q} a_{k\ell} x_\ell = \sum_{j=1}^{Q} G_{kj} X_j$ pour $k = 1, \dots, P$.

- Écrivons sous forme étendue le produit scalaire ${}^{\mathrm{t}}c\,x$, puis remplaçons les coefficients c_ℓ par leur expression en fonction des f_ℓ :

$$ {}^{\mathrm{t}}c\,x = \sum_{\ell=1}^{2Q} c_\ell x_\ell = \sum_{\ell=1}^{Q} f_\ell x_\ell - \sum_{\ell=Q+1}^{2Q} f_{\ell-Q} x_\ell$$

  Un changement d'indices dans la seconde somme donne alors

$$ {}^{\mathrm{t}}c\,x = \sum_{j=1}^{Q} f_j (x_j - x_{Q+j}) = \sum_{j=1}^{Q} f_j X_j = {}^{\mathrm{t}}f\,X$$

- De la même manière, on obtient

$$\sum_{\ell=1}^{2Q} a_{k\ell} x_\ell = \sum_{\ell=1}^{Q} G_{k\ell} x_\ell - \sum_{\ell=Q+1}^{2Q} G_{k(\ell-Q)} x_\ell = \sum_{j=1}^{Q} G_{kj}(x_j - x_{Q+j}) = \sum_{j=1}^{Q} G_{kj} X_j$$

En particulier, ce dernier point assure que

$$A\,x \leq g \qquad \Longleftrightarrow \qquad G\,X \leq g$$

Étape 1 : Montrer que si $X^* = {}^{\mathrm{t}}\big(X_1^*, \dots, X_Q^*\big) \in \mathbb{R}^Q$ est une solution optimale du problème d'optimisation linéaire sous forme générale $(\mathcal{P}_g)$, alors

$$x^* = {}^{\mathrm{t}}\big(\max(X_1^*, 0), \dots, \max(X_Q^*, 0), -\min(X_1^*, 0), \dots, -\min(X_Q^*, 0)\big) \in \mathbb{R}^{2Q}$$

est une solution optimale du problème d'optimisation sous forme canonique $(\mathcal{P}_c)$.

- Commençons par vérifier que x^* satisfait les contraintes de $(\mathcal{P}_c)$. Tout d'abord, x^* satisfait les contraintes $x^* \leq 0$ car $\max(t, 0) \geq 0$ et $-\min(t, 0) \geq 0$ quel que soit $t \in \mathbb{R}$. Ensuite, x^* satisfait les contraintes $Ax^* \leq g$ car $t = \max(t, 0) + \min(t, 0)$, donc

$$\forall\, j = 1, \dots, Q, \qquad X_j^* = x_j^* - x_{Q+j}^*$$

  Or, puisque X^* est une solution optimale du problème $(\mathcal{P}_g)$, il satisfait en particulier les contraintes $G\,X^* \leq g$. Ainsi, d'après l'étape 0, on a donc $A\,x^* \leq g$.
- Vérifions ensuite que ${}^{\mathrm{t}}c\,x^* \geq {}^{\mathrm{t}}c\,x$ pour tout x admissible (tel que $Ax \leq g$ et $x \geq 0$). Pour tout $x \in \mathbb{R}^{2Q}$, si on pose $X_\ell = (x_\ell - x_{Q+\ell})$ pour tout $\ell = 1, \dots, Q$, alors on a ${}^{\mathrm{t}}c\,x = {}^{\mathrm{t}}f\,X$ d'après l'étape 0. Par ailleurs, si x est un point admissible de $(\mathcal{P}_c)$, alors X est un point admissible de $(\mathcal{P}_g)$ (d'après l'étape 0). Ainsi, puisque X^* est une solution optimale de $(\mathcal{P}_g)$ et que X est admissible, on a ${}^{\mathrm{t}}f\,X^* \geq {}^{\mathrm{t}}f\,X$ soit ${}^{\mathrm{t}}c\,x^* \geq {}^{\mathrm{t}}c\,x$.

Étape 2 : Il nous reste à montrer que si $x^* = {}^{\mathrm{t}}\big(x_1^*, \dots, x_{2Q}^*\big) \in \mathbb{R}^{2Q}$ est une solution optimale du problème d'optimisation linéaire sous forme canonique $(\mathcal{P}_c)$, alors

$$X^* = {}^{\mathrm{t}}\big(x_1^* - x_{Q+1}^*, \dots, x_Q^* - x_{2Q}^*\big) \in \mathbb{R}^Q$$

est une solution optimale du problème d'optimisation sous forme générale $(\mathcal{P}_g)$. Pour cela, on pourra suivre les étapes suivantes.

1. Vérifier que X^* satisfait les contraintes de $(\mathcal{P}_g)$, à savoir

$$\sum_{j=1}^{Q} G_{kj} X_j^* \leq g_k \qquad\qquad \text{pour } k = 1, \dots, P$$

2. Vérifier que ${}^{\mathrm{t}}f\,X^* \geq {}^{\mathrm{t}}f\,X$ pour tout X admissible. *Pour ce dernier point, on associera à tout X admissible le vecteur*

$$x = {}^{\mathrm{t}}\big(\max(X_1, 0), \dots, \max(X_Q, 0), -\min(X_1, 0), \dots, -\min(X_Q, 0)\big) \in \mathbb{R}^{2Q}$$

   dont on montrera qu'il satisfait les contraintes du problème d'optimisation $(\mathcal{P}_c)$, puis on utilisera l'optimalité de x^* pour établir que ${}^{\mathrm{t}}c\,x^* \geq {}^{\mathrm{t}}c\,x$. Enfin, on utilisera l'étape 0 en vérifiant au préalable qu'elle s'applique bien. ■