Méthodes polyédriques et algorithmes

ENSIIE
Optimisation – Recherche opérationnelle

Alain FAYE

Préambule

Dans une première partie, on donne les notions essentielles relatives à la description des polyèdres: dimension, faces, facettes, points extrêmes, représentation minimale.

Dans une deuxième partie, on donne quelques méthodes pour résoudre les problèmes d'optimisation combinatoire décrits à l'aide de programmes linéaires de grande taille c'est-à-dire décrits comme la minimisation ou la maximisation d'une forme linéaire sur un polyèdre comportant un grand nombre de contraintes ou un grand nombre de variables.

A la fin, on propose des exercices parcourant les différents points abordés dans les parties précédentes.

Polyèdres

On considère l'espace $\mathbb{R}^d$, pour un entier d positif, muni du produit scalaire euclidien défini pour tout x, y de $\mathbb{R}^d$ par $xy = \sum_{i=1,d} x_i y_i$.

Espaces affines de $\mathbb{R}^d$

Définition: A, une partie de $\mathbb{R}^d$, est un espace affine si: $x, y \in A \Rightarrow (1-\lambda)x + \lambda y \in A \quad \forall \lambda \in \mathbb{R}$

Définition: une combinaison affine est un point de $\mathbb{R}^d$ de la forme $\sum_{i=1,p} \lambda_i x_i$ où $x_i \in \mathbb{R}^d, \lambda_i \in \mathbb{R}$ t.q. $\sum_{i=1,p} \lambda_i = 1$ et où p est un entier quelconque.

Remarquer que pour p=2 on retrouve une combinaison de la forme $(1-\lambda)x + \lambda y$.

Proposition: soit A un espace affine, A est stable par combinaison affine de ses points.
démonstration: récurrence sur p.

En translatant A d'un point quelconque x de A on obtient un sous-espace vectoriel comme le montre la proposition suivante.

Proposition: soit A un espace affine, il existe un unique sous-espace vectoriel L de $\mathbb{R}^d$ t.q. $A=L+x$ où x est un point de A quelconque.
démonstration : soit $x \in A$. Montrons que $L=A-x$ est un sous-espace de $\mathbb{R}^d$. Soient y_1, $y_2 \in A$, λ_1, λ_2 deux réels. On considère la combinaison linéaire z suivante :

$$z = \lambda_1(y_1 - x) + \lambda_2(y_2 - x) = \lambda_1 y_1 + \lambda_2 y_2 - (\lambda_1 + \lambda_2)x = \lambda_1 y_1 + \lambda_2 y_2 - (\lambda_1 + \lambda_2)x + x - x$$
$$= \lambda_1 y_1 + \lambda_2 y_2 - (\lambda_1 + \lambda_2 + 1)x - x$$

La somme des 3 premiers termes étant une combinaison affine de points de A, z est dans $L=A-x$.

Montrons que L est indépendant du choix de x. Soient $x \in A$, $x' \in A$ et $L=A-x$, $L'=A-x'$. Soit $y \in A$, considérons l'élément z de L' suivant : $z = y - x' = y - x' + x - x$. La somme des 3 premiers termes étant une combinaison affine de points de A, z est dans $L=A-x$. On a donc démontré que $L' \subset L$. L'inclusion inverse se montre de manière similaire.

♦

On dira que L est le sous-espace parallèle à A.

Une intersection d'espaces affines étant encore un espace affine, on définit l'enveloppe affine d'une partie quelconque de 3^d de la façon suivante.

Définition: étant donné M une partie de $\mathbb{R}^d$, Aff$M$, l'enveloppe affine de M, est le plus petit espace affine contenant M c'est-à-dire l'intersection de tous les espaces affines contenant M.

Proposition: Aff$M$ est l'ensemble des combinaisons affines de points de M.
démonstration: l'ensemble des combinaisons affines de M est un espace affine donc contient Aff$M$ par définition.
Aff$M$ étant un espace affine, une combinaison affine de ses points et en particulier des points de M est dans Aff$M$, ce qui montre l'inclusion inverse.

♦

Définition: la dimension d'un espace affine est la dimension de l'unique espace qui lui est parallèle.

On notera dim$A$ la dimension de l'espace affine A.

Définition: $x_1,\dots,x_p$ points de $\mathbb{R}^d$ sont affinement indépendants si $\left.\begin{array}{l}\lambda_1 x_1+\dots+\lambda_p x_p=0\\ \lambda_1+\dots+\lambda_p=0\end{array}\right\}\Rightarrow \lambda_1=\dots=\lambda_p=0$ ou de façon équivalente si $(x_2-x_1),\dots,(x_p-x_1)$ sont linéairement indépendants.

Exemple : les 3 points de $\mathbb{R}^2$ $\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}$ sont affinement indépendants. En effet, le système $\lambda_1\begin{pmatrix}1\\0\\1\end{pmatrix}+\lambda_2\begin{pmatrix}0\\1\\1\end{pmatrix}+\lambda_3\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$ admet une unique solution $\lambda_1=\lambda_2=\lambda_3=0$. Par ailleurs, si on retranche le premier point aux 2 autres on obtient $\begin{pmatrix}-1\\1\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}$ qui sont linéairement indépendants.

De ces deux définitions, il découle que la dimension d'un espace affine A est égal au nombre maximum de points affinement indépendants de A moins 1.

La proposition suivante montre que pour parcourir Aff$M$, les combinaisons affines d'une famille de points affinement indépendants de M suffit.

Proposition: si n+1 est le nombre maximum de points affinement indépendants de M, alors Aff$M$ est l'ensemble des combinaisons affines d'une famille de n+1 points de M affinement indépendants.
démonstration: M contient un maximum de n+1 points affinement indépendants, mettons $x_1,\dots,x_{n+1}$. Soit une combinaison affine $\sum_{i=1,n+1}\lambda_i x_i+\sum_{i=n+2,p}\lambda_i x_i$ de points de M (avec la somme des λ_i valant 1 et éventuellement des λ_i nuls pour i=1,…,n+1), alors un x_i $(n+2\le i\le p)$ peut s'exprimer comme combinaison affine des $x_1,\dots,x_{n+1}$. En remplaçant son expression on obtient une nouvelle combinaison affine où x_i ne figure plus.

♦

Corollaire : si n+1 est le nombre maximum de points affinement indépendants de M, alors la dimension de Aff$M$ est n.
démonstration : d'après la propriété précédente, Aff$M$ = Aff$\{x_1,\dots,x_{n+1}\}$ où $x_1,\dots,x_{n+1}$ sont des points de M affinement indépendants. Donc tout point de Aff$M$ est

combinaison affine de $x_1,\dots,x_{n+1}$ et le nombre maximum de points affinement indépendants de Aff$M$ est n+1 et dim(Aff$M$)=n.

♦

On pourra remarquer que l'écriture d'un point de Aff$M$ en fonction d'une famille fixée de points affinement indépendants est unique (c'est-à-dire les λ_i sont déterminés de façon unique).

Exemple : soit M l'ensemble des 3 points de $\mathbb{R}^2$ $\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\2\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}$ et A=Aff$M$. $\lambda_1\begin{pmatrix}1\\0\\1\end{pmatrix}+\lambda_2\begin{pmatrix}1\\2\\1\end{pmatrix}+\lambda_3\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$ admet pour solution $\lambda_1=\lambda_2=1,\lambda_3=-2$. Donc ces 3 points ne sont pas affinement indépendants. Par contre $\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}$ sont affinement indépendants. Donc $A=\text{Aff}\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}1\\1\end{pmatrix}\right\}$. Donc le nombre maximum de points affinement indépendants de A est 2 et dim$A$=1.

Définition: une application $\varphi:\mathbb{R}^d \mapsto \mathbb{R}^e$ est affine si $\varphi((1-\lambda)x+\lambda y)=(1-\lambda)\varphi(x)+\lambda\varphi(y)\ \ \forall\lambda\in\mathbb{R}$

Proposition: une application affine φ préserve les combinaisons affines c'est-à-dire $\varphi\left(\sum_{i=1,p}\lambda_i x_i\right)=\sum_{i=1,p}\lambda_i\varphi(x_i)$ lorsque $\sum_{i=1,p}\lambda_i=1$.
démonstration: récurrence sur p.

Proposition: une application affine φ est la somme d'une application linéaire et d'une application constante.
démonstration: poser $\varphi(x)=\phi(x)+\varphi(0)$. On vérifie que ϕ est linéaire.

En dimension finie, toute application linéaire étant continue, il en résulte la proposition suivante.

Proposition: une application affine est continue.

Lorsque la dimension de l'espace d'arrivée est e=1 on parle de fonction affine. D'après ce qui précède, une fonction affine est de la forme $\varphi(x)=a_0+\sum_{i=1,d}a_ix_i$ où $a_0,a_{i=1,d}$ sont des constantes réelles. Les fonctions affines permettent de définir des ensembles particuliers de $\mathbb{R}^d$.

Définition: si φ est une fonction affine non constante alors la partie de $\mathbb{R}^d$ $\varphi^{-1}(0)=\{x:\varphi(x)=0\}$ est appelée hyperplan de $\mathbb{R}^d$.

Il est facile de vérifier qu'un hyperplan est un espace affine et que donc l'intersection d'hyperplans est un espace affine. Réciproquement, tout espace affine propre (c'est-à-dire inclus strictement dans R^d) est l'intersection d'hyperplans comme le montre la proposition suivante.

Proposition: un espace affine propre ($\neq \mathsf{R}^d$) est l'intersection d'un nombre fini d'hyperplans.
démonstration: soit A un espace affine propre et L l'espace parallèle, $L \neq \mathsf{R}^d$ sinon A serait R^d . En considérant une base de l'orthogonal de L on établit que $L = \{x{:}\, Bx = 0\}$ où B est la matrice dont les lignes sont les vecteurs de base engendrant $L^{\perp} \neq \{0\}$. En translatant L d'un point a de A, on en déduit que $A = \{x + a{:}Bx = 0\} = \{y{:}By - Ba = 0\}$ et finalement $A = \{x{:}\, Bx - b = 0\}$ en posant b=Ba.

♦

Proposition: la dimension de l'espace affine $A = \{x{:}\, Bx - b = 0\}$ (avec b=Ba pour un a de A) est égale à d-rang$B$.
démonstration: l'espace parallèle à A est $L = \{x{:}\, Bx = 0\}$ et l'on sait que $d = \dim L + \mathrm{rang} B$ (L étant le noyau de l'application linéaire B).

♦

En particulier, un hyperplan est de dimension d-1.

Définition: si φ est une fonction affine non constante alors $\varphi^{-1}([0,+\infty[)=\{x : \varphi(x)\geq 0\}$ et $\varphi^{-1}(]-\infty,0])=\{x : \varphi(x)\leq 0\}$ sont appelés demi-espaces fermés de R^d . On définit de même demi-espaces ouverts en excluant des ensembles précédents $\varphi^{-1}(0)$ c'est-à-dire en considérant $\varphi^{-1}(]0,+\infty[)$ et $\varphi^{-1}(]-\infty,0[)$.

Ensembles convexes de R^d

Définition: C, une partie de R^d, est convexe si: $x, y \in C \Rightarrow (1-\lambda)x + \lambda y \in C \quad \forall \lambda \in [0,1]$

Définition: une combinaison convexe est un point de R^d de la forme $\sum_{i=1,p} \lambda_i x_i$ où $x_i \in R^d, \lambda_i \in R$ t.q. $\sum_{i=1,p} \lambda_i = 1$ et $\lambda_i \geq 0$ pour tout i (p étant un entier quelconque).

Remarquer que pour p=2 on retrouve une combinaison de la forme $(1-\lambda)x + \lambda y$ avec $\lambda \in [0,1]$.

Proposition: soit C un ensemble convexe, C est stable par combinaison convexe de ses points.
démonstration: récurrence sur p.

Une intersection d'ensembles convexes étant encore un ensemble convexe, on définit l'enveloppe convexe d'une partie quelconque de 3^d de la façon suivante.

Définition: étant donné M une partie de R^d, Conv$M$, l'enveloppe convexe de M, est le plus petit ensemble convexe contenant M c'est-à-dire l'intersection de tous les ensembles convexes contenant M.

Proposition: Conv$M$ est l'ensemble des combinaisons convexes de points de M.
démonstration: l'ensemble des combinaisons convexes de M est un ensemble convexe donc contient Conv$M$ par définition.
Conv$M$ étant un ensemble convexe, une combinaison convexe de ses points et en particulier des points de M est dans Conv$M$, ce qui montre l'inclusion inverse.

♦

Définition: la dimension d'un ensemble convexe est la dimension de son enveloppe affine.

On notera dim$C$ la dimension de l'ensemble convexe C.

Proposition : la dimension d'un convexe C est égale au nombre maximum de points affinement indépendants de C moins 1.
démonstration : la dimension de C est, par définition, la dimension de Aff$C$ et la dimension de Aff$C$ est, par le corollaire du chapitre précédent, le nombre maximum de points affinement indépendants de C moins 1.

♦

La proposition suivante montre que pour parcourir Conv$M$, il n'est pas nécessaire de considérer toutes les combinaisons convexes de points de M.

Proposition: étant donné M une partie de $\mathbb{R}^d$, Conv$M$ est l'ensemble des combinaisons convexes de points de M affinement indépendants.
démonstration : soit z la combinaison convexe de points de M suivante : $z = \sum_{i=1,\ldots,p} \lambda_i x_i$ avec $\sum_{i=1,\ldots,p} \lambda_i = 1$ et $\lambda_i>0$ pour tout i , p étant un entier quelconque (dans la somme on ne considère que les termes associés à un λ_i non nul). Si les x_i ne sont pas affinement indépendants cela implique qu'il existe un vecteur $\mu\neq 0$ t.q. $\sum_{i=1,\ldots,p} \mu_i x_i = 0$ et $\sum_{i=1,\ldots,p} \mu_i = 0$. Soit $\alpha>0$, en retranchant α fois le vecteur nul à z on obtient :
$z = \sum_{i=1,\ldots,p} \lambda_i x_i - \alpha \sum_{i=1,\ldots,p} \mu_i x_i = \sum_{i=1,\ldots,p} (\lambda_i - \alpha\mu_i) x_i$
avec $\sum_{i=1,\ldots,p} (\lambda_i - \alpha\mu_i) = \sum_{i=1,\ldots,p} \lambda_i - \alpha \sum_{i=1,\ldots,p} \mu_i = 1$. Si $\mu_i\leq 0$ alors $\lambda_i-\alpha\mu_i\geq 0$, si $\mu_i>0$ alors $\lambda_i-\alpha\mu_i$ peut être <0. Déterminons α pour que si $\mu_i>0$ alors $\lambda_i-\alpha\mu_i\geq 0$. Il faut $\alpha \leq \frac{\lambda_i}{\mu_i}$. En prenant $\alpha = \min_{i=1,\ldots,p \text{ t.q. } \mu_i>0} \frac{\lambda_i}{\mu_i}$, alors pour tout i $\lambda_i-\alpha\mu_i\geq 0$ et nul pour au moins un i celui réalisant le min dans la définition de α. Notons que $\mu\neq 0$ et $\sum_{i=1,\ldots,p} \mu_i = 0$ entraîne qu'il existe au moins un $\mu_i>0$. On a donc réécrit z sous forme d'une combinaison convexe avec au moins un terme de moins dans la somme. On peut réitérer jusqu'à ce qu'il n'y ait plus dans cette somme que des x_i affinement indépendants.

♦

Il faut remarquer que contrairement à l'enveloppe affine, l'enveloppe convexe n'est pas forcément obtenue avec une famille fixée de points de M affinement indépendants.
Par exemple, soit $M = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$. Le point $\begin{pmatrix} \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est dans Conv$M$ mais ne peut s'exprimer sous forme d'une combinaison convexe avec les 3 points affinement indépendants que sont $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

De la proposition précédente, on déduit le corollaire suivant connu sous le nom de théorème de Carathéodory.

Corollaire: si la dimension de Conv$M$ est n alors Conv$M$ est l'ensemble des combinaisons convexes d'au plus $n+1$ points de M .
démonstration: on a Aff(Conv$M$)=Aff$M$ et si dim(Conv$M$)=n alors M contient au plus $n+1$ points affinement indépendants.

♦

Lorsque M est de cardinal fini, l'enveloppe convexe de M donne lieu à des définitions d'ensembles particuliers.

Définition: un polytope est l'enveloppe convexe d'un nombre fini de points. Si les points sont affinement indépendants le polytope est appelé simplexe.

Intérieur relatif des convexes de $\mathbb{R}^d$

Dans $\mathbb{R}^3$ l'intérieur d'un disque inscrit dans un plan est vide. Par contre relativement au plan on peut dire que l'intérieur du disque n'est pas vide. C'est pourquoi il est plus intéressant de considérer la topologie d'un convexe relativement à son enveloppe affine. Etant donné un convexe C de $\mathbb{R}^d$, on muni l'espace affine Aff$C$ de la topologie de $\mathbb{R}^d$ induite sur Aff$C$ c'est-à-dire U est un ouvert de Aff$C$ ssi il existe U' un ouvert de $\mathbb{R}^d$ t.q. $U = U' \cap \mathrm{Aff}C$.

Définition: l'intérieur relatif d'un convexe C, noté ri$C$, est l'intérieur de C dans Aff$C$ c'est-à-dire le plus grand ouvert de Aff$C$ contenu dans C.

Remarquons qu'avec cette définition, l'inclusion d'un convexe dans un autre convexe n'entraîne pas (sauf si leurs enveloppes affines coïncident) l'inclusion de leurs intérieurs relatifs. Considérer par exemple un carré et ses côtés: l'intérieur relatif d'un côté est un segment ouvert qui n'est pas inclu dans l'intérieur du carré.

Dans notre exemple introductif nous avons vu que l'intérieur relatif du disque est non vide. Cette propriété se généralise. Démontrons d'abord le lemme suivant.

Lemme: soit S un simplexe de $\mathbb{R}^d$ alors ri$S$ est non vide.

démonstration: soit $S = \mathrm{Conv}\{x_1,\ldots,x_k\}$ avec $x_1,\ldots,x_k$ affinement indépendants. Un point x de Aff$S$ admet une décomposition unique $\sum_{i=1,k}\lambda_i x_i$ avec $\sum_{i=1,k}\lambda_i = 1$. Donc associer à x ses coordonnées $\lambda_1,\ldots,\lambda_k$ dans $x_1,\ldots,x_k$ est une application (l'image de x est unique) de Aff$S$ dans l'hyperplan de $\mathbb{R}^k$ d'équation $\sum_{i=1,k}\lambda_i = 1$. Notons φ cette application. On remarque que φ est affine donc continue. Notons H l'hyperplan de $\mathbb{R}^k$ d'équation $\sum_{i=1,k}\lambda_i = 1$ et considérons $O = \{(\lambda_1,\ldots,\lambda_k) \lambda_1 > 0,\ldots,\lambda_k > 0\} \cap H$ ouvert de H (comme intersection d'un ouvert de 3^k et de H). $\varphi^{-1}(O)$ est un ouvert de Aff$S$ et $\varphi^{-1}(O) = \{\sum_{i=1,k}\lambda_i x_i : \lambda_1 > 0,\ldots,\lambda_k > 0, \sum_{i=1,k}\lambda_i = 1\} \subset S$. De plus $\varphi^{-1}(O)$ est non vide puisqu'il contient le point $\sum_{i=1,k}\frac{1}{k}x_i$.

♦

Proposition: soit C un convexe (non vide) de $\mathbb{R}^d$ alors ri$C$ est non vide.
démonstration: soit n la dimension de C (C non vide contient au moins un point et est au moins de dimension 0) donc C contient n+1 points affinement indépendants et donc contient un simplexe S enveloppe convexe de $n+1 \geq 1$ points affinement indépendants de C. Alors Aff$S$=Aff$C$ donc ri$C$ contient ri$S$ qui est non vide d'après le lemme précédent et qui est un ouvert de Aff$C$.

♦

Avec la topologie induite, F est un fermé de Aff$C$ ssi il existe F' un fermé de $\mathbb{R}^d$ t.q. $F = F' \cap \mathrm{Aff}C$. Mais Aff$C$ étant un fermé (comme intersection d'hyperplans si propre ou $\mathbb{R}^d$ sinon) $F = F' \cap \mathrm{Aff}C$ est aussi un fermé de $\mathbb{R}^d$. Donc finalement les fermés de Aff$C$ sont les fermés de $\mathbb{R}^d$ inclus dans Aff$C$. Cela entraîne que les fermetures d'un convexe relativement à R^d et relativement à Aff$C$ sont les mêmes. On notera cl$C$ la fermeture du convexe C.

lemme (dit d'accessibilité): soit C un convexe de R^d alors pour tout x_0 de ri$C$ et tout x_1 de cl$C$ le segment ouvert $]x_0, x_1[$ est inclus dans ri$C$.
démonstration: par définition de ri$C$, il existe un ouvert de Aff$C$ inclus dans C et qui contient x_0, soit il existe $\varepsilon > 0$ t.q. $\{x : \|x - x_0\| < \varepsilon\} \cap \mathrm{Aff}C$ est inclus dans C.
Soit $x = (1-\lambda)x_0 + \lambda x_1$ avec $0 < \lambda < 1$ un point du segment ouvert $]x_0, x_1[$.
On va montrer que l'ouvert de Aff$C$, $O = \{z : \|z - x\| < (1-\lambda)\varepsilon\} \cap \mathrm{Aff}C$ est inclus dans C. Noter que x étant dans Aff$C$, O contient x.

Soit $z \in O$.
Soit l'ouvert de R^d $O_1 = \{z' : \|z' - x_1\| < \frac{1}{\lambda}((1-\lambda)\varepsilon - \|z - x\|)\}$. O_1 est un voisinage de x_1 et donc rencontre C car x_1 est dans cl$C$. Alors soit $z_1 \in O_1 \cap C$.
Considérons le point $z_0 = \frac{z - \lambda z_1}{1-\lambda}$. z_0 est dans Aff$C$ et satisfait $\|z_0 - x_0\| < \varepsilon$, donc z_0 est dans C et par suite $z=(1-\lambda)z_0+\lambda z_1$ est dans C par convexité de C.

On a montré que l'ouvert O de Aff$C$ qui contient x est inclus dans C et donc x est dans ri$C$. Finalement le segment $]x_0, x_1[$ est dans ri$C$.

♦

On peut donner maintenant une caractérisation géométrique de l'intérieur relatif.

Proposition: soit C un convexe (non vide) de R^d alors $x \in \mathrm{ri}C$ ssi $x \in C, \forall y \in C\, y \neq x\ \exists z \in C$ t.q. $x \in]y, z[$ (autrement dit tout segment $[y,x]$ peut être prolongé dans C au delà de x)
démonstration. Sens nécessaire: soit $\varepsilon > 0$ t.q. $\{z : \|z - x\| < \varepsilon\} \cap \mathrm{Aff}C \subset C$. La droite $(1-\lambda)x + \lambda y$ est dans Aff$C$. En prenant $\lambda = -\frac{\varepsilon}{2\|y-x\|}$ et $z = (1-\lambda)x + \lambda y$ alors $\|z - x\| < \varepsilon$ et donc z est dans C et de plus x est dans le segment $]y, z[$.
Sens suffisant: prenons y dans ri$C$ (non vide), si $y=x$ alors c'est terminé sinon il existe z de C t.q. $x \in]y, z[$. Par le lemme d'accessibilité le segment $]y, z[$ est dans ri$C$ et par conséquent x aussi.

♦

Définition: le bord relatif d'un convexe C, noté rb$C$, est le complémentaire de ri$C$ dans la fermeture de C c'est-à-dire rb$C$=cl$C$-ri$C$.

Faces d'un convexe de $\mathbb{R}^d$

On rappelle que pour $x,y\in\mathbb{R}^d$ le segment $[x,y]$ est l'ensemble des points z de la forme $z=(1-\lambda)x+\lambda y$ avec $0\leq\lambda\leq1$. Le segment ouvert $]x,y[$ est obtenu en restreignant λ à $0<\lambda<1$ (les extrémités x, y sont exclues).

Définition: soit C un convexe fermé de $\mathbb{R}^d$. Une partie F de C est une face de C si F est convexe et si pour tout x, y distincts de C t.q. le segment ouvert $]x,y[$ rencontre F alors x, y appartiennent à F, autrement dit : $x\neq y\in C$ et $]x,y[\cap F\neq\emptyset \Rightarrow x,y\in F$.

Dans la suite, on considérera toujours C un convexe fermé.

C lui-même est une face de C dite face impropre. Une face propre est une face non vide différente de C.

Définition: un point x de C t.q. $\{x\}$ est une face de C est appelé point extrême de C.

Un point extrême n'est dans l'intérieur relatif d'aucun segment de C. Ext$C$ désigne l'ensemble des points extrêmes de C.

Un point extrême est une face de dimension 0.

Définition: une facette de C est une face de C de dimension $\dim C-1\geq0$.

Proposition: une face est fermée.
démonstration : soit F une face de C et soit $x\in\mathrm{cl}F$. F est convexe, donc il existe $x_0\in\mathrm{ri}F$ (ri$F$ est non vide) et par le lemme d'accessibilité, le segment ouvert $]x_0,x[$ est dans ri$F$ donc dans F. Comme F est une face cela entraîne que $x\in F$. On a donc montré que $\mathrm{cl}C\subset C$. Comme la fermeture de C ,cl$C$,contient C, cela entraîne que $C=\mathrm{cl}C$ et que C est fermé.

♦

Cette dernière propriété permet de parler légitimement de face de face.

Proposition: soit F une face de C et G inclus dans F. G est une face de C ssi G est une face de F.
démonstration: remarquons d'abord que G face de F ou C entraîne que G est convexe.
Supposons G face de C, alors soit x,y de F t.q. $]x,y[$ rencontre G comme x,y sont dans C et que G est une face de C cela entraîne x,y appartiennent à G.
Supposons G face de F, alors soit x,y de C t.q. $]x,y[$ rencontre G , comme G est dans F le segment rencontre F et comme F est face de C cela implique que x,y sont dans F , maintenant comme G est une face de F cela implique que x,y sont dans G.

♦

En considérant des ensembles G réduits à un point, on établit le corollaire suivant.

Corollaire: soit F une face de C, alors $\mathrm{Ext}F = F\cap\mathrm{Ext}C$.

Proposition: soit F une face d'un convexe C t.q. $F \neq C$ alors F est incluse dans rb$C$.
démonstration : supposons que F ne soit pas incluse dans rb$C$. Cela signifie que F contient un point y de ri$C$. Soit $x \in C$, le segment de x à y peut être prolongé dans C au delà de y et il existe donc un point z de C t.q. $y \in]x,z[\subset C$ (caractéristique géométrique de l'intérieur relatif). Donc $]x,z[$ contient y un point de F et comme F est une face, cela entraîne que x, z sont dans F. On a montré que $C \subset F$ et donc $C=F$. Donc si $C \neq F$ alors F est incluse dans rb$C$.

♦

Proposition: soit F une face d'un convexe C alors $F = C \cap \mathrm{Aff}F$.
démonstration : de façon évidente $F \subset C \cap \mathrm{Aff}F$. Maintenant soit $x \in C \cap \mathrm{Aff}F$. $x = \sum_{i=1,\dots,p} \lambda_i x_i$ avec $\sum_{i=1,\dots,p} \lambda_i = 1$ et $x_i \in F$ pour $i=1,\dots,p$. Si $\lambda_i \geq 0$ $\forall i$ alors $x \in \mathrm{Conv}F=F$ et c'est terminé. Supposons qu'il y ait des $\lambda_i<0$. On partitionne les i en $I^+=\{i=1,\dots,p: \lambda_i \geq 0\}$ et $I^-=\{i=1,\dots,p: \lambda_i<0\}$. On pose $\alpha = \sum_{i\in I^+} \lambda_i, \beta = \sum_{i\in I^-} \lambda_i$. Alors $x = \alpha \sum_{i\in I^+} \frac{\lambda_i}{\alpha} x_i + \beta \sum_{i\in I^-} \frac{\lambda_i}{\beta} x_i = \alpha X + \beta Y$ avec $X = \sum_{i\in I^+} \frac{\lambda_i}{\alpha} x_i, Y = \sum_{i\in I^-} \frac{\lambda_i}{\beta} x_i$. X et Y sont combinaisons convexes de points de F et donc sont dans F. Par ailleurs $X = \frac{1}{\alpha} x - \frac{\beta}{\alpha} Y$ et apparaît donc comme une combinaison convexe de x et Y puisque $\frac{1}{\alpha} - \frac{\beta}{\alpha} = 1, \frac{1}{\alpha} > 0, \frac{-\beta}{\alpha} > 0$. Donc le segment ouvert $]x,Y[$ avec $x \in C$ et $Y \in C$ rencontre F en X. Comme F est une face de C cela entraîne que x, $Y \in F$. On a donc montré que $C \cap \mathrm{Aff}F \subset F$.

♦

De cette dernière proposition, on retrouve bien le fait qu'une face est fermée (comme intersection de deux fermés). Par ailleurs, cette proposition nous permet d'obtenir des résultats sur les dimensions des faces.

Corollaire1: soient G, F deux faces d'un convexe C t.q. G est strictement contenue dans F alors $\dim G < \dim F$.
démonstration : soit 2 faces de C t.q. $G \subset F$ alors $\mathrm{Aff}G \subset \mathrm{Aff}F$. Maintenant si $\dim G = \dim F$ alors $\mathrm{Aff}G = \mathrm{Aff}F$ et $F = C \cap \mathrm{Aff}F = C \cap \mathrm{Aff}G = G$.

♦

On en déduit une caractérisation des facettes.

Corollaire2: une facette est une face propre maximale (pour l'inclusion).
démonstration : soit F une facette de C et soit G une face de C t.q. $F \subset G \subset C$. Si F est strictement contenue dans G alors $\dim F < \dim G$ et donc $\dim C-1 < \dim G$ soit $\dim C \leq \dim G$ ce qui entraîne $\dim G = \dim C$ et finalement $G=C$.

♦

Etant donnée une fonction affine non constante φ t.q. $\varphi(x) \geq 0$ pour tout x appartenant à C, on considère l'ensemble $F = C \cap \varphi^{-1}(0)$ où $\varphi^{-1}(0) = \{x : \varphi(x)=0\}$.

Proposition: F est une face de C.
démonstration : on applique la définition d'une face.

Ce type de face est appelée <u>face exposée</u> de C. On dira que F est la face de C induite par l'inégalité $\varphi(x)\geq 0$.

Toutes les faces d'un convexe ne sont pas forcément exposées. Considérons par exemple l'ensemble C de $\mathbb{R}^2$ défini par l'ensemble des solutions du système d'inégalités suivant :

$$\begin{cases} x_2 - x_1^3 \geq 0 \\ x_2 \geq 0 \end{cases}$$

Le point (0,0) est un point extrême de C et on ne peut trouver une droite (i.e. un hyperplan de $\mathbb{R}^2$) t.q. C soit situé d'un coté de la droite et t.q. (0,0) soit l'intersection de cette droite avec C.

Polyèdres de R^d

Définition: un polyèdre est l'intersection d'un nombre fini de demi-espaces fermés.

Un polyèdre P de R^d est donc décrit par l'ensemble des solutions du système d'inégalités :

$$f_i(x) \le \alpha_i \quad i \in I$$

où $f_i(x)$ est une forme linéaire définie sur R^d, α_i une constante réelle et I un ensemble d'indices de cardinal fini.

Note : une forme linéaire f est une fonction de R^d dans R de la forme $f(x) = ax = \sum_{i=1,\ldots,d} a_i x_i$ avec $a_i \in$R pour i=1,..,d. On assimilera par la suite une forme linéaire f au vecteur $a = \begin{pmatrix} a_1 \\ \vdots \\ a_d \end{pmatrix}$ lorsque l'on parlera de rang de formes linéaires.

Un polyèdre étant un convexe, fermé, il possède des faces (par exemple lui-même). On va chercher une caractérisation de ses faces.

1. Caractérisation des faces d'un polyèdre

Soit J un sous-ensemble de I et considérons l'ensemble $F = P \cap \{x : f_i(x) = \alpha_i \ i \in J\}$ où les inégalités de J sont devenues des égalités. Il est facile de vérifier que F est une face de P.

La question que l'on se pose est: obtient-on de cette façon toutes les faces de P?

Proposition: soit F une face non vide d'un polyèdre P et soit $J \subset I$ t.q. $J = \{i \in I : f_i(x) = \alpha_i \ \forall x \in F\}$, alors Aff$F$ = $\{x : f_i(x) = \alpha_i \ i \in J\}$.

démonstration: • $F \subset \{x : f_i(x) = \alpha_i \ i \in J\}$ par définition de J, ce qui entraîne Aff$F \subset \{x : f_i(x) = \alpha_i \ i \in J\}$ puisque le membre de droite est un espace affine.

• Montrons l'inclusion inverse. Nous allons d'abord montrer qu'il existe $x_0 \in F$ t.q. $f_i(x_0) < \alpha_i \ \forall i \in I \setminus J$. Par définition de J, pour tout $i \in I \setminus J$ il existe $x^{(i)} \in F$ t.q. $f_i(x^{(i)}) < \alpha_i$. Si on note k=card($I \setminus J$), alors en prenant $x_0 = \sum_{i \in I \setminus J} \frac{1}{k} x^{(i)}$ on a la propriété voulue.

Soit $y \in \{x : f_i(x) = \alpha_i \ i \in J\}$. Si $y = x_0$ c'est terminé, on a $y \in$ Aff$F$.

Supposons $y \neq x_0$.

Soient $z^+ = x_0 + \varepsilon(y - x_0)$ et $z^- = x_0 - \varepsilon(y - x_0)$ avec $\varepsilon \neq 0$ donné par :

$$\varepsilon = \min\left\{\frac{\alpha_i - f_i(x_0)}{|f_i(y - x_0)|} : i \text{ t.q. } f_i(y - x_0) \neq 0\right\}$$

$(\varepsilon \neq 0$ car $f_i(y - x_0) \neq 0 \Rightarrow i \notin J \Rightarrow \alpha_i - f_i(x_0) > 0)$

ou $\varepsilon \neq 0$ quelconque si $f_i(y - x_0) = 0$ pour tout i.

Il est facile de vérifier que z^+ et z^- sont dans P (ils satisfont toutes les inégalités I). Alors x_0 étant dans $\left]z^-,z^+\right[$ (au milieu du segment) et F étant une face cela entraîne que z^+ et z^- sont dans F . Mais alors $y = \frac{1}{\varepsilon} z^+ + \frac{\varepsilon-1}{\varepsilon} x_0$ est combinaison affine de points de F et est donc dans Aff$F$.

♦

Sachant que si F est une face de P alors $F = P \cap \mathrm{Aff}F$ (cf. chapitre sur les faces), la proposition précédente entraîne qu'une face F quelconque de P s'obtient par l'intersection $F = P \cap \{x : f_i(x) = \alpha_i \ i \in J\}$ où J est l'ensemble des inégalités saturées par les points de F.

Une conséquence est que le nombre de faces d'un polyèdre est fini. Il est borné par le nombre de façon de choisir J dans I soit au plus $2^{|I|}$.

Une autre conséquence est que les polyèdres n'ont que des faces exposées contrairement au cas général des convexes quelconques.

Corollaire : toutes les faces d'un polyèdre P sont des faces exposées.
démonstration : soit F une face de P et soit $J = \{i \in I : f_i(x) = \alpha_i \ \forall x \in F\}$. En posant $f = \sum_{i \in J} f_i$, $\alpha = \sum_{i \in J} \alpha_i$ on obtient une inégalité valide pour P c'est-à-dire $f(x) \leq \alpha$ $\forall x \in P$ et de plus $F = P \cap \{x : f(x) = \alpha\}$. En effet, $f_i(x) \leq \alpha_i \forall i \in J$ et $\sum_{i \in J} f_i(x) = \sum_{i \in J} \alpha_i \Rightarrow f_i(x) = \alpha_i \forall i \in J$.

♦

On peut appliquer cette caractérisation des faces pour trouver les points extrêmes de P. Un point extrême étant une face de dimension 0, les points extrêmes de P sont les solutions (appartenant à P) des systèmes cramériens (solution unique) extraits du système $f_i(x) = \alpha_i \ i \in I$.

<u>Exemple</u> : soit P le polyèdre de $\mathbb{R}^2$ défini par le système suivant :

$$\begin{cases} x_1 + x_2 \leq 1 & (1) \\ x_1 - x_2 \leq 0 & (2) \\ x_1 \geq 0 & (3) \\ x_2 \geq 0 & (4) \end{cases}$$

En « serrant » (1) et (2) on obtient le système $\begin{cases} x_1 + x_2 = 1 \\ x_1 - x_2 = 0 \end{cases}$ ayant pour unique solution $x_1 = x_2 = {}^1/_2$. Ce point satisfait aussi (3) et (4) donc il est dans P. C'est donc une face de dimension 0 c'est-à-dire un point extrême de P.

Par contre, si l'on « serre » (1) et (4), on obtient le système $\begin{cases} x_1 + x_2 = 1 \\ x_2 = 0 \end{cases}$ ayant pour unique solution $x_1=1$, $x_2=0$. Mais ce point n'est pas dans P puisqu'il ne satisfait pas (2). Ce n'est donc pas un point extrême de P.

2. Description minimale d'un polyèdre

On partitionne I en $I^{\leq}$ et $I^{=}$ de la façon suivante: $\begin{cases} P \subset \{x: f_i(x) = \alpha_i\} & \forall i \in I^{=} \\ P \not\subset \{x: f_i(x) = \alpha_i\} & \forall i \in I^{\leq} \end{cases}$

Pour chaque i de $I^{\leq}$ il existe un point de P qui satisfait l'inégalité i strictement au contraire des inégalités i de $I^{=}$ qui sont satisfaites avec égalité par tous les points de P , les ensembles $I^{\leq}$ ou $I^{=}$ pouvant être éventuellement vides.

Alors on peut décrire P par l'ensemble des solutions du système $\begin{cases} f_i(x) = \alpha_i & i \in I^{=} \\ f_i(x) \leq \alpha_i & i \in I^{\leq} \end{cases}$

A noter que P étant une face de P , d'après le paragraphe précédent $\text{Aff}P = \{x: f_i(x) = \alpha_i \ i \in I^{=}\}$ et alors si P est non vide $\dim P = d - \text{rang}\{f_{i \in I^{=}}\}$.

<u>Exemple</u> : soit P le polyèdre de R^2 défini par le système d'inégalités suivant

$$\begin{cases} x_1 + x_2 \leq 1 & (1) \\ 2x_1 - x_2 \leq 1 & (2) \\ -x_1 - x_2 \leq -1 & (3) \\ -x_1 \leq 0 & (4) \\ -x_2 \leq 0 & (5) \end{cases}$$

(1) et (3) entraînent $x_1+x_2=1$ et donc (1) et (3) sont satisfaites à égalité par tous les points de P. $\begin{pmatrix} 0 \\ 1 \end{pmatrix} \in P$ et satisfait strictement (2) et (5), $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \in P$ et satisfait strictement (4). Donc $I^{=}=\{1,3\}$,$I^{\leq}=\{2,4,5\}$. Le rang des formes linéaires $f_1(x)=x_1+x_2$ et $f_3(x)=-x_1-x_2$ est donné par $\text{rang}\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}\right\} = 1$. Donc la dimension de P est $\dim P=2-1=1$.

Il est toujours intéressant d'obtenir une description de P comportant un minimum d'égalités et d'inégalités. On peut toujours éliminer une égalité qui soit combinaison linéaire des autres car cette égalité est redondante. Qu'en est-il des inégalités ?

Proposition: soit $j \in I^{\leq}$ l'indice d'une inégalité non redondante de P c'est-à-dire t.q. $P \neq P_j = \{x: f_i(x) = \alpha_i \ i \in I^{=},\ f_i(x) = \alpha_i \ i \in I^{\leq} \setminus j\}$ alors $F = \{x: f_j(x) = \alpha_j\} \cap P$ est une facette de P .

démonstration : soit $x \in P_j \setminus P$. Par définition, $f_j(x) > \alpha_j$. Nous allons d'abord montrer qu'il existe $x_0 \in P$ t.q. $f_i(x_0) < \alpha_i \ \forall i \in I^{\leq}$. Par définition de $I^{\leq}$, pour tout $i \in I^{\leq}$ il existe $x^{(i)} \in P$ t.q. $f_i(x^{(i)}) < \alpha_i$. Si on note $m=\text{card}(I^{\leq})$, alors en prenant $x_0 = \sum_{i \in I^{\leq}} \frac{1}{m} x^{(i)}$ on a la propriété voulue.

Maintenant, nous allons exhiber un point z de P sur le segment $]x_0,x[$ qui « sature » l'inégalité j uniquement. Pour cela, posons $\lambda = \frac{\alpha_j - f_j(x_0)}{f_j(x) - f_j(x_0)}$. Les propriétés de λ sont : $\lambda>0$, $1 - \lambda = \frac{f_j(x) - f_j(x_0)}{f_j(x) - f_j(x_0)} - \frac{\alpha_j - f_j(x_0)}{f_j(x) - f_j(x_0)} = \frac{f_j(x) - \alpha_j}{f_j(x) - f_j(x_0)} > 0$.

Soit $z=\lambda x-(1-\lambda)x_0$.

- x et x_0 satisfont les contraintes d'égalités de P, donc z aussi.
- x satisfait les contraintes d'inégalités $i \neq j$ et x_0 satisfait strictement les contraintes d'inégalité, donc z satisfait strictement les contraintes d'inégalités $i \neq j$.

- Il reste l'inégalité j. Par construction de λ, z « sature » cette inégalité : $f_j(z)=\lambda f_j(x)+(1-\lambda)f_j(x_0)=\frac{\alpha_j-f_j(x_0)}{f_j(x)-f_j(x_0)}f_j(x)+\frac{f_j(x)-\alpha_j}{f_j(x)-f_j(x_0)}f_j(x_0)=\alpha_j$.

Donc $z\in F$ et l'ensemble $J\subset I$ t.q. $J=\{i\in I : f_i(x)=\alpha_i \ \forall x\in F\}$ est $I^{=}\cup\{j\}$ car l'ajout de toute égalité construite en serrant une inégalité $i\in I^{\leq}\backslash j$ éliminerait z.
Donc $AffF=\{x : f_i(x)=\alpha_i \ i\in I^{=}\cup\{j\}\}$ et dim$F$=dim$P$-1.

♦

Ainsi on peut éliminer de $I^{\leq}$ les inégalités i t.q. $P\cap\{x: f_i(x)=\alpha_i\}$ n'est pas une facette de P.

Exemple :
Soit le polyèdre P de $\mathbb{R}^3$ défini par le système d'inégalités suivant :

$$\begin{cases} x_3\le x_1 & (1)\\ x_3\le x_2 & (2)\\ x_3\ge x_1+x_2-1 & (3)\\ x_1\le 1 & (4)\\ x_2\le 1 & (5)\end{cases}$$

Les 4 points suivants sont dans P : $\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}$. Ils sont affinement indépendants, on en déduit que dim$P$=3.
L'autre façon pour déterminer la dimension de P est de chercher $I^{=}$. Le point de coordonnées $x_1=x_2={}^1/_2$, $x_3={}^1/_4$ satisfait strictement toutes les contraintes (1) à (5) donc $I^{=}=\varnothing$ et P est de dimension 3 (contenu dans aucun hyperplan de $\mathbb{R}^3$).
Soit la face F de P obtenue en « serrant » l'inégalité (4) $F=P\cap\{x: x_1=1\}$.

$x_1=1,(3),(2)\Rightarrow x_3=x_2$. Donc $F\subset\{x: x_1=1, x_3=x_2, x_3=x_1+x_2-1\}$. Par contre $\begin{pmatrix}1\\0\\0\end{pmatrix}\in F$ et ne sature pas (1) et (5). Donc l'ensemble J des inégalités (1,…,5) qui une fois « serrées » contiennent F est J={2,3,4}. On en déduit que $\mathrm{Aff}F=\{x: x_1=1, x_3=x_2, x_3=x_1+x_2-1\}$. Le rang des formes linéaires $x_1, x_3-x_2, -x_3+x_1+x_2$ est déterminé par le rang des vecteurs $\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\-1\\1\end{pmatrix},\begin{pmatrix}1\\1\\-1\end{pmatrix}$ qui est 2. Donc dim$F$=3-2=dim$P$-2. F n'est donc pas une facette. On peut donc retirer l'inégalité (4) de la description de P. On peut d'ailleurs vérifier que (4) peut être obtenue par (2)+(3).

3. Caractérisation des facettes d'un polyèdre

Le problème qui se pose est comment reconnaître parmi les inégalités $I^{\leq}$ celles qui induisent les facettes de P. Ceci n'est pas chose facile mais on propose ici une caractérisation, bien utile, de ces inégalités.

proposition: soit F une face propre de P (c'est-à-dire non vide et $\neq P$) décrite par une inégalité $f(x)\le\alpha$ c'est-à-dire t.q. $F=P\cap\{x: f(x)=\alpha\}$ alors F est une facette de P

ssi pour tout couple (g,β) t.q. $F \subset \{x : g(x) = \beta\}$ il existe des réels $\mu, \lambda_{i \in I^=}$ t.q. $\begin{pmatrix} g \\ \beta \end{pmatrix} = \mu \begin{pmatrix} f \\ \alpha \end{pmatrix} + \sum_{i \in I^=} \lambda_i \begin{pmatrix} f_i \\ \alpha_i \end{pmatrix}$.

Exemple : reprenons le polyèdre précédent P défini par (1),…,(5). Montrons que (5) induit une facette de P c'est-à-dire $F=P\cap\{x : x_1+x_2-x_3=1\}$ est une facette de P.
Soit $g_1x_1+g_2x_2+g_3x_3=\beta$ l'équation d'un (hyper)plan H contenant F. Exhibons quelques points de F qui sont donc dans H et qui satisfont cette équation.
$(1\ 1\ 1)^T$, $(1\ 0\ 0)^T$, $(0\ 1\ 0)^T$ sont dans F. On a donc les relations :
$g_1+g_2+g_3=\beta$, $g_1=\beta$, $g_2=\beta$. On en déduit $g_3=-\beta$. L'équation de H s'écrit donc $\beta(x_1+x_2-x_3)=\beta$. Soit encore $(g_1\ g_2\ g_3\ \beta)^T=\beta(1\ 1\ {-1}\ 1)^T$. En posant $\mu=\beta$, on vérifie la partie suffisante de la proposition (rappelons que P n'est contenu dans aucun hyperplan et $I^=\!=\emptyset$). On en déduit que F est une facette de P.
On peut vérifier que F est bien une facette car les trois points exhibés sont affinement indépendants ce qui implique $\dim F \geq 2$. Comme par ailleurs F est contenue dans un hyperplan $\dim F \leq 3\text{-}1=2$ et $\dim F=2=\dim P\text{-}1$.

4. Projection d'un polyèdre

Dans ce paragraphe, pour plus de commodité, nous décrivons les polyèdres à l'aide de matrices et de vecteurs colonnes de dimensions adéquates.

On considère le polyèdre P défini par l'ensemble des vecteurs (x,y) solutions du système d'inégalités: $Ax+Gy\leq b$, $y \in \mathbb{R}_+^p$, $x\in\mathbb{R}^n$,
où

- A est une matrice m lignes et n colonnes
- G est une matrice m lignes et p colonnes
- b est un vecteur colonne de m lignes

La projection de P sur les variables x est l'ensemble P_x des vecteurs x de $\mathbb{R}^n$ t.q. pour tout $x\in P_x$, il existe $y\in\mathbb{R}_+^p$ t.q. $(x,y)\in P$.

Proposition : P_x est un polyèdre. Il est défini par l'ensemble des x vérifiant les inégalités de la forme $\nu(b-Ax)\geq 0$ où ν vecteur (ligne) de $\mathbb{R}_+^m$ parcourt l'ensemble des points extrêmes du polyèdre $\nu G\geq 0$, $\nu_1+...+\nu_m\leq 1$, $\nu \geq 0$.

On considère le polytope $Q=\text{Conv}M$ avec $M=\{x_1,\ldots,x_p\}$ un ensemble fini de points de $\mathbb{R}^d$. En utilisant la proposition précédente, nous allons montrer que Q est un polyèdre.

Théorème de Weyl : Q est un polyèdre de $\mathbb{R}^d$.
démonstration : par définition de Q, Q est la projection du polyèdre P suivant $P = \{(x,\lambda) : x = \lambda_1 x_1 + \ldots + \lambda_p x_p, \lambda_i \geq 0\ i = 1,\ldots,p, \lambda_1 + \ldots + \lambda_p = 1\}$ sur les variables x. Ce polyèdre comporte d+1 contraintes d'égalité et p contraintes d'inégalité (positivité des variables λ). On peut décomposer une égalité en deux inégalités. On a donc un polyèdre avec $2(d+1)+p$ inégalités. Par la proposition précédente, la projection de P sur les variables x est un polyèdre.

♦

<u>Exemple</u> : soit $M=\left\{\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$ un ensemble de 4 points de R^3. Ecrivons Conv*M* comme un polyèdre de R^3.

$\mathrm{Conv}M=\left\{\lambda_1\begin{pmatrix}1\\0\\0\end{pmatrix}+\lambda_2\begin{pmatrix}0\\1\\0\end{pmatrix}+\lambda_3\begin{pmatrix}0\\0\\1\end{pmatrix}+\lambda_4\begin{pmatrix}1\\1\\1\end{pmatrix}:\lambda_1+\ldots+\lambda_4=1,\lambda_1\geq 0,\ldots,\lambda_4\geq 0\right\}$. En notant x_1, x_2, x_3 les coordonnées des points, on obtient :

$$\begin{cases} x_1 = \lambda_1 + \lambda_4 & (1)\\ x_2 = \lambda_2 + \lambda_4 & (2)\\ x_3 = \lambda_3 + \lambda_4 & (3)\end{cases} \text{ avec } \lambda_1+\ldots+\lambda_4=1\,(4) \text{ et } \lambda_1\geq 0,\ldots,\lambda_4\geq 0$$

On obtient le résultat en éliminant les variables λ des équations (1), (2) et (3).

(1)+(2)+(3)-(4) donne $x_1+x_2+x_3=1+2\lambda_4$ (5). Ensuite $\lambda_4\geq 0$ donne $x_1+x_2+x_3\geq 1$.

(1)-½(5) donne $\frac{1}{2}x_1-\frac{1}{2}x_2-\frac{1}{3}x_3=\lambda_1-\frac{1}{2}$. Ensuite $\lambda_1\geq 0$ donne $x_1-x_2-x_3\geq -1$.

En éliminant de même $\lambda_2\geq 0$ et $\lambda_3\geq 0$, on obtient finalement :

Conv*M* est décrit par les solutions du système $\begin{cases} x_1+x_2+x_3\geq 1\\ x_1-x_2-x_3\geq -1\\ -x_1+x_2-x_3\geq -1\\ -x_1-x_2+x_3\geq -1\end{cases}$

Méthodes d’optimisation

Méthodes de coupes

1. Approche polyédrique

On considère le problème discret :

$$(\mathrm{P})\ \text{maximiser}\ f(x)\ \text{s.c.}\ x \in S$$

où f est linéaire et S est inclus dans $\mathbb{R}^n$ et de cardinal fini.

On se ramène au problème continu :

$$(\mathrm{PC})\ \text{maximiser}\ f(x)\ \text{s.c.}\ x \in \mathrm{Conv}S$$

Il est facile de montrer, utilisant la linéarité de f, qu'il existe une solution optimale de (PC) qui est dans S. Donc résoudre (PC) permet de résoudre (P).

Maintenant si on connaît l'expression de Conv$S$ comme polyèdre, le problème (PC) est un programme linéaire classique que l'on peut résoudre par l'algorithme du simplexe par exemple.

Noter que Conv$S$ est un polytope (cardinal de S fini) et que par théorème (Weyl) Conv$S$ est un polyèdre.

2. Algorithme de coupes

Supposons Conv$S$ décrit par un ensemble d'inégalités I partitionné en $I^=$ et $I^\leq$ (cf. chapitre sur les polyèdres) où $I^=$ sont les contraintes d'égalités décrivant l'enveloppe affine de Conv$S$ (c'est-à-dire l'enveloppe affine de S) et $I^\leq$ les inégalités décrivant les facettes de Conv$S$.
Le cardinal de $I^=$ ne peut excéder n (Aff$S \subset \mathbb{R}^n$) si $I^=$ ne contient pas d'égalités redondantes. Par contre le cardinal de $I^\leq$ peut être très grand devant n. C'est pourquoi on se doit de n'introduire que les inégalités vraiment utiles dans la résolution de (PC). C'est l'objet de l'algorithme de coupes.

Algorithme de coupes

0. choisir un polyèdre initial P contenant Conv$S$ et construit avec les égalités $I^=$ et quelques inégalités de $I^\leq$

1. soit x^* solution du problème : maximiser $f(x)$ s.c. $x \in P$

2. ajouter à P quelques inégalités de $I^\leq$ violées par x^*, s'il n'en existe pas STOP

3. retirer de P les inégalités largement satisfaites par x^* et aller en 1

La partie 1 consiste en la résolution d'un programme linéaire.

La partie 2 consiste à rechercher des inégalités de $I^{\leq}$ violées (non satisfaites) par x^*. Généralement on ajoute à P les inégalités les plus violées. La recherche d'inégalités violées est appelée <u>problème de séparation</u>. C'est le point essentiel de l'algorithme.
La partie 3 est destinée à retirer de P les inégalités qui a priori ne sont plus utiles. Par largement satisfaite on entend $f(x^*)+\varepsilon \leq \alpha$ si l'inégalité est $f(x) \leq \alpha$ où $\varepsilon > 0$ est un paramètre à choisir.

<u>Algorithme à 2 phases</u>

Si $I^{\leq}$ n'est pas connu exhaustivement ou si le problème de séparation n'est pas résolu de façon exacte (on ne détecte pas d'inégalités violées alors qu'il y en a) alors il se peut que l'algorithme de coupes s'arrête sans qu'on ait la solution de (P) c'est-à-dire avec $x^* \notin S$. Il est alors nécessaire de recourir à un algorithme de *Branch and Bound* en partant du dernier polyèdre P généré. C'est ce qu'on appelle un algorithme à 2 phases.

<u>Branch and Cut</u>

Il est possible dans cet algorithme de *Branch and Bound* de mettre en oeuvre l'algorithme de coupes en chacun des noeuds de l'arborescence d'exploration c'est ce qu'on appelle un algorithme *Branch and Cut*.

3. Exemple de problème de séparation

3.1. Problème du stable de poids maximum

Soit un graphe $G=(V,E)$ muni de poids c_i pour tout i de V. Le problème du stable de G de poids maximum s'énonce:

$$\max \sum_{i \in V} c_i x_i$$
$$\text{s.c.} \begin{cases} x_i + x_j \leq 1 & \forall (i,j) \in E \\ x_i \in \{0,1\} & \forall i \in V \end{cases}$$

avec la convention suivante $x_i = \begin{cases} 1 \text{ si } i \text{ appartient au stable cherch} \\ 0 \text{ sinon} \end{cases}$

L'ensemble des vecteurs x qui vérifient les contraintes du problème est l'ensemble des vecteurs caractéristiques des stables de G, noté S(G).

3.2. Inégalités valides pour le problème du stable de poids maximum

Soit $C \subset V$ l'ensemble de sommets d'un cycle de G t.q. $|C| = 2k+1$. Alors l'inégalité $\sum_{i \in C} x_i \leq k$ est valide pour S(G) c'est-à-dire vérifiée par tous les vecteurs de S(G). On appellera cette inégalite *inégalité de cycle impair*.

On se propose de mettre en oeuvre cette famille d'inégalités valides dans un algorithme de coupes avec comme polyèdre initial $P = \{x : x_i + x_j \leq 1 \ \forall (i,j) \in E,\ x_i \geq 0 \ \forall i \in V\}$

3.3. Un algorithme exact pour le problème de séparation des inégalités de cycle impair.

Soit x vérifiant les inégalités $x_i + x_j \leq 1\ \forall (i,j) \in E$. Le problème est de trouver une inégalité de cycle impair violée par x.

On construit le graphe biparti $B = (X \cup X', F)$ avec $X = V$ et $X' = \{i' : i \in V\}$, $F = \{(i,j') : (i,j) \in E\} \cup \{(i',j) : (i,j) \in E\}$.
On munit les arêtes de B de poids de la façon suivante: $(i,j'),(i',j) \in F$ sont munies du poids $1 - x_i - x_j \geq 0$.
On remarque que dans B un chemin de i à i' correspond à un cycle C impair ($|C| = 2k+1$) dans G passant par i. De plus la longueur du chemin est $2k+1-\sum_{j\in C} 2x_j$.
Par exemple soit un chemin 1, 2', 3, 4',..., $(2k)$', $2k$+1, 1' dans B. La valeur du chemin est
$(1-x_1-x_2)+(1-x_2-x_3)+(1-x_3-x_4)+\ldots+(1-x_{2k}-x_{2k+1})+(1-x_{2k+1}-x_1)=$
$2k+1-2x_1-2x_2-2x_3-\ldots-2x_{2k+1}$.

Si la valeur du plus court chemin de i à i' est ≥ 1 on obtient $2k+1-\sum_{j\in C} 2x_j \geq 1$ soit $k \geq \sum_{j\in C} x_j$. On en déduit qu'il n'existe pas d'inégalité cycle violée par x (pour tout cycle passant par i). Par contre si la valeur du plus court chemin est <1 on obtient $k < \sum_{j\in C} x_j$ et on a mis en évidence une inégalité cycle violée par x.

D'où l'algorithme suivant:

Pour tout $i \in V$, chercher le plus court chemin de i à i' dans B.
Considérer le plus court de ces plus courts chemins et reconstituer le cycle C correspondant dans G. Si la valeur du plus court chemin est ≥ 1 alors il n'y a pas d'inégalité cycle violée par x sinon l'inégalité $\sum_{j\in C} x_j \leq \frac{|C|-1}{2}$ est violée par x.

Noter que les poids étant positifs ou nuls, on peut utiliser l'algorithme de Dijkstra pour calculer les plus courts chemins dans B.

Méthode de décomposition

1. Décomposition

On considère le problème (Pb):

$$\max_{x,y}\ cx+hy \ \text{ s.c. } Ax+Gy\leq b\ ,\ y\in \mathbb{R}_+^p\ ,\ x\in X\subseteq \mathbb{R}^n$$

où

A est une matrice m lignes n colonnes
G est une matrice m lignes p colonnes
b est un vecteur colonne de m lignes
c est un vecteur ligne de n colonnes
h est un vecteur ligne de p colonnes

et

X est une partie de $\mathbb{R}^n$

On réécrit (Pb) de la façon suivante:

$$\max_{x,z,y}\ cx+z \ \text{ s.c. } Ax+Gy\leq b\ ,\ z=hy\ ,\ y\in 3_+^p\ ,\ x\in X\subseteq \mathbb{R}^n\ ,\ z\in \mathbb{R}$$

On introduit alors l'ensemble P des vecteurs (x,z,y) vérifiant $Ax+Gy\leq b$, $z=hy$, $y\in \mathbb{R}_+^p$, $x\in X\subseteq \mathbb{R}^n$, $z\in \mathbb{R}$,

et on réécrit le problème:

$$\max_{x,z}\ cx+z \ \text{ s.c. } (x,z)\in P_{x,z} \qquad \text{(Pb')}$$

où $P_{x,z}$ est la projection de P sur l'espace des variables x,z

Par définition de la projection, $(x,z)\in P_{x,z}$ ssi $\exists y$ t.q. $(x,z,y)\in P$. Pour $x\in X$ et z fixés, cette dernière condition revient à dire qu'il existe $y\in \mathbb{R}_+^p$ solution du système $Gy\leq b-Ax$, $hy=z$. Par le théorème de Farkas cette dernière condition est équivalente à:
$\forall \upsilon\in \mathbb{R}_+^m$, $\forall \mu\in \mathbb{R}$ t.q. $\upsilon G+\mu h\geq 0$ on a $\upsilon(b-Ax)+\mu z\geq 0$

2. Algorithme de coupes

On s'intéresse maintenant à la résolution du problème écrit sous la forme (Pb').
On va décrire un algorithme qui génère de façon dynamique les inégalités utiles décrivant la projection de P sur les variables x, z.

<u>algorithme</u>:
soit I un ensemble de quelques inégalités valides pour $P_{x,z}$
1. maximiser $cx+z$ s.c. $x\in X$ et (x,z) vérifie les inégalités I
soit $(\bar{x},\bar{z})$ une solution

2. Si $(\bar{x},\bar{z}) \in P_{x,z}$

alors STOP on a résolu (Pb').

sinon

générer une inégalité séparant $(\bar{x},\bar{z})$ et $P_{x,z}$ (c'est-à-dire violée par $(\bar{x},\bar{z})$ et valide pour $P_{x,z}$)

l'ajouter à I et aller en 1.

Le point crucial est le problème de séparation (étape 2 de l'algorithme).
On cherche y t.q. $(\bar{x},\bar{z},y) \in P$ et s'il n'en existe pas on doit exhiber une inégalité valide pour $P_{x,z}$ et violée par $(\bar{x},\bar{z})$.

Pour cela on résout le problème suivant:

$$\min_{s,y} s \equiv \max_{s,y} -s \quad \text{s.c.} \begin{cases} Gy - s\mathbf{1} \le b - A\bar{x} \\ hy - s \le \bar{z} \\ hy + s \ge \bar{z} \end{cases}, \ y \in \mathbb{R}_+^p,\ s \in \mathbb{R}_+ \quad \text{(SP)}$$

où $\mathbf{1}$ est le vecteur colonne constitué de m coordonnées valant 1.

Si la solution de (SP) vérifie s=0 alors il existe y t.q. $(\bar{x},\bar{z},y) \in P$ c'est-à-dire $(\bar{x},\bar{z}) \in P_{x,z}$.
Sinon (s>0) considérons le dual de (SP), υ, μ_1,μ_2 ses variables duales et l'inégalité $\upsilon(b - Ax) + (\mu_1 - \mu_2)z \ge 0$ (i)
Les variables duales vérifient $\upsilon G + (\mu_1 - \mu_2)h \ge 0$ et $\upsilon \in \mathbb{R}_+^m$
alors d'après le théorème de Farkas (cf. premier paragraphe) l'inégalité (i) est valide pour $P_{x,z}$.
De plus la fonction objectif du dual est $\upsilon(b - A\bar{x}) + (\mu_1 - \mu_2)\bar{z} = -s < 0$ et l'inégalité (i) est donc violée par $(\bar{x},\bar{z})$.

3. Cas particuliers

Si h=0 alors on peut poser z=0 et le problème (Pb') revient à maximiser cx sur la projection de l'ensemble P des vecteurs (x,y) vérifiant $Ax+Gy \le b$, $y \in \mathbb{R}_+^p$, $x \in X$, sur les variables x.

<u>Méthode de génération de colonnes</u>

On considère le problème (Pb) suivant:

(Pb) $\max_x f(x)$ s.c. $g_j(x) \geq 0 \;\; j \in J,\, x \in X$

où $X = \{x_1, \ldots, x_p\}$ est une partie finie de $\mathbb{R}^n$, J est l'ensemble (fini) des indices des contraintes d'inégalités.
On suppose f linéaire, g_j affine pour tout j.

On suppose également savoir résoudre efficacement le problème d'optimisation d'une forme linéaire sur X même lorsque p est grand.

On considère la relaxation suivante de (Pb) où $x \in X$ est relaxé à $x \in \mathrm{Conv}X$:

(P) $\max_x f(x)$ s.c. $g_j(x) \geq 0 \;\; j \in J,\, x \in \mathrm{Conv}X$

1. Génération de colonnes

Sachant que x dans l'enveloppe convexe de X s'écrit comme combinaison convexe des éléments de X, par linéarité de f et des g_j , (P) se réécrit:

(P′) $\max_\lambda \sum_{i=1}^{p} \lambda_i f(x_i)$ s.c. $\sum_{i=1}^{p} \lambda_i g_j(x_i) \geq 0 \;\; j \in J,\; \sum_{i=1}^{p} \lambda_i = 1$, $\lambda_i \geq 0 \;\; i = 1, \ldots, p$

(P′) est un programme linéaire en les variables (λ_i). Lorsque p est grand, le nombre de variables de ce problème est grand d'où la nécessité de n'utiliser que les variables (colonnes) utiles.

<u>algorithme de génération de colonnes</u>:

Soit I un sous-ensemble de $\{1,\ldots,p\}$ contenant quelques indices t.q. pour au moins un i de I, x_i élément de X, vérifie les contraintes $g_j(x) \geq 0 \;\; j \in J$.

1. résoudre $\max_\lambda \sum_{i \in I} \lambda_i f(x_i)$ s.c. $\sum_{i \in I} \lambda_i g_j(x_i) \geq 0 \;\; j \in J,\; \sum_{i \in I} \lambda_i = 1,\; \lambda_i \geq 0 \;\; i \in I$
Soit $\mu_j \;\; j \in J,\; \eta$ les variables duales optimales

2. recherche d'une variable entrante de coût réduit positif
$\bar{\gamma} = \max_{x \in X} f(x) + \sum_{j \in J} \mu_j g_j(x) - \eta$
Soit $x_i \in X$ réalisant le max
Si $\bar{\gamma} > 0$
alors ajouter l'indice i à I et aller en 1.
sinon tous les coûts réduits sont négatifs ou nuls STOP on a résolu (P′)

Commentaires.

A chaque itération de l'algorithme, par dualité η est égal à la valeur du problème résolu en 1.

Soit $x^* = \sum_{i=1}^{p} \lambda_i^* x_i$ solution de (P') , on a:

$$\eta \leq f(x^*) \leq f(x^*) + \sum_{j \in J} \mu_j g_j(x^*) = \sum_{i=1}^{p} \lambda_i^* \left(f(x_i) + \sum_{j \in J} \mu_j g_j(x_i) \right) \leq \max_{x \in X} f(x) + \sum_{j \in J} \mu_j g_j(x)$$

(la deuxième inégalité étant due au fait que $g_j(x^*) \geq 0$ et $\mu_j \geq 0$ pour tout $j \in J$)

et finalement $\eta \leq f(x^*) \leq \eta + \bar{\gamma}$

Donc à chaque itération, on a un encadrement de la valeur optimale de (P′).

En particulier lorsque $\bar{\gamma} \leq 0$, η est ègal à $f(x^*)$ et (P′) est résolu.

D'autre part si l'on peut se contenter d'une solution approchée, on peut arrêter l'algorithme lorsque $\bar{\gamma}$ atteint une valeur suffisamment petite.

Soit (λ_i^*) la solution trouvée à la fin de l'algorithme (implicitement $\lambda_i^* = 0 \quad i \notin I$).

Si $\lambda_i^* = 1$ pour un i (et donc 0 pour les autres) alors on a résolu (Pb) le problème sur X (non relaxé). Sinon on peut toujours rajouter à J des inégalités valides et violées par $x^* = \sum_{i=1}^{p} \lambda_i^* x_i$ la solution trouvée et réitérer l'algorithme.

2. Génération alternative de colonnes et de coupes

Supposons connu un ensemble d'inégalités valides pour l'ensemble $X \cap \{x : \; g_j(x) \geq 0 \quad j \in J\}$

algorithme de génération de colonnes et de coupes:

1. Résoudre (P′) par l'algorithme de génération de colonnes du paragraphe précédent
Soit $x^* = \sum_{i=1}^{p} \lambda_i^* x_i$ la solution trouvée

2. S'il existe une inégalité valide violée par x^*
   alors l'ajouter à J et aller en 1.
   sinon STOP

Noter que d'une itération à l'autre quand on réapplique l'algorithme de génération de colonnes (étape 1), on peut repartir du dernier ensemble de colonnes I trouvé.

3. Dualité lagrangienne

On considère ici (Pb) et son problème dual lagrangien (D) obtenu par relaxation des contraintes $g_j(x)\geq 0$ $j\in J$.

(D) $\min_{\alpha\geq 0} \max_{x\in X} f(x) + \sum_{j\in J} \alpha_j g_j(x)$

Les contraintes de (Pb) sont relâchées et injectées dans la fonction objectif via des multiplicateurs α positifs ou nuls. Pour x satisfaisant les contraintes de (Pb), $\alpha_j g_j(x)\geq 0$ et donc $\max_{x\in X} f(x) + \sum_{j\in J} \alpha_j g_j(x) \geq f(x) + \sum_{j\in J} \alpha_j g_j(x) \geq f(x)$ et ceci $\forall\alpha\geq 0$. La valeur de (D) est donc supérieure ou égale à la valeur de (Pb).

En introduisant une variable supplémentaire z, (D) se met sous la forme d'un programme linéaire avec autant de contraintes qu'il y a d'éléments dans X.

(D) $\min_{\alpha\geq 0, z} z$ s.c. $z \geq f(x) + \sum_{j\in J} \alpha_j g_j(x) \quad \forall x \in X$

Théorème
Le problème (D) est le dual de (P′).
démonstration : écrire le dual du programme linéaire (D) avec les variables $\lambda_i\geq 0$ $i=1,\ldots,p$ associées à chacune des contraintes de (D).

Il résulte du théorème précédent que les problèmes (D) et (P′), et donc (P), donnent la même valeur. Donc faire la relaxation lagrangienne de (Pb) ou relaxer $x\in X$ en $x\in \mathrm{Conv}X$ sont deux méthodes équivalentes dans le sens où elles donnent la même valeur.

Références bibliographiques

An Introduction to Convex Polytopes. Arne Brondsted. Graduate Texts in Mathematics, Springer Verlag (1983)

Integer and Combinatorial Optimization. George L. Nemhauser, Laurence A. Wolsey. Wiley (1988)

Etude Géométrique des Espaces Vectoriels II, Polyèdres et Polytopes Convexes. Jacques Bair, René Fourneau. Lecture Notes in Mathematics, Springer Verlag (1980)

Etude Géométrique des Espaces Vectoriels, Une Introduction. Jacques Bair, René Fourneau. Lecture Notes in Mathematics, Springer Verlag (1975)

Programmation mathématique (tomes 1 et 2). M. Minoux. Dunod (1983)

Optimisation mathématique. A. Faye. Cours IIE 1ère année

Convex Analysis. R.T. Rockafellar. Princeton University Press (1970)

Anciennes démonstrations

1. Caractérisation des faces d'un polyèdre

Soit J un sous-ensemble de I et considérons l'ensemble $F = P \cap \{x: f_i(x) = \alpha_i \ i \in J\}$ où les inégalités de J sont devenus des égalités. Il est facile de vérifier que F est une face de P.

La question que l'on se pose est: obtient-on de cette façon toutes les faces de P?

Proposition: soit F une face non vide d'un polyèdre P et soit $x_0 \in \mathrm{ri}F$, alors $\mathrm{Aff}F = \{x: f_i(x) = \alpha_i \ i \in J\}$ où $J = \{i \in I: f_i(x_0) = \alpha_i\}$ (J est l'ensemble des inégalités "saturées" par x_0).

démonstration: par la caractérisation géométrique de l'intérieur relatif on a $\forall y \in F (y \neq x_0)\ \exists z \in F$ t.q. $x_0 \in]y,z[$ c'est-à-dire $x_0 = (1-\lambda)y + \lambda z$ avec $0 < \lambda < 1$. Il est facile alors de vérifier que $f_i(x_0) = \alpha_i \Rightarrow f_i(y) = f_i(z) = \alpha_i$ et donc finalement tout y de F satisfait avec égalité les inégalités de J c'est-à-dire $F \subset \{x: f_i(x) = \alpha_i \ i \in J\}$ ce qui entraine $\mathrm{Aff}F \subset \{x: f_i(x) = \alpha_i \ i \in J\}$ puisque le membre de droite est un espace affine.
Montrons l'inclusion inverse. Soit $y \in \{x: f_i(x) = \alpha_i \ i \in J\}$. Si $y=x_0$ c'est terminé.
Supposons $y \neq x_0$.
Soient $z^+ = x_0 + \varepsilon(y - x_0)$ et $z^- = x_0 - \varepsilon(y - x_0)$ avec $\varepsilon \neq 0$ donné par $\varepsilon = \min\left\{\frac{\alpha_i - f_i(x_0)}{|f_i(y - x_0)|}: i \text{ t.q. } f_i(y - x_0) \neq 0\right\}$ ($\varepsilon \neq 0$ car $f_i(y - x_0) \neq 0 \Rightarrow i \notin J \Rightarrow \alpha_i - f_i(x_0) > 0$) et quelconque si $f_i(y - x_0) = 0$ pour tout i. Il est facile de vérifier que z^+ et z^- sont dans P (ils satisfont toutes les inégalités I). Alors x_0 étant dans $]z^-, z^+[$ (au milieu du segment) et F étant une face cela entraine que z^+ et z^- sont dans F . Mais alors $y = \frac{1}{\varepsilon} z^+ + \frac{\varepsilon - 1}{\varepsilon} x_0$ est combinaison affine de points de F et est donc dans $\mathrm{Aff}F$.