

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTRE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
ECOLE NATIONALE SUPERIEURE DE LAGHOUAT



DEPARTEMENT : MATHEMATIQUES

MEMOIRE DE MAGISTER

En vue de l'obtention du diplôme de Magister en Mathématiques

Option : Analyse Fonctionnelle Et EDP

THEME DU PROJET:

Application de la théorie des opérateurs pour la résolution d'un problème d'optimisation quadratique

Présenté par :

M. Lakehal Ismail.

Dirigé par :

Pr. Mokhtari Abdelkader

DEVANT LES MEMBRES DU JURY

Mme Bendaoud Zohra ,	M.C.A, Université de Laghouat,	Présidente.
M. Bentobache Mohand,	M.C.A, Université de Laghouat,	Examineur.
Mme. Yamna Boukhatem,	M.C.A, Université de Laghouat,	Examineur.
M. Mokhtari Abdelkader,	Pr, Université de Laghouat,	Encadreur.

ANNÉE UNIVERSITAIRE : 2017-2018

Résumé

Le travail présenté dans ce mémoire porte sur les applications de la théorie des opérateurs. En particulier sur la projection orthogonale, théorème de Riesz et les conditions de K.K.T. Le but essentiel est la résolution d'un problème d'optimisation quadratique sous contraintes linéaires, l'étude du problème dual, et la décomposition de la fonction-objectif.

Abstract

This work deals with the applications of the theory of operators. Particularly, the orthogonal projection, theorem of Riesz and the conditions of K.K.T.

The essential goal is the resolution of the problem of quadratic optimization with linear constraints, the study of dual problem, and the decomposition of the objective function.

Remerciements

J'aimerais en premier lieu remercier mon Dieu Allah qui m'a donné la volonté et le courage pour la réalisation de ce travail.

Je tiens remercier vivement mon Encadreur de mémoire, Pr **Mokhtari Abdelkader**, pour ces conseils, son encouragement et sa disponibilité.

Je remercie également Madame **Bendaoud Zohra**, qui m'a fait l'honneur de présider le jury de ce memoire.

Je remercie, Dr **Bentobache Mohand**, qui a accepté d'examiner ce memoire et faire partie du jury.

Je tiens à remercier aussi Madame **Boukhatem Yamna**, qui a accepté de juger ce travail. J'adresse un grand merci à toute ma famille qui a toujours été présente lorsque j'en ai eu besoin, en particulier à ma mère .

Enfin, je remercie toutes les personnes qui ont participé de près ou de loin à la réalisation de ce memoire .

Dédicace

Je dédie ce modeste travail :

-A mes parents,

-A mes frères,

-A mes soeures,

-A toute la famille

Je tiens à remercier l'ensemble de tous les étudiants et étudiantes de ma promotion.

En fin je dédie ce memoire à tous mes collègues.

Table des matières

Introduction	6
1 Calcul différentiel	9
1.1 Applications continument différentiable	9
1.2 Formule des accroissements finis	12
1.3 Problème d'extremum	13
1.4 Condition pour l'existence d'un extremum	13
1.4.1 Différentielles d'ordre supérieur	13
1.5 Maximum et minimum relatifs	15
1.5.1 Notion sur les formes quadratiques	16
1.5.2 Maximum et minimum relatifs liés	19
2 Rappel d'analyse convexe	22
2.1 Les ensembles convexes	22
2.1.1 Caractéristiques des ensembles convexes	24
2.1.2 Ensembles convexes polyédriques	26
2.1.3 Enveloppes conique	31
2.1.4 Structure des ensembles admissibles	34
2.2 Fonctions convexes	39
2.2.1 L'hyperplan tangent à l'hypersurface	41
2.2.2 Minimum et maximum d'une fonction convexe	44
3 Optimisation de la fonction concave	48
3.1 Décomposition de la fonction-objectif	48
3.2 Optimisation de la fonction concave	49

4	Programmation linéaire et non-linéaire	57
4.1	Dual d'un programme linéaire	57
4.1.1	Problème dual d'un problème de programmation linéaire (P . L)	57
4.1.2	Construction du dual	64
4.2	Introduction à la programmation non linéaire	70
4.2.1	Lemme de Farkas- Minkowski	71
4.2.2	Les relations de Kuhn et Tucker	73
4.2.3	Lagrangien et points -selles	81
4.3	Points selles d'une forme quadratique	90
	Conclusion	96
	Bibliographie	97

la programmation quadratique, sujet de notre travail. Cependant, un problème de programmation quadratique diffère d'un problème de programmation linéaire seulement dans le fait que la fonction-objectif contient en plus les termes x_j^2 et $x_j x_i$ ($j \neq i$). La programmation quadratique convexe est ainsi considérée comme une transition naturelle de la programmation linéaire vers la programmation non-linéaire.

La linéarité des dérivées partielles de la fonction-objectif quadratiques convexe a permis d'établir une théorie très élégante, importante par elle-même. Les programmes quadratiques peuvent apparaître dans divers domaines :

- **Régression** : Trouver la meilleure approximation par moindres carrés dans le cas où on sait à priori que certains paramètres satisfont des contraintes d'inégalités.
- **Production** : Maximisation des profits dans le cas de fonctions de production linéaires et de variation linéaire des coûts marginaux.
- **Variance minimum** : Résoudre un programme linéaire comportant des coefficients de coût variable avec espérances de coût données et variance minimum.
- **Programmes convexes** : Minimiser une fonction convexe quelconque sous des contraintes linéaires.

Le sujet de ce mémoire a été consacré à l'optimisation des fonctions quadratiques quelconques sous contraintes linéaires, on s'intéresse à l'illustration de quelques méthodes. Le contenu de ce mémoire est réparti en quatre chapitres :

Au premier et deuxième chapitres, nous avons rappelé brièvement les définitions de quelques notions importantes, ainsi que quelques propriétés, comme la notion de convexité et le calcul différentiel qui permettent d'obtenir des résultats d'existence et d'unicité de la solution du problème d'optimisation. Nous envisageons particulièrement présenter les théorèmes les plus utiles qui synthétisent les propriétés de minimisation (ou maximisation) d'une fonction convexe (ou concave).

Dans le troisième chapitre, on s'intéressera à l'optimisation de la fonction concave :

$$\varphi(x) = \sum_{i=1}^n \alpha_i x_i + \sum_{i=1}^n \beta_i x_i^2 \quad \text{avec } \beta_i < 0, \forall i = 1 \dots n .$$

Pour calculer l'optimum, on a utilisé la méthode présentée dans [6].

Dans le quatrième chapitre, on a étudié les points suivants :

- Problème dual d'un problème de programmation linéaire.
- Introduction à la programmation non-linéaire.
- Lemme de Farkas-Minkowski.

- Relations de KKT.
- Les points-selles.
- Points-selles d'une forme quadratique.

Chapitre 1

Calcul différentiel

Dans ce chapitre, nous allons étudier quelques rappels sur le calcul différentiel, nous allons nous intéresser plus particulièrement aux dérivées.

1.1 Applications continument différentiable

Dérivée suivant un vecteur

Proposition 1.1.1. Soit f une application d'un ouvert $U \subset \mathbb{R}^n$ dans \mathbb{R} , si $u \in U$ et h est un vecteur non nul de \mathbb{R}^n , il existe un réel $\delta > 0$ tel que pour tout $t \in]-\delta, \delta[$, $u + th \in U$. On peut donc définir l'application $f_h :]-\delta, \delta[\rightarrow \mathbb{R}$ par

$$t \longrightarrow f(u + th).$$

Car f est définie sur U dans \mathbb{R} .

Définition 1.1.1. (Dérivée en un point suivant un vecteur)

Avec les hypothèses et les notations de la proposition (3.1.1), on dit que f admet une dérivée au point u suivant le vecteur h si la fonction f_h admet une dérivée en 0, on la note

$$\lim_{t \rightarrow 0} \frac{f(u + th) - f(u)}{t} = f'_h(0) = D_h f(u).$$

Définition 1.1.2. (Dérivées partielles)

Avec les hypothèses et les notations de la proposition (3.1.1), on dit que f admet une dérivée

partielle au point u suivant la variable d'indice j , si elle admet une dérivée au point u suivant le vecteur e_j , laquelle est notée par :

$$\lim_{t \rightarrow 0} \frac{f(u + te_j) - f(u)}{t} = D_{e_j} f(u) = D_j f(u) = f'_{e_j}(0).$$

Définition 1.1.3. On dit que la fonctionnelle f définie sur un voisinage de $u \in \mathbb{R}^n$ est dérivable (ou différentiable) en u au sens de Fréchet, s'il existe une forme linéaire $df(u)$ continue sur \mathbb{R}^n , qui vérifie :

$$\forall h \in \mathbb{R}^n, f(u + h) = f(u) + df(u).h + \|h\|\varepsilon(h),$$

où $\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}$ avec $\varepsilon(h) \rightarrow 0$ quand $\|h\| \rightarrow 0$.

L'opérateur linéaire $df(u)$ lui-même s'appelle dérivée forte. $df(u)$ est un opérateur linéaire continu.

Remarque 1.1.1.

1) D'après le théorème de représentation de Riesz, $df(u)$ est une forme linéaire sur \mathbb{R}^n , donc il existe $P \in \mathbb{R}^n$ tel que

$$df(u).h = \langle P, h \rangle \quad \forall h \in \mathbb{R}^n.$$

On appelle P le gradient de f en u et on note $P = \nabla f(u)$. Alors on a

$$df(u).h = \nabla f(u).h.$$

2) Si de plus $u \rightarrow df(u)$ est continue, alors on dit que f est de classe C^1 .

Définition 1.1.4. On dit que la fonctionnelle f , définie sur un voisinage de $u \in \mathbb{R}^n$ est dérivable (ou différentiable) au sens de Gâteaux en u , s'il existe $P \in \mathbb{R}^n$ tel que

$$\forall d \in \mathbb{R}^n \quad \lim_{t \rightarrow 0} \frac{f(u + t.d) - f(u)}{t} = \langle P, d \rangle = Df(u, d).$$

On l'appelle différentielle de Gâteaux ou différentielle faible.

Remarque 1.1.2. [15].

Une fonction dérivable au sens de Fréchet l'est aussi au sens de Gâteaux, mais la réciproque est fausse.

Soit f un fonction dérivable au sens de Fréchet .Alors il existe une forme linéaire $df(u)$ continue sur \mathbb{R}^n , qui vérifie, $\forall h \in \mathbb{R}^n$, $f(u+h) = f(u) + df(u).h + \|h\|\varepsilon(h)$. Donc si $t \in \mathbb{R}^*$, on a $f(u+th) - f(u) = df(u).th + \|th\|\varepsilon(th)$, avec $\varepsilon(h) \rightarrow 0$

$$\frac{f(u+th) - f(u)}{t} = df(u).h + \|h\|\varepsilon(th).$$

Mais on a

$$\lim_{t \rightarrow 0^+} \|h\|\varepsilon(th) = 0.$$

Par suite, on a

$$\lim_{t \rightarrow 0} \frac{f(u+th) - f(u)}{t} = df(u).h.$$

Alors f est dérivable au sens de Gâteaux.

Mais la réciproque est fausse

Exemple

Soit la fonction de deux variables

$$f(x_1, x_2) = \begin{cases} \frac{x_1^3 \cdot x_2}{x_1^4 + x_2^2} & \text{si } (x_1, x_2) \neq (0, 0), \\ 0 & \text{si } (x_1, x_2) = (0, 0). \end{cases}$$

Cette fonction est continue partout sur le plan car

$$\begin{aligned} \lim_{(x_1, x_2) \rightarrow (0, 0)} f(x_1, x_2) &= \lim_{r \rightarrow 0} \frac{r^3 \cos^3(\theta) \cdot r \sin(\theta)}{r^4 \cos^4(\theta) + r^2 \sin^2(\theta)} = \lim_{r \rightarrow 0} \frac{r^4 \cos^3(\theta) \cdot \sin(\theta)}{r^2 (r^2 \cos^4(\theta) + \sin^2(\theta))} \\ &= \lim_{r \rightarrow 0} \frac{r^2 \cos^3(\theta) \cdot \sin(\theta)}{r^2 \cos^4(\theta) + \sin^2(\theta)} = 0 \end{aligned}$$

et on a

$$\lim_{t \rightarrow 0} \frac{f(0+th) - f(0)}{t} = \lim_{t \rightarrow 0} \frac{t^4 h_1^3 \cdot h_2}{t(t^4 h_1^4 + t^2 h_2^2)} = \lim_{t \rightarrow 0} \frac{t^4 h_1^3 \cdot h_2}{t^3 (t^2 h_1^4 + h_2^2)} = \lim_{t \rightarrow 0} \frac{t h_1^3 \cdot h_2}{t^2 h_1^4 + h_2^2} = 0.$$

Néanmoins, cette différentielle n'est pas la partie principale de l'accroissement de la fonction f au point $(0,0)$.

En effet, si l'on pose $h_2 = h_1^2$ on obtient.

$$\lim_{\|h\| \rightarrow 0} \frac{f(h_1; h_2) - f(0; 0)}{\|h\|} = \lim_{h_1 \rightarrow 0} \frac{h_1^5}{2h_1^4 \cdot \sqrt{h_1^2 + h_1^4}} = \lim_{h_1 \rightarrow 0} \frac{1}{2 \cdot \sqrt{1 + h_1^2}} = \frac{1}{2} \neq 0.$$

Car ($f(0+h) - f(0) = df(0).h + O(\|h\|)$).

1.2 Formule des accroissements finis

Proposition 1.2.1. Soit O un ouvert de \mathbb{R}^n , et le segment $[x_0, x] \subset O$

$$f : O \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$$

ayant en tout point du segment $[x_0, x]$ une dérivée faible f' . Alors on a

$$|f(x) - f(x_0)| \leq \sup_{0 \leq \theta \leq 1} |f'(x_0 + \theta \Delta x)| \cdot \|\Delta x\| : (\Delta x = x - x_0)$$

Preuve. Considérons la fonction $g(t) = \varphi(f(x_0 + t\Delta x))$ définie pour $0 \leq t \leq 1$ où φ est une fonction arbitraire de l'espace dual de \mathbb{R}^n ($\varphi \in \mathcal{L}(\mathbb{R}^n)$).

Cette fonction est différentiable par rapport à t .

En effet,

$$\begin{aligned} \frac{g(t + \Delta t) - g(t)}{\Delta t} &= \frac{\varphi(f(x_0 + t\Delta x + \Delta t\Delta x)) - \varphi(f(x_0 + t\Delta x))}{\Delta t} \\ &= \varphi\left(\frac{f(x_0 + t\Delta x + \Delta t\Delta x) - f(x_0 + t\Delta x)}{\Delta t}\right) : (\varphi \in \mathcal{L}(\mathbb{R}^n) \text{ et } \Delta t \in \mathbb{R}). \end{aligned}$$

En passant à limite, on obtient

$$g'(t) = \varphi(f'(x_0 + t\Delta x)\Delta x)$$

et on a $g(1) - g(0) = g'(\theta)$ où $0 \leq \theta \leq 1$, (la formule des accroissements finis sur \mathbb{R}).

Alors

$$\varphi(f(x) - f(x_0)) = \varphi(f'(x_0 + \theta\Delta x)\Delta x) : \forall \varphi \in \mathcal{L}(\mathbb{R}^n) \text{ (car : } g(1) = \varphi(f(x)) \text{ et } g(0) = \varphi(f(x_0))\text{)}.$$

On déduit que

$$|\varphi(f(x) - f(x_0))| \leq \|\varphi\| \cdot \sup_{0 \leq \theta \leq 1} |f'(x_0 + \theta\Delta x)| \cdot \|\Delta x\|.$$

D'après la théorème de Hahn Banach on choisit $\varphi \in \mathcal{L}(\mathbb{R}^n)$ qui vérifie

$$|\varphi(f(x) - f(x_0))| = \|\varphi\| \cdot |f(x) - f(x_0)|, \quad [4].$$

Alors on obtient

$$\begin{aligned} \|\varphi\| \cdot |f(x) - f(x_0)| &\leq \|\varphi\| \cdot \sup_{0 \leq \theta \leq 1} |f'(x_0 + \theta\Delta x)| \cdot \|\Delta x\| \\ \implies |f(x) - f(x_0)| &\leq \sup_{0 \leq \theta \leq 1} |f'(x_0 + \theta\Delta x)| \cdot \|\Delta x\|. \end{aligned}$$

■

1.3 Problème d'extremum

Définition 1.3.1. (Point extremum). Soit f un fonction réelle définie sur \mathbb{R}^n .

On dit que la fonction f admet au point x_0 un minimum, si pour tout les x assez voisins de x_0 on a $f(x) - f(x_0) \geq 0$.

- On défini de manière analogue le maximum d'une fonction.
- Si la fonction f admet au point x_0 un minimum ou maximum, alors on dit que f admet en ce point un extrémum.

Théorème 1.3.1. Pour qu'une fonction différentiable f présente au point x_0 un extremum, il faut que sa différentielle en ce point soit nulle pour tous les h , $f'(x_0).h \equiv 0$.

Preuve. f est différentiable alors, $f(x_0 + h) - f(x_0) = f'(x_0).h + O(h)$.

On suppose que il existe un h tel que $f'(x_0).h \neq 0$, alors pour des valeurs réelles assez petites de λ , le signe de l'expression $f'(x_0).(\lambda h) + O(\lambda h)$ coïncide avec le signe de sa partie principale $f'(x_0).(\lambda h)$ mais $f'(x_0)$ est une fonction linéaire, on a donc

$$f'(x_0).(\lambda h) = \lambda f'(x_0).h.$$

- Si x_0 est un minimum alors $f(x_0 + \lambda h) - f(x_0) \geq 0$ ($\lambda \neq 0$ est assez petit), mais si λ vérifie $f'(x_0).(\lambda h) = \lambda f'(x_0).h < 0$. Alors on a $f(x_0 + \lambda h) - f(x_0) < 0$ contradiction.
- Si x_0 est un maximum alors $f(x_0 + \lambda h) - f(x_0) \leq 0$ ($\lambda \neq 0$ est assez petit), mais si λ vérifie $f'(x_0).(\lambda h) = \lambda f'(x_0).h > 0$, alors on a $f(x_0 + \lambda h) - f(x_0) > 0$, contradiction.

De sorte que la fonction f ne peut pas a d'extremum au point x_0 . ■

Remarque 1.3.1.

Dans le cas des fonctions de n variables, pour l'existence d'un extremum on a la condition nécessaire suivante: lorsque une fonction différentiable f admet au point $x_0 = (x_1^0, x_2^0, \dots, x_n^0)$ un extremum, alors en ce point on a $df = 0$ c'est à dire.

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \dots = \frac{\partial f}{\partial x_n} = 0.$$

1.4 Condition pour l'existence d'un extremum

1.4.1 Différentielles d'ordre supérieur

On définit la différentielle d'ordre deux de la fonction f par la formule $d^2 f = f''(x)(h, h)$ c-à-d comme l'expression quadratique qui correspond à l'application $f''(x) \in B(\mathbb{R}^2, \mathbb{R})$,

d'une façon générale on appelle différentiel d'ordre n de f l'expression, $d^n f = f^{(n)}(x)(h, h, \dots, h)$ c-à-d l'image dans \mathbb{R} d'un élément $(h, h, \dots, h) \in \mathbb{R}^n$ par l'application $f^{(n)}(x)$.

Voyons maintenant, dans quelle mesure ces résultats peuvent être étendus aux fonctions définies sur \mathbb{R}^n .

Théorème 1.4.1. *Soit f une fonction réelle, définie sur \mathbb{R}^n et admet au voisinage d'un point x_0 une dérivée seconde continue. Si cette fonction admet au point x_0 un minimum, on a $d^2 f(x_0) \geq 0$.*

Preuve.

D'après la formule de Taylor on a

$$f(x_0 + h) - f(x_0) = \frac{1}{2} f''(x_0)(h; h) + O(\|h\|^2)$$

car x_0 est un minimum de f , alors $f'(x_0) = 0$.

On suppose qu'il existe un élément h tel que

$$f''(x_0)(h; h) < 0 \tag{1.1}$$

alors, puisque $f''(x_0)(\varepsilon h, \varepsilon h) = \varepsilon^2 f''(x_0)(h, h)$, il existe des h vérifiant la condition (1.1) et ayant la norme aussi petite que l'on veut. Or, pour des h suffisamment petits en norme le signe de l'expression $\frac{1}{2} f''(x_0)(h, h) + O(\|h\|^2)$ est défini par le signe du terme $\frac{1}{2} f''(x_0)(h, h)$. Par conséquent,

$$f(x_0 + h) - f(x_0) = \frac{1}{2} f''(x_0)(h; h) + O(\|h\|^2) < 0 \implies f(x_0 + h) < f(x_0)$$

c-à-d la fonction f n'a pas de minimum en x_0 . ■

Définition 1.4.1. *Une fonction quadratique g est dite fortement positive, s'il existe une constante $c > 0$ telle que $g(x, x) \geq c\|x\|^2$ pour tous les x , où*

$$g : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

$$(x, y) \longrightarrow g(x, y).$$

Théorème 1.4.2. *Si une fonction f , définie sur \mathbb{R}^n , vérifie les conditions*

1) $df(x_0) = 0$.

2) $d^2 f(x_0)$ est une fonction quadratique fortement positive, elle admet au point x_0 un minimum.

Preuve.

Choisissons $\varepsilon > 0$ de façon que pour $\|h\|^2 < \varepsilon$ et vérifie la condition

$$\frac{-C}{4}\|h\|^2 \leq O(\|h\|^2) \leq \frac{C}{4}\|h\|^2.$$

Alors

$$f(x_0 + h) - f(x_0) = \frac{1}{2}f''(x_0)(h; h) + O(\|h\|^2) > \frac{1}{2}c\|h\|^2 - \frac{1}{4}c\|h\|^2 = \frac{C}{4}\|h\|^2.$$

$$\Rightarrow f(x_0 + h) - f(x_0) > \frac{c}{4}\|h\|^2 \text{ pour } \|h\|^2 < \varepsilon.$$

Mais $\frac{c}{4}\|h\|^2 > 0$, donc $f(x_0 + h) - f(x_0) > 0$. D'où

$$f(x_0) < f(x_0 + h) \quad \forall h : \|h\| < \varepsilon.$$

■

1.5 Maximum et minimum relatifs

Définition 1.5.1. On dit que la fonction f admet un minimum relatif (resp-un minimum relatif strict) en un point $a \in A$ (partie de \mathbb{R}^n), s'il existe un voisinage V de a dans \mathbb{R}^n tel que pour tout $x \in V \cap A$, $f(x) \geq f(a)$ (resp. $f(x) > f(a)$).

On dit que f admet un maximum relatif. (resp, un maximum relatif strict) au point a si la fonction $-f$ admet un minimum relatif (resp, un minimum relatif strict) au point a .

Théorème 1.5.1. Si la fonction f admet en un point a intérieur à A un minimum relatif, ou un maximum relatif, et si f est différentiable en a , sa différentielle en ce point nulle.

Preuve. Supposons $f'(a) \neq 0$, alors $\exists v \in \mathbb{R}^n$ tel que $f'(a)(v) \neq 0$ donc $f'(a)(v) > 0$

(si $f'(a)(v) < 0$ en remplaçant v par $-v$, $f'(a)(-v) = -f'(a)(v) > 0$).

Posons $w = \frac{v}{\|v\|}$ et $M = f'(a)(w)$, alors on remarque que $M > 0$

(car $M = f'(a)(w) = f'(a)\left(\frac{v}{\|v\|}\right) = \frac{1}{\|v\|}f'(a)(v) > 0$) et on a f différentiable, alors

$f(a + x) - f(a) - f'(a)(x) = o(\|x\|)$, donc $\exists \eta > 0$ tel que pour $x \in \mathbb{R}^n$, vérifiant $\|x\| \leq \eta$ avec $a + x$ soit intérieur à A , on a

$$|f(a + x) - f(a) - f'(a)(x)| \leq \frac{M}{2}\|x\|, \text{ donc}$$

$$f'(a)(x) - \frac{M}{2}\|x\| \leq f(a+x) - f(a) \leq f'(a)(x) + \frac{M}{2}\|x\|.$$

Soit V un voisinage de a dans \mathbb{R}^n , le point a étant intérieur à A , alors il existe $\theta > 0$ tel que pour $t \in \mathbb{R}$, et $|t| \leq \theta$, on obtient $a + tw$ soit intérieur à A et élément de $A \cap V$. Imposons de plus $\theta \leq \eta$, nous avons alors, dans les mêmes conditions, $\|tw\| = |t| \leq \eta$ et, d'après la double inégalité précédente (on pose $x = tw$). On a alors

$$f'(a)(x) = f'(a)(tw) = t.f'(a)(w) = tM.$$

Alors

$$\begin{aligned} tM - \frac{M}{2}|t|\|w\| &\leq f(a+tw) - f(a) \leq tM + \frac{M}{2}|t|\|w\| \\ \Rightarrow M(t - \frac{|t|}{2}) &\leq f(a+tw) - f(a) \leq M(t + \frac{|t|}{2}). \end{aligned}$$

Ce qui prouve que $f(a+tw) - f(a)$ est de même signe que t , le point $a+tw$ étant élément de $V \cap A$, ceci prouve que f n'admet en a ni minimum relatif, ni maximum relatif.

Le théorème ci-dessus donne, sous certaines hypothèses, une condition nécessaire pour que la fonction f admette un maximum relatif ou un minimum relatif en un point a . ■

1.5.1 Notion sur les formes quadratiques

Soit $\tilde{\varphi}$ une forme bilinéaire sur \mathbb{R}^n , nous supposons $\tilde{\varphi}$ symétrique, i.e.

$$\forall x, y \in \mathbb{R}^n : \tilde{\varphi}(x, y) = \tilde{\varphi}(y, x).$$

(a) On appelle forme quadratique associée à $\tilde{\varphi}$ l'application $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ définie par $\varphi(x) = \tilde{\varphi}(x, x)$. La connaissance de φ détermine $\tilde{\varphi}$ puisque,

$$\forall x, y : \tilde{\varphi}(x, y) = \frac{1}{2}(\varphi(x+y) - \varphi(x) - \varphi(y)).$$

Et on dit que la forme quadratique φ est positive si $\varphi(x) > 0$, $\forall x \in \mathbb{R}^n$ et on dit aussi que la forme bilinéaire symétrique $\tilde{\varphi}$ est positive si la forme quadratique associée φ est positive. (φ est négative si $-\varphi$ est positive).

(b) On dit que la forme quadratique φ est coercive s'il existe $\lambda > 0$ tel que

$$\forall x \in \mathbb{R}^n, \varphi(x) \geq \lambda\|x\|^2$$

et on dit la forme bilinéaire symétrique $\tilde{\varphi}$ est coercive si φ est coercive.

- Une forme quadratique coercive est évidemment positive.

Théorème 1.5.2. Soit f une fonction définie sur une partie A de \mathbb{R}^n . Si f admet en un point $a \in A$ un minimum relatif (resp. un maximum relatif), et si f est deux fois différentiable en a , la forme quadratique associée à sa différentielle seconde au point a est positive (resp. négative).

Preuve. Supposons que f admette en a un minimum relatif, et que $f''(a)$ ne soit pas positive. Alors $\exists v \in \mathbb{R}^n : f''(a)(v, v) < 0$. On pose $w = \frac{v}{\|v\|}$, et on a : $f''(a)(w, w) = -M < 0$. D'autre part, d'après la théorème (3.3.1) $f'(a) = 0$. Alors la formule de Taylor au point a à l'ordre 2 s'écrit

$$f(a+x) - f(a) - \frac{1}{2}f''(a)(x, x) = O(\|x\|^2).$$

Alors $\exists \eta > 0, \forall x \in \mathbb{R}^n$ vérifiant $\|x\| \leq \eta$ avec $a+x \in A$ tel que

$$f(a+x) - f(a) - \frac{1}{2}f''(a)(x, x) \leq \frac{M}{4}\|x\|^2 \Rightarrow f(a+x) - f(a) \leq \frac{1}{2}f''(a)(x, x) + \frac{M}{4}\|x\|^2.$$

Soit V un voisinage de a dans \mathbb{R}^n , comme $a \in A$, $\exists \theta > 0$ et $\forall t \in \mathbb{R}$ vérifiant $|t| \leq \theta$, $a+tw \in A$ et $a+tw \in A \cap V$. Imposons de plus $\theta \leq \eta$. Dans les mêmes conditions, nous avons $\|tw\| = |t| \leq \eta$ et alors on pose $x = tw$, alors

$$\begin{aligned} f(a+tw) - f(a) &\leq \frac{1}{2}t^2 f''(a)(w; w) + \frac{M}{4}\|tw\|^2 \\ \Rightarrow f(a+tw) - f(a) &\leq \left(-\frac{M}{2} + \frac{M}{4}\right)t^2 = -\frac{M}{4}t^2 < 0, \end{aligned}$$

donc $f(a+tw) - f(a) < 0$ et $a+tw \in V \cap A$ ceci contredit l'hypothèse selon laquelle f admet un minimum relatif en a . ■

Le théorème suivant donne une condition suffisante pour que f admette en a un minimum relatif strict.

Théorème 1.5.3. Soit f définie sur une partie A de \mathbb{R}^n . Si la fonction f est deux fois différentiable en un point $a \in A$ et si $f'(a) = 0$ et sa différentielle seconde en ce point est coercive, alors f admet en a un minimum relatif strict.

Preuve. On a $f''(a)$ est coercive, alors

$$\exists \lambda > 0, \forall x \in \mathbb{R}^n : f''(a)(x, x) \geq \lambda\|x\|^2.$$

La formule de Taylor à l'ordre 2 en a nous permet d'écrire

$$f(a+x) - f(a) - \frac{1}{2}f''(a)(x, x) = O(\|x\|^2).$$

Alors $\exists \eta > 0$, $\forall x \in \mathbb{R}^n$, $\|x\| \leq \eta$ et tel que $a+x \in A$,

$$\begin{aligned} |f(a+x) - f(a) - \frac{1}{2}f''(a)(x, x)| &\leq \frac{\lambda}{4}\|x\|^2 \implies f(a+x) - f(a) \geq \frac{1}{2}f''(a)(x, x) - \frac{\lambda}{4}\|x\|^2 \\ &\implies f(a+x) - f(a) \geq \left(\frac{\lambda}{2} - \frac{\lambda}{4}\right)\|x\|^2 \\ &\implies f(a+x) - f(a) \geq \frac{\lambda}{4}\|x\|^2 > 0. \end{aligned}$$

Donc f admet en a un minimum relatif strict. ■

La proposition suivante indique des conditions dans lesquelles on peut affirmer que $f''(a)$ est coercive.

Proposition 1.5.1. Soit $\bar{\varphi}$ une forme bilinéaire continue et symétrique sur \mathbb{R}^n , φ la forme quadratique associée. On note $\hat{\varphi}$ l'application linéaire continue de \mathbb{R}^n dans son dual topologique qui associe à chaque élément x de \mathbb{R}^n , la forme linéaire $\hat{\varphi}$ définie par

$$\hat{\varphi}(x)(y) = \bar{\varphi}(x, y), \quad y \in \mathbb{R}^n$$

Si φ est positive et si $\hat{\varphi}$ est un isomorphisme de \mathbb{R}^n sur son dual, alors φ est coercive.

Preuve. Soit $x \in \mathbb{R}^n, x \neq 0$. La norme M^{-1} de $\hat{\varphi}^{-1}$ est strictement positive, puisque $\hat{\varphi} : \mathbb{R}^n \rightarrow \mathcal{L}(\mathbb{R}^n)$ est un isomorphisme, c'est-à-dire est linéaire, bijective, continue et d'inverse continue. Nous avons donc

$$\|\hat{\varphi}(x)\| \geq M\|x\|.$$

D'après la définition de la norme d'une application linéaire, il existe $y \in \mathbb{R}^n$ vérifiant $\|y\| = 1$ tel que

$$\begin{aligned} |\hat{\varphi}(x)(y)| &= |\bar{\varphi}(x, y)| \geq \frac{M}{2}\|x\|, \\ \exists y : \|y\| = 1 \text{ et } \|\hat{\varphi}(x)\| &= |\hat{\varphi}(x)(y)| \geq \frac{M}{2}\|x\|. \end{aligned}$$

Compte tenu de l'inégalité de Schwartz, nous en déduisons

$$\bar{\varphi}(x, y) \leq \bar{\varphi}^{\frac{1}{2}}(x, x) \cdot \bar{\varphi}^{\frac{1}{2}}(y, y) \implies |\bar{\varphi}(x, y)|^2 \leq \varphi(x) \cdot \varphi(y)$$

et on a $|\bar{\varphi}(x, y)|^2 \geq \frac{M^2}{4} \|x\|^2$. Alors on obtient

$$\varphi(x) \cdot \varphi(y) \geq \frac{M^2}{4} \|x\|^2.$$

Mais si $\|\bar{\varphi}\| = k > 0$, puisque φ est positive et que $\|y\| = 1$,

$$\varphi(y) = \bar{\varphi}(y, y) \leq k \cdot \|y\|^2 = k,$$

mais on a

$$\varphi(x) \cdot k \geq \varphi(x) \cdot \varphi(y) \geq \frac{M^2}{4} \cdot \|x\|^2 \implies \varphi(x) \geq \frac{M^2}{4k} \cdot \|x\|^2.$$

Ce qui prouve que φ est coercive. ■

1.5.2 Maximum et minimum relatifs liés

Soit U un ouvert de \mathbb{R}^n , et $g : U \rightarrow F$ une application différentiable de classe C^p ($p \geq 1$) de U dans partie de \mathbb{R}^n ($F \subseteq \mathbb{R}^n$).

Soit A la partie de U : $A = g^{-1}(0) = \{x \in U ; g(x) = 0\}$.

Soit d'autre part $f : U \rightarrow R$. Nous nous intéressons aux points de A où la fonction $f|_A$ (restriction de f à A) présente un minimum relatif ou un maximum relatif. Ces points sont traditionnellement appelés minima ou maxima liés de la fonction f , car la condition $g(x) = 0$, exprimant que les points $x \in U$ auxquels on s'intéresse doivent appartenir à A , est souvent appelée une liaison.

Théorème 1.5.4. [16]

Soit a un point de A où la fonction $f|_A$ admet un minimum relatif ou un maximum relatif. On suppose que les trois conditions suivantes sont satisfaites :

- (1) la différentielle $g'(a)$ de g au point a est surjective,
- (2) il existe un sous-espace vectoriel fermé G de \mathbb{R}^n supplémentaire de $\ker g'(a)$,
- (3) la fonction f est différentiable au point a .

Alors on a l'inclusion $\ker g'(a) \subset \ker f'(a)$.

Preuve. Posons pour alléger l'écriture, $K = \ker g'(a)$. C'est un sous-espace vectoriel fermé de \mathbb{R}^n . Par hypothèse, il existe un autre sous-espace vectoriel fermé G de \mathbb{R}^n supplémentaire de K . L'application de $K \times G$ dans \mathbb{R}^n , $(x, y) \rightarrow x + y$, avec $x \in K$ et $y \in G$, est linéaire, continue et bijective. D'après le théorème de Banach, son inverse est continue. Nous pouvons donc identifier, au moyen de cette application, l'espace \mathbb{R}^n au produit $K \times G$, donc

considérer U comme un ouvert de $K \times G$, et $A = g^{-1}(0)$ comme une partie de $U (A \subset U)$, donc de $K \times G$. Le point a de A s'écrit donc $a = (a_K, a_G)$, avec $a_K \in K, a_G \in G$.

L'application $g : U \rightarrow F$ s'écrit comme $(x, y) \rightarrow g(x, y)$, $x \in K, y \in G$. Ses différentielles partielles $g'_x(a_K, a_G)$ et $g'_y(a_K, a_G)$ sont les restrictions de $g'(a)$, respectivement à K et à G , cela résulte en effet de l'identification que nous avons faite de \mathbb{R}^n avec $K \times G$. Comme K est le noyau de $g'(a)$, on a $g'_x(a_K, a_G) = 0$. L'autre différentielle partielle $g'_y(a_K, a_G)$ est linéaire continue et bijective de G sur F . D'après le théorème de Banach, son inverse est continue, et nous pouvons affirmer que $g'_y(a_K, a_G)$ est un isomorphisme de G sur F . Le théorème des fonctions implicites nous permet d'affirmer qu'il existe un voisinage ouvert V de (a_K, a_G) dans $K \times G$, $V \subset U$, un voisinage ouvert W de a_K dans K et une application différentiable h , de classe C^p , de W dans G , tels que les assertions suivantes :

$$a) (x, y) \in V \text{ et } g(x, y) = 0$$

$$b) x \in W \text{ et } y = h(x)$$

soient équivalentes.

Nous voyons ainsi que $V \cap A = V \cap g^{-1}(0)$ est le graphe de l'application h :

$$V \cap A = \{ (x, y) \in K \times G, x \in W, y = h(x) \}.$$

L'application $x \rightarrow (x, h(x))$ est un homéomorphisme de l'ouvert W de K sur l'ouvert $V \cap A$ de A . La fonction f admettant au point a de $V \cap A$ un minimum relatif ou un maximum relatif, sa composée $f \circ (id_W, h)$ avec l'homéomorphisme (id_W, h) admet au point a_K de W un maximum relatif ou un minimum relatif. Mais la fonction $f \circ (id_W, h)$ est différentiable au point a_K . Alors

$$\forall u \in K : D(f \circ (id_W, h))(a_K)(u) = 0 \tag{1.2}$$

Donc

$$D(f \circ (id_W, h))(a_K)(u) = f'(a)(u, h'(a_K)(u)) \tag{1.3}$$

Calculons $h'(a_K)$ on a

$$\forall x \in W, g(x, h(x)) = 0.$$

En différentiant cette expression au point a_K , nous obtenons, pour tout $u \in K$:

$$g'_x(a_K, a_G)(u) + g'_y(a_K, a_G)(h'(a_K)(u)) = 0.$$

Puisque $g'_y(a_K, a_G)$ est un isomorphisme et que $g'_x(a_K, a_G) = 0$, cette égalité prouve que

$$h'(a_K) = 0 \tag{1.4}$$

La condition (1.2), compte tenu de (1.3) et de (1.4), équivaut donc à

$$\forall u \in \text{Ker } g'(a), f'(a)(u) = 0.$$

Alors on obtient

$$\text{Ker } g'(a) \subset \text{Ker } f'(a).$$

■

Chapitre 2

Rappel d'analyse convexe

2.1 Les ensembles convexes

dans ce chapitre, nous allons donner les définitions de les ensembles convexes, et leurs propriétés fondamentale.

Notions fondamentales

Définition 2.1.1. *Un ensemble $X \subseteq \mathbb{R}^n$ est convexe si, chaque fois qu'il contient deux points $x_1; x_2$ quelconques, il contient également tout les points de la forme*

$$x = \alpha x_1 + (1 - \alpha) x_2 : 0 \leq \alpha \leq 1.$$

La droite passant par les point x_1, x_2 , a pour équation paramétrique :

$$x(\alpha) = x_2 + (x_1 - x_2)\alpha \text{ car } \alpha x_1 + (1 - \alpha)x_2 = x_2 + (x_1 - x_2)\alpha, \alpha \in \mathbb{R}.$$

Exemple 2.1.1. *Le triangle X définie par : $X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 + x_2 \leq 1 : x_1 \geq 0 \text{ et } x_2 \geq 0\}$ est convexe .*

Soient $(x_1, x_2), (x'_1, x'_2)$ deux points du triangle X , et α un nombre arbitraire tel que $0 \leq \alpha \leq 1$. Vérifions que le point de coordonnées $(\alpha x_1 + (1 - \alpha)x'_1, \alpha x_2 + (1 - \alpha)x'_2)$ appartient à X . En d'autres termes , il faut tester les inégalités :

$$\alpha x_1 + (1 - \alpha)x'_1 + \alpha x_2 + (1 - \alpha)x'_2 \leq 1 \text{ et } \alpha x_1 + (1 - \alpha)x'_1 \geq 0 \text{ et } \alpha x_2 + (1 - \alpha)x'_2 \geq 0 .$$

Comme $x_1 \geq 0$ et $x'_1 \geq 0$ et on a $0 \leq \alpha \leq 1$, alors $\alpha x_1 + (1 - \alpha)x'_1 \geq 0$ et on a $x_1 + x_2 \leq 1$ et $x'_1 + x'_2 \leq 1$, on les multiplie par les nombres non négatifs α et $1 - \alpha$ respectivement on fait la somme, il vient

$$\alpha x_1 + (1 - \alpha)x'_1 + \alpha x_2 + (1 - \alpha)x'_2 = \alpha(x_1 + x_2) + (1 - \alpha)(x'_1 + x'_2) \leq \alpha \cdot 1 + (1 - \alpha) \cdot 1 = \alpha + 1 - \alpha = 1$$

Alors $\alpha x_1 + (1 - \alpha)x'_1 + \alpha x_2 + (1 - \alpha)x'_2 \leq 1$.

Donc X est convexe. (Fig 1)

Exemple 2.1.2. Le disque X définie par : $X = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq r^2\}$ est convexe.

Soit $(x_1, x_2) \in X, (x'_1, x'_2) \in X$ et $0 \leq \alpha \leq 1$ vérifions que

$(\alpha x_1 + (1 - \alpha)x'_1, \alpha x_2 + (1 - \alpha)x'_2)$ est un élément de X .

On effectue les transformations simples suivantes :

$$(\alpha x_1 + (1 - \alpha)x'_1)^2 + (\alpha x_2 + (1 - \alpha)x'_2)^2 = \alpha^2(x_1^2 + x_2^2) + (1 - \alpha)^2((x'_1)^2 + (x'_2)^2) + 2\alpha(1 - \alpha)(x_1 x'_1 + x_2 x'_2)$$

et on utilise l'inégalité fondamentale du produit scalaire

$$x_1 x'_1 + x_2 x'_2 \leq \sqrt{x_1^2 + x_2^2} \cdot \sqrt{(x'_1)^2 + (x'_2)^2}. \text{ On a } x_1^2 + x_2^2 \leq r^2 \text{ et } (x'_1)^2 + (x'_2)^2 \leq r^2.$$

$$\text{Alors } (\alpha x_1 + (1 - \alpha)x'_1)^2 + (\alpha x_2 + (1 - \alpha)x'_2)^2 \leq r^2(\alpha^2 + (1 - \alpha)^2 + 2\alpha(1 - \alpha)) = r^2.$$

Donc, le disque est une ensemble convexe. (Fig 2)

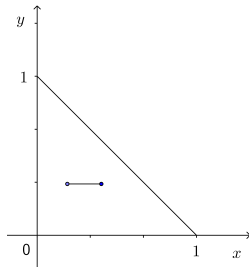


Fig 1

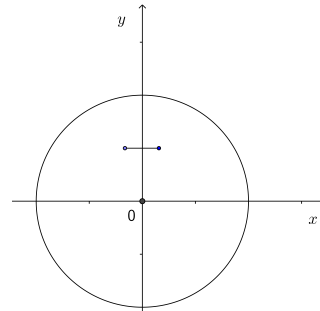


Fig 2

Théorème 2.1.1. L'intersection finie d'ensembles convexes est convexe.

Preuve.

Soit $X = \bigcap_{i=1}^n X_i$ avec X_i est ensemble convexe pour toute $i = 1, 2, \dots, n$ et soient x, x' deux points quelconques de X . Comme $X \subseteq X_i, \forall i = 1; 2, \dots, n$, alors $x, x' \in X_i$, mais X_i étant convexe il en résulte que le segment $[x, x']$ tout entier est dans $X_i, \forall i = 1, 2, \dots, n$.

Donc $[x, x'] \subset \bigcap_{i=1}^n X_i = X$, alors X est un ensemble convexe. ■

Définition 2.1.2. On appelle somme de deux convexes X_1 et X_2 l'ensemble défini par

$$X = \{x = x_1 + x_2 / x_1 \in X_1, x_2 \in X_2\}$$

Théorème 2.1.2. La somme d'ensembles convexes est convexe.

Preuve.

Soient $x, x' \in X$ où $X = X_1 + X_2$ avec X_1 et X_2 deux ensembles convexes .

Il existe dans X_1 et X_2 des éléments tels que $x = x_1 + x_2$, $x' = x'_1 + x'_2$ avec $x_1, x'_1 \in X_1$ et $x_2, x'_2 \in X_2$.

Comme X_1 et X_2 sont convexes, alors pour tout $0 \leq \alpha \leq 1$ on a $\alpha x_1 + (1 - \alpha)x'_1 \in X_1$ et $\alpha x_2 + (1 - \alpha)x'_2 \in X_2$.

De la définition de X , il résulte que la somme de ces éléments est dans X :

$$\alpha x_1 + (1 - \alpha)x'_1 + \alpha x_2 + (1 - \alpha)x'_2 = \alpha(x_1 + x_2) + (1 - \alpha)(x'_1 + x'_2) = \alpha x + (1 - \alpha)x' \in X .$$

Donc X est convexe . ■

Exemple dans l'espace \mathbb{R}^n , $\Omega = \{x \in \mathbb{R}^n : Ax = b, A \text{ une matrice } n \times m, b \in \mathbb{R}^m\}$.

2.1.1 Caractéristiques des ensembles convexes

Théorème 2.1.3. (Théorème de projection) . [8]

Soit un espace de Hilbert H et soit $C \subset H$ une partie convexe fermée non vide alors pour tout $x_0 \in H \setminus C$ il existe un unique $y_0 \in C$ tel que $\|x_0 - y_0\| = \inf_{y \in C} \|x_0 - y\|$.

Preuve. Posons $d = \inf_{y \in C} \|x_0 - y\|$. Alors pour tout $z \in C$ on a : $\|x_0 - z\| \geq d$ (qui existe, puisque c'est l'inf d'une partie non vide minorée par 0). Par définition, il existe $(y_n)_{n \in \mathbb{N}} \subseteq C$ tel que $d \leq \|x_0 - y_n\| \leq d + \frac{1}{n}$, $\forall n \in \mathbb{N}$.

Montrons que $(y_n)_{n \in \mathbb{N}}$ est une suite de Cauchy. Ainsi, H étant Hilbertien c'est un espace complet, donc on pourra conclure quant a la convergence de la suite (y_n) .

Soit $n, m \in \mathbb{N}$: $\|y_n - y_m\|^2 = \|y_n - x_0 + x_0 - y_m\|^2 = \|y_n - x_0 - (y_m - x_0)\|^2$.

D'après la loi du parallélogramme on a

$$\|(y_n - x_0) - (y_m - x_0)\|^2 + \|(y_n - x_0) + (y_m - x_0)\|^2 = 2\|y_n - x_0\|^2 + 2\|y_m - x_0\|^2 .$$

Alors $\|(y_n - x_0) - (y_m - x_0)\|^2 = 2\|y_n - x_0\|^2 + 2\|y_m - x_0\|^2 - 4\|\frac{y_n + y_m}{2} - x_0\|^2$.

Mais $\frac{y_n + y_m}{2} \in C$ car C est convexe, donc $4\|\frac{y_n + y_m}{2} - x_0\|^2 \geq 4d^2$ d'où

$$\|y_n - y_m\|^2 \leq 2(d + \frac{1}{n})^2 + 2(d + \frac{1}{m})^2 - 4d^2 = \frac{4d}{n} + \frac{4d}{m} + \frac{2}{n^2} + \frac{2}{m^2} \longrightarrow_{n, m \rightarrow \infty} 0 .$$

Donc (y_n) est de Cauchy, alors $y_n \rightarrow y_0 : y_0 \in C$ car C est fermé et on a

$$d \leq \|x_0 - y_0\| = \|x_0 - y_n + y_n - y_0\| \leq \|x_0 - y_n\| + \|y_n - y_0\| \leq d + \frac{1}{n} + \epsilon$$

car $y_n \rightarrow y_0 \implies d \leq \|x_0 - y_0\| \leq d + \frac{1}{n} + \epsilon, \forall n \geq n_0, \forall \epsilon \geq 0$. D'où $\|x_0 - y_0\| = d$

• **Unicité de y_0 .** Supposons qu'il existe $z_0 \neq y_0$ et $z_0 \in C$, $d = \|x_0 - z_0\|$.

$$\begin{aligned} \|z_0 - y_0\|^2 &= \|(z_0 - x_0) - (y_0 - x_0)\|^2 = 2\|z_0 - x_0\|^2 + 2\|y_0 - x_0\|^2 - 4\left\|\frac{z_0 + y_0}{2} - x_0\right\|^2 \\ &= 2d^2 + 2d^2 - 4\left\|\frac{z_0 + y_0}{2} - x_0\right\|^2 \leq 4d^2 - 4d^2 = 0. \end{aligned}$$

Alors $\|z_0 - y_0\|^2 = 0 \implies z_0 = y_0$. Contradiction, alors on a unicité. ■

Définition 2.1.3. Un ensemble de points $x = (x_1, x_2, \dots, x_n)$ de coordonnées telles que $p_1x_1 + p_2x_2 + \dots + p_nx_n = c$, s'appelle hyperplan dans \mathbb{R}^n .

Où le vecteur $p = (p_1, p_2, \dots, p_n) \in \mathbb{R}^n$ est la normale à cet hyperplan (avec $p \neq 0$) et $c \in \mathbb{R}$.

Définition 2.1.4. On dit que l'hyperplan $\langle p, x \rangle = c$ est un hyperplan d'appui dans l'ensemble X en un point $b \in X$ si l'on a pour tout élément $x \in X$ les relations $\langle p, x \rangle \leq c$ et $\langle p, b \rangle = c$.

Théorème 2.1.4. La fermeture d'un ensemble convexe est convexe.

Preuve. Soient $x, x' \in \overline{X}$. Alors il existe une suite $(x_k)_{k \in \mathbb{N}} \subset X$ tel que $x = \lim_{k \rightarrow \infty} x_k$ et il existe une suite $(x'_k)_{k \in \mathbb{N}} \subset X$ tel que $x' = \lim_{k \rightarrow \infty} x'_k$.

Mais X est convexe, alors $\forall \alpha \in [0, 1]$ on a $\alpha x_k + (1 - \alpha)x'_k \in X$, et on a

$$\lim_{k \rightarrow \infty} \alpha x_k + (1 - \alpha)x'_k = \alpha x + (1 - \alpha)x' \in \overline{X}.$$

Donc \overline{X} est convexe. ■

Définition 2.1.5. Un ensemble $K \subseteq \mathbb{R}^n$ s'appelle cône convexe si, toutes les fois qu'il contient des points x', x'' quelconques, il contient tout les points de la forme

$$x = \alpha x' + \beta x'', \alpha \geq 0 \text{ et } \beta \geq 0.$$

• Il en résulte de suite que, premièrement, le cône convexe K est un ensemble convexe et deuxièmement, si K contient le vecteur x , il contient également la demi-droite αx , $\alpha \geq 0$.

2.1.2 Ensembles convexes polyédriques

Combinaison linéaire convexe

Définition 2.1.6. [2].

Soient x_1, x_2, \dots, x_k des points quelconques de \mathbb{R}^n .

On appelle combinaison linéaire convexe de ces points la somme $a_1x_1 + a_2x_2 + \dots + a_kx_k$ pour $a_i, i = 1, 2, \dots, k$ des nombres non négatifs arbitraires dont la somme est 1 ou $a_i \geq 0$ et

$$\sum_{i=1}^k a_i = 1.$$

Remarque 2.1.1. Un ensemble X est convexe s'il contient la combinaison linéaire convexe de deux points quelconques.

Théorème 2.1.5. [11].

Soient X un ensemble convexe et x_1, x_2, \dots, x_k , k points quelconques de X , alors X contient toute combinaison linéaire convexe de ces points, / i.e /

si $(x_i)_{i=1}^k \subset X$, $\sum_{i=1}^k \lambda_i = 1$, $\lambda_i \geq 0$ $i = 1, 2, \dots, k$ et $\lambda_i \in \mathbb{R}$.

Alors $\sum_{i=1}^k \lambda_i x_i \in X$.

Preuve. On démontre par récurrence sur le nombre k de point.

- Si $k = 2$ comme X est convexe donc $\forall x_1, x_2 \in X$ et $\lambda_1 + \lambda_2 = 1$ avec $\lambda_1 \geq 0$ et $\lambda_2 \geq 0$ on a $\lambda_1 x_1 + \lambda_2 x_2 \in X$ ($\lambda_2 = 1 - \lambda_1$).

- On suppose que

$$\forall \lambda_i \in]0,1[\text{ et } \sum_{i=1}^{k-1} \lambda_i = 1 \text{ on a } \sum_{i=1}^{k-1} \lambda_i x_i \in X \text{ et } (x_i)_{i=1}^k \subset X.$$

Soient x_1, x_2, \dots, x_k . k points quelconques de X , et soient $(\lambda_i)_{i=1}^k \subset]0,1[$ avec

$$\sum_{i=1}^k \lambda_i = 1, \text{ alors on a } \sum_{i=1}^{k-1} \lambda_i = 1 - \lambda_k \implies \sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} = 1.$$

Donc

$$\sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i \in X. \text{ En posant } \left(\sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i = y \right), \text{ on a}$$

$$\sum_{i=1}^k \lambda_i x_i = \sum_{i=1}^{k-1} \lambda_i x_i + \lambda_k x_k = (1 - \lambda_k) \sum_{i=1}^{k-1} \frac{\lambda_i}{1 - \lambda_k} x_i + \lambda_k x_k = (1 - \lambda_k) y + \lambda_k x_k \in X$$

(car X est convexe avec $y, x_k \in X$) ■

Définition 2.1.7. Soit M un ensemble de points (fini ou infini) de \mathbb{R}^n .

On appelle enveloppe convexe de M , et on note $C(M)$, l'ensemble de toutes combinaisons linéaires convexes des points de M .

Théorème 2.1.6. L'enveloppe convexe $C(M)$ de l'ensemble M coïncide avec le plus petit ensemble convexe contenant M .

Preuve.

D'une part, $C(M)$ est un ensemble convexe et d'autre part. Soit X étant un convexe contenant M contient toutes les combinaisons linéaires convexes des points de M , i.e.

X convexe et $M \subseteq X$. Alors

$$\forall (x_i)_{i=1}^k \subset M \text{ et } \sum_{i=1}^k \lambda_i = 1 : \lambda_i \geq 0 \implies (x_i)_{i=1}^k \subset X \text{ et } \sum_{i=1}^k \lambda_i = 1 \text{ Alors s}$$

$$\sum_{i=1}^k \lambda_i x_i \in X \implies C(M) \subseteq X.$$

Comme X quelconque, on obtient $C(M)$ est le plus petit ensemble convexe contenant M .

■

Définition 2.1.8 (Point extrême). On dit qu'un point x d'un convexe $C \subseteq \mathbb{R}^n$ est extrême s'il n'existe pas $x_1, x_2 \in C$ avec

$$x_1 \neq x_2 \text{ et } x = \frac{1}{2}x_1 + \frac{1}{2}x_2.$$

On note que tout ensemble convexe ne possède pas de points extrêmes.

Ainsi le demi-plan supérieur fermé est convexe, a des points frontières (qui sont tous les points de la droite définissant le demi-plan), mais il ne possède pas de points extrêmes.

Définition 2.1.9 (Point extrême). Un ensemble $C \subseteq \mathbb{R}^n$ est convexe, non vide, un point $x \in C$ est appelé point extrême de C , si $x = \lambda x_1 + (1 - \lambda)x_2$ avec $x_1, x_2 \in C$ et $\lambda \in]0,1[$

$$\implies x = x_1 = x_2.$$

Définition 2.1.10 (Direction de C). Soit $C \subseteq \mathbb{R}^n$ ($C \neq \emptyset$) convexe et fermé.

Un vecteur non-nul $d \in \mathbb{R}^n$ est appelé une direction du convexe C si pour tout $x \in C$ on a

$$x + \theta d \in C, \theta \in \mathbb{R}.$$

- Deux direction d_1 et d_2 sont distinctes si $d_1 \neq \alpha d_2$ pour $\alpha > 0$.
- Une direction de C est dite direction extrême si elle ne peut être exprimée comme une combinaison linéaire positive de deux directions distinctes ,i.e. Si $d = \lambda_1 d_1 + \lambda_2 d_2$ pour $\lambda_1, \lambda_2 \geq 0$, alors $d_1 = \alpha d_2$, pour $\alpha > 0$.

Exemple 2.1.3. Soit $E = \{(x; y) \in \mathbb{R}^2 : y \geq |x|\}$ on a $d_1 = (1, 1)$ et $d_2 = (-1, 1)$ sont deux directions extrêmes de E . (Fig 3)

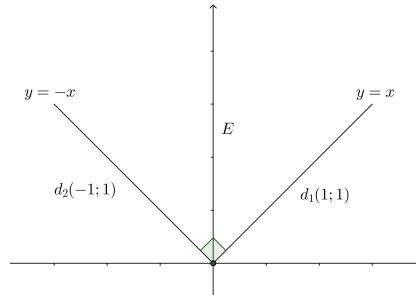


Fig 3

Proposition 2.1.1. [13]

Soit $C \subseteq \mathbb{R}^n$ un convexe non vide alors .

- (a) Si $x \in \text{Ir}(C)$ (les points intérieure à C), alors x n'est jamais extrême.
 (b) Soit $C \subseteq \mathbb{R}^n$ un convexe compact et soit $x \notin C$, soit $z \in C$ tel que

$$\|x - z\| = \sup_{y \in C} \|x - y\| \quad (\text{norme euclidienne}).$$

Si z existe ,alors z est un point extrême de l'ensemble C .

Preuve. (a) . Si $x \in \text{Ir}(C)$ et $u = z - x$ avec $z \in C \setminus \{x\}$, alors il existe $\alpha \geq 0$ tel que $x + [-\alpha, \alpha]u \subset C$ (car $x \in \text{Ir}(C) \implies \exists O$ ouvert tel que $x \in O \subset C$. Alors $\exists \alpha > 0 : x + [-\alpha, \alpha]u \subset O \subset C$). (Fig 4)

Donc $x = \frac{x - \alpha u}{2} + \frac{x + \alpha u}{2}$ avec $x - \alpha u \in C$ et $x + \alpha u \in C$.

Alors x n'est pas extrême. ■

Exemple 2.1.4. Si B est la boule euclidienne unité, alors tous les points de la sphère unité $S(B)$ sont extrémaux.

En effet, si $\|x\| = 1$ et $x = \frac{x_1}{2} + \frac{x_2}{2}$ avec $x_1 \leq 1$ et $x_2 \leq 1$ alors : $\langle x_1, x_2 \rangle = 1$ car

$$\|x\|^2 = \left\| \frac{x_1}{2} + \frac{x_2}{2} \right\|^2 = \frac{1}{4} (\|x_1\|^2 + \|x_2\|^2 + 2 \langle x_1, x_2 \rangle)$$

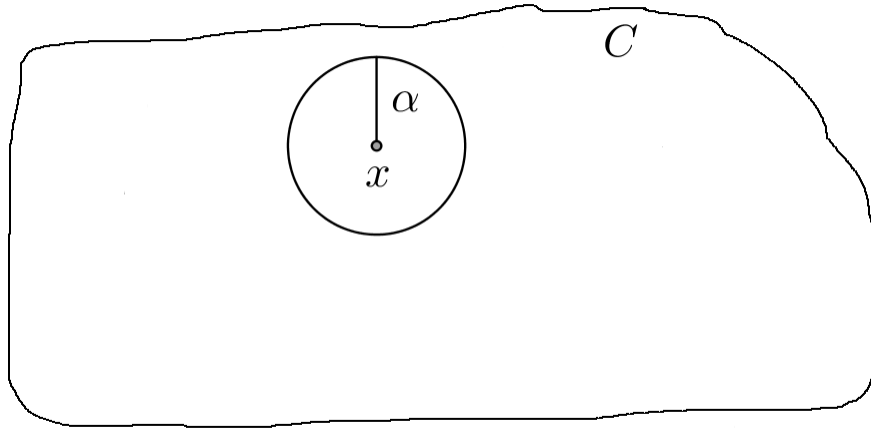


Fig 4

$$= \frac{1}{4} \|x_1\|^2 + \frac{1}{4} \|x_2\|^2 + \frac{1}{2} \langle x_1, x_2 \rangle \leq \frac{1}{4} + \frac{1}{4} + \frac{1}{2} \langle x_1, x_2 \rangle .$$

Alors

$$1 \leq \frac{1}{4} + \frac{1}{4} + \frac{1}{2} \langle x_1, x_2 \rangle = \frac{1}{2} + \frac{1}{2} \langle x_1, x_2 \rangle \implies \frac{1}{2} \langle x_1, x_2 \rangle \geq \frac{1}{2} \\ \implies \langle x_1, x_2 \rangle \geq 1$$

et de l'inégalité de Cauchy - Schwarz on a

$$\langle x_1, x_2 \rangle \leq \|x_1\| \|x_2\| \leq 1 \cdot 1 = 1 \implies \langle x_1, x_2 \rangle \leq 1, \text{ alors}$$

$$\langle x_1, x_2 \rangle = 1 = \|x_1\| \cdot \|x_2\|$$

Ce qui implique l'existence de $\lambda > 0$ tel que $x_2 = \lambda x_1$, d'où

$$1 \leq \frac{\|x_1\|}{2} + \frac{\|x_2\|}{2} = \frac{\|x_1\|}{2} + \frac{\lambda \|x_1\|}{2} = \frac{1+\lambda}{2} \|x_1\| \leq \frac{1+\lambda}{2}$$

$$\implies \frac{1+\lambda}{2} \geq 1, \text{ donc}$$

$$\frac{\lambda}{2} \geq \frac{1}{2} \implies \lambda \geq 1.$$

Comme $x_2 = \mu \cdot x_1$ avec $\mu = \frac{1}{\lambda}$, on obtient donc aussi $\lambda \leq 1$ d'où $\lambda = 1$ et $x_1 = x_2$ alors $x = x_1 = x_2$.

Donc x est un point extrême.

(b) La fonction $y \rightarrow \|x - y\|$ étant continue sur le compact C , elle atteint son maximum sur ce compact. Soit donc $z \in C$ réalisant ce maximum (on remarque que $z \neq x$), et soient $x_1, x_2 \in C$ tels que $z = \frac{x_1 + x_2}{2}$. On a alors

$$\|x - z\| = \left\| \frac{x - x_1}{2} + \frac{x - x_2}{2} \right\| \leq \frac{\|x - x_1\|}{2} + \frac{\|x - x_2\|}{2} \leq \frac{\|x - z\|}{2} + \frac{\|x - z\|}{2} = \|x - z\|,$$

car $x_1, x_2 \in C$ et $\|x - z\| = \sup_{y \in C} \|x - y\|$. On est donc dans un cas où la norme de la somme de deux vecteurs est égal à la somme des normes, i.e.

$$\begin{aligned} \|(x - x_1) + (x - x_2)\| &= \|x - x_1\| + \|x - x_2\| \\ \|(x - x_1) + (x - x_2)\|^2 &= (\|x - x_1\| + \|x - x_2\|)^2 \\ \implies \|x - x_1\|^2 + \|x - x_2\|^2 + 2 \langle x - x_1, x - x_2 \rangle &= \|x - x_1\|^2 + \|x - x_2\|^2 + 2\|x - x_1\| \cdot \|x - x_2\| \\ \implies \langle x - x_1, x - x_2 \rangle &= \|x - x_1\| \|x - x_2\|. \end{aligned}$$

Ce qui implique donc l'existence de $\lambda > 0$ tel que

$$x - x_1 = \lambda(x - x_2).$$

Donc

$$\begin{aligned} \|x - z\| &= \frac{\|x - x_1\|}{2} + \frac{\lambda \|x - x_1\|}{2} = \frac{1 + \lambda}{2} \|x - x_1\| \leq \frac{1 + \lambda}{2} \|x - z\| \quad (\|x - z\| \neq 0) \\ \implies \frac{1 + \lambda}{2} &\geq 1 \implies \lambda \geq 1. \end{aligned}$$

De même, comme $x - x_2 = \lambda^{-1}(x - x_1)$ le même raisonnement montre que $\lambda^{-1} \geq 1$, donc $\lambda = 1$. Alors $x - x_1 = x - x_2 \implies x_1 = x_2 \implies x_1 = x_2 = z$. D'où z est un point extrême du convexe compact C .

Théorème 2.1.7. [1] et [6]

Tout ensemble compact convexe de \mathbb{R}^n est l'enveloppe convexe de ses points extrêmes.

Définition 2.1.11. *Un ensemble X s'appelle polyèdre convexe s'il est l'enveloppe convexe des points en nombres finis.*

Autrement dit un ensemble X est un polyèdre convexe s'il contient un nombre fini de points x_1, x_2, \dots, x_k tels que tout élément x de X s'écrive

$$x = \sum_{i=1}^k \alpha_i x_i : 0 \leq \alpha_i \leq 1, \sum_{i=1}^k \alpha_i = 1$$

et que tous les points de cette forme soient dans X .

Théorème 2.1.8. [2]

Un polyèdre convexe est compact.

Preuve. Un polyèdre convexe X est l'enveloppe convexe de ses points x_1, x_2, \dots, x_k , i.e.

$$X = C(x_1, x_2, \dots, x_k).$$

On utilise le théorème de Weierstrass si de toute suite infinie $(x^n)_{n \in \mathbb{N}}$ d'éléments de X , on peut extraire une sous-suite infinie $(x^{n_l})_{l \in \mathbb{N}}$ qui converge vers un élément de X .

• Soient $(x^n)_{n \in \mathbb{N}} \subset X$, si bien que

$$x^n = \sum_{i=1}^k \alpha_i^n x_i \text{ avec } 0 \leq \alpha_i^n \leq 1 \text{ et } \sum_{i=1}^k \alpha_i^n = 1.$$

Chaque suite $(\alpha_i^n)_{n \in \mathbb{N}}$ est bornée de \mathbb{R} on en choisit donc une sous-suite convergente.

Alors pour tout $i = 1, 2, \dots, k$ il existe une sous-suite $(\alpha_i^{n_l})_{l \in \mathbb{N}}$ convergent.

Soit

$$\lim_{l \rightarrow +\infty} \alpha_i^{n_l} = \alpha_i : \alpha_i \geq 0 \text{ (car } 0 \leq \alpha_i^{n_l} \leq 1) \text{ et}$$

$$\lim_{l \rightarrow +\infty} \sum_{i=1}^k \alpha_i^{n_l} = \sum_{i=1}^k \alpha_i = 1$$

et on a

$$\lim_{l \rightarrow +\infty} \sum_{i=1}^k \alpha_i^{n_l} x_i = \sum_{i=1}^k \alpha_i x_i = x' \in X \text{ (car } X = C(x_1, x_2, \dots, x_k)).$$

Alors

$$\exists (x^{n_l})_{l \in \mathbb{N}} \text{ tel que } \lim_{l \rightarrow \infty} x^{n_l} = x' \in X.$$

Donc X est compact. ■

2.1.3 Enveloppes conique .

Définition 2.1.12. Soit $M \subseteq \mathbb{R}^n$ un ensemble quelconque. On appelle enveloppe conique, et on note $CO(M)$, l'ensemble de toutes les combinaisons linéaires non négatives

$$\sum_{i=1}^k \alpha_i x_i : \alpha_i \geq 0 \text{ de points de } M.$$

• On montre que $CO(M)$ est le plus petit cône convexe contenant M

Définition 2.1.13. [6]. *Un cône convexe K est dit polyédrique s'il est l'enveloppe conique d'un nombre fini de ses points.*

• *Le cône polyédrique est fermé, et la démonstration de cette propriété provient du lemme suivant :*

Lemme 2.1.1. *On suppose qu'un vecteur x non nul est combinaison linéaire non négative des vecteurs*

$$x_1, x_2, \dots, x_k: x = \sum_{i=1}^k \alpha_i x_i \text{ avec } \alpha_i \geq 0 \quad i = 1; 2, \dots, k.$$

Le vecteur x s'écrit comme combinaison linéaire non négative d'un certain nombre de vecteurs linéairement indépendants de l'ensemble $\{x_1, x_2, \dots, x_k\}$.

Preuve. On raisonne par récurrence sur k

- L'affirmation est triviale pour $k = 1$.
- Supposons qu'elle soit vraie pour les $k - 1$ vecteurs.

S'il existe $i_0 \in \{1, 2, \dots, k\}$ telle que $\alpha_{i_0} = 0$, on applique l'hypothèse de récurrence.

Si pour tout $i \in \{1, 2, \dots, k\}$ on a $\alpha_i > 0$. On suppose que x_1, x_2, \dots, x_k sont linéairement indépendants, i.e.

$\exists (\lambda_i)_{i=1}^k$ sont non tous nuls tel que

$$\sum_{i=1}^k \lambda_i x_i = 0.$$

Alors il existe au moins un certain λ_i positif (dans le cas contraire, on multiplie tous par -1).

Soit

$$\theta = \max_{1 \leq i \leq k} \left(\frac{\lambda_i}{\alpha_i} \right).$$

Il est clair que $\theta > 0$. Alors il existe i_0 tel que $\theta = \frac{\lambda_{i_0}}{\alpha_{i_0}}$ pour fixer les idées, auquel cas,

$$x = x - \frac{1}{\theta} \sum_{i=1}^k \lambda_i x_i = \sum_{i=1}^k \alpha_i x_i - \frac{1}{\theta} \sum_{i=1}^k \lambda_i x_i = \sum_{i=1}^k \left(\alpha_i - \frac{1}{\theta} \lambda_i \right) x_i.$$

et on a $\theta \geq \frac{\lambda_i}{\alpha_i} \implies \alpha_i \geq \frac{\lambda_i}{\theta} \implies \alpha_i - \frac{\lambda_i}{\theta} \geq 0 \quad i = 1, 2, \dots, k$. Mais on a $\alpha_{i_0} - \frac{\lambda_{i_0}}{\theta} = 0$. Par conséquent, le vecteur x est combinaison linéaire non négative des $k - 1$ vecteurs. ■

Théorème 2.1.9. [6]. *Un cône polyédrique convexe est fermé.*

Preuve. Soit $K = CO(x_1, x_2, \dots, x_k)$.

Plaçons-nous dans le cas où le système de vecteurs $x_i, i = 1, 2, \dots, k$, est linéairement indépendant.

Soit L un sous-espace vectoriel k - dimensionnell de \mathbb{R}^n engendré par ces vecteurs. Alors $\{x_1, x_2, \dots, x_k\}$ est une base de L .

On munit L d'un produit scalaire de façon que la base donnée soit orthonormée, ce qui permet d'identifier L et \mathbb{R}^k , d'une part, K et \mathbb{R}_+^k , de l'autre.

Il est évident que \mathbb{R}_+^k est un ensemble fermé, ce qui *entraîne* la propriété de K d'être fermé pour la norme associée au produit scalaire choisi.

Mais les normes sont équivalentes. Soient le cas où les vecteurs, x_1, x_2, \dots, x_k , sont linéairement indépendants.

On considère un sous-système libre quelconque $\{x_{i_1}, x_{i_2}, \dots, x_{i_s}\}$ de cet ensemble et le cône $K_\sigma = CO(x_{i_1}, x_{i_2}, \dots, x_{i_s})$, avec $\sigma = \{i_1, i_2, \dots, i_s\}$.

Comme nous l'avons vu , K_σ est fermé. On démontre que $K = \bigcup_\sigma K_\sigma$, la réunion étant prise par rapport a tous les jeux σ correspondant aux sous-systèmes libres de vecteurs x_1, x_2, \dots, x_k .

En effet, on a évidemment

$$K \subseteq \bigcup_\sigma K_\sigma \subseteq K \implies K = \bigcup_\sigma K_\sigma$$

et l'inclusion inverse découle du lemme.

La propriété, pour la réunion d'ensembles fermés en nombre fini, d'être fermée démontre le théorème. ■

2.1.4 Structure des ensembles admissibles

On se propose d'étudier d'une manière systématique la structure des ensembles admissibles des problèmes de programmation linéaire. [2].

Un ensemble admissible est décrit par le système d'inéquations linéaires de la forme :

$$(P) : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b_2 \\ \cdot \quad \cdot \quad \quad \quad \cdot \\ \cdot \quad \cdot \quad \quad \quad \cdot \\ \cdot \quad \cdot \quad \quad \quad \cdot \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq b_m. \end{cases}$$

On n'écrit pas les contraintes $x_i \geq 0$ car ces contraintes, si elles existent, sont incluses dans le système (P) sous forme $-x_i \leq 0$. Le symbole a_i désignera dans la suite le vecteur $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$ (la ligne i de la matrice A des coefficients de (P)), et a^j désignera le vecteur $a^j = (a_{1j}, a_{2j}, \dots, a_{mj})$ (la colonne j de la matrice A).

$$\langle a_1, x \rangle \leq b_1, \langle a_2, x \rangle \leq b_2, \dots, \langle a_m, x \rangle \leq b_m$$

x étant comme toujours $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

Si l'on fait recours à la notation

$$A = (a_{ij}) : i = 1, 2, \dots, m, j = 1, 2, \dots, n,$$

Notre système se réécrit

$$Ax \leq b, \quad b = (b_1, b_2, \dots, b_m).$$

Et on désignera par la suite l'ensemble des vecteurs réalisables (c'est-à-dire l'ensemble des vecteurs x vérifiant (P)) par

$$\Omega = \{ x \in \mathbb{R}^n : Ax \leq b \}$$

Remarque 2.1.2. Il se peut que Ω soit vide.

Exemple 2.1.5. L'ensemble Ω défini par les contraintes, $x \leq 2$ et $-x \leq -3$ et Ω est vide.

Théorème 2.1.10. Tout ensemble Ω des vecteurs réalisables est convexe et fermé.

Preuve. On montre que le demi-espace Ω_i défini par $\langle a_i, x \rangle \leq b_i$ est convexe.

En effet, soit $x, x' \in \Omega_i$. On a $\langle a_i, x \rangle \leq b_i$, $\langle a_i, x' \rangle \leq b_i$. Soit α un nombre quelconque tel que $0 \leq \alpha \leq 1$ auquel

$$\alpha \langle a_i, x \rangle \leq \alpha b_i, \quad (1 - \alpha) \langle a_i, x' \rangle \leq (1 - \alpha) b_i.$$

On fait la somme et on utilise la linéarité du produit scalaire, il vient

$$\langle a_i, \alpha x + (1 - \alpha)x' \rangle \leq \alpha b_i + (1 - \alpha)b_i = b_i \implies \alpha x + (1 - \alpha)x' \in \Omega_i.$$

D'où la convexité de Ω_i .

On constate que l'ensemble Ω est l'intersection de demi-espaces Ω_i .i.e

$x \in \Omega$ si et seulement si x dans de tous les Ω_i , $i = 1, 2, \dots, m$.

Il ne reste qu'à appliquer le théorème (l'intersection d'ensemble finis convexes est convexe).

On peut utiliser l'écriture matricielle ($Ax \leq b$), pour démontrer le théorème.

Si $x, x' \in \Omega$, alors $Ax \leq b, Ax' \leq b$, dans ce cas

$$\forall \lambda \in [0, 1], \lambda Ax \leq \lambda b \text{ et } (1 - \lambda)Ax' \leq (1 - \lambda)b.$$

D'où

$$\lambda Ax + (1 - \lambda)Ax' \leq \lambda b + (1 - \lambda)b = b \implies A(\lambda x + (1 - \lambda)x') \leq b.$$

Alors on obtient $\lambda x + (1 - \lambda)x' \in \Omega$. D'où la convexité de Ω .

• La fermeture de Ω .

Soit $(x_n)_{n \in \mathbb{N}} \subset \Omega$, une suite convergente telle que

$$\lim_{n \rightarrow \infty} x_n = \bar{x}.$$

On a $Ax_n \leq b$, alors la limite pour passage à la limite, on obtient

$$\lim_{n \rightarrow \infty} Ax_n \leq b \implies A\bar{x} \leq b, \text{ donc } \bar{x} \in \Omega.$$

■

Définition 2.1.14. Un point $x_0 \in \Omega$ est dit intérieur s'il vérifie strictement toutes les inéquations du (P) ($\langle a_i, x_0 \rangle < b_i$ pour $i = 1, 2, \dots, m$).

Tous les autres points de Ω seront qualifiés de point frontières.

Remarque 2.1.3. Cette définition est parfaitement cohérente avec la notion usuelle d'un point intérieur.

En effet, si l'on utilise la continuité des fonction $\langle a_i, x \rangle$ et le fait qu'elles sont en nombre fini dans notre cas ($i = 1, 2, \dots, m$), on trouve facilement un nombre ε tel que si la longueur du vecteur $\delta = (\delta_1, \delta_2, \dots, \delta_n)$ est plus petite que ε , alors on ait les inégalités

$\langle a_i, x_0 + \delta \rangle \leq b_i$, $i = 1, 2, \dots, m$ c'est-à-dire le point $(x_0 + \delta)$ appartient lui aussi à Ω .

En d'autres termes, tant que X contient un point intérieur x_0 il contient un ε -voisinage de ce point.

Définition 2.1.15. [14].

Le support d'un point frontière x_0 est l'ensemble des contraintes (P) vérifiées par x_0 sous forme d'égalités.

Autrement dit, le support de x_0 est l'ensemble des hyperplans $\langle a_i, x \rangle = b_i$ qui définissent Ω et qui contiennent x_0 .

Soient $\sigma = (i_1, i_2, \dots, i_k)$ $1 \leq i_j \leq m$ $j = 1, 2, \dots, k$ les numéros des vecteurs-lignes a_i qui définissent le support de x_0 et $\bar{\sigma}$ l'ensemble des numéros $i \in \{1, 2, \dots, m\}$ restants.

On note A_σ la matrice de vecteurs-lignes a_i , $i \in \sigma$ et b_σ le vecteur formé de coordonnées de b indicés par les éléments de σ .

On introduit de même les notations $A_{\bar{\sigma}}$ et $b_{\bar{\sigma}}$.

Si on prend comme exemple

$$A = \begin{pmatrix} 3 & 5 \\ -1 & 1 \\ 10 & 7 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 15 \\ 2 \\ 35 \\ -1 \\ 0 \\ 0 \end{pmatrix} \quad \text{et } x = \left(\frac{70}{29}, \frac{45}{29} \right).$$

On a

$$Ax = \begin{pmatrix} 3 & 5 \\ -1 & 1 \\ 10 & 7 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{70}{29} \\ \frac{45}{29} \end{pmatrix} = \begin{pmatrix} 15 \\ \frac{-25}{29} \\ 35 \\ \frac{115}{29} \\ \frac{-70}{29} \\ \frac{45}{29} \end{pmatrix}.$$

Alors

$$\sigma = \{1, 3\}, \bar{\sigma} = \{2, 4, 5, 6\} \quad A_\sigma = \begin{pmatrix} 3 & 5 \\ 10 & 7 \end{pmatrix} \text{ et } b_\sigma = \begin{pmatrix} 15 \\ 35 \end{pmatrix}$$

$$A_{\bar{\sigma}} = \begin{pmatrix} -1 & 1 \\ -1 & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \text{ et } b_{\bar{\sigma}} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix} .$$

Les contraintes (P) ($Ax \leq b$) deviennent pour x_0 .

$$A_\sigma x_0 = b_\sigma, \quad A_{\bar{\sigma}} x_0 < b_{\bar{\sigma}}.$$

On rappelle que le rang d'une matrice est le nombre maximum des lignes (ou colonnes) linéairement indépendantes .

Théorème 2.1.11. *Un point $x_0 \in \Omega$ est un point extrême si et seulement si le rang de A_σ est n .*

Preuve. On montre en premier lieu que si x_0 est un point extrême, alors le rang de A_σ vaut n .

Supposons que ce n'est pas le cas. Le rang de A_σ est alors inférieur à n , si bien que $A_\sigma x = 0$ a pour solution $\bar{x} \neq 0$.

Comme $A_{\bar{\sigma}} x_0 < b_{\bar{\sigma}}$, on trouve $\alpha \neq 0$ suffisamment petit tel que

$$A_{\bar{\sigma}} (x_0 + \alpha \bar{x}) < b_{\bar{\sigma}}; \quad A_{\bar{\sigma}} (x_0 - \alpha \bar{x}) < b_{\bar{\sigma}}$$

et

$$A_\sigma (x_0 + \alpha \bar{x}) = A_\sigma x_0 + \alpha A_\sigma \bar{x} = A_\sigma x_0 = b_\sigma.$$

De même $A_\sigma (x_0 - \alpha \bar{x}) = A_\sigma x_0 - \alpha A_\sigma \bar{x} = A_\sigma x_0 = b_\sigma$.

Il en résulte que $(x_0 + \alpha \bar{x}), (x_0 - \alpha \bar{x}) \in \Omega$, ce qui contredit l'hypothèse x_0 est un point extrême (en effet, $x_0 = \frac{1}{2}(x_0 + \alpha \bar{x}) + \frac{1}{2}(x_0 - \alpha \bar{x})$).

Inversement, soit n le rang de A_σ .

On suppose que x_0 n'est pas un point extrême de Ω . Alors il existe deux points

$x', x'' \in \Omega$ avec $x' \neq x''$ et $x_0 = \frac{1}{2}x' + \frac{1}{2}x''$. Alors on a $b_\sigma = A_\sigma x_0 = \frac{1}{2}A_\sigma x' + \frac{1}{2}A_\sigma x''$.

Du moment que $A_\sigma x' \leq b_\sigma, A_\sigma x'' \leq b_\sigma$, la dernière égalité entraîne $A_\sigma x' = A_\sigma x'' = b_\sigma$.

Car on suppose que

$$A_\sigma x' < b_\sigma \implies \frac{1}{2}A_\sigma x' < \frac{1}{2}b_\sigma .$$

Donc

$$b_\sigma = \frac{1}{2}A_\sigma x' + \frac{1}{2}A_\sigma x'' < \frac{1}{2}b_\sigma + \frac{1}{2}b_\sigma = b_\sigma \text{ contredit.}$$

Alors $A_\sigma x' = A_\sigma x'' = b_\sigma$, d'où $A_\sigma(x' - x'') = 0$, ce qui contredit l'hypothèse faite sur le rang de A_σ (car $x' - x'' \neq 0$).

Donc le théorème est démontré. ■

Ce théorème nous donne deux résultats importants.

Théorème 2.1.12. *Les point extrême de l'ensemble Ω , s'il en existe, sont en nombre fini.*

Preuve. D'après le théorème précédent, à tout point extrême il correspond un matrice carrée non dégénérée A_σ qui le définit complètement : $x_0 = A_\sigma^{-1}b_\sigma$.

Les points extrêmes de l'ensemble Ω sont donc en nombre fini (car il en est de même des sous-matrices de A). ■

Théorème 2.1.13. *Si l'ensemble Ω est borné, alors c'est un polyèdre convexe.*

Preuve. On a Ω est borné fermé alors Ω est compact et on a Ω convexe donc Ω est l'enveloppe convexe de ses points extrêmes mais le nombre des points extrêmes de Ω est finie. Alors Ω un polyèdre convexe. ■

Définition 2.1.16. *Si le rang de la matrice A est égal à n , l'ensemble $\Omega = \{ x \mid Ax \leq b \}$ est dit régulier.*

Théorème 2.1.14. [1].

L'ensemble Ω possède des points extrêmes si et seulement s'il est régulier.

2.2 Fonctions convexes

Soit Ω un ouvert de \mathbb{R}^n , a, b deux points de Ω tel que le segment $((1-t)a + tb) \in \Omega$ pour tout $t \in [0, 1]$, on suppose que la fonction

$$\begin{aligned} f : \Omega \subseteq \mathbb{R}^n &\longrightarrow \mathbb{R} \\ x &\longrightarrow f(x). \end{aligned}$$

admet des dérivées partielles continues dans Ω , c.à.d, f est de classe C^1 .

Alors la fonction

$$\begin{aligned} g : [0, 1] &\longrightarrow \mathbb{R} \\ t &\longrightarrow g(t) = f((1-t)a + tb) \end{aligned}$$

est dérivable dans $[0, 1]$. De plus,

$$\forall t \in [0, 1], \quad g'(t) = \sum_{i=1}^n (b_i - a_i) \frac{\partial f}{\partial x_i}((1-t)a + tb).$$

Car $(\sum_{i=1}^n (b_i - a_i) \frac{\partial f}{\partial x_i}((1-t)a + tb) = \langle b - a, \nabla f((1-t)a + tb) \rangle = (b - a) \cdot \nabla f((1-t)a + tb)$ où $b = (b_i)_{i=1}^n$ et $a = (a_i)_{i=1}^n$ en vertu du théorème de Lagrange, on a $g(1) - g(0) = (1-0)g'(t_0)$ où $t_0 \in]0, 1[$, mais $g(1) = f(b)$ et $g(0) = f(a)$, et en posant $t_0 = \varepsilon$ on obtient

$$f(b) - f(a) = \langle b - a, \nabla f((1-\varepsilon)a + \varepsilon b) \rangle = \sum_{i=1}^n (b_i - a_i) \frac{\partial f}{\partial x_i}((1-\varepsilon)a + \varepsilon b) \quad \text{où } \varepsilon \in]0, 1[.$$

Soit

$$\begin{aligned} f : \Omega &\longrightarrow \mathbb{R}, & \varphi : I \subset \mathbb{R} &\longrightarrow \Omega, & \Omega &\text{ un ouvert de } \mathbb{R}^n \\ t &\longrightarrow \varphi(t) \end{aligned}$$

Considérons l'hypersurface $f(x) = b$, et une courbe φ telle que $\varphi(t) = x$ et $f(x) = b$.

φ est une courbe qui appartient à l'hypersurface $f(x) = b$.

Soit $t_0 \in I : \varphi(t_0) = x_0, f(x_0) = b$.

Définition 2.2.1. On dit qu'une droite est tangente à une surface en un point x_0 si elle est tangente à n'importe quelle courbe φ passant par x_0 .

Définition 2.2.2. Un point x_0 de la surface $f(x) = b$ est un point singulier si

$$\frac{\partial f}{\partial x_i}(x_0) = 0, \quad \forall i = 1, 2, \dots, n.$$

S'il existe $i_0 \in \{1, 2, \dots, n\}$ tel que $\frac{\partial f}{\partial x_{i_0}}(x_0) \neq 0$ on dira que c'est un point simple de la surface.

Toutes les droites tangentes à la surface en un point simple x_0 appartiennent à un même plan, appelé plan tangent à la surface au point x_0 .

On a : $\varphi(t) = (\varphi_1(t), \dots, \varphi_n(t))$, et

$$\begin{aligned} x_1 &= \varphi_1(t), & \varphi_1'(t) &= \frac{dx_1}{dt} \\ & \cdot & \cdot & \\ & \cdot & \cdot & \\ & \cdot & \cdot & \\ x_n &= \varphi_n(t), & \varphi_n'(t) &= \frac{dx_n}{dt} \end{aligned}$$

$(\varphi_1'(t_0), \dots, \varphi_n'(t_0))$ est un vecteur tangent à la surface $f(x) = b$ au point x_0 ($\varphi(t_0) = x_0$). Mais $f(x) = b$, alors

$$f(x) - b = 0 \Rightarrow \frac{\partial f}{\partial x_1} \cdot \frac{\partial x_1}{\partial t} + \frac{\partial f}{\partial x_2} \cdot \frac{\partial x_2}{\partial t} + \dots + \frac{\partial f}{\partial x_n} \cdot \frac{\partial x_n}{\partial t} = 0.$$

$N = \left(\frac{\partial f(x_0)}{\partial x_1}, \dots, \frac{\partial f(x_0)}{\partial x_n} \right)$ est le vecteur normal à la surface au point x_0 .

Soit x_0 un point, $f(x) = b$, et soit x un point du plan tangent à la surface au point x_0 .

Alors $(x - x_0)$ est un vecteur du plan tangent et ce vecteur est perpendiculaire à N , donc

$\left\langle \left(\frac{\partial f(x_0)}{\partial x_i} \right)_i, x - x_0 \right\rangle = 0$ d'ou $\langle \nabla f(x_0), x - x_0 \rangle = 0$. Ce qui implique que l'équation du plan tangent à la surface $f(x) = b$ est : donnée par

$$\langle \nabla f(x_0), x \rangle = \langle \nabla f(x_0), x_0 \rangle.$$

Définition 2.2.3. Soient Ω un convexe de \mathbb{R}^n non vide, et $f : \Omega \rightarrow \mathbb{R}$. On dit que f est convexe si pour tout x_1, x_2 de Ω et pour tout $\lambda : 0 \leq \lambda \leq 1$, on a

$$f(\lambda x_2 + (1 - \lambda)x_1) \leq \lambda f(x_2) + (1 - \lambda)f(x_1).$$

On dira que f est concave si $(-f)$ est convexe. On dira que f est strictement convexe si pour toutes x_1, x_2 de Ω on a $f(\lambda x_2 + (1 - \lambda)x_1) < \lambda f(x_2) + (1 - \lambda)f(x_1)$.

Théorème 2.2.1. *Soit f est fonction de classe C^1 sur un ouvert convexe $\Omega \subset \mathbb{R}^n$. Si f est convexe dans Ω , alors on a*

$$\langle (x_2 - x_1), \nabla f(x_1) \rangle \leq f(x_2) - f(x_1) \text{ pour tout } x_1, x_2 \in \Omega.$$

Preuve. Soit $x_1, x_2 \in \Omega$ et $\lambda \in]0, 1]$. Comme f est convexe alors

$$f(\lambda x_2 + (1 - \lambda)x_1) \leq \lambda f(x_2) + (1 - \lambda)f(x_1) \implies \frac{f(\lambda x_2 + (1 - \lambda)x_1) - f(x_1)}{\lambda} \leq f(x_2) - f(x_1).$$

Mais d'après l'égalité

$$f(b) - f(a) = \langle b - a, \nabla f((1 - \varepsilon)a + \varepsilon b) \rangle,$$

et on posent $x_\lambda = \lambda x_2 + (1 - \lambda)x_1$, on obtient

$$\begin{aligned} f(x_\lambda) - f(x_1) &= \langle x_\lambda - x_1, \nabla f((1 - \varepsilon)x_1 + \varepsilon x_\lambda) \rangle \\ &= \langle \lambda x_2 + (1 - \lambda)x_1 - x_1, \nabla f(x_1 - \varepsilon x_1 + \varepsilon(\lambda x_2 + x_1 - \lambda x_1)) \rangle \\ &= \langle \lambda x_2 + x_1 - \lambda x_1 - x_1, \nabla f(x_1 - \varepsilon x_1 + \varepsilon \lambda x_2 + \varepsilon x_1 - \varepsilon \lambda x_1) \rangle \\ &= \langle \lambda(x_2 - x_1), \nabla f(x_1 + \lambda \varepsilon(x_2 - x_1)) \rangle = \lambda \langle x_2 - x_1, \nabla f(x_1 + \lambda \varepsilon(x_2 - x_1)) \rangle \\ &\leq \langle (x_2 - x_1), \nabla f(x_1 + \lambda \varepsilon(x_2 - x_1)) \rangle \leq f(x_2) - f(x_1). \end{aligned}$$

Par passage à la limite quand $\lambda \rightarrow 0$, on obtient

$$\langle x_2 - x_1, \nabla f(x_1) \rangle \leq f(x_2) - f(x_1).$$

■

2.2.1 L'hyperplangent à l'hypersurface

Théorème 2.2.2. *Soit f une fonction convexe alors l'ensemble défini par $\{x : f(x) \leq b\}$ est convexe.*

Preuve. Posons $\Omega = \{x : f(x) \leq b\}$

Si Ω contient un point le théorème est évident .

Sinon, soient x_1, x_2 deux points de Ω alors

$$\lambda x_1 + (1 - \lambda)x_2 \in \Omega, \forall \lambda \in [0, 1].$$

Car

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) \leq \lambda b + (1 - \lambda)b = b \implies \lambda x_1 + (1 - \lambda)x_2 \in \Omega.$$

■

Remarque 2.2.1. L'ensemble $\{x, f(x) = b\}$ n'est pas en général convexe, car

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) = b.$$

Cependant, l'application $x \rightarrow \langle a, x \rangle = ax$ est convexe et $\Omega = \{x : \langle a, x \rangle = b, x \geq 0\}$ convexe.

Théorème 2.2.3. [9] et [18].

Soit f une fonction convexe de classe C^1 pour tout $x \geq 0$.

Alors l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geq 0$ est l'hyperplan, d'appui à $\Omega = \{x \geq 0, f(x) \leq b\}$ au point x_0 .

Preuve. Ω est convexe, comme f admet des dérivées partielles continues, elle est continue, donc Ω est convexe fermé.

On sait qu'on peut construire un hyperplan d'appui à Ω au point x_0

Ici, on veut montrer que cet hyperplan à Ω au point x_0 , ($f(x_0) = b$) est le plan tangent à la surface $f(x) = b$ au point x_0 .

Mais on sait que l'équation de l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geq 0$ est

$$\langle \nabla f(x_0), x \rangle = \langle \nabla f(x_0), x_0 \rangle.$$

Il est évident que l'hyperplan $\langle \nabla f(x_0), x \rangle = \langle \nabla f(x_0), x_0 \rangle$ contient le vecteur x_0 .

Il rest à montrer qu'il sépare le convexe fermé $\Omega = \{x \geq 0 : f(x) \leq b\}$.

Soit $x_1 \in \Omega$, $x_1 \geq 0 : f(x_1) \leq b$

$$\langle \nabla f(x_0), x_0 \rangle - \langle \nabla f(x_0), x_1 \rangle = \langle \nabla f(x_0), -x_1 + x_0 \rangle.$$

Mais

$$\begin{aligned} \langle \nabla f(x_0), -x_1 + x_0 \rangle &= -\langle x_1 - x_0, \nabla f(x_0) \rangle \geq -(f(x_1) - f(x_0)) \\ &\geq f(x_0) - f(x_1) = b - f(x_1) \\ &\geq 0. \end{aligned}$$

Alors

$$\langle \nabla f(x_0), x_0 \rangle - \langle \nabla f(x_0), x_1 \rangle \geq 0 \Rightarrow \langle \nabla f(x_0), x_0 \rangle \geq \langle \nabla f(x_0), x_1 \rangle.$$

Mais on a

$$\langle \nabla f(x_0), x_0 \rangle = \langle \nabla f(x_0), x \rangle \quad (f(x) = b).$$

On obtient donc

$$\langle \nabla f(x_0), x \rangle \geq \langle \nabla f(x_0), x_1 \rangle, \quad \forall x_1 \in \Omega \text{ et } f(x) = b.$$

■

On peut de la *même* manière démontrer le résultat suivant pour le cas d'une fonction concave.

Théorème 2.2.4. *Soit f est fonction de classe C^1 sur un ouvert convexe $\Omega \subset \mathbb{R}^n$, si f concave alors*

$$\langle \nabla f(x_1), x_2 - x_1 \rangle \geq f(x_2) - f(x_1).$$

Preuve. On a $(-f)$ est convexe alors on obtien $\forall x_1, x_2 \in \Omega$,

$$\begin{aligned} \langle \nabla(-f)(x_1), x_2 - x_1 \rangle &\leq (-f)(x_2) - (-f)(x_1) \\ \Rightarrow \langle -\nabla f(x_1), x_2 - x_1 \rangle &\leq -f(x_2) + f(x_1) \\ \Rightarrow -\langle \nabla f(x_1), x_2 - x_1 \rangle &\leq -(f(x_2) - f(x_1)) \\ \Rightarrow \langle \nabla f(x_1), x_2 - x_1 \rangle &\geq f(x_2) - f(x_1). \end{aligned}$$

■

Théorème 2.2.5. *Si f est une fonction concave alors l'ensemble $\Omega = \{x : f(x) \geq b\}$ est convexe.*

Preuve. Soit $x_1, x_2 \in \Omega$ alors $f(x_1) \geq b$ et $f(x_2) \geq b$. Comme f est concave, $\forall \lambda \in [0, 1]$

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)x_2) &\geq \lambda f(x_1) + (1 - \lambda)f(x_2) \\ &\geq \lambda b + (1 - \lambda)b = b. \end{aligned}$$

Alors $\lambda x_1 + (1 - \lambda)x_2 \in \Omega$.

■

Théorème 2.2.6. *Si f est une fonction concave de classe C^1 pour tout $x \geq 0$. Alors l'hyperplan tangent à l'hypersurface $f(x) = b$ au point $x_0 \geq 0$ est l'hyperplan d'appui à*

$$\Omega = \{x \geq 0, f(x) \geq b\}.$$

2.2.2 Minimum et maximum d'une fonction convexe

Définition 2.2.4. On dit que x^* est un point de minimum local de f sur l'ensemble Ω si et seulement si

$$\exists \eta > 0, \forall x \in \Omega, \|x - x^*\| \leq \eta \Rightarrow f(x^*) \leq f(x).$$

Et on dit que x^* est un point de minimum global de f sur Ω si et seulement si

$$\forall x \in \Omega : f(x^*) \leq f(x).$$

Théorème 2.2.7. Soit f une fonction convexe sur un convexe fermé Ω de \mathbb{R}^n , alors le minimum local de f est un minimum global de f sur Ω .

Preuve. Soit x_0 un minimum local de f sur Ω , donc

$$\exists \eta > 0, \forall y \in \Omega : \|y - x_0\| \leq \eta \Rightarrow f(x_0) \leq f(y).$$

Soit $x \in \Omega$, pour $\theta \in]0, 1[$ suffisamment petit $x_\theta = \theta x + (1 - \theta)x_0$ vérifie $\|x_\theta - x_0\| \leq \eta$ et on a $x_\theta \in \Omega$ car Ω est convexe. On a f convexe, donc

$$\begin{aligned} f(x_\theta) &= f(\theta x + (1 - \theta)x_0) \leq \theta f(x) + (1 - \theta)f(x_0) \\ \Rightarrow f(x_0) &\leq f(x_\theta) \leq \theta f(x) + f(x_0) - \theta f(x_0) \quad / \quad f(x_0) \leq \theta f(x) + f(x_0) - \theta f(x_0) \\ &\Rightarrow \theta f(x_0) \leq \theta f(x) \quad \text{mais } \theta > 0 \text{ alors } f(x_0) \leq f(x), \forall x \in \Omega. \end{aligned}$$

Ce qui montre bien que x_0 est un minimum global sur Ω . ■

Remarque 2.2.2. Ici, on suppose que le minimum local existe, f est localement bornée, donc bornée sur le fermé Ω .

Théorème 2.2.8. Soit f une fonction convexe sur un convexe fermé Ω de \mathbb{R}^n . Alors l'ensemble des points minimaux locaux de f est convexe.

Preuve. Si x_1 et x_2 sont deux minimum locaux tel que $x_1 \neq x_2$, et si $\theta \in [0, 1]$, alors

$$x = \theta x_1 + (1 - \theta)x_2 \in \Omega, \text{ et on a } f(x) \leq \theta f(x_1) + (1 - \theta)f(x_2).$$

$$\text{Mais } f(x_1) = f(x_2) = \inf_{y \in \Omega} f(y),$$

$$\text{alors } f(x) \leq \theta \inf_{y \in \Omega} f(y) + (1 - \theta) \inf_{y \in \Omega} f(y) = \inf_{y \in \Omega} f(y)$$

$\Rightarrow f(x) \leq \inf_{y \in \Omega} f(y)$ alors $f(x) = \inf_{y \in \Omega} f(y)$ car $x \in \Omega$, donc x est encore un minimum de f sur Ω . ■

Remarque 2.2.3. [11] et [19].

$f(x) = \langle a, x \rangle$, f est convexe. Si x_1 est un minimum de f , x_2 un autre alors le segment deivent un minimum.

Théorème 2.2.9. Si f est strictement convexe sur un convexe fermé Ω de \mathbb{R}^n , alors le minimum locale qui est global est unique.

Preuve. On suppose qu'il exist x_1 et x_2 deux minimum. Si $\theta \in]0, 1[$, alors $y = \theta x_1 + (1 - \theta)x_2 \in \Omega$. Donc

$$f(y) < \theta f(x_1) + (1 - \theta)f(x_2) = \theta \inf_{z \in \Omega} f(z) + (1 - \theta) \inf_{z \in \Omega} f(z) = \inf_{z \in \Omega} f(z) \Rightarrow f(y) < \inf_{z \in \Omega} f(z)$$

contradiction. ■

Théorème 2.2.10. Soit f une fonction de classe C^1 sur un convexe Ω de \mathbb{R}^n . Si f est convexe, et si $\nabla f(x_0) = 0$, avec $x_0 \in \overset{0}{\Omega}$ (intérieur de Ω), alors le minimum global est $x = x_0$.

Preuve. On a $\langle x - x_0, \nabla f(x_0) \rangle \leq f(x) - f(x_0)$, mais $\nabla f(x_0) = 0$, alors

$f(x) - f(x_0) \geq 0 \Rightarrow f(x) \geq f(x_0)$ pour tout $x \in \Omega$. Donc x_0 est le minimum global de f . ■

Théorème 2.2.11. [20].

Soit f une fonction convexe sur un convexe fermé borné. Si le maximum global de f est fini, alors il est atteint en un ou plusieurs sommets.

Preuve.

i) Supposons que $\max_{x \in \Omega} f(x) = f(x^*)$ et que $x^* \in \overset{0}{\Omega}$. Si f n'est pas constante, $\exists x_1 \in \Omega$ tel que $f(x_1) < f(x^*)$, mais $x^* \in \overset{0}{\Omega}$, donc

$$\exists x_2 \in \Omega \text{ et } \lambda \in]0, 1[\text{ tel que } x^* = \lambda x_2 + (1 - \lambda)x_1,$$

comme f convexe alors $f(x^*) \leq \lambda f(x_2) + (1 - \lambda)f(x_1)$.

Mais $f(x_1) < f(x^*)$, donc $f(x^*) < \lambda f(x_2) + (1 - \lambda)f(x^*)$

$$\Rightarrow f(x^*) < \lambda(f(x_2) - f(x^*)) + f(x^*),$$

d'où $\lambda(f(x_2) - f(x^*)) > 0 \Rightarrow f(x^*) < f(x_2)$, contraction. Donc f est une constante sur Ω .

Mais un ensemble Ω de \mathbb{R}^n fermé et borné possède au moins un point extrême \bar{x} et

$f(\bar{x}) = c^{st}$. D'où $\max_{x \in \Omega} f(x) = f(\bar{x}) = c^{st}$.

ii) Supposons que x^* est un point frontière et que le maximum de f n'est pas un point intérieur de Ω . On doit montrer que x^* est un point extrême. On sait qu'il existe un hyperplan H , qui sépare x^* et le convexe fermé Ω , x^* est un point frontière donc l'hyperplan d'appui à Ω au point x^* est $H = H_1$.

Posons $T_1 = \Omega \cap H_1 \subset \Omega$, T_1 est fermé, convexe et borné (car $T_1 \subset \Omega$), donc possède au moins un point extrême, mais $T_1 \subset H_1$ (hyperplan), donc

$$\dim T_1 = n - 1, x^* \in H_1, x^* \in \Omega \Rightarrow x^* \in T_1.$$

Si x^* est un point intérieur à T_1 d'après (i) $\Rightarrow f = c^{st}$ et on aura une contradiction.

Sinon $x^* \in T_1$ est un point frontière de T_1 d'après (ii), posons $T_2 = T_1 \cap H_2, \dots$, on arrive alors à $T_n = T_{n-1} \cap H_n$, et maintenant $\dim T_n = 0 \Rightarrow T_n$ contient un point unique.

Ce point unique est bien évidemment x^* , $T_n = \{x^*\}$.

Un singleton est un point extrême car T_n est fermé et borné, donc il doit contenir un point extrême. ■

D'une manière analogue, on peut démontrer les théorèmes suivants.

Théorème 2.2.12. Soit f une fonction concave définie sur un convexe fermé de \mathbb{R}^n , alors le maximum local de f est global (car, $\max_{x \in \Omega} f(x) = - \inf_{x \in \Omega} (-f(x))$).

Théorème 2.2.13. Soit f une fonction concave définie sur un convexe fermé de \mathbb{R}^n , alors l'ensemble des maximum globaux est un convexe.

Théorème 2.2.14. Soit f une fonction concave définie sur un convexe fermé de \mathbb{R}^n , si x_1 et x_2 sont deux maximum alors $(\lambda x_1 + (1 - \lambda)x_2)$ est aussi un maximum global pour tout $\lambda \in [0, 1]$.

Théorème 2.2.15. Soit f une fonction de classe C^1 dans un ouvert convexe Ω de \mathbb{R}^n , f est concave, et $\nabla f(x_0) = 0$, avec $x_0 \in \Omega$ alors f atteint son maximum global au point x_0 .

Théorème 2.2.16. Soit f une fonction de classe C^1 , sur un convexe fermé et borné et si f est concave sur Ω , et si le minimum global de f existe, alors il est atteint en un point extrême du convexe Ω .

Définition 2.2.5. L'épigraphe de la fonction f , définie sur un sous-ensemble non vide Ω de \mathbb{R}^n est

$$\text{epi}(f) = \{(t, x) \in \mathbb{R}^{n+1}, x \in \Omega, f(x) \leq t\}.$$

Géométriquement, pour définir l'épigraphe il faut prendre le graphe de la fonction la surface $\{t = f(x), x \in \Omega\}$ dans \mathbb{R}^{n+1} et ajouter à cette surface tous les points qui sont au-dessus.

La définition géométrique (équivalente) d'une fonction convexe est donnée par :

Proposition 2.2.1. (Définition de la convexité en termes d'épigraphe)

La fonction f définie sur un sous-ensemble de \mathbb{R}^n est convexe si et seulement si son épigraphe est un ensemble convexe non vide dans \mathbb{R}^{n+1} .

Proposition 2.2.2. (L'inégalité de Jensen)

Soit f une fonction convexe et soit Ω le domaine de f alors pour n'importe quelle combinaison convexe, $\sum_{i=1}^n \lambda_i = 1, \lambda_i \in \mathbb{R}_+$ et $\sum_{i=1}^n \lambda_i x_i$ des points de Ω , on a

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Preuve. On a les points

$$(f(x_i), x_i) \in \text{epi}(f) \text{ car } (f(x_i), x_i) \in \mathbb{R}^{n+1} \text{ et } f(x_i) \leq f(x_i).$$

Comme f est convexe alors son épigraphe est un ensemble convexe de sorte que la combinaison convexe

$$\sum_{i=1}^n \lambda_i (f(x_i), x_i) = \left(\sum_{i=1}^n \lambda_i f(x_i), \sum_{i=1}^n \lambda_i x_i \right).$$

Par la définition de l'épigraphe, ça implique que $f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i)$.

■

Notez que la définition de la convexité d'une fonction f est exactement la condition que f satisfait l'inégalité de Jensen dans le cas de $n = 2$.

Chapitre 3

Optimisation de la fonction concave

3.1 Décomposition de la fonction-objectif

Considérons le problème suivant :

$$(P) : \begin{cases} x \geq 0 \\ Ax \leq b \\ \max_{x \in \Omega} f(x) \\ \text{avec } f(x) = \sum_{i=1}^n \alpha_i x_i + \sum_{i,j} \gamma_{ij} x_i x_j \end{cases}$$

où $\Omega = \{x \in \mathbb{R}^n / x \geq 0 \text{ et } Ax \leq b\}$, A une matrice de dimension $m \times n$, b un vecteur de \mathbb{R}^m tel que $b_i \geq 0 \quad i = 1, 2, \dots, m$.

Il est facile de qu'on peut se restreindre à l'étude du cas suivant

$$\sum_{i=1}^n (\alpha_i x_i + \beta_i x_i^2) \text{ avec } \alpha_i \in \mathbb{R} \text{ et } \beta_i \in \mathbb{R},$$

où α_i et β_i sont des coefficients réels quelconques, et la fonction f est quelconque.

Commençons par comment optimiser la fonction f .

Soit $I_1 = \{i / \beta_i < 0\}$ et $I_2 = \{i / \beta_i \geq 0\}$, on a donc

$$f(x) = \sum_{i=1}^n (\alpha_i x_i + \beta_i x_i^2) = \sum_{i \in I_1} (\alpha_i x_i + \beta_i x_i^2) + \sum_{i \in I_2} (\alpha_i x_i + \beta_i x_i^2).$$

On pose

$$\varphi(x) = \sum_{i \in I_1} (\alpha_i x_i + \beta_i x_i^2) \text{ et } \psi(x) = \sum_{i \in I_2} (\alpha_i x_i + \beta_i x_i^2),$$

ce qui implique que la fonction f devient :

$$f(x) = \varphi(x) + \psi(x) .$$

Nous allons commencer par la fonction concave φ .

Si $\varphi(x) = \sum_{i=1}^n (\alpha_i x_i + \beta_i x_i^2)$, on a donc

$$\frac{\partial \varphi}{\partial x_i} = \alpha_i + 2\beta_i x_i$$

alors

$$\frac{\partial \varphi}{\partial x_i}(x^*) = 0 \implies x^* = \left(\frac{-\alpha_i}{2\beta_i} \right)_i .$$

On va que la solution optimale est la projection du point x^* sur un nouveau convexe Ω' . Si le point critique x^* est à l'intérieur du convexe Ω , alors la solution optimale du problème (P) se trouve exactement en ce point. Sinon, la projection directe du point x^* sur Ω donne seulement une solution approchée.

Nous montrons dans le théorème 4.1.1, que

$$\max_{x \in \Omega} \varphi(x) = \varphi(\bar{x})$$

où \bar{x} est la projection du point x^* sur le nouveau convexe Ω' .

3.2 Optimisation de la fonction concave

Soit φ est un fonction concave définie par

$$\varphi(x) = \sum_{i=1}^n (\alpha_i x_i + \beta_i x_i^2) \text{ avec } \alpha_i \in \mathbb{R} \text{ et } \beta_i < 0 \text{ pour tout } i = 1, 2, \dots, n.$$

φ définie sur l'ensemble convexe fermé borné de \mathbb{R}^n , et posons :

$$x^* = (x_i^*)_i = \left(\frac{-\alpha_i}{2\beta_i} \right)_i, \quad y^* = (y_i^*)_i = \left(\frac{\alpha_i}{2\sqrt{-\beta_i}} \right)_i$$

$$y = (y_i)_i = \left(\sqrt{-\beta_i x_i} \right)_i, \quad \max_{x \in \Omega} \varphi(x) = \varphi(\bar{x}), \quad \bar{x} = (\bar{x}_i)_i \in \Omega$$

Théorème 3.2.1. [6].

Il existe un ensemble convexe fermé borné Ω' de \mathbb{R}^n , et un vecteur $y_0 = (y_{0i})_{i \in \Omega'}$, tel que les condition suivantes soient satisfaites :

$$\max_{x \in \Omega} \varphi(x) = \varphi(x^*) - \|y^* - y_0\|^2 \quad (3.1)$$

$$\|y^* - y_0\| = \inf_{y \in \Omega'} \|y^* - y\| \quad (3.2)$$

$$\bar{x}_i = \frac{y_{0i}}{\sqrt{-\beta_i}} \text{ pour tout } i, i = 1, 2, \dots, n. \quad (3.3)$$

Lemme 3.2.1. .

Si φ un fonction continue sur un ensemble borné on a donc,

$$\inf_{x \in \Omega} (b - \varphi(x)) = b - \max_{x \in \Omega} \varphi(x) \quad b \in \mathbb{R}^n .$$

Preuve. (lemme 4.1.1).

On a toujours

$$\inf_{x \in \Omega} (b - \varphi(x)) \leq b - \varphi(x), \forall x \in \Omega \implies -b + \inf_{x \in \Omega} (b - \varphi(x)) \leq -\varphi(x) \quad \forall x \in \Omega.$$

Alors

$$b - \inf_{x \in \Omega} (b - \varphi(x)) \geq \varphi(x), \forall x \in \Omega,$$

on aura donc

$$b - \inf_{x \in \Omega} (b - \varphi(x)) \geq \max_{x \in \Omega} \varphi(x) .$$

Alors on obtient

$$- \inf_{x \in \Omega} (b - \varphi(x)) \geq -b + \max_{x \in \Omega} \varphi(x) \implies \inf_{x \in \Omega} (b - \varphi(x)) \leq b - \max_{x \in \Omega} \varphi(x).$$

D'un autre coté on a

$$\varphi(x) \leq \max_{x \in \Omega} \varphi(x) \implies b - \varphi(x) \geq b - \max_{x \in \Omega} \varphi(x), \forall x \in \Omega.$$

Alors comme $b - \max_{x \in \Omega} \varphi(x)$ est constant on obtient donc

$$\inf_{x \in \Omega} (b - \varphi(x)) \geq b - \max_{x \in \Omega} \varphi(x).$$

On aura,

$$\inf_{x \in \Omega} (b - \varphi(x)) = b - \max_{x \in \Omega} \varphi(x).$$

■

Preuve. (Théorème 4.1.1)

Pour tout $x \in \Omega$ on posons $\Delta\varphi_i = (\alpha_i x_i^* + \beta_i x_i^{*2}) - (\alpha_i x_i + \beta_i x_i^2)$. Alors

$$\begin{aligned} \Delta\varphi_i &= \left(\alpha_i \frac{-\alpha_i}{2\beta_i} + \beta_i \left(\frac{-\alpha_i}{2\beta_i} \right)^2 \right) - \alpha_i x_i - \beta_i x_i^2 = \frac{-\alpha_i^2}{2\beta_i} + \frac{\alpha_i^2}{4\beta_i} - \alpha_i x_i - \beta_i x_i^2 \\ &= \frac{-2\alpha_i^2 + \alpha_i^2}{4\beta_i} - \alpha_i x_i - \beta_i x_i^2 = \frac{-\alpha_i^2}{4\beta_i} - \alpha_i x_i - \beta_i x_i^2 = -\beta_i \left(\left(\frac{\alpha_i}{2\beta_i} \right)^2 + \frac{\alpha_i}{\beta_i} x_i + x_i^2 \right) \\ &= -\beta_i \left(x_i + \frac{\alpha_i}{2\beta_i} \right)^2 = -\beta_i (x_i - x_i^*)^2. \end{aligned}$$

Mais

$$\varphi(x^*) - \varphi(x) = \sum_{i=1}^n \Delta\varphi_i.$$

Alors

$$\varphi(x^*) - \varphi(x) = \sum_{i=1}^n -\beta_i (x_i - x_i^*)^2 \text{ pour tout } x \in \Omega.$$

cela donne ,

$$\inf_{x \in \Omega} (\varphi(x^*) - \varphi(x)) = \inf_{x \in \Omega} \left(\sum_{i=1}^n -\beta_i (x_i - x_i^*)^2 \right)$$

peut être écrite sous la forme suivante :

$$\varphi(x^*) - \max_{x \in \Omega} \varphi(x) = \inf_{x \in \Omega} \left(\sum_{i=1}^n (\sqrt{-\beta_i} (x_i - x_i^*))^2 \right) = \inf_{y \in \Omega'} \left(\sum_{i=1}^n (y_i - y_i^*)^2 \right)$$

avec

$$\Omega' = \left\{ y = (y_i)_i \in \mathbb{R}^n : y_i = \sqrt{-\beta_i} x_i / x = (x_i)_i \in \Omega \right\}.$$

Parce que

$$\inf_{y \in \Omega'} \left(\sum_{i=1}^n (y_i^* - y_i)^2 \right) = \|y^* - y_0\|^2 : y_0 \in \Omega'$$

(car Ω' fermé et borné) donc : $\max_{x \in \Omega} \varphi(x) = \varphi(x^*) - \|y^* - y_0\|^2$ alors on obtiens (3.1) .

Puisque $\inf_{y \in \Omega'} \|y^* - y\|^2 = \|y^* - y_0\|^2$, alors $\|y^* - y_0\|^2 \leq \|y^* - y\|^2$ pour tout $y \in \Omega'$,

ceci implique que $\|y^* - y_0\| \leq \|y^* - y\|$ pour tout $y \in \Omega'$.

Nous avons par conséquent, $\|y^* - y_0\| \leq \inf_{y \in \Omega'} \|y^* - y\|$, mais $y_0 \in \Omega'$, alors on a :

$$\|y^* - y_0\| \geq \inf_{y \in \Omega'} \|y^* - y\| \text{ donc } \|y^* - y_0\| = \inf_{y \in \Omega'} \|y^* - y\|$$

Donc on obtient (3.2).

On a le vecteur y_0 est la projection du vecteur y^* sur le nouveau convexe Ω'

on a :

$$\begin{aligned} \max_{x \in \Omega} \varphi(x) &= \varphi(\bar{x}) = \varphi(x^*) - \|y^* - y_0\|^2 \\ &= \varphi(x^*) - \inf_{x \in \Omega} \left(\sum_{i=1}^n (y_i^* - \sqrt{-\beta_i} x_i)^2 \right) \\ &= \varphi(x^*) - \sum_{i=1}^n (y_i^* - \sqrt{-\beta_i} \bar{x}_i)^2. \end{aligned}$$

Alors

$$\sum_{i=1}^n (y_i^* - \sqrt{-\beta_i} \bar{x}_i)^2 = \sum_{i=1}^n (y_i^* - y_{i0})^2 \implies \bar{x}_i = \frac{y_{i0}}{\sqrt{-\beta_i}}.$$

Donc on obtient (3.3) ■

Remarque 3.2.1. Si $x^* \in \Omega$ on a $\inf_{x \in \Omega} \sum_{i=1}^n (-\beta_i (x_i^* - x_i)^2) = 0$. Alors $\max_{x \in \Omega} \varphi(x) = \varphi(x^*)$.

Remarque 3.2.2. Soit Ω un ensemble, et soit $T : \Omega \subset \mathbb{R}^n \longrightarrow \Omega' \subset \mathbb{R}^n$ l'application qui transforme le convexe Ω en un convexe Ω' .

$T(x) = vx$ tel que $v = (\sqrt{-\beta_1}, \sqrt{-\beta_2}, \dots, \sqrt{-\beta_n})$, on a $\beta_i < 0$ pour tout i . Alors la matrice de cette transformation :

$$\begin{pmatrix} \sqrt{-\beta_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{-\beta_n} \end{pmatrix}$$

Son déterminant est $\prod_{i=1}^n \sqrt{-\beta_i} \neq 0$. Donc elle est conforme.

Remarque 3.2.3. On a Ω est convexe borné, on peut se restreindre au cas $\alpha_i \geq 0$ pour tout i .

En effet, supposons que nous avons $\alpha_i x_i + \beta_i x_i^2$ avec $\alpha_i < 0$

on pose : $\delta^* = \max(x_i)$, $y_i = \delta^* - x_i \geq 0 \implies x_i = \delta^* - y_i$ par conséquent

$$\alpha_i x_i + \beta_i x_i^2 = \alpha_i (\delta^* - y_i) + \beta_i (\delta^* - y_i)^2$$

$$\begin{aligned}
&= \alpha_i \delta^* - \alpha_i y_i + \beta_i \delta^{*2} - 2\beta_i \delta^* y_i + \beta_i y_i^2 \\
&= \beta_i y_i^2 + (-\alpha_i - 2\beta_i \delta^*) y_i + \alpha_i \delta^* + \beta_i \delta^{*2}
\end{aligned}$$

on pose, $\alpha'_i = -\alpha_i - 2\beta_i \delta^* \geq 0$ et $k_i = \alpha_i \delta^* + \beta_i \delta^{*2}$. Alors on obtient

$$\alpha_i x_i + \beta_i x_i^2 = \alpha'_i y_i + \beta_i y_i^2 + k_i, \text{ où } \alpha'_i \geq 0 \text{ et } \beta_i \leq 0,$$

par conséquent il suffit de remplacer: x_i par $\delta^* - y_i$ et $\alpha_i x_i + \beta_i x_i^2$ par $\alpha'_i y_i + \beta_i y_i^2 + k_i$.

Exemple 3.2.1.

$$\begin{cases} x_1, x_2 \geq 0 \\ 3x_1 + x_2 \leq 9 \\ x_1 + 2x_2 \leq 8 \\ \max \varphi(x_1, x_2), \text{ avec : } \varphi(x_1, x_2) = 5x_1 + 5x_2 - x_1^2 - x_2^2. \end{cases}$$

On remarque que $\beta_1 = \beta_2 = -1$ alors on a $y_i = \sqrt{-\beta_i} x_i \implies y_i = x_i$, $\Omega = \Omega'$ et on a

$$\begin{aligned}
\frac{\partial \varphi}{\partial x_1}(x_1, x_2) &= 5 - 2x_1, \text{ alors } \frac{\partial \varphi}{\partial x_1}(x_1, x_2) = 0 \implies 5 - 2x_1^* = 0 \implies x_1^* = \frac{5}{2} \\
\frac{\partial \varphi}{\partial x_2}(x_1, x_2) &= 5 - 2x_2, \text{ alors } \frac{\partial \varphi}{\partial x_2}(x_1, x_2) = 0 \implies 5 - 2x_2^* = 0 \implies x_2^* = \frac{5}{2}.
\end{aligned}$$

Alors le point critique $x^* = (\frac{5}{2}, \frac{5}{2}) \notin \Omega$ car $3\frac{5}{2} + \frac{5}{2} = 10 > 9$ donc le projection de x^* est \bar{x} tel que :

$$a(\bar{x} - x^*) = a\bar{x} - ax^* = b - \langle a, x^* \rangle$$

car \bar{x} est un point frontière de Ω

$$\bar{x} = x^* - \frac{\langle x^*, a \rangle - b}{\|a\|^2} a$$

et on a $a = (3, 1), b = 9$, alors

$$\langle x^*, a \rangle = \frac{5}{2} \cdot 3 + \frac{5}{2} = 10, \|a\|^2 = 10.$$

Donc

$$\bar{x} = \left(\frac{5}{2}, \frac{5}{2}\right) - \frac{1}{10}(3, 1) = \left(\frac{22}{10}, \frac{24}{10}\right) \implies \max_{(x_1, x_2) \in \Omega} \varphi(x_1, x_2) = \varphi(\bar{x}) = 12.40.$$

Ici

$$\Omega = \left\{ (x_1, x_2) \in \mathbb{R}^2 : \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 9 \\ 8 \end{pmatrix} \text{ et } x_1 \geq 0, x_2 \geq 0 \right\}.$$

Exemple 3.2.2.

$$\begin{cases} x_1, x_2 \geq 0 \\ 3x_1 + 2x_2 \leq 6 \\ \max \varphi(x_1, x_2), \text{ avec : } \varphi(x_1, x_2) = 5x_1 + 8x_2 - x_1^2 - 2x_2^2. \end{cases}$$

Dans ce cas $\beta_1 = -1$ et $\beta_2 = -2$ alors on remarque que $\Omega \neq \Omega'$ et on a

$$\frac{\partial \varphi}{\partial x_1}(x_1, x_2) = 5 - 2x_1, \quad \frac{\partial \varphi}{\partial x_2}(x_1, x_2) = 8 - 4x_2.$$

Alors on a

$$\frac{\partial \varphi}{\partial x_1}(x_1, x_2) = 0 \implies 5 - 2x_1 = 0 \implies x_1 = \frac{5}{2}$$

et

$$\frac{\partial \varphi}{\partial x_2}(x_1, x_2) = 0 \implies 8 - 4x_2 = 0 \implies x_2 = 2.$$

On obtient le point critique $x^* = (2.5, 2) \notin \Omega$ car $3 \times \frac{5}{2} + 2 \times 2 = 7.5 + 4 = 11.5 > 6$, alors on a

$$x_0 = P_{\Omega}(x^*) = x^* - \frac{\langle x^*, a \rangle - b}{\|a\|^2} a \text{ avec } a = (3, 2) \text{ et } b = 6$$

$\langle x^*, a \rangle = \frac{5}{2} \times 3 + 2 \times 2 = \frac{23}{2}$, $\|a\|^2 = 13$, on obtient donc.

$$x_0 = \left(\frac{5}{2}, 2\right) - \frac{\frac{23}{2} - 6}{13} (3, 2) = \left(\frac{5}{2}, 2\right) - \frac{11}{26} (3, 2) = \left(\frac{16}{13}, \frac{15}{13}\right)$$

et on a $y_1 = x_1$ et $y_2 = \sqrt{2}x_2$, alors

$$\Omega' = \left\{ (y_1, y_2) \in \mathbb{R}^2 / 3y_1 + \frac{2}{\sqrt{2}}y_2 \leq 6, y_1 \geq 0, y_2 \geq 0 \right\}.$$

De plus $y^* = \left(\frac{5}{2}, 2\sqrt{2}\right)$, alors la projection de ce point sur Ω' donne le point suivant y_0 :

$$y_0 = y^* - \frac{\langle y^*, a' \rangle - b}{\|a'\|^2} a' \text{ tel que } a' = (3, \sqrt{2}) \text{ et } b = 6.$$

Alors

$$\begin{aligned} y_0 &= \left(\frac{5}{2}, 2\sqrt{2}\right) - \frac{3 \times \frac{5}{2} + 2\sqrt{2} \times \sqrt{2} - 6}{11} (3, \sqrt{2}) = \left(\frac{5}{2}, 2\sqrt{2}\right) - \frac{\frac{15}{2} + 4 - 6}{11} (3, \sqrt{2}) \\ &= \left(\frac{5}{2}, 2\sqrt{2}\right) - \frac{11}{22} (3, \sqrt{2}) = \left(\frac{5}{2}, 2\sqrt{2}\right) - \left(\frac{3}{2}, \frac{\sqrt{2}}{2}\right) \end{aligned}$$

$$= \left(\frac{5-3}{2}, 2\sqrt{2} - \frac{\sqrt{2}}{2} \right) = \left(1, \frac{3}{\sqrt{2}} \right).$$

Alors $y_0 = \left(1, \frac{3}{\sqrt{2}} \right) \implies \bar{x} = \left(1, \frac{3}{2} \right)$ (car $x_1 = y_{1,0}$ et $x_2 = \frac{y_{2,0}}{\sqrt{2}}$)

et $\max_{x \in \Omega} \varphi(x) = \varphi(\bar{x}) = \varphi\left(1, \frac{3}{2}\right) = 11,50$.

Notons que la projection du point x^* sur Ω donne le point $\left(\frac{16}{13}, \frac{15}{13}\right)$ et $\varphi\left(\frac{16}{13}, \frac{15}{13}\right) = 11,207$.

Évidemment $\max_{x \in \Omega} \varphi(x) \leq \varphi(x^*) = \varphi\left(\frac{5}{2}, 2\right) = 14,25$.

Alors on remarque que la projection du point x^* est le plus proche point du point critique x^* , mais n'est pas la solution optimale de φ .

Exemple 3.2.3.

$$\begin{cases} x_1, x_2, x_3 \geq 0 \\ x_1 + 3x_2 \leq 18 \\ x_1 + x_2 + x_3 \leq 8 \\ 2x_1 + x_2 + 2x_3 \leq 14 \\ \max f(x_1, x_2, x_3), \text{ avec : } f(x_1, x_2, x_3) = 2x_1 + 3x_2 + 3x_3 + x_1^2 + 2x_2^2 - x_3^2. \end{cases}$$

On pose $f(x_1, x_2, x_3) = \psi(x_1, x_2) + \varphi(x_3)$ tel que

$$\varphi(x_3) = 3x_3 - x_3^2; \quad \psi(x_1, x_2) = 2x_1 + 3x_2 + x_1^2 + 2x_2^2.$$

On a $\varphi(x_3) = 3x_3 - x_3^2$ alors le point critique est $x^* = \left(\frac{3}{2}\right)$ et $\max_{x_3 \in \Omega} \varphi(x_3) = \varphi\left(\frac{3}{2}\right) = \frac{9}{4}$.

• ψ est une fonction convexe, et sa solution optimale est un point extrême (sommet de Ω).

$$\max_{(x_1, x_2) \in \Omega} \psi(x_1, x_2) = \psi(0, 6) = 90.$$

Mais

$$\max_{x \in \Omega} f(x) \leq \max_{x \in \Omega} \psi(x) + \max_{x \in \Omega} \varphi(x).$$

Par conséquent $\max_{x \in \Omega} f(x) \leq 92,25$.

Mais on a le point, $(0, 6, \frac{3}{2}) \in \Omega$. Alors $f\left((0, 6, \frac{3}{2})\right) \leq \max_{x \in \Omega} f(x) \implies 92,25 \leq \max_{x \in \Omega} f(x)$.

Alors on obtient $\max_{x \in \Omega} f(x) = 92,25$.

Remarque 3.2.4. Si nous posons $\psi(x) = \sum_{\beta_i > 0} (\alpha_i x_i + \beta_i x_i^2)$, alors

$$\psi(x^*) - \psi(x) = \sum_{i=1}^n [\alpha_i x_i^* + \beta_i (x_i^*)^2 - \alpha_i x_i - \beta_i x_i^2] = \sum_{i=1}^n -\beta_i \left[\left(\frac{-\alpha_i}{\beta_i} \right) x_i^* - (x_i^*)^2 + \frac{\alpha_i}{\beta_i} x_i + x_i^2 \right]$$

$$\begin{aligned}
&= \sum_{i=1}^n -\beta_i \left[2 \left(\frac{-\alpha_i}{2\beta_i} \right) x_i^* - (x_i^*)^2 - 2 \left(\frac{-\alpha_i}{2\beta_i} \right) x_i + x_i^2 \right] = \sum_{i=1}^n -\beta_i (2(x_i^*)^2 - (x_i^*)^2 - 2x_i^* x_i + x_i^2) \\
&= \sum_{i=1}^n -\beta_i (x_i^* - x_i)^2 \text{ pour tout } x \in \Omega \text{ et}
\end{aligned}$$

$$\inf_{x \in \Omega} (\psi(x^*) - \psi(x)) = \inf_{x \in \Omega} \left(- \sum_{i=1}^n \beta_i (x_i^* - x_i)^2 \right) \implies \psi(x^*) - \max_{x \in \Omega} \psi(x) = - \max_{x \in \Omega} \left(\sum_{i=1}^n \beta_i (x_i^* - x_i)^2 \right),$$

donc

$$\begin{aligned}
\max_{x \in \Omega} \psi(x) &= \psi(x^*) + \max_{x \in \Omega} \left(\sum_{i=1}^n \beta_i (x_i^* - x_i)^2 \right) \\
&= \psi(x^*) + \max_{x \in \Omega} \left(\sum_{i=1}^n \left(\sqrt{\beta_i} x_i^* - \sqrt{\beta_i} x_i \right)^2 \right).
\end{aligned}$$

Soit $y^* = (y_i^*)_i = \left(\frac{\alpha_i}{2\sqrt{\beta_i}} \right)_i$ et $y = (y_i)_i = (\sqrt{\beta_i} x_i)_i$

Notons que y_0 est le point le plus éloigné du convexe

$$\Omega' = \left\{ y \in \mathbb{R}^n \mid y_i = \sqrt{\beta_i} x_i, x = (x_i)_i \in \Omega \right\}.$$

Chapitre 4

Programmation linéaire et non-linéaire

4.1 Dual d'un programme linéaire

4.1.1 Problème dual d'un problème de programmation linéaire (P . L)

Considérons le problème de (P L) suivant

$$(P) \begin{cases} x \geq 0 \\ Ax = b \\ \max Z = C^T .x \end{cases}$$

Soit x^0 une solution optimale de base non dégénérée du problème (P) et A_B sa matrice de base associée .

- Cette solution correspond au vecteur des multiplicateurs suivants,

$$\pi^T = C_B^T A_B^{-1} \quad (4.1)$$

- De l'optimalité de x^0 , le vecteur

$$\Delta_N = A_N^T \pi - C_N \geq 0 \quad (4.2)$$

car

$x^0 = (x_B^0, x_N^0)$ avec $x_N^0 = 0$ alors on a $A_B x_B^0 + A_N x_N^0 = b$ ($A_N x_N^0 = 0$), donc $x_B^0 = A_B^{-1} b$ et $Z(x^0) = C^T x^0 = C_B^T x_B^0 + C_N^T x_N^0 = C_B^T A_B^{-1} b$.

Comme x^0 solution optimale alors $Z(x^0) - Z(x) \geq 0 : \forall x \geq 0$

$Z(x^0) - Z(x) = C_B^T A_B^{-1} b - (C_B^T x_B + C_N^T x_N)$, mais $x_B = A_B^{-1} b - A_B^{-1} A_N x_N$, alors
 $Z(x^0) - Z(x) = C_B^T A_B^{-1} b - (C_B^T A_B^{-1} b + (C_N^T - C_B^T A_B^{-1} A_N) x_N) = (C_B^T A_B^{-1} A_N - C_N^T) x_N$.
 Mais on a $Z(x^0) - Z(x) \geq 0$ et $x_N \geq 0$, donc

$$C_B^T A_B^{-1} A_N - C_N^T \geq 0 \Rightarrow \pi^T A_N - C_N^T \geq 0 \Rightarrow A_N^T \pi - C_N \geq 0.$$

Donc (4.2) vérifie .

D'ou le vecteur π verifie l'ingalité

$$A^T \pi \geq C \quad (4.3)$$

(car $A_B^T \pi - C_B = 0$) on a

$$\pi^T A_B - C_B^T = C_B^T A_B^{-1} A_B - C_B^T = C_B^T - C_B^T = 0.$$

• Soit y un vecteur vérifiant , $A^T y \geq C$ ou $y^T A \geq C^T$, comme $x^0 \geq 0$, en multipliant le dernière inégalité par x^0 on aura , $y^T A x^0 \geq C^T x^0 \Rightarrow y^T b \geq C^T x^0 \Rightarrow b^T y \geq C^T x^0, \forall y \in \mathbb{R}^m$, vérifiant

$$A^T y \geq C \quad (4.4)$$

De (4.3) et (4.4) on déduit que le vecteur π est une solution optimale du problém de

programation linéaire (PL) suivant : $(P^*) \begin{cases} y \in \mathbb{R}^n \\ A^T y \geq C \\ \min w = b^T y . \end{cases}$

Car , si $y^0 = \pi$ alors $b^T y^0 = b^T \pi = C_B^T A_B^{-1} b$, mais

$$\begin{aligned} A_B^{-1} . b = x_B^0 \Rightarrow b^T y^0 &= C_B^T x_B^0 \\ \Rightarrow b^T y^0 &= C^T x^0 \quad (C_N^T x_N^0 = 0) , \end{aligned}$$

donc $\min w = \max z$.

On déduit que le problemé (P) et (P*) sont équivalent.

• Le problème (P) est dit problème primal et le problème (P*) est appelé dual du problème (P). [17]

Définitions

a) Cas d'un (P.L) canonique de type max

Définition 4.1.1. On appelle programme linéaire dual du (P.L) canonique de type max (P)

$$\begin{cases} x \geq 0 \\ Ax \leq b \\ C^T x = z \quad (\max z) \end{cases} , \text{ le (P.L) canonique de type } \min(P^*) \begin{cases} y \geq 0 \\ A^T y \geq C \\ b^T y = w \quad (\min w) \end{cases} .$$

$$\text{Exemple 4.1.1. } (P) : \begin{cases} x \geq 0 \\ \begin{pmatrix} 3 & 2 & -1 \\ 5 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} 2 \\ 3 \end{pmatrix} \\ z = 2x_1 - x_2 + x_3 \text{ (max } z) \end{cases}$$

$$\text{a pour dual } (P^*) : \begin{cases} y \geq 0 \\ \begin{pmatrix} 3 & 5 \\ 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \geq \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \\ w = 2y_1 + 3y_2 \text{ (min } w) \end{cases} .$$

$$\text{C'est -a-dire, en " développé " } (P^*) : \begin{cases} y \geq 0 \\ 3y_1 + 5y_2 \geq 2 \\ 2y_1 - y_2 \geq -1 \\ -y_1 + y_2 \geq 1 \\ w = 2y_1 + 3y_2 \text{ (min } w) \end{cases} .$$

Remarque 4.1.1. Nous aurions écrit (P^*) sous la forme équivalente suivante, obtenue par

$$\text{simple transposition de l'écriture } A^T y \geq C : (P^*) : \begin{cases} y \geq 0 \\ y^T A \geq C^T \\ y^T b = w \text{ (min } w) \end{cases}$$

$$\text{ainsi } (P^*) : \begin{cases} y \geq 0 \\ \begin{pmatrix} y_1 & , y_2 \end{pmatrix} \begin{pmatrix} 3 & 2 & -1 \\ 5 & -1 & 1 \end{pmatrix} \geq \begin{pmatrix} 2, & -1, & 1 \end{pmatrix} \\ \begin{pmatrix} y_1 & , y_2 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = w \text{ (min } w) \end{cases}$$

b) Cas d'un (P.L) canonique de type min

$$(P) : \begin{cases} x \geq 0 \\ Ax \geq b \\ C^T x = z \text{ (min } z) \end{cases} \Leftrightarrow (P) : \begin{cases} x \geq 0 \\ (-A)x \leq -b \\ -C^T x = z' \text{ (max } z') \text{ (où } z' = -z) \end{cases}$$

sous cette forme , la définition 5.1.1 ,s'applique á (P) et l'on a

$$(P^*) : \begin{cases} y \geq 0 \\ (-A)^T y \geq -C \\ (-b)^T y = w' \text{ (min } w') \end{cases} \Leftrightarrow (P^*) : \begin{cases} y \geq 0 \\ A^T y \leq C \\ b^T y = w \text{ (max } w) \text{ (où } w = -w') \end{cases}$$

C'est cette dernière forme qu'on appellera dual de (P).

Définition 4.1.2. On appelle programmation linéaire dual du (P.L) canonique de type min,

$$(P) : \begin{cases} x \geq 0 \\ Ax \geq b \\ C^T x = z \text{ (min } z) \end{cases} \text{ le (P.L) canonique de type max, } (P^*) : \begin{cases} y \geq 0 \\ A^T y \leq C \\ b^T y = w \text{ (max } w) . \end{cases}$$

c) Le cas général

Généralisons maintenant le travail précédent nous avons introduit pour tout (P.L) de type max (respectivement min) de manière Systématique , un (P.L) canonique de type max (respectivement min) équivalent.

Définition 4.1.3. On appelle programme linéaire dual d'un programme quelconque (P) le dual (P*) de son programme canonique équivalent .

C'est en fait une forme équivalente de (P*), que nous considérons comme dual de (P).

Commençant par préciser cela sur quelques exemples.

Exemple 1 (Un type max avec contraintes supérieur et égale)

$$(P) : \begin{cases} x \geq 0 \\ x_1 - x_2 + x_3 \geq 6 \\ x_1 + x_2 + x_3 \leq 2 \\ 2x_1 + 2x_2 + x_3 \leq 4 \\ \dots\dots\dots \\ 4x_1 + 2x_2 + 2x_3 = z \text{ (max } z) \end{cases} \Leftrightarrow \text{(ligne1) } (P) : \begin{cases} x \geq 0 \\ -x_1 + x_2 - x_3 \leq -6 \\ x_1 + x_2 + x_3 \leq 2 \\ 2x_1 + 2x_2 + x_3 \leq 4 \\ \dots\dots\dots \\ 4x_1 + 2x_2 + 2x_3 = z \text{ (max } z) . \end{cases}$$

Alors le problème dual de (P).

$$(P^*) : \begin{cases} {}^t y' = (y'_1; y_2; y_3) \geq 0 \\ -y'_1 + y_2 + 2y_3 \geq 4 \\ y'_1 + y_2 + 2y_3 \geq 2 \\ -y'_1 + y_2 + y_3 \geq 2 \\ \dots\dots\dots \\ -6y'_1 + 2y_2 + 4y_3 = w \text{ (min } w) \end{cases}$$

on posant $y_1 = -y'_1$ alors on obtient

$$(P^*) : \begin{cases} y_1 \leq 0, y_2 \geq 0, y_3 \geq 0 \\ y_1 + y_2 + 2y_3 \geq 4 \\ -y_1 + y_2 + 2y_3 \geq 2 \\ y_1 + y_2 + y_3 \geq 2 \\ \dots\dots\dots \\ 6y_1 + 2y_2 + 4y_3 = w \text{ (min } w) . \end{cases}$$

Ainsi , après calculs , la présence de \geq dans la (ligne 1) ne se traduit dans (P^*) que par une condition $y_1 \leq 0$. Tout se passe comme si , considérant (P_1) identique à (P) , sauf $x_1 - x_2 + x_3 \leq 6$, on formait P_1^* et on changeait $y_1 \geq 0$ en $y_1 \leq 0$.

Il ne faut toutefois . Changes les signes dans la première ligne de A et b_1 , équivaut à changer le signe de la première variable y_1 de y .

Exemple 2 (un type max avec constraints égale)

$$(P) : \begin{cases} x \geq 0 \\ x_1 - x_2 \leq 2 \\ 5x_1 + 2x_2 = 4 \\ \dots\dots\dots \\ 2x_1 + x_2 = z \text{ (max } z) \end{cases} \iff (P) : \begin{cases} x \geq 0 \\ x_1 - x_2 \leq 2 \\ 5x_1 + 2x_2 \leq 4 \\ -5x_1 - 2x_2 \leq -4 \\ \dots\dots\dots \\ 2x_1 + x_2 = z \text{ (max } z) . \end{cases}$$

Alors on a le problème dual de (P)

$$(P^*) : \begin{cases} {}^t y' = (y_1, y_2', y_3') \geq 0 \\ y_1 + 5y_2' - 5y_3' \geq 2 \\ -y_1 + 2y_2' - 2y_3' \geq 1 \\ \dots\dots\dots \\ 2y_1 + 4y_2' - 4y_3' = w \text{ (min } w) \end{cases} .$$

Mais il est aisé de constater que y_2' et y_3' n'interviennent que sous la forme "groupée" $y = y_2' - y_3' \in \mathbb{R}$, donc on obtient

$$(P^*) : \begin{cases} y_1 \geq 0, y_2 \in \mathbb{R} \\ y_1 + 5y_2 \geq 2 \\ -y_1 + 2y_2 \geq 1 \\ \dots\dots\dots \\ 2y_1 + 4y_2 = w \text{ (min } w) \end{cases} .$$

Nous notons y_2 pour signifier que cette variable est non astreinte.

Dans ce cas encore, par rapport à un programme (P_1) qui aurait comporté comme ligne(2), la seule modification de (P^*) par rapport à (P_1^*) consiste à remplacer $y_2 \geq 0$ par $y_2 \in \mathbb{R}$.

Exemple 3 (Un peu de tout ce qui précède . . . et une variable non astreinte).

Considérons d'abord (P_1) :

$$(P_1) : \begin{cases} x \geq 0 \\ x_1 + x_2 + x_3 \leq 9 \\ 2x_1 + 3x_2 \geq 3 \\ x_1 - x_2 - x_3 = 1 \\ \dots\dots\dots \\ x_1 - 2x_2 + x_3 = z \text{ (max } z) \end{cases} .$$

Des exemples antérieurs certainement que les variables duales vérifieront

$y_1 \geq 0, y_2 \leq 0, y_3 \in \mathbb{R}$, si bien que l'on obtiendra :

$$(P^*) : \begin{cases} y_1 \geq 0, y_2 \leq 0, y_3 \in \mathbb{R} \\ y_1 + 2y_2 + y_3 \geq 1 \\ y_1 + 3y_2 - y_3 \geq -2 \\ y_1 - y_3 \geq 1 \\ \dots\dots\dots \\ 9y_1 + 3y_2 + y_3 = w \text{ (min } w\text{)}. \end{cases}$$

Mais que devient - il si nous modifions (P_1) en (P) par $x_3 \in \mathbb{R}$ (non astreinte) au lieu $x_3 \geq 0$. On pose alors $x_3 = x'_3 - x''_3$ avec $x'_3 \geq 0, x''_3 \geq 0$ et on écrit :

$$(P) : \begin{cases} x = (x_1, x_2, x'_3, x''_3) \geq 0 \\ x_1 + x_2 + x'_3 - x''_3 \leq 9 \\ -2x_1 - 3x_2 \leq -3 \\ x_1 - x_2 - x'_3 + x''_3 \leq 1 \\ -x_1 + x_2 + x'_3 - x''_3 \leq -1 \\ \dots\dots\dots \\ x_1 - 2x_2 + x'_3 - x''_3 = z \text{ (max } z\text{)}. \end{cases}$$

D'où

$$(P^*) : \begin{cases} {}^t y' = (y'_1, y'_2, y'_3, y'_4) \geq 0 \\ y'_1 - 2y'_2 + y'_3 - y'_4 \geq 1 \\ y'_1 - 3y'_2 - y'_3 + y'_4 \geq -2 \\ y'_1 - y'_3 + y'_4 \geq 1 \\ -y'_1 + y'_3 - y'_4 \geq -1 \\ \dots\dots\dots \\ 9y'_1 - 3y'_2 + y'_3 - y'_4 = w \text{ (min } w\text{)} . \end{cases}$$

Soit , on posant $y_1 = y'_1 (\geq 0)$, $y_2 = -y'_2 (\leq 0)$ et $y_3 = y'_3 - y'_4 \in \mathbb{R}$

$$(P^*) : \begin{cases} y_1 \geq 0, y_2 \leq 0, y_3 \in \mathbb{R} \\ y_1 + 2y_2 + y_3 \geq 1 \\ y_1 + 3y_2 - y_3 \geq -2 \\ y_1 - y_3 \geq 1 \\ -y_1 + y_3 \geq -1 \\ \dots\dots\dots \\ 9y_1 + 3y_2 + y_3 = w \text{ (min } w) \end{cases} \iff (P^*) : \begin{cases} y_1 \geq 0, y_2 \leq 0, y_3 \in \mathbb{R} \\ y_1 + 2y_2 + y_3 \geq 1 \\ y_1 + 3y_2 - y_3 \geq -2 \\ y_1 - y_3 = 1 \dots\dots\dots (4) \\ \dots\dots\dots \\ 9y_1 + 3y_2 + y_3 = w \text{ (min } w) \end{cases} .$$

Ainsi (P_1^*) ne diffère-t-il de (P^*) que par égale au lieu de supérieur et égale dans la troisième contrainte! Le rôle de " x_3 non astreinte " par rapport à $x_3 \geq 0$ peut être " suivi à la trace " dans les lignes (4) dans (P^*) et colons (4) dans (P) .

4.1.2 Construction du dual

Primal (P)	Dual(P*)
$max z = c^T x$	$min w = b^T y$
$A_i x = b_i$	$y_i \in \mathbb{R}$
$A_i x \leq b_i$	$y_i \geq 0$
$A_i x \geq b_i$	$y_i \leq 0$
$x_j \geq 0$	$A_j^T y \geq c_j$
$x_j \leq 0$	$A_j^T y \leq c_j$
$x_j \in R$	$A_j^T y = c_j$

Proposition 4.1.1. (*Dual d'un programme standard*)

$$(P) : \begin{cases} x \geq 0 \\ Ax = b \\ \dots\dots\dots \\ c^T x = z \text{ (max } z) \end{cases} \quad a \text{ pour dual } (P^*) : \begin{cases} y \text{ non astriente} \\ y^T A \geq c^T \\ \dots\dots\dots \\ y^T b = w \text{ (min } w) \end{cases} .$$

Preuve.

$$(P) : \begin{cases} x \geq 0 \\ Ax \leq b \\ Ax \geq b \\ \dots\dots\dots \\ c^T x = z \text{ (max } z) \end{cases} \iff (P) : \begin{cases} x \geq 0 \\ Ax \leq b \\ -Ax \leq -b \\ \dots\dots\dots \\ c^T x = z \text{ (max } z) \end{cases} .$$

Alors on a

$$(P) : \begin{cases} x \geq 0 \\ \begin{pmatrix} A \\ -A \end{pmatrix} . x \leq \begin{pmatrix} b \\ -b \end{pmatrix} \\ \dots\dots\dots \\ c^T x = z \text{ (max } z) \end{cases}$$

avec, si $A \in M_{m;n}(\mathbb{R})$, $b \in M_{m,1}(\mathbb{R})$. Alors on a

$$\begin{pmatrix} A \\ -A \end{pmatrix} \in M_{2m;n}(\mathbb{R}) \text{ et } \begin{pmatrix} b \\ -b \end{pmatrix} \in M_{2m,1}(\mathbb{R}) .$$

Il y a donc $2m$ variables duales, séparons les en deux blocs.

Correspondant aux m premières contraintes (y') et aux m dernières (y''):

$$(P^*) : \begin{cases} y' \geq 0, y'' \geq 0 \\ (y', y'') \begin{pmatrix} A \\ -A \end{pmatrix} \geq c^T \\ \dots\dots\dots \\ (y'^T, y''^T) \begin{pmatrix} b \\ -b \end{pmatrix} = w \text{ (min } w) \end{cases} \iff (P^*) : \begin{cases} y' \geq 0, y'' \geq 0 \\ (y' - y'') A \geq c^T \\ \dots\dots\dots \\ (y'^T - y''^T) b = w \text{ (min } w) \end{cases}$$

et il n'y a plus qu'à poser $y = y' - y''$ par construction aucune variable n'est astreinte .

$$(P^*) : \begin{cases} y \in \mathbb{R} \\ y^T A \geq c^T \\ \dots\dots\dots \\ y^T . b = w \text{ (min } w \text{)}. \end{cases}$$

■

Théorème 4.1.1. Soit (P) un programme linéaire, et (P^*) son dual . Alors $(P^*)^* = (P)$.

• compte tenu de la définition 5.1.3 , il suffit de prouver le résultat pour un problème canonique de type max, (C) . Or (C^*) est du type " fourniture " F .

$$(C) : \begin{cases} x \geq 0 \\ A.x \leq b \\ \dots\dots\dots \\ c^T .x = z \text{ (max } z \text{)} \end{cases} \implies (C^*) = (F) : \begin{cases} y \geq 0 \\ y^T .A \geq c^T \\ \dots\dots\dots \\ y^T .b = w \text{ (min } w \text{)} \end{cases}$$

$$\implies (C^*)^* = (F^*) : \begin{cases} x \geq 0 \\ (A^T)^T .x \leq (b^T)^T \\ \dots\dots\dots \\ c^T .x = z \text{ (max } z \text{)} \end{cases} \iff (C^*)^* : \begin{cases} x \geq 0 \\ A.x \leq b \\ \dots\dots\dots \\ c^T .x = z \text{ (max } z \text{)} \end{cases} .$$

Ce résultat est la justification du mot dualité : (P) et (P^*) jouent des rôles symétriques, chacun étant le dual de l'autre.

Théorèmes de la dualité.

Étant donné un problème de (P.L) primal .

$$(P) : \begin{cases} x \geq 0 \\ A.x \leq b \\ \dots\dots\dots \\ c^T .x = z \text{ (max } z \text{)} \end{cases} \quad \text{et son dual : } (P^*) : \begin{cases} y \geq 0 \\ y^T .A \geq c^T \\ \dots\dots\dots \\ y^T .b = w \text{ (min } w \text{)} \end{cases} .$$

Théorème 4.1.2. (Théorème faible de le dualité). [5]

Si x et y sont deux solutions réalisables du problème primal et du problème dual respectivement. Alors $c^T x \leq b^T y$.

Preuve. On a x est solution réalisable pour le problème primal(P) donc $A.x \leq b$ après multiplication de cette équation par $y^T \geq 0$, on obtient $y^T .A.x \leq y^T .b$.

Comme y est un solution réalisable du problème dual (P^*) donc $y^T .A \geq c^T$ et en multipliant cette dernière équation par $x(x \geq 0)$, on obtient $y^T .A.x \geq c^T .x$.

D'on

$$c^T .x \leq y^T .A.x \leq y^T .b \implies c^T .x \leq y^T .b .$$

On notera le rôle essentiel de $x \geq 0$ et $y \geq 0$ dans cette démonstration .

■

Corollaire 4.1.1. Si les problèmes (P) et (P^*) admettent des solutions optimales : \bar{x} et \bar{y} respectivement, elles vérifient $c^T .\bar{x} \leq \bar{y}^T .b$.

Preuve. \bar{x} et \bar{y} sont en particulier réalisable, elles vérifient donc l'inégalité .

■

Exemple 1 . Soit le problème (PL) suivant (P) et le duale de (P)

$$(P) : \begin{cases} x \geq 0 \\ x_1 + x_2 \leq 1 \\ -3x_1 + x_2 \leq -3 \\ \dots\dots\dots \\ 3x_1 + x_2 = z \text{ (max } z) \end{cases} .$$

D'où

$$(P^*) : \begin{cases} y \geq 0 \\ y_1 - 3y_2 \geq 3 \\ y_1 + y_2 \geq 1 \\ \dots\dots\dots \\ y_1 - 3y_2 = w \text{ (min } w) . \end{cases}$$

On a $\Omega = \{A(1,0)\}$ (fig5) et on a $c^T .x = 3x_1 + x_2 \implies x = A \implies c^T .x = 3$.

Alors $\max z = 3$.

et on a les points extrémaux du Ω' , $\{A'(3;0)\}$ (fig6).

Alors on a

$$y^T .b = y_1 - 3y_2 \text{ donc } y = A' \implies y^T .b = 3, \text{ alors } \min w = 3$$

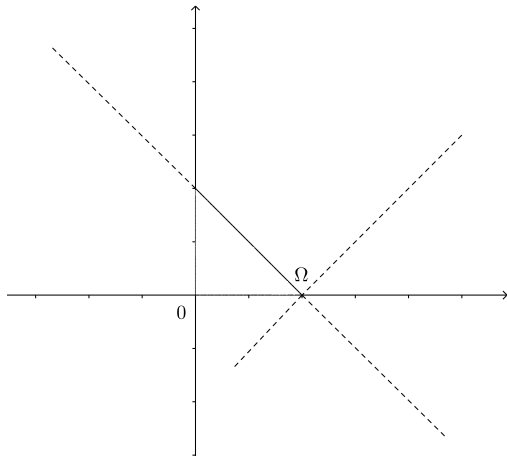


Fig 5

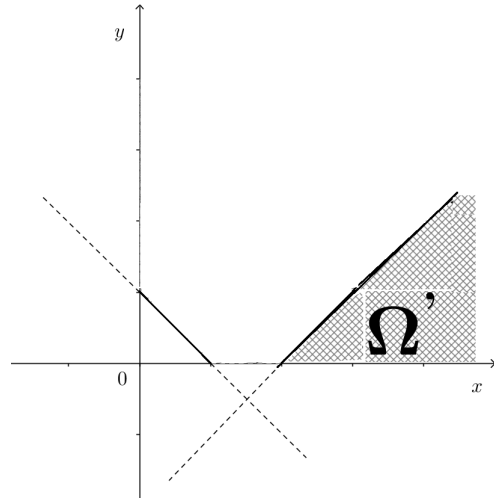


Fig 6

ou on a : $c^T \cdot \bar{x} = (3, 1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3$ et $\bar{y}^T \cdot b = (3, 0) \begin{pmatrix} 1 \\ -3 \end{pmatrix} = 3$, alors $c^T \cdot \bar{x} = \bar{y}^T \cdot b$.

Il y a donc égalité à l'optimum.

Exemple 2.

$$(P) : \begin{cases} x \geq 0 \\ x_1 + 2x_2 + x_3 \leq 5 \\ x_2 + x_3 = 2 \\ \dots\dots\dots \\ x_1 + 4x_2 + \frac{5}{2}x_3 = z \text{ (max } z\text{)}. \end{cases}$$

D'où

$$(P^*) : \begin{cases} y_1 \geq 0, y_2 \in \mathbb{R} \\ y_1 \geq 1 \\ 2y_1 + y_2 \geq 4 \\ y_1 + y_2 \geq \frac{5}{2} \\ \dots\dots\dots \\ 5y_1 + 2y_2 = w \text{ (min } w\text{)}. \end{cases}$$

Si on pose $x_3 = 2 - x_2$, alors comme $x_3 \geq 0$ on a $x_2 \leq 2$ donc on obtient

$$(P) : \begin{cases} x_1, x_2 \geq 0 \\ x_2 \leq 2 \\ x_1 + 2x_2 + 2 - x_2 \leq 5 \\ \dots\dots\dots \\ x_1 + 4x_2 + \frac{5}{2}(2 - x_2) = z \text{ (max } z) \end{cases} \iff (P) : \begin{cases} x_1, x_2 \geq 0 \\ x_2 \leq 2 \\ x_1 + x_2 \leq 3 \\ \dots\dots\dots \\ x_1 + \frac{3}{5}x_2 + 5 = z \text{ (max } z) \end{cases}$$

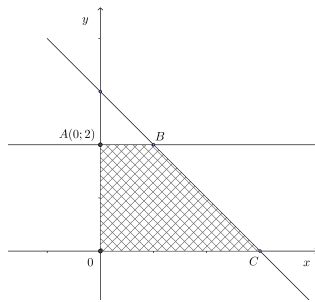


Fig 7

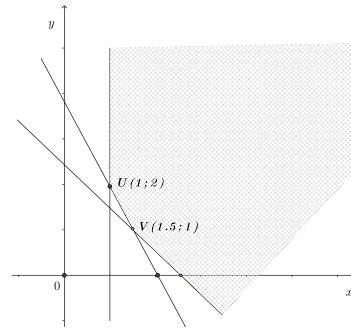


Fig 8

La résolution graphique est aisée. comparons le valeurs de z et w .

Point (x_1, x_2)	$z = c^T \cdot x$
$O(0,0)$	5
$A(0,2)$	8
$B(1,2)$	9
$C(3,0)$	8

Point (y_1, y_2)	$w = y^T \cdot b$
$U(1,2)$	9
$V(1.5,1)$	9.5

On vérifie d'abord, pour tout les couples de sommets, l'inégalité fondamentale ($c^T \cdot x \leq y^T \cdot b$). Mais on observe également que, de nouveau, ($c^T \cdot \bar{x} = \bar{y}^T \cdot b = 9$).

Corollaire 4.1.2. . [6]

Si les problèmes (P) et (P^*) possèdent des solutions réalisables x^0 et y^0 (respectivement) et vérifiant de plus, $c^T \cdot x^0 = y^0 \cdot b$. Alors x^0 est solution optimale de (P) et y^0 solution optimale de (P^*)

Preuve. On a pour tout x réalisable de (P) , et pour tout y réalisable de (P^*) $c^T \cdot x \leq y^T \cdot b$, alors si $y = y^0$ on a

$$c^T \cdot x \leq y^{0T} \cdot b \implies c^T \cdot x \leq c^T \cdot x^0 \text{ (car } c^T \cdot x^0 = y^{0T} \cdot b \text{)}.$$

Donc pour tout x réalisable de (P) : $c^T \cdot x \leq c^T \cdot x^0$, alors x^0 est optimale de (P) . Démonstration identique concernant y^0 .

■

4.2 Introduction à la programmation non linéaire

Considérons le problème de programmation non linéaire :
trouver u tel que

$$(P) : \begin{cases} u \in U = \{v \in \mathbb{R}^n : \varphi_i(v) \leq 0, 1 \leq i \leq m\} \\ J(u) = \inf_{v \in U} J(v) \end{cases},$$

où J et φ_i , $1 \leq i \leq m$ des fonctions définies sur un ensemble ouvert de \mathbb{R}^n , dans \mathbb{R} .

Et considérons la fonction L tel que : $L(v; \mu) = J(v) + \sum_{i=1}^m \mu_i \varphi_i(v)$ pour $v \in \mathbb{R}^n$ et $\mu \in \mathbb{R}_+^m$. Appelée le Lagrangien associé au problème (P) , et on montre que, si u désigne la solution du problème (P) , il existe dans certains cas un élément $\lambda \in \mathbb{R}_+^m$ vérifiant

$$L(u, \lambda) = \inf_{v \in \mathbb{R}^n} L(v, \lambda) = \sup_{\mu \in \mathbb{R}_+^m} \inf_{v \in \mathbb{R}^n} L(v, \mu)$$

soit encore $L(u, \lambda) = \sup_{\mu \in \mathbb{R}_+^m} L(u, \mu)$ avec u_μ est solution du problème sont contraintes :

$$(P_\mu) : u_\mu \in \mathbb{R}^n, L(u_\mu, \mu) = \inf_{v \in \mathbb{R}^n} L(v, \mu).$$

On établit également que

$$L(u, \lambda) = \sup_{\mu \in \mathbb{R}_+^m} L(u, \mu) = \inf_{v \in \mathbb{R}^n} \sup_{\mu \in \mathbb{R}_+^m} L(v, \mu)$$

les points (u, λ) apparaissant ainsi comme des points-selles du Lagrangien L . Dans cette direction, une relation entre l'ensemble des solutions du problème (P) et l'ensemble des premiers arguments des points-selles du Lagrangien L .

A l'aide de la fonction :

$$G : \mu \in \mathbb{R}_+^m \longrightarrow G(\mu) = \inf_{v \in \mathbb{R}^n} L(v; \mu).$$

On appelle problème dual du problème primal (P) : trouver λ tel que

$$(Q) : \lambda \in \mathbb{R}_+^m ; G(\lambda) = \sup_{\mu \in \mathbb{R}_+^m} G(\mu)$$

- Tout les espace vectoriels considérés sont réels.

4.2.1 Lemme de Farkas- Minkowski

Théorème 4.2.1. (*Lemme de Farkas- Minkowski*)

Soient $a_i : i \in I$, où I est un ensemble fini d'indices, et b des éléments d'un espace de Hilbert V , de produit scalaire $\langle \cdot, \cdot \rangle$. Alors l'inclusion

$$\{w \in V : \langle a_i, w \rangle \geq 0, i \in I\} \subseteq \{w \in V : \langle b, w \rangle \geq 0\}$$

est satisfaite si et seulement si (\Leftrightarrow)

$$\exists \lambda_i \geq 0, i \in I, \text{ tel que } b = \sum_{i \in I} \lambda_i a_i.$$

Preuve. La condition est évidemment suffisante. Sa nécessité est établie en trois étapes :

(i) L'ensemble

$$C = \left\{ \sum_{i \in I} \lambda_i a_i \in V : \lambda_i \geq 0, i \in I \right\}$$

est un cône de sommet l'origine, convexe, fermé dans V . La convexité résulte de l'égalité :

$$\theta \sum_{i \in I} \lambda_i a_i + (1 - \theta) \sum_{i \in I} \mu_i a_i = \sum_{i \in I} (\theta \lambda_i + (1 - \theta) \mu_i) a_i.$$

Supposons les vecteurs $a_i, i \in I$, linéairement indépendants. Toute suite $(v_k)_{k \geq 0}$ de points $v_k = \sum_{i \in I} \lambda_i^k a_i$ de C est en particulier une suite de points du sous-espace vectoriel U de dimension finie, donc fermé, engendré par les vecteurs a_i . Si elle converge dans V , elle converge donc dans U , et comme la convergence dans U équivaut à la convergence des composantes, on déduit que

$$\lim_{k \rightarrow \infty} v_k = \sum_{i \in I} \{ \lim_{k \rightarrow \infty} \lambda_i^k \} a_i \in C \left(\text{car} : \lambda_i^k \geq 0 \implies \lim_{k \rightarrow \infty} \lambda_i^k \geq 0 \right)$$

ce qui montre que l'ensemble C est fermé.

Si $a_i, i \in I$ sont linéairement dépendants, alors $\exists \mu_i, i \in I$, tels que $\sum_{i \in I} \mu_i a_i = 0$ et

$J = \{i \in I, \mu_i < 0\} \neq \emptyset$, alors $v = \sum_{i \in I} \lambda_i a_i \in C$ peut aussi s'écrire $v = \sum_{i \in I} (\lambda_i + t\mu_i) a_i \in C$

où $t = \min_{i \in J} \{-\frac{\lambda_i}{\mu_i}\} \geq 0$.

Les nombres $(\lambda_i + t\mu_i) \geq 0, i \in I$, mais l'un d'eux au moins est nul : on a donc démontré l'égalité :

$$C = \bigcup_{j \in I} \left\{ v = \sum_{i \in (I - \{j\})} \lambda_i a_i, \lambda_i \geq 0, i \in (I - \{j\}) \right\}.$$

Et il suffit de recommencer ce raisonnement jusqu'à ce que le cône C soit écrit comme une réunion finie de cônes associés à des vecteurs a_i linéairement indépendants, donc fermés d'après la première partie du raisonnement. on obtient C est fermé.

(ii) Soit C une partie convexe fermée non vide d'un espace de Hilbert V , et $b \in V$ avec $b \notin C$. Alors on peut trouver, des vecteurs $u \in V$ et des nombres réels α tels que

$$\begin{cases} \langle v, u \rangle > \alpha \\ \langle b, u \rangle < \alpha \end{cases} : \forall v \in C.$$

D'après le théorème de projection, $\exists c \in C$ et un seul tel que

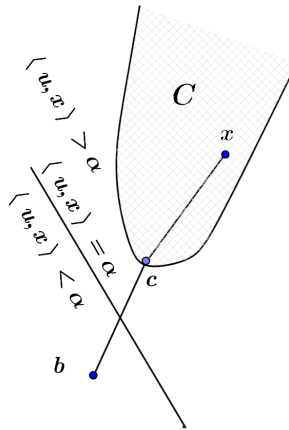


Fig 9

$$\begin{cases} \|b - c\|_V = \inf_{v \in C} \|b - v\|_V > 0 \\ \langle v - c, c - b \rangle \geq 0, \forall v \in C \end{cases}$$

et on a $\langle c - b, c - b \rangle > 0 \implies \langle c, c - b \rangle > \langle b, c - b \rangle$

donc $\langle v, c - b \rangle \geq \langle c, c - b \rangle > \langle b, c - b \rangle$.

Il suffit de choisir $u = (c - b)$ et n'importe quel nombre α vérifiant $\langle c, c - b \rangle > \alpha > \langle b, c - b \rangle$.

(iii) Si le point b n'appartient pas à l'ensemble

$$C = \{v = \sum_{i \in I} \lambda_i a_i, \lambda_i \geq 0, i \in I\}.$$

Alors, $\exists u \in V : \langle a_i, u \rangle \geq 0, i \in I$ et $\langle b, u \rangle < 0$.

D'après l'étape (i), l'ensemble C est une partie convexe fermée non vide de V . Si donc $b \notin C$, d'après l'étape (ii), il existe un vecteur $u \in V$ et un nombre réel α tels que

$$\begin{cases} \langle v, u \rangle > \alpha \\ \langle b, u \rangle < \alpha \end{cases}, \forall v \in C.$$

Comme C est un cône de sommet o . $o \in C \implies \alpha < 0$, et par suite

$$\langle v, u \rangle > \frac{\alpha}{\lambda}, \forall \lambda > 0 \implies \langle v, u \rangle \geq 0, \forall v \in C \text{ (car : } v \in C \implies \lambda v \in C, \forall \lambda > 0).$$

On a donc établi que $\langle a_i, u \rangle \geq 0, i \in I$ (puisque $a_i \in C$), et $\langle b, u \rangle < \alpha < 0$.

Ce qui donne la contradiction car, $\forall u \in V$, si $\langle a_i, u \rangle_{i \in I} \geq 0 \implies \langle b, u \rangle \geq 0$. ce qui démontre l'assertion.

(\Leftarrow) Si $b = \sum_{i \in I} \lambda_i a_i, \lambda_i \geq 0$. Soit $u \in \{w \in V : \langle a_i, w \rangle \geq 0, i \in I\}$. Alors $\langle a_i, u \rangle \geq 0, i \in I$.

Comme $\lambda_i \geq 0 \implies \langle \lambda_i a_i, u \rangle \geq 0, i \in I$, donc $\langle \sum_{i \in I} \lambda_i a_i, u \rangle = \langle b, u \rangle \geq 0$.

Alors $u \in \{w \in V, \langle b, w \rangle \geq 0\}$, donc

$$\{w \in V : \langle a_i, w \rangle, i \in I\} \subseteq \{w \in V, \langle b, w \rangle \geq 0\}.$$

■

4.2.2 Les relations de Kuhn et Tucker

Soit V un espace vectoriel normé et U une partie non vide de V .

En tout point $u \in U$, le cône $C(u)$ des directions admissibles est la réunion de $\{o\}$ et de l'ensemble des vecteurs $w \in V$ pour lesquels il existe (au moins) une suite de points $(u_k)_{k \geq 0}$ telle que

$$\begin{cases} u_k \in U \text{ et } u_k \neq u : \forall k > 0, \lim_{k \rightarrow \infty} u_k = u \\ \lim_{k \rightarrow \infty} \frac{u_k - u}{\|u_k - u\|} = \frac{w}{\|w\|}, w \neq 0. \end{cases}$$

La dernière condition s'écrivant de façon équivalente

$$u_k = u + \|u_k - u\| \frac{w}{\|w\|} + \|u_k - u\| \delta_k, \quad \lim_{k \rightarrow \infty} \delta_k = 0; \quad w \neq 0.$$

Il est clair que l'ensemble $C(u)$ est un cône, non nécessairement convexe.

Si une telle courbe γ est représentée par une application $t \rightarrow u(t)$ avec $u(0) = u$, dérivable en 0 et si $u'(0) \neq 0$ on peut en effet écrire

$$u_k - u = t_k u'(0) + t_k \varepsilon_k; \quad \lim_{k \rightarrow \infty} \varepsilon_k = 0 \text{ avec } : u_k = u(t_k), \quad t_k > 0; \quad \lim_{k \rightarrow \infty} t_k = 0,$$

de sorte que

$$\lim_{k \rightarrow \infty} \frac{u_k - u}{\|u_k - u\|} = \frac{u'(0)}{\|u'(0)\|}.$$

Théorème 4.2.2. Soit U une partie quelconque non vide d'un espace vectoriel normé V .

(1) En tout point $u \in U$, le cône $C(u)$ des directions admissibles est fermé.

(2) Soit $J : \Omega \subset V \rightarrow \mathbb{R}$ une fonction définie sur un ouvert $\Omega \subset U$. Si la fonction J admet en un point $u \in U$ un minimum relatif par rapport à l'ensemble U , et si elle est dérivable en u , alors

$$J'(u)(v - u) \geq 0, \quad \forall v \in \{u + C(u)\}.$$

Preuve.

(1) Soit $(w_n)_{n \geq 0} \subseteq C(u)$ et $w_n \xrightarrow{n \rightarrow +\infty} w$ vers $w \in V$.

Si $w = 0$ on a $0 \in C(u)$ stop. sinon,

on suppose que $w \neq 0$ et donc supposez $w_n \neq 0$ pour tout n . Par définition, $\forall n \in \mathbb{N} : \exists (u_k^n)_{k \geq 0}$ tel que

$$\begin{cases} u_k^n \in U \text{ et } u_k^n \neq u : \forall k > 0, \quad \lim_{k \rightarrow \infty} u_k^n = u \\ u_k^n = u + \|u_k^n - u\| \cdot \frac{w_n}{\|w_n\|} + \|u_k^n - u\| \cdot \delta_k^n, \quad \lim_{k \rightarrow +\infty} \delta_k^n = 0 \end{cases}.$$

Soit $(\varepsilon_n)_{n \geq 0}$ une suite de nombres $\varepsilon_n > 0$ telle que $\lim_{n \rightarrow +\infty} \varepsilon_n = 0$. Alors il existe un entier $k(n)$ tel que

$$\|u_{k(n)}^n - u\| \leq \varepsilon_n, \quad \|\delta_{k(n)}^n\| \leq \varepsilon_n.$$

Considérons alors la suite $(u_{k(n)}^n)_n$. On a d'une part

$$u_{k(n)}^n \in U, \quad u_{k(n)}^n \neq u, \quad \forall n \geq 0. \quad \lim_{n \rightarrow +\infty} u_{k(n)}^n = u$$

et l'on peut écrire , d'autre part

$$u_{k(n)}^n = u + \|u_{k(n)}^n - u\| \cdot \frac{w}{\|w\|} + \|u_{k(n)}^n - u\| \cdot \{\delta_{k(n)}^n + (\frac{w_n}{\|w_n\|} - \frac{w}{\|w\|})\}.$$

Comme $\lim_{n \rightarrow +\infty} \frac{w_n}{\|w_n\|} = \frac{w}{\|w\|}$ l'assertion est démontrée . ($w \in C(u)$) donc $C(u)$ fermé .

(2) Soit $w = (v - u)$ un vecteur non nul du cône $C(u)$, et soit $(u_k)_k$ une suite de points de $U - \{u\}$ telle que

$$\begin{cases} \lim_{k \rightarrow +\infty} u_k = u \\ u_k - u = \|u_k - u\| \cdot \frac{w}{\|w\|} + \|u_k - u\| \cdot \delta_k , \quad \lim_{k \rightarrow +\infty} \delta_k = 0 \\ J(u) \leq J(u_k) , \quad \forall k \end{cases} .$$

Comme J dérivable on a

$$\begin{aligned} 0 \leq J(u_k) - J(u) &= J'(u)(u_k - u) + (u_k - u)\varepsilon_k = J'(u)(u_k - u) + \|u_k - u\| \frac{u_k - u}{\|u_k - u\|} \varepsilon_k \\ &= J'(u)(u_k - u) + \|u_k - u\| \left(\frac{w}{\|w\|} + \delta_k \right) \varepsilon_k \\ &= J'(u)(u_k - u) + \|u_k - u\| \varepsilon'_k ; \quad \lim_{k \rightarrow +\infty} \varepsilon'_k = 0, \end{aligned}$$

on a

$$\begin{aligned} 0 \leq J'(u) \left(\|u_k - u\| \cdot \frac{w}{\|w\|} + \|u_k - u\| \cdot \delta_k \right) + \|u_k - u\| \varepsilon'_k &= \frac{\|u_k - u\|}{\|w\|} (J'(u)w + \|w\|J'(u)\delta_k + \|w\| \varepsilon'_k) \\ &= \frac{\|u_k - u\|}{\|w\|} (J'(u)w + \eta_k) , \quad \lim_{k \rightarrow +\infty} \eta_k = 0. \end{aligned}$$

Alors

$$\frac{\|u_k - u\|}{\|w\|} (J'(u)w + \eta_k) \geq 0 \quad (\text{avec } \lim_{k \rightarrow +\infty} \eta_k = 0).$$

Il en résulte que le nombre $J'(u).w \geq 0$. ■

Nous considérons les ensembles U particuliers, de la forme

$$U = \{v \in \Omega : \varphi_i(v) \leq 0, 1 \leq i \leq m\},$$

avec Ω est un ouvert de V ,et considérons l'ensemble d'indices

$$I(u) = \{1 \leq i \leq m, \varphi_i(u) = 0\}$$

et

$$C^*(u) = \{w \in V, \varphi'_i(u)w \leq 0, i \in I(u)\}.$$

Définition 4.2.1. On dit que les contraintes sont qualifiées en un point $u \in U$ si

- ou bien toutes les fonctions φ_i , $i \in I(u)$ sont affines.
- ou bien il existe un vecteur $\bar{w} \in V$ tel que

$$\begin{cases} \varphi'_i(u)\bar{w} \leq 0 \\ \varphi'_i(u)\bar{w} < 0 \text{ si } \varphi_i \text{ n'est pas affine} \end{cases} \quad \text{pour tout } i \in I(u)$$

Théorème 4.2.3. Soit u un point de l'ensemble

$$U = \{ v \in \Omega, \varphi_i(v) \leq 0, 1 \leq i \leq m \},$$

et les fonctions $\varphi_i : \Omega \subseteq V \rightarrow \mathbb{R}$, $i \in I(u)$ sont dérivables. Alors :

- (1) Le cône des directions admissibles en ce point vérifie toujours l'inclusion $C(u) \subseteq C^*(u)$
- (2) Si les contraintes sont qualifiées en u et si les fonctions φ_i ; $i \notin I(u)$ sont continues en u , on a l'égalité $C(u) = C^*(u)$

Preuve.

(1) On a $\varphi_i(v) \leq 0 : \forall v \in U$ et on a $\varphi_i(u) = 0 : i \in I(u)$, alors chaque fonction $\varphi_i : i \in I(u)$ admet en u un maximum relatif par rapport à l'ensemble U , Une application du théorème (5.3.2) montre alors que :

$$\varphi'_i(u)w \leq 0, \forall w \in C(u), i \in I(u) \implies \forall w \in C(u) \text{ et } i \in I(u) \text{ on a } \varphi'_i(u)w \leq 0.$$

Alors $C(u) \subseteq C^*(u)$.

(2) Supposons d'abord toutes les fonctions φ_i ; $i \in I(u)$ affines.

Étant donné un vecteur w non nul de $C^*(u)$, considérons la suite (u_k) définie par

$$u_k = u + \varepsilon_k w \text{ où } \varepsilon_k > 0 \text{ et } \lim_{k \rightarrow +\infty} \varepsilon_k = 0.$$

Alors $u_k - u = \varepsilon_k w \neq 0$ pour tout $k \geq 0$ et $\lim_{k \rightarrow +\infty} u_k = u$ par ailleurs

$$0 > \varphi_i(u) = \lim_{k \rightarrow +\infty} \varphi_i(u_k) \text{ pour } i \notin I(u) \text{ (par continuité)}$$

$$\varphi_i(u_k) = \varphi_i(u_k) - \varphi_i(u) = \varphi'_i(u)(u_k - u) = \varepsilon_k \varphi'_i(u)w \leq 0, \text{ pour } i \in I(u).$$

Ce qui montre que les points u_k sont dans l'ensemble U pour k suffisamment grand.

Comme enfin :

$$\frac{u_k - u}{\|u_k - u\|} = \frac{w}{\|w\|} \text{ pour tout } k.$$

On a bien démontré que $w \in C(u)$.

Passons au cas général : il existe un vecteur $\bar{w} \neq 0$ tel que

$$\begin{cases} \varphi'_i(u)\bar{w} \leq 0 \\ \varphi'_i(u)\bar{w} < 0 \text{ si } \varphi_i \text{ n'est pas affine} \end{cases} \quad \text{pour tout } i \in I(u) .$$

Étant donné un vecteur w non nul de $C^*(u)$. Soit un nombre $\delta > 0$ tel que le vecteur $(w + \delta\bar{w})$ soit non nul. Considérons la suite (u_k) définie par

$$u_k = u + \varepsilon_k(w + \delta\bar{w}), \quad \varepsilon_k > 0, \quad \lim_{k \rightarrow +\infty} \varepsilon_k = 0.$$

Alors $u_k - u = \varepsilon_k(w + \delta\bar{w}) \neq 0$ pour tout k et $\lim_{k \rightarrow +\infty} u_k = u$. Par ailleurs

$$0 > \varphi_i(u) = \lim_{k \rightarrow +\infty} \varphi_i(u_k), \quad \text{pour } i \notin I(u) \text{ (par continuité)}$$

$$\varphi_i(u_k) = \varphi'_i(u)(u_k - u) = \varphi'_i(u)\varepsilon_k(w + \delta\bar{w}) = \varepsilon_k(\varphi'_i(u)w + \delta\varphi'_i(u)\bar{w}) \leq 0,$$

pour $i \in I(u)$ et φ_i affine.

$$\varphi_i(u_k) = \varphi'_i(u)(u_k - u) + \delta_k : \lim_{k \rightarrow +\infty} \delta_k = 0 \quad (\varphi_i(u) = 0)$$

$$\begin{cases} \varphi_i(u_k) = \varepsilon_k(\varphi'_i(u)w + \delta\varphi'_i(u)\bar{w} + \alpha_k), \quad \lim_{k \rightarrow +\infty} \alpha_k = 0 \text{ et donc} \\ \varphi_i(u_k) \leq \varepsilon_k(\delta\varphi'_i(u)\bar{w} + \alpha_k), \text{ pour } i \in I(u) \text{ et } \varphi_i \text{ non affine} \end{cases}$$

(car $\varphi'_i(u)w \leq 0$ $w \in C^*(u)$). Ce qui montre que les points $u_k \in U$ pour k suffisamment grand (ce qui explique l'inégalité stricte $\varphi'_i(u)\bar{w} < 0$ d'une part, et l'introduction des vecteurs $(w + \delta\bar{w})$ d'autre part). Comme

$$\frac{u_k - u}{\|u_k - u\|} = \frac{w + \delta\bar{w}}{\|w + \delta\bar{w}\|} \quad \text{pour tout } k.$$

On a montré que $(w + \delta\bar{w}) \in C(u)$ pour tout nombre $\delta > 0$, suffisamment petit. La suite (w_n) définie par $w_n = w + \varepsilon_n\bar{w}$ avec $\lim_{n \rightarrow +\infty} \varepsilon_n = 0$, est une suite de points de $C(u)$ (pour n assez grand). Comme l'ensemble $C(u)$ est fermé (théorème 5.3.2). On conclut que $\lim_{n \rightarrow +\infty} w_n = w \in C(u)$. Alors $C^*(u) \subseteq C(u)$ on par suite $C^*(u) = C(u)$. ■

Exemple 4.2.1. Soit $V = \mathbb{R}^n$ et toutes les fonctions φ_i sont affines, l'ensemble U se présentant alors sous la forme :

$$U = \left\{ v \in \mathbb{R}^n, \sum_{j=1}^n c_{ij}v_j \leq d_i, 1 \leq i \leq m \right\} = \{v \in \mathbb{R}^n, c.v \leq d\}$$

où $c = (c_{ij}) \in M_{m;n}(\mathbb{R})$ est une matrice donnée et $d \in \mathbb{R}^m$ un vecteur donné, le cône des directions admissibles en un point $u \in U$ est l'ensemble

$$C^*(u) = C(u) \implies C(u) = \left\{ w \in \mathbb{R}^n, \sum_{j=1}^n c_{ij} w_j \leq 0, i \in I(u) \right\}.$$

- Revenant à la situation générale, nous obtenons un des résultats les plus importants de

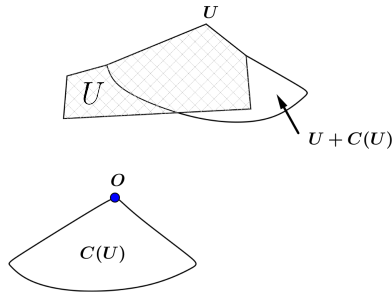


Fig 10

l'optimisation.

Théorème 4.2.4. (Condition nécessaire de minimum en programmation non linéaire).

Soit $\varphi_i : \Omega \subseteq V \rightarrow \mathbb{R}, 1 \leq i \leq m$ avec Ω est un ouvert d'un espace de Hilbert V et

$U = \{v \in \Omega, \varphi_i(v) \leq 0, 1 \leq i \leq m\}$ et $I(u) = \{1 \leq i \leq m, \varphi_i(u) = 0\}$.

On suppose les fonctions $\varphi_i, i \in I(u)$ dérivables en u , et $\varphi_i, i \notin I(u)$ continues en u .

Enfin, on se donne une fonction $J : \Omega \subseteq V \rightarrow \mathbb{R}$ dérivable en u .

Si J admet en u un minimum relatif par rapport à l'ensemble U , et si les contraintes sont qualifiées au point u , alors il existe des nombres $\lambda_i(u), i \in I(u)$ tels que

$$J'(u) + \sum_{i \in I(u)} \lambda_i(u) \varphi'_i(u) = 0 \text{ et } \lambda_i(u) \geq 0; i \in I(u).$$

Preuve.

D'après le théorème.5.3.2,

$$J'(u)w \geq 0 \text{ pour tout } w \in C(u).$$

D'après le théorème .5.3.3,

$$C(u) = \{w \in V : \varphi'_i(u)w \leq 0, i \in I(u)\}.$$

Désignant par $\langle \cdot, \cdot \rangle$ le produit scalaire de l'espace V et par τ l'isométrie canonique de V sur V' , nous sommes donc dans la situation suivante :

$$\{w \in V, \langle -\tau\varphi'_i(u), w \rangle \geq 0, i \in I(u)\} \subseteq \{w \in V, \langle \tau J'(u), w \rangle \geq 0\}$$

car

$$\begin{aligned} w \in V, \langle -\tau\varphi'_i(u), w \rangle \geq 0, i \in I(u) &\implies w \in V, \varphi'_i(u)w \leq 0, i \in I(u) \\ &\implies w \in C(u), \text{ alors } J'(u)w \geq 0 \\ &\implies w \in V \text{ et } \langle \tau J'(u); w \rangle \geq 0, \end{aligned}$$

c'est-à-dire exactement les conditions d'application du lemme de Farkas-Minkowski (théorème .5.3.1) : on obtient

$$\begin{aligned} \exists \lambda_i(u) \geq 0, i \in I(u) \text{ tel que } \tau J'(u) &= \sum_{i \in I(u)} \lambda_i(u)(-\tau\varphi'_i(u)) \\ &\implies \tau J'(u) + \tau \sum_{i \in I(u)} \lambda_i(u)\varphi'_i(u) = 0 \\ &\implies J'(u) + \sum_{i \in I(u)} \lambda_i(u)\varphi'_i(u) = 0 \dots (\star). \end{aligned}$$

Alors les relations (\star) , s'applent les relations de Kuhn et Tucker

- Si on poser $\lambda_i(u) = 0, \forall i \notin I(u)$. Donc

$$\begin{cases} J'(u) + \sum_{i=1}^m \lambda_i(u)\varphi'_i(u) = 0 \\ \lambda_i(u) \geq 0, 1 \leq i \leq m, \sum_{i=1}^m \lambda_i(u)\varphi_i(u) = 0. \end{cases}$$

On appelle le vecteur $\lambda(u) = (\lambda_i(u)) \in \mathbb{R}_+^m$ un multiplicateur de Lagrange généralisé, associé au minimum relatif u . ■

Remarque 4.2.1. Si $I(u) = \emptyset$, alors la relation de Kuhn et Tucker se réduisent à $J'(u) = 0$.

Définition 4.2.2. On dit que des contraintes convexes $\varphi_i : \Omega \subseteq V \longrightarrow \mathbb{R}, 1 \leq i \leq m$ sont qualifiées si :

- ou bien toutes les fonctions $\varphi_i, 1 \leq i \leq m$ sont affines, et l'ensemble (convexe si Ω est convexe)

$$U = \{v \in \Omega, \varphi_i(v) \leq 0, 1 \leq i \leq m\}$$

est non vide,

- ou bien il existe un point $\bar{v} \in \Omega$ tel que

$$\begin{cases} \varphi_i(\bar{v}) \leq 0 \\ \varphi_i(\bar{v}) < 0 \text{ si } \varphi_i \text{ n'est pas affine} \end{cases} : 1 \leq i \leq m$$

Théorème 4.2.5. (Conditions nécessaires et suffisantes de minimum en programmation convexe). Soit $J : \Omega \subseteq V \rightarrow \mathbb{R}$ une fonction définie sur un ouvert convexe Ω d'un espace de Hilbert V , et

$$U = \{v \in \Omega : \varphi_i(v) \leq 0, 1 \leq i \leq m\}$$

une partie de Ω , les contraintes $\varphi_i : \Omega \subseteq V \rightarrow \mathbb{R}, 1 \leq i \leq m$ étant supposées convexes.

Soit $u \in U$ un point en lequel les fonctions $\varphi_i, 1 \leq i \leq m$ et J sont dérivables.

(1) Si la fonction J admet en u un minimum relatif par rapport à l'ensemble U et si les contraintes sont qualifiées, alors il existe des nombres $\lambda_i(u), 1 \leq i \leq m$ tels que les relations de Kuhn et Tucker :

$$\begin{cases} J'(u) + \sum_{i=1}^m \lambda_i(u) \varphi'_i(u) = 0 \\ \lambda_i(u) \geq 0, 1 \leq i \leq m, \sum_{i=1}^m \lambda_i(u) \varphi_i(u) = 0 \end{cases}$$

soient satisfaites.

(2) Réciproquement, si la fonction $J : U \rightarrow \mathbb{R}$ est convexe et s'il existe des nombres $\lambda_i, 1 \leq i \leq m$ tels que les relations de Kuhn et Tucker soient satisfaites, alors la fonction J admet en u un minimum par rapport à l'ensemble U .

Preuve.

(1) Il suffit naturellement de montrer que, si des contraintes convexes φ_i sont qualifiées au sens ci-dessus, elles le sont aussi en tout point $u \in U$ au sens entendu antérieurement, ce qui permet d'appliquer ensuite le (théorème .5.3.4).

Or, si $u \neq \bar{v}$ le vecteur $\bar{w} = \bar{v} - u$ répond à la question, puisque

$$\varphi'_i(u) \bar{w} = \varphi_i(u) + \varphi'_i(u) (\bar{v} - u) \leq \varphi_i(\bar{v}), \forall i \in I(u) \quad (\varphi_i(u) = 0, i \in I(u)).$$

Les fonctions φ_i étant convexes : si $u = \bar{v}$ l'ensemble $I(u)$ ne peut contenir que des indices i pour lesquels les fonctions φ_i correspondantes sont affines, et le vecteur $\bar{w} = 0$

répond à la question.

(2) Soit $v \in U$ (convexe). Alors

$$\begin{aligned}
 J(u) &\leq J(u) - \sum_{i=1}^m \lambda_i \varphi_i(v) \quad (\lambda_i \geq 0, \varphi_i(v) \leq 0) \\
 &= J(u) - \sum_{i \in I(u)} \lambda_i (\varphi_i(v) - \varphi_i(u)) \quad (\lambda_i = 0 \text{ si } i \notin I(u), \varphi_i(u) = 0, i \in I(u)) \\
 &\leq J(u) - \sum_{i \in I(u)} \lambda_i \cdot \varphi_i'(u)(v - u) \quad (\varphi_i \text{ convexe}) \\
 &\leq J(u) + (v - u) \left(- \sum_{i \in I(u)} \lambda_i \cdot \varphi_i'(u) \right) \\
 &\leq J(u) + J'(u)(v - u) \quad (\text{Kuhn et Tucker}) \\
 &\leq J(v) \quad (J \text{ convexe}).
 \end{aligned}$$

■

4.2.3 Lagrangien et points -selles

Soient V et M deux ensembles quelconques et

$$L: V \times M \longrightarrow \mathbb{R}$$

une fonction. On dit qu'un point $(u, \lambda) \in V \times M$ est un point - selle de la fonction L si le point u un minimum pour la fonction $L(\cdot, \lambda) : v \in V \longrightarrow L(v, \lambda) \in \mathbb{R}$ et si le point λ est un maximum pour la fonction $L(u, \cdot) : \mu \in M \longrightarrow L(u, \mu) \in \mathbb{R}$, autrement dit, si

$$\sup_{\mu \in M} L(u, \mu) = L(u, \lambda) = \inf_{v \in V} L(v, \lambda).$$

Remarques

- 1) On notera que la définition fait intervenir de façon essentielle (mais évidemment arbitraire) l'ordre dans lequel les deux variables interviennent.
- 2) Dans l'usage que nous ferons de cette notion, les variables $\mu \in M$ décrivent un espace de " multiplicateurs de Lagrange généralisés.
 - On a tendance à se représenter un point-selle comme " col ", c'est-à-dire un point d'une surface ayant la forme d'une selle.

Théorème 4.2.6. Si (u, λ) est un point-selle d'une fonction $L : V \times M \longrightarrow \mathbb{R}$, alors :

$$\sup_{\mu \in M} \inf_{v \in V} L(v, \mu) = L(u, \lambda) = \inf_{v \in V} \sup_{\mu \in M} L(v, \mu).$$

Preuve. Montrons d'abord que :

$$\sup_{\mu \in M} \inf_{v \in V} L(v, \mu) \leq \inf_{v \in V} \sup_{\mu \in M} L(v, \mu),$$

étant donné des éléments quelconque $\bar{v} \in V$ et $\bar{\mu} \in M$ on a sûrement

$$\inf_{v \in V} L(v, \bar{\mu}) \leq L(\bar{v}, \bar{\mu}) \leq \sup_{\mu \in M} L(\bar{v}, \mu).$$

(on n'exclut par les valeur $-\infty$ pour $\inf\{.\}$ et $+\infty$ pour $\sup\{.\}$). Comme $\inf_{v \in V} L(v, \bar{\mu})$ est fonction de la seul variable $\bar{\mu} \in M$, et comme $\sup_{\mu \in M} L(\bar{v}, \mu)$ est fonction de la seul variable $\bar{v} \in V$, alors l'inégalité suivant :

$$\sup_{\mu \in M} \inf_{v \in V} L(v, \mu) \leq \inf_{v \in V} \sup_{\mu \in M} L(v, \mu) \tag{4.5}$$

• pour établir notre l'inégalité, exprimons que (u, λ) est un point-selle, on écrit

$$\inf_{v \in V} \sup_{\mu \in M} L(v, \mu) \leq \sup_{\mu \in M} L(u, \mu) = L(u, \lambda) = \inf_{v \in V} L(v, \lambda) \leq \sup_{\mu \in M} \inf_{v \in V} L(v, \mu).$$

Alors

$$\inf_{v \in V} \sup_{\mu \in M} L(v, \mu) \leq \sup_{\mu \in M} \inf_{v \in V} L(v, \mu) \tag{4.6}$$

De (4.5) et (4.6), on obtient donc

$$\inf_{v \in V} \sup_{\mu \in M} L(v, \mu) = L(u, \lambda) = \sup_{\mu \in M} \inf_{v \in V} L(v, \mu).$$

Reprenons le problème général de la programmation non linéaire, où nous supposons les diverses fonctions définies sur l'espace V tout entier, dans le seul but d'alléger l'écriture : étant donné des fonctions $J : V \longrightarrow \mathbb{R}$ et $\varphi_i : V \longrightarrow \mathbb{R}$, $1 \leq i \leq m$, définies sur un espace de Hilbert V , on définit l'ensemble

$$U = \{ v \in V, \varphi_i(v) \leq 0, 1 \leq i \leq m \}$$

et il s'agit de trouver le, où les éléments, u vérifiant

$$u \in U : J(u) = \inf_{v \in U} J(v)$$

c'est ce que nous appellerons dorénavant le problème (P) .

L'intérêt fondamental de la notion de point-selle est le suivant (c'est ce que nous établissons au prochain théorème) sous certaines conditions, toute solution u du problème (P) est également premier argument d'un point-selle (u, λ) d'une fonction convenable $L : V \times M \rightarrow \mathbb{R}$ (l'espace M étant donc à déterminer), appelée le lagrangien associé au problème (P) considéré, et réciproquement, si (u, λ) est un point-selle de ce même lagrangien L alors u est solution du problème (P) .

On appelle le second argument λ d'un tel point-selle (u, λ) un multiplicateur de Lagrange généralisé associé à la solution u du problème (P) .

De façon plus précise, définissons le Lagrangien associé au problème (P) ci-dessus comme étant la fonction :

$$L : (v, \mu) \in V \times \mathbb{R}^m \longrightarrow L(v, \mu) = J(v) + \sum_{i=1}^m \mu_i \varphi_i(v)$$

le " deuxième " espace M étant donc \mathbb{R}_+^m .

■

Théorème 4.2.7. (1) Si $(u, \lambda) \in V \times \mathbb{R}_+^m$ est point-selle du Lagrangien L , le point u , qui appartient à l'ensemble U , est solution du problème (P) .

(2) On suppose les fonctions J et φ_i , $1 \leq i \leq m$, convexes, dérivables en un point $u \in U$, et les contraintes qualifiées. Alors, si u est solution du problème (P) , il existe au moins un vecteur $\lambda \in \mathbb{R}_+^m$ tel que le couple $(u, \lambda) \in V \times \mathbb{R}_+^m$ soit point-selle du Lagrangien L .

Preuve.

(1) Les inégalités $L(u, \mu) \leq L(u, \lambda) : \forall \mu \in \mathbb{R}_+^m$ s'écrivent : $L(u, \mu) - L(u, \lambda) \leq 0$. Alors

$$J(u) + \sum_{i=1}^m \mu_i \varphi_i(u) - J(u) - \sum_{i=1}^m \lambda_i \varphi_i(u) \leq 0 \implies \sum_{i=1}^m (\mu_i - \lambda_i) \varphi_i(u) \leq 0 : \forall \mu \geq 0$$

ce qui entraîne $\varphi_i(u) \leq 0$, $1 \leq i \leq m$ (faire tendre chaque μ_i vers $+\infty$).

Si $\mu = 0$ on obtient $\sum_{i=1}^m -\lambda_i \varphi_i(u) \leq 0 \implies \sum_{i=1}^m \lambda_i \varphi_i(u) \geq 0$ et comme $\lambda_i \geq 0$ et $\varphi_i(u) \leq 0$

on obtient $\sum_{i=1}^m \lambda_i \varphi_i(u) \leq 0$, on a donc déjà montré que $u \in U$ et $\sum_{i=1}^m \lambda_i \varphi_i(u) = 0$, mais on a les inégalités $L(u, \lambda) \leq L(v, \lambda) : \forall v \in V$, on obtient donc

$$J(u) \leq J(v) + \sum_{i=1}^m \lambda_i \varphi_i(v) : \forall v \in V$$

comme $\lambda_i \geq 0$, et $\varphi_i(u) \leq 0$, et si $v \in U \subset V$, donc

$$J(u) \leq J(v) + \sum_{i=1}^m \lambda_i \varphi_i(v) \implies J(u) \leq J(v) : \forall v \in U.$$

Alors comme $J(u) \leq J(v) : \forall v \in U$, et $u \in U$ donc

$$J(u) = \inf_{v \in U} J(v)$$

a par suite u est un solution de (P).

(2) On est exactement dans les conditions d'application du théorème . 5.3 .5 :

Si u est solution du problème (P), il existe un vecteur $\lambda \in R_+^m$ tel que $\sum_{i=1}^m \lambda_i \varphi_i(u) = 0$ et

$J'(u) + \sum_{i=1}^m \lambda_i \varphi_i'(u) = 0$ (relations de Kuhn et Tuckes). La première égalité donne

$$L(u, \mu) = J(u) + \sum_{i=1}^m \mu_i \varphi_i(u) \leq J(u) = L(u, \lambda) : \forall \mu \in R_+^m \implies L(u, \lambda) = \sup_{\mu \in R_+^m} L(u, \mu),$$

(car $\mu_i \geq 0$ et $\varphi_i(u) \leq 0$),

et la seconde égalité est une condition suffisante de minimum pour la fonction convexe (somme de fonctions convexes)

$$L(., \lambda) : v \in V \longrightarrow J(v) + \sum_{i=1}^m \lambda_i \varphi_i(v)$$

comme J convexe et la relations de Kuhn et Tuckes, on a

$$J(u) - J(v) \leq (u - v)J'(u) \implies J(u) - J(v) \leq (u - v) \left(- \sum_{i=1}^m \lambda_i \varphi_i'(u) \right)$$

$$\implies J(u) - J(v) \leq \sum_{i=1}^m \lambda_i (\varphi_i'(u)(v - u))$$

et on a comme φ_i convexe pour tout i , on a

$$\varphi_i(u) - \varphi_i(v) \leq \varphi_i'(u)(u - v) \implies \varphi_i(u) - \varphi_i(v) \leq -\varphi_i'(u)(v - u)$$

$$\implies \varphi'_i(u)(v - u) \leq \varphi_i(v) - \varphi_i(u).$$

Alors

$$\begin{aligned} J(u) - J(v) &\leq \sum_{i=1}^m \lambda_i \varphi_i(v) - \sum_{i=1}^m \lambda_i \varphi_i(u) \\ \implies J(u) + \sum_{i=1}^m \lambda_i \varphi_i(u) &\leq J(v) + \sum_{i=1}^m \lambda_i \varphi_i(v) \\ L(u, \lambda) &\leq L(v, \lambda) \end{aligned}$$

par suite $L(u, \lambda) \leq L(v, \lambda)$ pour tout $v \in V$, et on a donc établi que le point $(u; \lambda)$ est un point-selle du Lagrangien L .

- L'ensemble des solutions du problème (P) : trouver u tel que

$$(P) : \begin{cases} u \in U = \{v \in V : \varphi_i(v) \leq 0 : 1 \leq i \leq m\} \\ J(u) = \inf_{v \in U} J(v) \end{cases}$$

coïncide avec l'ensemble des premiers arguments des points-selles du Lagrangien

$$L : (v, \mu) \in V \times \mathbb{R}_+^m \longrightarrow L(v, \mu) = J(v) + \sum_{i=1}^m \mu_i \varphi_i(v)$$

disons λ , des seconds arguments de ces points-selles. le problème (P) avec contraintes serait remplacé par le problème sans contraintes (P_λ) : trouver (u_λ) tel que :

$$(P_\lambda) : u_\lambda \in V : L(u_\lambda, \lambda) = \inf_{v \in V} L(v, \lambda).$$

Comment trouver un tel élément $\lambda \in \mathbb{R}_+^m$. Si l'on se rappelle que (théorème . 5.3 .6)

$$L(u_\lambda, \lambda) = \inf_{v \in V} L(v, \lambda) = \sup_{\mu \in \mathbb{R}_+^m} \inf_{v \in V} L(v, \mu)$$

on est naturellement conduit à chercher λ comme solution du problème (Q) : trouver λ tel que

$$\lambda \in \mathbb{R}_+^m : G(\lambda) = \sup_{\mu \in \mathbb{R}_+^m} G(\mu)$$

où l'application, $G : \mathbb{R}_+^m \longrightarrow \mathbb{R}$ est définie par

$$G : \mu \in \mathbb{R}_+^m \longrightarrow G(\mu) = \inf_{v \in V} L(v; \mu).$$

On appelle (Q) le problème dual du problème (P) qui, de ce point de vue, devient le problème primal. De la même façon, on appelle $\mu \in \mathbb{R}_+^m$ la variable duale, de la variable

primale $v \in U \subseteq V$.

Le problème dual apparaît donc à nouveau comme un problème d'optimisation avec contraintes, mais les contraintes $\mu_i \geq 0, 1 \leq i \leq m$, sont aisément prises en compte puisqu'on connaît l'opérateur de projection associé, alors que les contraintes

$\varphi_i(u) \leq 0, 1 \leq i \leq m$, sont en général impossibles à prendre en compte numériquement.

Alors qu'au théorème précédent, nous avons établi le lien existant entre le problème primal (P) et un problème primal-dual, qui faisait intervenir à la fois les variables primale et duale, de recherche de point-selle, il nous reste à établir le lien existant entre les solutions du problème primal (P) et les solutions du problème dual (Q) , c'est l'objet du résultat qui suit. ■

Théorème 4.2.8. *On suppose que les fonctions $\varphi_i : V \rightarrow \mathbb{R}, 1 \leq i \leq m$, sont continues et que, pour tout $\mu \in \mathbb{R}_+^m$, le problème (P_μ) : trouve $u_\mu \in V$ tel que*

$$(P_\mu)_{\mu \in \mathbb{R}_+^m} : u_\mu \in V, L(u_\mu, \mu) = \inf_{v \in V} L(v, \mu)$$

à une solution et une seule u_μ dépendant continûment de $\mu \in \mathbb{R}_+^m$. Alors, si λ est une solution quelconque du problème (Q) , la solution u_λ du problème (P_λ) correspondant est solution du problème (P) .

Preuve. Soit λ une solution quelconque du problème (Q) . On a

$$\lambda \in \mathbb{R}_+^m, G(\lambda) = L(u_\lambda, \lambda) = \inf_{v \in V} L(v, \lambda)$$

et nous allons établir la relation

$$\sup_{\mu \in \mathbb{R}_+^m} L(u_\lambda, \mu) = L(u_\lambda, \lambda)$$

Ces deux relations constituant précisément la définition d'un point-selle (u_λ, λ) du Lagrangien L , on déduira du théorème 5.3.7 que u_λ est solution du problème (P) . Considérons deux points μ et $(\mu + \xi)$ de l'ensemble \mathbb{R}_+^m , de composantes μ_i et $(\mu_i + \xi_i), 1 \leq i \leq m$, les inégalités

$$L(u_\mu, \mu) \leq L(u_{\mu+\xi}, \mu) \text{ et } L(u_{\mu+\xi}, \mu + \xi) \leq L(u_\mu, \mu + \xi)$$

s'écrivent encore :

$$\sum_{i=1}^m \xi_i \varphi_i(u_{\mu+\xi}) \leq G(\mu + \xi) - G(\mu) \leq \sum_{i=1}^m \xi_i \varphi_i(u_\mu)$$

car

$$\begin{aligned} G(\mu + \xi) - G(\mu) &= \inf_{v \in V} L(v, \mu + \xi) - \inf_{v \in V} L(v, \mu), \quad (\inf_{v \in V} L(v, \mu) = L(u_\mu, \mu)) \\ &\leq L(u_\mu, \mu + \xi) - L(u_\mu, \mu) \end{aligned}$$

mais

$$L(u_\mu, \mu + \xi) - L(u_\mu, \mu) = J(u_\mu) + \sum_{i=1}^m (\mu_i + \xi_i) \varphi_i(u_\mu) - J(u_\mu) - \sum_{i=1}^m \mu_i \varphi_i(u_\mu) = \sum_{i=1}^m \xi_i \varphi_i(u_\mu).$$

Alors on obtient

$$G(\mu + \xi) - G(\mu) \leq \sum_{i=1}^m \xi_i \varphi_i(u_\mu) \quad (4.7)$$

Respectivement on a

$$L(u_\mu, \mu) \leq L(u_{\mu+\xi}, \mu) \implies G(\mu) \leq L(u_{\mu+\xi}, \mu).$$

Alors

$$-L(u_{\mu+\xi}, \mu) \leq -G(\mu)$$

par suite

$$L(u_{\mu+\xi}, \mu + \xi) - L(u_{\mu+\xi}, \mu) \leq G(\mu + \xi) - G(\mu), \quad (G(\mu + \xi) = \inf_{v \in V} L(v, \mu + \xi) = L(u_{\mu+\xi}, \mu + \xi))$$

et on a

$$\begin{aligned} L(u_{\mu+\xi}, \mu + \xi) - L(u_{\mu+\xi}, \mu) &= J(u_{\mu+\xi}) + \sum_{i=1}^m (\mu_i + \xi_i) \varphi_i(u_{\mu+\xi}) - J(u_{\mu+\xi}) - \sum_{i=1}^m \mu_i \varphi_i(u_{\mu+\xi}) \\ &= \sum_{i=1}^m \xi_i \varphi_i(u_{\mu+\xi}). \end{aligned}$$

Alors on obtient

$$\sum_{i=1}^m \xi_i \varphi_i(u_{\mu+\xi}) \leq G(\mu + \xi) - G(\mu) \quad (4.8)$$

De (4.7) et (4.8) on obtient donc

$$\sum_{i=1}^m \xi_i \varphi_i(u_{\mu+\xi}) \leq G(\mu + \xi) - G(\mu) \leq \sum_{i=1}^m \xi_i \varphi_i(u_\mu)$$

il existe $\theta \in [0, 1]$ tel que

$$G(\mu + \xi) - G(\mu) = (1 - \theta) \sum_{i=1}^m \xi_i \varphi_i(u_\mu) + \theta \sum_{i=1}^m \xi_i \varphi_i(u_{\mu+\xi})$$

$$= \sum_{i=1}^m \xi_i \varphi_i(u_\mu) + \theta \sum_{i=1}^m \xi_i (\varphi_i(u_{\mu+\xi}) - \varphi_i(u_\mu)).$$

Les fonctions $\mu \in \mathbb{R}_+^m \longrightarrow u_\mu \in V$, et $\varphi_i : V \longrightarrow \mathbb{R}$, étant continues par hypothèse, on peut écrire, en tout point $\mu \in \mathbb{R}_+^m$

$$\begin{aligned} \theta \sum_{i=1}^m \xi_i (\varphi_i(u_{\mu+\xi}) - \varphi_i(u_\mu)) &= \theta \xi \cdot (\varphi(u_{\mu+\xi}) - \varphi(u_\mu)) \\ &= \|\xi\| \theta \frac{\xi}{\|\xi\|} \cdot (\varphi(u_{\mu+\xi}) - \varphi(u_\mu)) \\ &= \|\xi\| \epsilon(\xi) \text{ avec } \epsilon(\xi) = \theta \frac{\xi}{\|\xi\|} \cdot (\varphi(u_{\mu+\xi}) - \varphi(u_\mu)). \end{aligned}$$

Alors

$$G(\mu + \xi) - G(\mu) = \sum_{i=1}^m \xi_i \varphi_i(u_\mu) + \|\xi\| \cdot \epsilon(\xi) \text{ avec } \lim_{\xi \rightarrow 0} \epsilon(\xi) = 0.$$

($\|\cdot\|$: norme vectorielle quelconque sur \mathbb{R}^m), ce qui établit la dérivabilité de la fonction $G : \mathbb{R}^m \longrightarrow \mathbb{R}$, d'une part, et qui, d'autre part, donne l'expression de la dérivée, on a en effet

$$G(\mu + \xi) - G(\mu) = \xi G'(\mu) + \|\xi\| \epsilon(\xi) \implies \xi \cdot G'(\mu) = \sum_{i=1}^m \xi_i \varphi_i(u_\mu) : \forall \xi \in \mathbb{R}^m.$$

Puisque la fonction G admet un maximum au point λ de l'ensemble convexe \mathbb{R}_+^m alors on a :

$$G'(\lambda)(\mu - \lambda) \leq 0 : \forall \mu \in \mathbb{R}_+^m$$

soit encore :

$$\sum_{i=1}^m (\mu_i - \lambda_i) \varphi_i(u_\lambda) \leq 0 \implies \sum_{i=1}^m \mu_i \varphi_i(u_\lambda) \leq \sum_{i=1}^m \lambda_i \varphi_i(u_\lambda) : \forall \mu \in \mathbb{R}_+^m$$

par suite :

$$L(u_\lambda, \mu) = J(u_\lambda) + \sum_{i=1}^m \mu_i \varphi_i(u_\lambda) \leq J(u_\lambda) + \sum_{i=1}^m \lambda_i \varphi_i(u_\lambda) = L(u_\lambda, \lambda) : \forall \mu \in \mathbb{R}_+^m$$

ce qui est précisément la deuxième inégalité caractérisant un point-selle.

■

Théorème 4.2.9. [7]

On suppose que le problème (P) a au moins une solution u , que les fonctions J et φ_i , $1 \leq i \leq m$, sont convexes, dérivables en u , et enfin que les contraintes sont qualifiées. Alors le problème (Q) a au moins une solution.

Preuve. D'après le théorème (5.3.7), il existe (au moins) un vecteur $\lambda \in \mathbb{R}_+^m$. tel que le couple (u, λ) soit un point-selle du Lagrangien L . et Le théorème (5.3.6) entraîne alors.

$$L(u, \lambda) = \inf_{v \in V} L(v, \lambda) = \sup_{\mu \in \mathbb{R}_+^m} \inf_{v \in V} L(v, \mu)$$

ce qui peut encore s'écrire :

$$G(\lambda) = \sup_{\mu \in \mathbb{R}_+^m} G(\mu).$$

■

Remarque 4.2.2. *Si (u, λ) est un point-selle du Lagrangien L sur $V \times \mathbb{R}_+^m$, il est clair d'après le théorème. 5.3.6 que λ est solution du problème (Q) par définition de ce dernier, réciproquement, avec les hypothèses du théorème précédent, si λ est solution de (Q), le couple (u_λ, λ) est point-selle du Lagrangien L : il y a donc coïncidence de l'ensemble des solutions du problème dual (Q) avec celui des seconds arguments λ des points-selles (u_λ, λ) du Lagrangien L .*

Notre contribution est dans la suite :

4.3 Points selles d'une forme quadratique

Considérons le problème suivant :

$$(P) : \begin{cases} f(x) = \sum_{i=1}^p (\alpha_i x_i + \beta_i x_i^2) + \sum_{i=p+1}^n (\alpha_i x_i + \beta_i x_i^2) = \varphi(x) + \psi(x) \\ \Omega = \{x \in \mathbb{R}_+^n, Ax \leq b\}, \max_{x \in \Omega} f(x), b = (b_i)_{i=1}^m \in \mathbb{R}_+^m. \end{cases}$$

Notations et rappels :

$$\max_{x \in \Omega} f(x) = f(\bar{x}).$$

Soit a un vecteur de \mathbb{R}^n

$$[a] = \{x \in \mathbb{R}^n, x = \lambda a : \lambda \in \mathbb{R}\}.$$

$[a]$ est le sous-espace vectoriel de \mathbb{R}^n engendré par $a (a \neq 0)$ c'est le plus petit sous-espace vectoriel contenant a il est fermé, d'après le théorème de Riesz, $\mathbb{R}^n = [a] \oplus [a]^\perp$.

Algorithme de recherche de \bar{x} : On a deux possibilités

(i) \bar{x} est un sommet de Ω , Dans ce cas on a un algorithme de recherche de \bar{x} , [3]

(ii) Si \bar{x} n'est pas un sommet de Ω ,

on sait que \bar{x} est un point frontière de Ω , $\left(\text{sinon } \bar{x} = \left(\frac{-\alpha_i}{2\beta_i} \right)_i \right)$.

Alors il existe un hyperplan $ax = b$ contentent \bar{x}

$$(a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i) \quad (a\bar{x} = b).$$

Dans le cas où $f'(\bar{x}) \in [a]$, on peut déterminer la valeur de \bar{x}

$f'(\bar{x}) \in [a] \Rightarrow f'(\bar{x}) = k.a$ avec $k \in \mathbb{R}$ (nombre réel). Cette equation s'ecrit :

$$(\alpha_i + 2\beta_i \bar{x}_i)_i = k.a$$

avec

$$a = (a_{i1}, a_{i2}, \dots, a_{in}) = (a_1, a_2, \dots, a_n)$$

$$\Rightarrow \alpha_i + 2\beta_i \bar{x}_i = k.a_i, \quad i = 1, 2, \dots, n$$

$$\bar{x}_i = \frac{-\alpha_i}{2\beta_i} + k \cdot \frac{\alpha_i}{2\beta_i} \implies \bar{x} = x^* + k \cdot \left(\frac{\alpha_i}{2\beta_i} \right)_i \left(\text{si on pose } x^* = \left(\frac{-\alpha_i}{2\beta_i} \right)_i \right)$$

– On a la transformation suivante

$$K \longrightarrow 2k \implies \bar{x} = x^* + k \cdot \left(\frac{\alpha_i}{\beta_i} \right)_i$$

calcul de k

$$i = 1 \left\{ \begin{array}{l} \mathbf{a)} \ a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \text{ "contenant } \bar{x} \text{ " } a\bar{x} = b_1 \iff ax^* + ka \left(\frac{\alpha_i}{2\beta_i} \right)_i = b_1 \\ \implies \text{valeur de } k \\ \mathbf{b)} \ \text{Calcule de } \bar{x} = x^* + k \left(\frac{\alpha_i}{2\beta_i} \right)_i \\ \mathbf{c)} \ \text{Si } \bar{x} \in \Omega \\ \mathbf{d)} \ \text{Stop } x = \bar{x} \text{ la solution optimale du problème (P)} \\ \mathbf{e)} \ \text{Si } A.\bar{x} > b, \text{ on passe á } i = 2 \dots m \end{array} \right.$$

Remarque importante. On a les deux possibilités suivantes :

(★) \bar{x} est un sommet.

(★★) \bar{x} n'est pas un sommet et soit l'hyperplan contenant \bar{x} : $ax = b$.

On veut chercher k tel que $\bar{x} = x^* + k \left(\frac{\alpha_i}{\beta_i} \right)_i$.

Montrons que si $f(x) = \sum_{i=1}^p \alpha_i x_i + \beta_i x_i^2 = \varphi(x)$, et si $\bar{x} \neq \left(\frac{-\alpha_i}{2\beta_i} \right)_i$, alors $\bar{x} = P_{\Omega'}(x^*)$.

Sans restreindre la généralisation, on peut supposer que $\beta_i = -1, \forall i$.

$$\bar{x} = x^* - k(a_{1i})_i \implies a\bar{x} = b \text{ et } a\bar{x} = b = a.x^* - ka.(a_{1i})_i = a.x^* - k\|a\|^2.$$

D'où

$$k = -\frac{b - a.x^*}{\|a\|^2} = \frac{a.x^* - b}{\|a\|^2}.$$

Alors

$$\bar{x} = x^* - \frac{a.x^* - b}{\|a\|^2} . a.$$

Théorème 4.3.1. Si $\beta_i \neq 0, \forall i = 1 \dots n$.

Alors il existe un convexe fermé et borné Ω' de R^n , tel que

$$f(x) = \phi(y)$$

avec

$$\phi(y) = \sum_{i=1}^p (\alpha'_i y_i - y_i^2) + \sum_{i=p+1}^n (\alpha'_i y_i + y_i^2), \quad \beta'_i = \pm 1$$

pour tout $y \in \Omega'$.

Preuve.

pose $\Omega' = \{ y = (y_i)_i \in \mathbb{R}^n : y_i = \sqrt{|\beta_i|} \cdot x_i ; x = (x_i)_i \in \Omega \}$

le transformé Ω' de Ω devient

$$f(x) = \phi(y) = \sum_{i=1; \beta_i < 0}^p \left(\frac{\alpha_i}{\sqrt{-\beta_i}} y_i - y_i^2 \right) + \sum_{i=p+1; \beta_i > 0}^n \left(\frac{\alpha_i}{\sqrt{\beta_i}} y_i + y_i^2 \right)$$

$$Ax \leq b \left\{ \begin{array}{l} a_{11} \frac{y_1}{\sqrt{-\beta_1}} + \dots + a_{1p} \frac{y_p}{\sqrt{-\beta_p}} + a_{1p+1} \frac{y_{p+1}}{\sqrt{\beta_{p+1}}} + \dots + a_{1n} \frac{y_n}{\sqrt{\beta_n}} \leq b_1 \\ \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\ \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\ \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \quad \quad \quad \cdot \\ a_{n1} \frac{y_1}{\sqrt{-\beta_1}} + \dots + a_{np} \frac{y_p}{\sqrt{-\beta_p}} + a_{np+1} \frac{y_{p+1}}{\sqrt{\beta_{p+1}}} + \dots + a_{nn} \frac{y_n}{\sqrt{\beta_n}} \leq b_n \end{array} \right.$$

on pose : $\hat{a}_{1i} = \frac{a_{1i}}{\sqrt{|\beta_i|}}$ donc on obtient $\hat{A}y \leq b \implies \bar{y} = y^* + k(\hat{a}_{1i})$ où

$$\hat{a}_{1i} = \left(\frac{-a_{11}}{\sqrt{-\beta_1}}, \dots, \frac{-a_{1p}}{\sqrt{-\beta_p}}, \frac{a_{1p+1}}{\sqrt{\beta_{p+1}}}, \dots, \frac{a_{1n}}{\sqrt{\beta_n}} \right).$$

Sans restreindre la généralisation, on peut écrire f sous la forme :

$$f(x) = \sum_{i=1}^p (\alpha_i x_i - x_i^2) + \sum_{i=p+1}^n (\alpha_i x_i + x_i^2).$$

(Il suffit de poser $y_i = \sqrt{|\beta_i|} x_i$, ici on suppose que $\beta_i \neq 0$).

Considérons le cas où \bar{x} n'est pas un sommet de Ω .

Soit $\bar{x} = x^* + k \left(\frac{a_{1i}}{\beta_i} \right)_i ; k \in \mathbb{R}$.

On pose

$$\Omega_1 = \{ x \in \mathbb{R}_+^n : a_{11} x_1 + \dots + a_{1p} x_p = b_1 \text{ et } x_i = 0 : \forall i = p+1, \dots, n \}$$

$$\Omega_2 = \{ x \in \mathbb{R}_+^n : a_{1p+1} x_{p+1} + \dots + a_{1n} x_n = b_2 \text{ et } x_i = 0 : \forall i = 1, \dots, p \}$$

avec $b = b_1 + b_2$ et $x^* = (x_1^*, x_2^*, \dots, x_n^*)$.

On pose

$$\underline{x}^* = (x_1^*; x_2^*; \dots; x_p^*, 0, \dots, 0) \text{ et } \bar{x}^* = (0, \dots, 0, x_{p+1}^*, x_{p+2}^*, \dots, x_n^*)$$

$$x^* = \bar{x}^* + \underline{x}^*, \quad a_\varphi = (a_{11}, \dots, a_{1p}, 0, \dots, 0) \text{ et } a_\psi = (0, \dots, 0, a_{1p+1}, \dots, a_{1n})$$

$$\begin{cases} p_{\Omega_1}(\underline{x}^*) = \underline{x}^* - \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2} a_\varphi \\ p_{\Omega_2}(\bar{x}^*) = \bar{x}^* - \frac{a_\psi \bar{x}^* - b_2}{\|a_\psi\|^2} a_\psi \end{cases} \implies$$

$$\begin{aligned} p_{\Omega_1}(\underline{x}^*) + p_{\Omega_2}(\bar{x}^*) &= \underline{x}^* - \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2} a_\varphi + \bar{x}^* - \frac{a_\psi \bar{x}^* - b_2}{\|a_\psi\|^2} a_\psi \\ &= x^* - \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2} a_\varphi - \frac{a_\psi \bar{x}^* - b_2}{\|a_\psi\|^2} a_\psi. \end{aligned}$$

Mais

$$\bar{x} = x^* + k \begin{pmatrix} a_{1i} \\ \beta_i \end{pmatrix}_i \quad \begin{cases} \beta_i = 1 & \text{si } i = p+1, \dots, n \\ \beta_i = -1 & \text{si } i = 1, \dots, p \end{cases}$$

donne

$$a\bar{x} = ax^* + ka \begin{pmatrix} a_{1i} \\ \beta_i \end{pmatrix}_i \text{ avec } a = (a_{11}, \dots, a_{1p}, \dots, a_{1n})$$

$$\begin{aligned} a\bar{x} &= ax^* + k \left(- \sum_{i=1}^p a_{1i}^2 + \sum_{i=p+1}^n a_{1i}^2 \right) \\ &= ax^* - k \sum_{i=1}^p a_{1i}^2 + k \sum_{i=p+1}^n a_{1i}^2 \\ &= a\underline{x}^* + a\bar{x}^* - k \sum_{i=1}^p a_{1i}^2 + k \sum_{i=p+1}^n a_{1i}^2. \end{aligned}$$

Alors on a

$$a.\bar{x} = \left(a\underline{x}^* - k \sum_{i=1}^p a_{1i}^2 \right) + \left(a\bar{x}^* + k \sum_{i=p+1}^n a_{1i}^2 \right).$$

Posons alors $a\underline{x}^* - k \sum_{i=1}^p a_{1i}^2 = b_1$ et $a\bar{x}^* + k \sum_{i=p+1}^n a_{1i}^2 = b_2$ avec $b = b_1 + b_2$

$$\implies k = - \frac{b_1 - a\underline{x}^*}{\|a_\varphi\|^2} = - \frac{b_1 - a_\varphi \underline{x}^*}{\|a_\varphi\|^2}$$

$$= \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2},$$

car $a \underline{x}^* = (a_{11}, \dots, a_{1p}, \dots, a_{1n}) \cdot (x_1^*, \dots, x_p^*, 0, \dots, 0) = a_\varphi \underline{x}^*$.

Par suite $a \underline{x}^* - k \sum_{i=1}^p a_{1i}^2 = a_\varphi \underline{x}^* - k \sum_{i=1}^p a_{1i}^2 = a_\varphi \underline{x}^* - k \|a_\varphi\|^2 = b_1$.

De même la manière $k = \frac{b_2 - a_\psi \bar{x}^*}{\|a_\psi\|^2}$, $a \bar{x}^* = a_\psi \bar{x}^*$

$$\begin{aligned} \bar{x} &= \underline{x}^* + \bar{x}^* + k \left(\frac{a_{11}}{-1}, \dots, \frac{a_{1p}}{-1}, \frac{a_{1p+1}}{1}, \dots, \frac{a_{1n}}{1} \right) \\ &= \underline{x}^* + k(-a_{11}, \dots, -a_{1p}, 0, \dots, 0) + \bar{x}^* + k(0, \dots, 0, a_{1p+1}, \dots, a_{1n}) \end{aligned}$$

$$= \underline{x}^* + \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2} (-1) a_\varphi + \bar{x}^* - \frac{a_\psi \bar{x}^* - b_2}{\|a_\psi\|^2} (+1) a_\psi$$

$$= \left(\underline{x}^* - \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2} a_\varphi \right) + \left(\bar{x}^* - \frac{a_\psi \bar{x}^* - b_2}{\|a_\psi\|^2} a_\psi \right)$$

$$\Omega_1 = \{x \in \mathbb{R}^n : a_{11}x_1 + \dots + a_{1p}x_p = b_1\}$$

$$\Omega_2 = \{x \in \mathbb{R}^n : a_{1p+1}x_{p+1} + \dots + a_{1n}x_n = b_2\}$$

$$p_{\Omega_1}(\underline{x}^*) = \underline{x}^* - \frac{a_\varphi \underline{x}^* - b_1}{\|a_\varphi\|^2} a_\varphi ; p_{\Omega_2}(\bar{x}^*) = \bar{x}^* - \frac{a_\psi \bar{x}^* - b_2}{\|a_\psi\|^2} a_\psi.$$

Alors

$$\bar{x} = p_{\Omega_1}(\underline{x}^*) + p_{\Omega_2}(\bar{x}^*)$$

Notons par : $\bar{y} = \bar{x}_\varphi$ et $\bar{x} = \bar{x}_\psi$ donc $a_\varphi \bar{y} + a_\psi \bar{x} = b$

$$\Rightarrow \begin{cases} a_\varphi \bar{y} = b_1, \\ a_\psi \bar{x} = b_2. \end{cases}$$

Si $\bar{x} = 0$, on a donc $b_1 = b$.

Si $\bar{y} = 0$, on a donc $b_2 = b$.

Supposons qu'on sait \bar{x}_ψ .

$$\begin{cases} \max \varphi(x) \\ A_\varphi x_\varphi + A_\psi \bar{x} \leq b. \end{cases}$$

Si \bar{x}_φ n'est pas un sommet, $A_\varphi x + A_\psi \bar{x} \leq b$, on a

$$\bar{x}_\varphi = \underline{x}^* - \frac{a_\varphi \underline{x}^* - b + a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} a_\varphi.$$

$$\text{Car : } \bar{x}_\varphi a_\varphi + \bar{x}_\psi a_\psi = b \implies \bar{x}_\varphi a_\varphi = b - \bar{x}_\psi a_\psi$$

$$\bar{x}_\varphi = \frac{b - a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} \cdot a_\varphi \implies \bar{x}_\varphi = \underline{x}^* - \frac{a_\varphi \underline{x}^*}{\|a_\varphi\|^2} \cdot a_\varphi - \frac{-b + a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} \cdot a_\varphi \implies \bar{x}_\varphi = \underline{x}^* - \frac{a_\varphi \underline{x}^* - b + a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} \cdot a_\varphi$$

où

$$\bar{x}_\varphi = (\bar{x}_1, \dots, \bar{x}_p, 0, \dots, 0), \quad \bar{x}_\psi = (0, \dots, 0, \bar{x}_{p+1}, \dots, \bar{x}_n)$$

$$\bar{x} = \bar{x}_\varphi + \bar{x}_\psi$$

$$\bar{x}_\varphi = \underline{x}^* - \frac{a_\varphi \underline{x}^* - b}{\|a_\varphi\|^2} a_\varphi - \frac{a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} a_\varphi.$$

$$\bar{x}_\varphi = (\underline{x}^* - \frac{a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} a_\varphi) - \frac{a_\varphi \underline{x}^* - b}{\|a_\varphi\|^2} a_\varphi.$$

On pose $\hat{x}^* = \underline{x}^* - \frac{a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} a_\varphi$. Alors

$$\bar{x}_\varphi = \hat{x}^* - \frac{a_\varphi \underline{x}^* - b}{\|a_\varphi\|^2} a_\varphi.$$

Mais on a

$$a_\varphi \hat{x}^* = a_\varphi \underline{x}^* - \frac{a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} \|a_\varphi\|^2 = a_\varphi \underline{x}^* - a_\psi \bar{x}_\psi$$

$$\implies a_\varphi \underline{x}^* = a_\varphi \hat{x}^* + a_\psi \bar{x}_\psi$$

$$P_{\Omega_1}(\hat{x}^*) = \hat{x}^* - \frac{a_\varphi \hat{x}^* - b + a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} a_\varphi = \hat{x}^* - \frac{[a_\varphi \underline{x}^* - a_\psi \bar{x}_\psi] - b + a_\psi \bar{x}_\psi}{\|a_\varphi\|^2} a_\varphi$$

$$= \hat{x}^* - \frac{a_\varphi \underline{x}^* - b}{\|a_\varphi\|^2} a_\varphi = \bar{x}_\varphi.$$

D'où $\bar{x}_\varphi = P_{\Omega_1}(\hat{x}^*)$.

Donc le vecteur $\bar{x}_\varphi = (\bar{x}_1; \dots; \bar{x}_p, 0, \dots, 0)$, est projection orthogonale du vecteur \hat{x}^* sur Ω_1 .

■

Conclusion

Les algorithmes de résolution des problèmes d'optimisation quadratique sont très complexes et très lourds à mettre en œuvre pour les problèmes de grande taille ou pour la résolution en temps réel.

Des formes particulières de ce problème ont été étudiées pour réduire le temps du calcul. L'une de ces cas est l'application des opérateurs pour la résolution d'un problème d'optimisation quadratique [8].

En développant plusieurs exemples à l'aide de la méthode [7], nous avons montré la simplicité des calculs et surtout la précision des résultats obtenus après un nombre d'itérations très limité.

Nous avons explicité les conditions d'utilisation de cette méthode en démontrant l'existence des solutions et les techniques permettant de les calculer :

1. Décomposition de la fonction f en deux fonctions, φ et ψ ,
2. Optimisation de la fonction concave φ . \bar{x} est la projection orthogonale de x^* sur un nouveau convexe Ω' ;
3. L'extrémum de la fonction convexe ψ est un sommet ;
4. Si \bar{x} est un sommet de Ω , dans ce cas on a un algorithme de recherche de \bar{x} [3], et si \bar{x} n'est pas sommet de Ω on sait que \bar{x} est un point frontière. Alors il existe un hyperplan $a.x = b$ contenant \bar{x} donc si $\phi'(\bar{x}) \in [a]$, on peut calculer \bar{x} ;
5. On a rappelé le cas linéaire, en particulier la méthode du simplexe, et le problème dual ainsi que le rajout d'une contrainte.

Notons qu'il est parfois difficile de trouver la projection sur un polyèdre.

Bibliographie

- [1] S. Achmanov, Programmation linéaire, Edition Mir, Moscou, 1984.
- [2] A. Auslender, Optimisation Méthodes numériques, Masson, 1979.
- [3] A. Belabbaci, la programmation quadratique et son application à la programmation séparable, Mémoire de magister, Laghouat 2007.
- [4] H. Brezis, Analyse fonctionnelle, Théorie et applications, Dunod, 1999.
- [5] P. Caron, A. Juhel, Programmation linéaire méthodes et application, Dunod, 1988.
- [6] A. Chikhaoui, B. Djebbar, A. Belabbaci, A. Mokhtari, optimisation of a quadratic function under its canonical form, Journal of Applied Sciences, 444-510 ISSN 1996-3343, 2009.
- [7] P.G. Ciarlet, Introduction à l'analyse numérique matricielle et à l'optimisation , Dunod, 1998.
- [8] F. DEMENGEL, G. DEMENGEL, Convexité dans les espaces fonctionnels, ellipses, 2004.
- [9] J. Dieudonné, Algèbre linéaire et géométrie élémentaire, Herman, 1964
- [10] J. Dieudonné, Calcul infinitésimal, Hermene, 1997.
- [11] P.D. Bertsekas, Convex analysis and optimization, Athenta Scientific, 2002.

- [12] M. Dozzi, Méthodes non linéaires et stochastiques de la recherche opérationnelle, 2004.
- [13] J-B. Hiriart-Urruty, Analyse variationnelle et optimisation, cepadues, 2010.
- [14] M. Kadda, Optimisation d'une Fonction Quadratique Sous Contraintes Linéaires, Mémoire de magister, Tiaret 2010.
- [15] A. Kolmogorov, S. Fomine, Éléments de la théorie des fonctions et de l'analyse fonctionnelle, Moscou, 1973
- [16] C.M. Marle, P. Pilibossian, Calcul différentiel, ellipses, 1997.
- [17] M. Minoux, Programmation mathématique théorie et algorithmes.
- [18] M. Simmonard, Programmation linéaire. Vol 1 Fondements, Dunod, Paris, 1972.
- [19] M. Simmonard, Programmation linéaire. Vol 2 Extensions, Dunod, Paris, 1973.
- [20] G. R. Walsh, Methods of Optimization, Printed page Bros, Ltd Mile crois Lavie, Norwich, 1979.