

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique



UNIVERSITÉ D'EL-OUED

FACULTÉ DES SCIENCES ET DE TECHNOLOGIE

Mémoire de fin d'étude

LICENCE ACADEMIQUE

Domaine: Mathématiques et Informatique

Filière: Mathématiques

Spécialité: Modélisation mathématiques & simulation
numérique

Présenté par : Chihi Sabrina

Guediri Ouidad

Meguerhi Soumaya

Thème

Programmation Linéaire

Soutenu le ...juin 2014

Devant le jury composé de:

Mr. Nisse khadidja
Mr. Naoire aziza souad
Mlle. Miloudi Majda

MC (B) Univ. El Oued Président
MA (B) Univ. El Oued examinateur
MA(A) Univ. El Oued Rapporteur

Année universitaire 2013 – 2014

Remerciements

La louange est à Allah, qui nous a facilité l'accomplissement de ce travail de recherche chose ne peut être qu'avec la volonté de Dieu -à lui la toute puissance et la Majesté- et que la louange initiale et finale appartient à Allah, Seigneur des mondes

Aussi, il nous fait plaisir que nous, au commencement de ce travail, présentons nos grands remerciements, estimations et reconnaissances à notre encadreur puissant "Miloudi Majda" de nous avoir encourager moralement la durée de recherche et que ce travail est le fruit de ces encouragements

En fin, nous présentons nos veridiques remerciements à toute personne, du proche ou du loin, qui nous a donné un coup de main, à fin de terminer ce travail de recherche "Que la Grace et la paix soient sur notre prophète Muhammad ainsi que sur sa famille et ses compagnons".

Table des matières

Introduction générale	1
Notations et conventions	3
1 Généralités sur l'optimisation	4
1.1 Rappels et notations de calcul différentiel	4
1.1.1 Différentielle d'ordre 2, matrice hessienne	6
1.2 Optimisation sans contrainte	6
1.2.1 Conditions d'optimalité	7
1.2.2 Résultats d'existence et d'unicité	8
1.3 Optimisation sous contraintes	9
1.3.1 Existence-Unicité-Conditions d'optimalité simple	10
1.3.2 Conditions d'optimalité dans le cas de contraintes égalités	11
1.3.3 Conditions d'optimalité dans le cas de contraintes inégalités	13
2 Programmation linéaire	15
2.1 Problème de programmation linéaire	15
2.2 Généralités sur la programmation linéaire	15
2.3 Exemples de problèmes de programmation linéaire	18
2.4 La méthode du simplexe	19
2.5 Méthodes de dualité	27
2.6 Dualité et programmation linéaire	31
Bibliographie	34

Introduction générale

Nous étudions dans ce travail les problèmes de programmation linéaire. Un problème de programmation linéaire est un problème de minimisation d'une fonctionnelle linéaire sur un polyèdre de $\mathbb{R}^N$.

D'une façon plus précise, dans le premier chapitre. On trouvera:

-Généralités sur l'optimisation : la théorie qui concerne les problèmes d'optimisation sans et avec contraintes, les résultats d'existence et d'unicité de solutions de ces problèmes, caractérisation éventuelle d'une solution de ces problèmes, c'est-à-dire des conditions nécessaires et suffisantes dans certains cas.

Les conditions nécessaires font généralement intervenir la différentielle d'ordre 1 et 2. Pour la recherche des minimum locaux, on a des autres conditions nécessaires et suffisantes, on obtient aussi dans ce chapitre des conditions suffisantes dans les deux cas convenables, le cas des contraintes égalités, on a l'existence des multiplicateurs de Lagrange qui permet de déterminer le minimum d'une fonctionnelle sous ces contraintes. Aussi pour le cas des contraintes inégalités. On obtient une condition suffisante qui permet de déterminer le minimum d'une fonction sur ces contraintes (théorème de Khun-Tucker).

Dans le deuxième chapitre. Nous considérons un problème de programmation linéaire. En ce qui concerne l'existence et l'unicité d'une solution, les résultats antérieurement établis ne donnent aucun renseignement. En effet, la fonctionnelle f étant convexe, tout au plus peut-on-affirmer l'existence (mais non l'unicité) d'une solution dans le cas où l'ensemble $\mathbb{k}$ est borné.

Après avoir passé rapidement en revue au paragraphe 2.3 quelques exemples de problème de programmation linéaire (d'origine essentiellement « économique »), nous décrivons en détail au paragraphe 2.4 la méthode du simplexe. Cette méthode, d'emploi universel, est

tout à fait remarquable: Dans presque tout les cas, elle permet en effet à l'aide d'un nombre fini d'opérations élémentaires. Il peut apparaître néanmoins un phénomène de cyclage dans certains cas exceptionnels.

Au paragraphe 2.5, nous commençons par montrer que les résultats du chapitre précédent concernant la dualité sont d'une portée limitée lorsqu'on les applique au problème (p1); cette observation nous conduit ensuite à transformer ce dernier sous la forme équivalente (p2) puis à définir un nouveau type de problème dual (2.6.3). L'introduction de ces forme ((p2), (2.6.3)) de problèmes nous permet d'établir le résultat fondamental concernant la dualité en programmation linéaire (théorème 2.6.2). Nous précisons enfin (théorème 2.6.3) le lien, à priori inattendu mais pourtant réel, qui existe entre la méthode du simplexe et la dualité.

Notations et conventions

$C(\mathbb{R}^N, \mathbb{R}) :$	l'ensemble des fonctions continues de $\mathbb{R}^N$ dans $\mathbb{R}$.
$C^1(E, F) :$	l'ensemble des fonctions continument différentiable d'ordre 1.
$C^2(E, F) :$	l'ensemble des fonctions continument différentiable d'ordre 2.
$\mathcal{L}(E, F) :$	l'ensemble des applications linéaires de E dans F .
$E, F :$	deux espaces vectoriels normés.
$T_x :$	la différentielle de f au point x .
$\nabla f(x) :$	le gradient de f .
$\partial f(x) :$	la dérivée partielle de f .
$H_f(x) :$	la matrice hessienne de f en x .
$J_f(x) :$	la matrice jacobienne de f en x .
$M_{p,N}(\mathbb{R}) :$	l'espace des matrices de p lignes et N colonnes.
$\ \cdot\ :$	la norme.
$ \cdot :$	la valeur absolue.
$\mathbb{k} :$	sous ensemble de $\mathbb{R}^N$.
$B(\bar{x}, \alpha) :$	la boule ouverte de centre $\bar{x}$ et de rayon α .
$\text{rang}(C) :$	rang de la matrice C .
$\text{card}(I^*(\bar{x})) :$	cardinal d'une partie $I^*(\bar{x})$.
$(\cdot, \cdot)_N :$	produit scalaire eüclidien sur $\mathbb{R}^N$.
$\inf_{x \in \mathbb{k}} f(x) :$	la borne inferieur de fonction f sur l'ensemble $\mathbb{k}$.
$\sup_{x \in \mathbb{k}} f(x) :$	la borne supérieur de fonction f sur l'ensemble $\mathbb{k}$.
$\min :$	le minimum.
$\max :$	le maximum.

Chapitre 1

Généralités sur l'optimisation

L'objectif de ce chapitre est de rechercher des minima ou des maxima d'une fonction continue avec ou sans contrainte.

Pour commencer l'étude, on passe aux principales définitions et les résultats fondamentaux du calcul différentiel dans les espaces vectoriels normés.

Après avoir rappelé des conditions nécessaires et suffisantes d'extremum relatif ou local.

On donne aussi quelques résultats d'existence et d'unicité d'extremum.

1.1 Rappels et notations de calcul différentiel

Définition 1.1.1 (*Application différentiable*). Soient E et F deux espaces vectoriels normés, f une application de E dans F et $x \in E$. On rappelle que f est différentiable en x s'il existe $T_x \in \mathcal{L}(E, F)$ (où $\mathcal{L}(E, F)$ est l'ensemble des applications linéaires de E dans F) telle que:

$$f(x + h) = f(x) + T_x(h) + \|h\| \varepsilon(h) \text{ avec } \varepsilon(h) \rightarrow 0 \text{ quand } \|h\| \rightarrow 0 \quad (1.1.1)$$

L'application T_x est alors unique et on note $Df(x) = T_x \in \mathcal{L}(E, F)$ la différentielle de f au point x . Si f est différentiable en tout point de E , alors on appelle différentielle de f l'application $Df : E \rightarrow \mathcal{L}(E, F)$ qui à $x \in E$ associe l'application linéaire $Df(x)$, de E dans F .

Remarque 1.1.1 Remarquons tout de suite que si f est une application linéaire de E dans F , alors f est différentiable, et $Df = f$. En effet, si f est linéaire, $f(x+h) - f(x) = f(h)$, et donc l'égalité (1.1.1) est vérifiée avec $T_x = f$ et $\varepsilon = 0$. Voyons maintenant quelques cas particuliers d'espaces E et F :

Cas où $E = \mathbb{R}$ et $F = \mathbb{R}$. Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, dire que f est différentiable en x revient à dire que f est dérivable en x . En effet, dire que f est dérivable en x revient à dire que: $\lim_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h}$ existe, et $\lim_{h \rightarrow 0} \frac{f(x+h)-f(x)}{h} = f'(x)$, ce qui s'écrit encore

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \varepsilon(h),$$

avec $\varepsilon(h) \rightarrow 0$ lorsque $h \rightarrow 0$. C'est-à-dire $f(x+h) - f(x) = T_x(h) + h\varepsilon(h)$; avec $T_x(h) = f'(x)h$, ce qui revient à dire que f est différentiable en x , et que sa différentielle en x est l'application linéaire $T_x : \mathbb{R} \rightarrow \mathbb{R}$, qui à h associe $f'(x)h$. On a ainsi vérifié que pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$, la notion de différentielle coïncide avec celle de dérivée.

Cas où $E = \mathbb{R}^N$ et $F = \mathbb{R}^p$ Soit $f : \mathbb{R}^N \rightarrow \mathbb{R}^p$, $x \in \mathbb{R}^N$ et supposons que f est différentiable en x ; alors $Df(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^p)$, et il existe une unique matrice $J_f(x) \in M_{p,N}(\mathbb{R})$ telle que: $\underbrace{Df(x)}_{\in \mathbb{R}^p}(y) = \underbrace{J_f(x)}_{\in \mathbb{R}^p}y, \forall y \in \mathbb{R}^N$. On confond alors l'application linéaire $Df(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^p)$ avec la matrice $J_f(x) \in M_{p,N}(\mathbb{R})$ qui la représente. On écrit donc: $J_f(x) = Df(x) = (a_{i,j})_{1 \leq i \leq p, 1 \leq j \leq N}$ où $a_{i,j} = \partial_j f_i(x)$. ∂_j désignant la dérivée partielle par rapport à la j -ème variable.

La matrice $J_f(x)$ s'appelle la matrice jacobienne de f au point x (d'où la notation J_f). Notons que si $N = p = 1$, la fonction f est de $\mathbb{R}$ dans $\mathbb{R}$ et la matrice jacobienne en x n'est autre que la dérivée en x : $J_f(x) = f'(x)$. Voir [4]

Définition 1.1.2 (Point critique). Soit $f : \mathbb{R}^N \rightarrow \mathbb{R}$ une fonction différentiable. Tout vecteur $x \in \mathbb{R}^N$ tel que $\nabla f(x) = 0$ est appelé point critique ou point stationnaire de f .

Définition 1.1.3 (Extremum d'une fonction). Soit E un espace vectoriel et $f : E \rightarrow \mathbb{R}$. On dit que $\bar{x}$ est un minimum local de f , s'il existe un voisinage V de $\bar{x}$ tel que:

$$f(\bar{x}) \leq f(x), \forall x \in V.$$

De même, on dit que $\bar{x}$ est un maximum local de f s'il existe un voisinage V de $\bar{x}$ tel que:

$$f(x) \leq f(\bar{x}), \forall x \in V.$$

On dit que $\bar{x}$ est un minimum global de f si

$$f(x) \geq f(\bar{x}), \forall x \in E.$$

De même, on dit que $\bar{x}$ est un maximum global de f si

$$f(x) \leq f(\bar{x}), \forall x \in E.$$

On dit que $\bar{x}$ est un extremum local (resp. global) de f si c'est un minimum local (resp. global) ou un maximum local (resp. global). Voir [4]

1.1.1 Différentielle d'ordre 2, matrice hessienne

Revenons maintenant au cas général de deux espaces vectoriels normés E et F , et supposons maintenant que $f \in C^2(E, F)$. Le fait que $f \in C^2(E, F)$ signifie que: $Df \in C^1(E, \mathcal{L}(E, F))$. Par définition, on a $D^2f(x) \in \mathcal{L}(E, \mathcal{L}(E, F))$ et donc pour $y \in E$, $D^2f(x)(y) \in \mathcal{L}(E, F)$; en particulier, pour $z \in E$, $D^2f(x)(y)(z) \in F$. Considérons maintenant le cas particulier $E = \mathbb{R}^N$ et $F = \mathbb{R}$. On a:

$$f \in C^2(\mathbb{R}^N, \mathbb{R}) \iff [f \in C^1(\mathbb{R}^N, \mathbb{R}) \text{ et } \nabla f \in C^1(\mathbb{R}^N, \mathbb{R}^N)].$$

Soit $g = \nabla f \in C^1(\mathbb{R}^N, \mathbb{R}^N)$, et $x \in \mathbb{R}^N$ alors $Dg(x) \in M_N(\mathbb{R})$ et on peut définir la matrice hessienne de f , qu'on note H_f , par:

$$H_f(x) = Dg(x) = D(Df)(x) = (b_{i,j})_{i,j=1\dots N} \in M_N(\mathbb{R}).$$

où $b_{i,j} = \partial^2_{i,j} f(x)$ où $\partial^2_{i,j}$ désigne la dérivée partielle par rapport à la variable i de la dérivée partielle par rapport à la variable j . Notons que par définition, $Dg(x)$ est la matrice jacobienne de g en x . Voir [3]

1.2 Optimisation sans contrainte

Définition 1.2.1 Soit $f \in C(E, \mathbb{R})$ et E un espace vectoriel normé (le plus souvent $\mathbb{R}^N$).

On cherche soit un minimum global de f , ce type de problème s'écrit:

$$\begin{cases} \bar{x} \in E \\ f(\bar{x}) = \inf_E f \end{cases} \quad (1.2.1)$$

1.2.1 Conditions d'optimalité

Condition nécessaire d'optimalité du 1^{er} ordre

Théorème 1.2.1 *soit E un espace vectoriel normé et soint $f \in C(E, \mathbb{R})$ et $\bar{x} \in E$ telle que f est différentiable en $\bar{x}$. Si $\bar{x}$ est solution de (1.2.1), alors $Df(\bar{x}) = 0$.*

Preuve. Supposons qu'il existe $\alpha > 0$, tel que $f(\bar{x}) \leq f(y)$, pour tout $y \in B(\bar{x}, \alpha)$ soit $z \in E \setminus \{0\}$ alors si $|t| < \frac{\alpha}{\|z\|}$ on a: $\bar{x} + tz \in B(\bar{x}, \alpha)$ (où $B(\bar{x}, \alpha)$ est la boule ouverte de centre $\bar{x}$ et de rayon α) et on a: $f(\bar{x}) \leq f(\bar{x} + tz)$ comme f est différentiable en $\bar{x}$. donc: $f(\bar{x} + tz) = f(\bar{x}) + Df(\bar{x})(tz) + |t| \varepsilon_z(t)$ où $\varepsilon_z(t) \rightarrow 0$ lorsque $t \rightarrow 0$, on a donc: $f(\bar{x}) + tDf(\bar{x})(z) + |t| \varepsilon_z(t) \geq f(\bar{x})$, pour $0 < t < \frac{\alpha}{\|z\|}$: $Df(\bar{x})(z) + \varepsilon_z(t) \geq 0$. On prend $t \rightarrow 0$ alors $Df(\bar{x})(z) \geq 0, \forall z \in E$, on a aussi $Df(\bar{x})(-z) \geq 0 \implies -Df(\bar{x})(z) \geq 0 \implies Df(\bar{x})(z) \leq 0, \forall z \in E \implies Df(\bar{x})(z) = 0, \forall z \in E \implies Df(\bar{x}) = 0$. ■

Remarque 1.2.1 *Attention, la proposition précédente donne une condition nécessaire mais non suffisante. En effet, $Df(\bar{x}) = 0$ n'entraîne pas que f atteigne un minimum (ou un maximum) même local, en $\bar{x}$. Prendre par exemple $E = \mathbb{R}$, $\bar{x} = 0$ et la fonction f définie par: $f(x) = x^3$ pour s'en convaincre. Voir [4].*

Condition nécessaire d'optimalité du 2^{ème} ordre

Théorème 1.2.2 *Soit E un espace vectoriel normé et soint $f \in C(E, \mathbb{R})$, f est deux fois différentiable en $\bar{x}$ et $\bar{x}$ est solution de (1.2.1), alors:*

$$D^2f(\bar{x})x \cdot x \geq 0, \forall x \in E.$$

($D^2f(\bar{x})$ est semi défini positive).

Preuve. On prend la preuve précédente on a $Df(\bar{x}) = 0$, $Df(\bar{x})(z) + \varepsilon_z(t) \geq 0$ et la formule de Taylor à l'ordre 2 permet d'écrire pour tout $x \in E$: $D^2f(\bar{x})x \cdot x \geq 0$ (pour le passage à la limite de $t \rightarrow 0$). ■

Condition suffisantes d'optimalité

Théorème 1.2.3 *Soit une fonction $f : \mathbb{R}^N \rightarrow \mathbb{R}$, deux fois différentiable dans un sous-ensemble ouvert V de $\mathbb{R}^N$, et soit $\bar{x} \in V$ qui vérifie les conditions $\nabla f(\bar{x}) = 0$, et $\nabla^2 f(\bar{x})$ est définie positive. Alors, $\bar{x}$ est un minimum local strict de f . Voir [?]*

1.2.2 Résultats d'existence et d'unicité

Théorème 1.2.4 (Existence). Soit $E = \mathbb{R}^N$ et $f : E \rightarrow \mathbb{R}$ une application telle que

(i) f est continue.

(ii) $f(x) \rightarrow +\infty$ quand $\|x\| \rightarrow +\infty$.

Alors il existe $\bar{x} \in \mathbb{R}^N$ tel que $f(\bar{x}) \leq f(y)$ pour tout $y \in \mathbb{R}^N$.

Preuve. La condition (ii) peut encore s'écrire

$$\forall A \in \mathbb{R}, \exists R \in \mathbb{R}; \|x\| \geq R \implies f(x) > A. \quad (1.2.2)$$

On écrit (1.2.2) avec $A = f(0)$. On obtient alors: $\exists R \in \mathbb{R}$ tel que $\|x\| \geq R \implies f(x) \geq f(0)$

On en déduit que $\inf_{\mathbb{R}^N} f = \inf_{B_R} f$, où $B_R = \{x \in \mathbb{R}^N; \|x\| \leq R\}$. Or, B_R est un compact de $\mathbb{R}^N$ et f est continue donc il existe $\bar{x} \in B_R$ tel que $f(\bar{x}) = \inf_{B_R} f$ et donc $f(\bar{x}) = \inf_{\mathbb{R}^N} f$. ■

Remarque 1.2.2 1. Le théorème est faux si E est de dimension infinie (i.e. si E est espace de Banach au lieu de $E = \mathbb{R}^N$), car si E est de dimension infinie, B_R n'est pas compacte.

2. L'hypothèse (ii) du théorème peut être remplacée par

$$(ii) \exists b \in \mathbb{R}^N, \exists R > 0 \text{ tel que } \|x\| \geq R \implies f(x) \geq f(b).$$

3. Sous les hypothèses du théorème il n'y a pas toujours unicité de $\bar{x}$ même dans le cas $N = 1$, prendre pour s'en convaincre la fonction f définie de $\mathbb{R}$ dans $\mathbb{R}$ par $f(x) = x^2(x-1)(x+1)$.

Convexité

Définition 1.2.2 Soit E un espace vectoriel et $f : E \rightarrow \mathbb{R}$ On dit que f est convexe si: $\forall (x, y) \in E^2$ tel que: $x \neq y$ et $t \in [0, 1]$,

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

On dit que f est strictement convexe si $\forall (x, y) \in E^2$ tel que: $x \neq y$ et $\forall t \in]0, 1[$:

$$f(tx + (1-t)y) < tf(x) + (1-t)f(y).$$

Voir [4]

Condition suffisante d'unicité

Théorème 1.2.5 Soit E un espace vectoriel normé et $f : E \rightarrow \mathbb{R}$ strictement convexe, alors il existe au plus un $\bar{x} \in E$ tel que: $f(\bar{x}) \leq f(y), \forall y \in E$.

Preuve. Soit f strictement convexe, supposons qu'il existe $\bar{x}$ et $\bar{\bar{x}} \in E$ tels que $f(\bar{x}) = f(\bar{\bar{x}}) = \inf_{\mathbb{R}^N} f$. Comme f est strictement convexe, si $\bar{x} \neq \bar{\bar{x}}$ alors $f(\frac{1}{2}\bar{x} + \frac{1}{2}\bar{\bar{x}}) < \frac{1}{2}f(\bar{x}) + \frac{1}{2}f(\bar{\bar{x}}) = \inf_{\mathbb{R}^N} f$, ce qui est impossible ; donc $\bar{x} = \bar{\bar{x}}$. ■

Remarque 1.2.3 Ce théorème ne donne pas l'existence. Par exemple dans le cas $N = 1$ la fonction f définie par $f(x) = \exp(x)$ n'atteint pas son minimum; en effet, $\inf_{\mathbb{R}} f = 0$ et $f(x) \neq 0$ pour tout $x \in \mathbb{R}$, et pourtant f est strictement convexe. Voir [4]

1.3 Optimisation sous contraintes

Définition 1.3.1 Soit $E = \mathbb{R}^N$, soit $f \in C(E, \mathbb{R})$, et soit $\mathbb{k}$ un sous ensemble de E . On s'intéresse à la recherche de $\bar{x} \in \mathbb{k}$ tel que:

$$\begin{cases} \bar{x} \in \mathbb{k} \\ f(\bar{x}) = \inf_{\mathbb{k}} f \end{cases} \quad (1.3.1)$$

Ce problème est un problème de minimisation avec contrainte (ou "sous contrainte") au sens où l'on cherche $\bar{x}$ qui minimise f en astreignant $\bar{x}$ à être dans $\mathbb{k}$. Voyons quelques exemples de ces contraintes (définies par l'ensemble $\mathbb{k}$), qu'on va expliciter à l'aide des p fonctions continues, $g_i \in C(E, \mathbb{R}), i = 1, \dots, p$.

1. Contraintes égalités: On pose $\mathbb{k} = \{x \in E, g_i(x) = 0, i = 1, \dots, p\}$. On verra plus loin que le problème de minimisation de f peut alors être résolu grâce au théorème des "multiplicateurs de Lagrange".

2. Contraintes inégalités: On pose $\mathbb{k} = \{x \in E, g_i(x) \leq 0, i = 1, \dots, p\}$. On verra plus loin que le problème de minimisation de f peut alors être résolu grâce au théorème de "Kuhn-Tucker".

– Programmation linéaire: Avec un tel ensemble de contraintes $\mathbb{k}$, si de plus f est linéaire, c'est-à-dire qu'il existe $b \in \mathbb{R}^N$ tel que $f(x) = b \cdot x$, et les fonctions g_i sont affines, c'est-à-dire

qu'il existe $b_i \in \mathbb{R}^N$ et $C_i \in \mathbb{R}$ tels que $g_i(x) = b_i \cdot x + C_i$, alors on dit qu'on a affaire un problème de "programmation linéaire". Ces problèmes sont souvent résolus numériquement à l'aide de l'algorithme de Dantzig, inventé vers 1950.

– Programmation quadratique: Avec le même ensemble de contraintes $\mathbb{k}$, si de plus f est quadratique, c'est-à-dire si f est de la forme: $f(x) = \frac{1}{2}Ax \cdot x - b \cdot x$, et les fonctions g_i sont affines, alors on dit qu'on a affaire un problème de "programmation quadratique".

– Programmation convexe: Dans le cas où f est convexe et $\mathbb{k}$ est convexe, on dit qu'on a affaire un problème de "programmation convexe".

1.3.1 Existence-Unicité-Conditions d'optimalité simple

Théorème 1.3.1 (Existence). Soit $E = \mathbb{R}^N$ et $f \in C(E, \mathbb{R})$:

1. Si $\mathbb{k}$ est un sous-ensemble fermé borné de E , alors il existe $\bar{x} \in \mathbb{k}$ tel que $f(\bar{x}) = \inf_{\mathbb{k}} f$.
2. Si $\mathbb{k}$ est un sous-ensemble fermé de E , et si f est croissante à l'infini, c'est-à-dire que $f(x) \rightarrow +\infty$ quand $\|x\| \rightarrow +\infty$, alors $\exists \bar{x} \in \mathbb{k}$ tel que: $f(\bar{x}) = \inf_{\mathbb{k}} f$.

Preuve. 1. Si $\mathbb{k}$ est un sous-ensemble fermé borné de E , comme f est continue, elle atteint ses bornes sur $\mathbb{k}$, d'où l'existence de $\bar{x}$.

2. Si f est croissante à l'infini, alors il existe $R > 0$ tel que si $\|x\| > R$ alors $f(x) > f(0)$; donc $\inf_{\mathbb{k}} f = \inf_{\mathbb{k} \cap B_R} f$, où B_R désigne la boule de centre 0 et de rayon R . L'ensemble $\mathbb{k} \cap B_R$ est compact, car intersection d'un fermé et d'un compact. Donc, par ce qui précède, il existe $\bar{x} \in \mathbb{k}$ tel que $f(\bar{x}) = \inf_{\mathbb{k} \cap B_R} f = \inf_{B_R} f$. ■

Théorème 1.3.2 (Unicité). Soit $E = \mathbb{R}^N$ et $f \in C(E, \mathbb{R})$. On suppose que f est strictement convexe et que $\mathbb{k}$ est convexe. Alors il existe au plus un élément $\bar{x}$ de $\mathbb{k}$ tel que $f(\bar{x}) = \inf_{\mathbb{k}} f$.

Preuve. Supposons que $\bar{x}$ et $\bar{\bar{x}}$ soient deux solutions du problème (1.3.1), avec $\bar{x} \neq \bar{\bar{x}}$. Alors $f\left(\frac{1}{2}\bar{x} + \frac{1}{2}\bar{\bar{x}}\right) < \frac{1}{2}f(\bar{x}) + \frac{1}{2}f(\bar{\bar{x}}) = \inf_{\mathbb{k}} f$. On aboutit donc à une contradiction. Voir [4] ■

Remarque 1.3.1 $\mathbb{k}$ est convexe $\Leftrightarrow \forall x, y \in \mathbb{k}, \forall t \in [0, 1], tx + (1 - t)y \in \mathbb{k}$.

Exemple 1.3.1 (Sur l'existence et l'unicité)

soit $E = \mathbb{R}$, $f : \mathbb{R} \longrightarrow \mathbb{R}$ est définie par $f(x) = x^2$, et pour les deux ensembles $\mathbb{k}$ suivants:

(i) $\mathbb{k} = \{|x| \leq 1\}$, (ii) $\mathbb{k} = \{|x| = 1\}$. la fonction f est continue, strictement convexe, et croissante à l'infini. Etudions maintenant les propriétés de $\mathbb{k}$ dans les deux cas proposés:

(i) L'ensemble $\mathbb{k} = \{|x| \leq 1\}$, est fermé borné et convexe. On peut donc appliquer le théorème d'existence et le théorème d'unicité. En remarquant que $f(x) \geq 0$ pour tout $x \in \mathbb{R}$ et que $f(0) = 0$, on en déduit que l'unique solution du problème (1.3.1) est donc $\bar{x} = 0$.

(ii) L'ensemble $\mathbb{k} = \{|x| = 1\}$, est fermé borné mais non convexe. De fait, on peut remarquer que $\mathbb{k} = \{-1, 1\}$, et donc $\{f(x), x \in \mathbb{k}\} = \{1\}$. Il existe donc deux solutions du problème (1.3.1): $\bar{x}_1 = 1$ et $\bar{x}_2 = -1$.

Condition simple d'optimalité

Proposition 1.3.1 Soient $E = \mathbb{R}^N$, $f \in C(E, \mathbb{R})$ et $\bar{x} \in \mathbb{k}$ tel que $f(\bar{x}) = \inf_{\mathbb{k}} f$. On suppose que f est différentiable en $\bar{x}$

1. Si $\bar{x} \in \overset{\circ}{\mathbb{k}}$, alors $\nabla f(\bar{x}) = 0$.
2. Si $\mathbb{k}$ est convexe, alors $\nabla f(\bar{x}) \cdot (x - \bar{x}) \geq 0$ pour tout $x \in \mathbb{k}$.

Preuve. 1. Si $\bar{x} \in \overset{\circ}{\mathbb{k}}$, alors il existe $\varepsilon > 0$, tel que $B(\bar{x}, \varepsilon) \subset \mathbb{k}$ et $f(\bar{x}) \leq f(x), \forall x \in B(\bar{x}, \varepsilon)$. Alors on a déjà vu que ceci implique $\nabla f(\bar{x}) = 0$.

2. Soit $x \in \mathbb{k}$. Comme $\bar{x}$ réalise le minimum de f sur $\mathbb{k}$, on a:

$$f(\bar{x} + t(x - \bar{x})) = f(tx + (1 - t)\bar{x}) \geq f(\bar{x}),$$

pour tout $t \in]0, 1]$, par convexité de $\mathbb{k}$. On en déduit que

$$\frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \geq 0, \text{ pour tout } t \in]0, 1].$$

.En passant à la limite lorsque t tend vers 0 dans cette dernière inégalité, on obtient: $\nabla f(\bar{x}) \cdot (x - \bar{x}) \geq 0$. Voir [4] ■

1.3.2 Conditions d'optimalité dans le cas de contraintes égalités

Dans tout ce paragraphe, on considèrera les hypothèses et notations suivantes:

$$\begin{aligned} f &\in C(\mathbb{R}^N, \mathbb{R}), g_i \in C^1(\mathbb{R}^N, \mathbb{R}), i = 1, \dots, p; \\ \mathbb{k} &= \{x \in \mathbb{R}^N, g_i(x) = 0, \forall i = 1, \dots, p\}; \\ g &= (g_1, \dots, g_p)^t \in C^1(\mathbb{R}^N, \mathbb{R}^p). \end{aligned} \quad (1.3.2)$$

Remarque 1.3.2 (Quelques rappels de calcul différentiel)

Comme $g \in C^1(\mathbb{R}^N, \mathbb{R}^p)$, si $x \in \mathbb{R}^N$, alors $Dg(x) \in \mathcal{L}(\mathbb{R}^N, \mathbb{R}^p)$, ce qui revient à dire, en confondant l'application linéaire $Dg(x)$ avec sa matrice, que $Dg(x) \in M_{p,N}(\mathbb{R})$. Par définition, $\text{Im}(Dg(x)) = \{Dg(x) \cdot z, z \in \mathbb{R}^N\} \subset \mathbb{R}^p$, et $\text{rang}(Dg(x)) = \dim(\text{Im}(Dg(x))) \leq p$. On rappelle de plus que

$$Dg(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}, & \dots, & \frac{\partial g_1}{\partial x_N} \\ \dots, & \dots, & \dots \\ \frac{\partial g_p}{\partial x_1}, & \dots, & \frac{\partial g_p}{\partial x_N} \end{pmatrix}$$

et que $\text{rang}(Dg(x)) \leq \min(N, p)$. De plus, si $\text{rang}(Dg(x)) = p$, alors les vecteurs $Dg_i(x)$, $i = 1, \dots, p$, sont linéairement indépendants dans $\mathbb{R}^N$.

Théorème 1.3.3 (Multiplicateurs de Lagrange). Soit $\bar{x} \in \mathbb{K}$ tel que $f(\bar{x}) = \inf_{\mathbb{K}} f$.

On suppose que f est différentiable en $\bar{x}$ et $\dim(\text{Im}(Dg(\bar{x}))) = p$ (ou $\text{rang}(Dg(\bar{x})) = p$), alors: il existe $(\lambda_1, \dots, \lambda_p)^t \in \mathbb{R}^p$ tels que:

$$\nabla f(\bar{x}) + \sum_{i=1}^p \lambda_i \nabla g_i(\bar{x}) = 0.$$

(Cette dernière égalité a lieu dans $\mathbb{R}^N$).

Remarque 1.3.3 (Utilisation pratique du théorème de Lagrange) Soit $f \in C^1(\mathbb{R}^N, \mathbb{R})$, $g = (g_1, \dots, g_p)^t$ avec $g_i \in C^1(\mathbb{R}^N, \mathbb{R})$ pour: $i = 1, \dots, p$, et soit

$$\mathbb{K} = \{x \in \mathbb{R}^N, g_i(x) = 0, i = 1, \dots, p\}.$$

Le problème qu'on cherche à résoudre est le problème de minimisation (1.3.1) qu'on rappelle ici:

$$\begin{cases} \bar{x} \in \mathbb{K} \\ f(\bar{x}) = \inf_{\mathbb{K}} f \end{cases}$$

D'après le théorème des multiplicateurs de Lagrange, si $\bar{x}$ est solution de (1.3.1) et $\text{Im}(Dg(\bar{x})) = \mathbb{R}^p$, alors il existe $(\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$, tel que $\bar{x}$ est solution du problème:

$$\begin{cases} \frac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^p \lambda_i \frac{\partial g_i(\bar{x})}{\partial x_j} = 0, j = 1, \dots, N \\ g_i(\bar{x}) = 0, i = 1, \dots, p \end{cases}. \quad (1.3.3)$$

Le système (1.3.3) est un système non linéaire de $(N + p)$ équations et à $(N + p)$ inconnues $(\bar{x}_1, \dots, \bar{x}_N, \lambda_1, \dots, \lambda_p)$. Ce système sera résolu par une méthode de résolution de système non linéaire (Newton par exemple). Voir [4]

Remarque 1.3.4 On vient de montrer que si $\bar{x}$ solution de (1.3.1) et $\text{Im}(Dg(\bar{x})) = \mathbb{R}^p$, alors $\bar{x}$ solution de (1.3.3). Par contre, si $\bar{x}$ est solution de (1.3.3), ceci n'entraîne pas que $\bar{x}$ est solution de (1.3.1).

1.3.3 Conditions d'optimalité dans le cas de contraintes inégalités

Soit $f \in C(\mathbb{R}^N, \mathbb{R})$ et $g_i \in C^1(\mathbb{R}^N, \mathbb{R})$ $i = 1, \dots, p$, on considère maintenant un ensemble $\mathbb{k}$ de la forme:

$$\mathbb{k} = \{x \in \mathbb{R}^N, g_i(x) \leq 0, \forall i = 1, \dots, p\},$$

et on cherche à résoudre le problème de minimisation (1.3.1) qui s'écrit:

$$\begin{cases} \bar{x} \in \mathbb{k} \\ f(\bar{x}) \leq f(x), \forall x \in \mathbb{k} \end{cases}$$

Remarque 1.3.5 Soit $\bar{x}$ une solution de (1.3.1) et supposons que $g_i(\bar{x}) < 0$, pour tout $i \in \{1, \dots, p\}$. Il existe alors: $\varepsilon > 0$ tel que si $x \in B(\bar{x}, \varepsilon)$ alors $g_i(x) < 0$, pour tout $i = 1, \dots, p$. On a donc $f(\bar{x}) \leq f(x), \forall x \in B(\bar{x}, \varepsilon)$. On est alors ramené à un problème de minimisation sans contrainte, et ssi f est différentiable en $\bar{x}$, on a donc $\nabla f(\bar{x}) = 0$. On donne maintenant sans démonstration le théorème de Kuhn-Tucker qui donne une caractérisation de la solution du problème (1.3.1).

Théorème 1.3.4 (Kuhn-Tucker). Soit $f \in C(\mathbb{R}^N, \mathbb{R})$, soit $g_i \in C^1(\mathbb{R}^N, \mathbb{R})$; pour $i = 1, \dots, p$, et soit $\mathbb{k} = \{x \in \mathbb{R}^N, g_i(x) \leq 0, \forall i = 1, \dots, p\}$. On suppose qu'il existe $\bar{x}$ solution de (1.3.1), et on pose $I(\bar{x}) = \{i \in \{1, \dots, p\}; g_i(\bar{x}) = 0\}$. On suppose que f est différentiable en $\bar{x}$ et que la famille $(\text{de } \mathbb{R}^N) \{\nabla g_i(\bar{x}), i \in I(\bar{x})\}$ est libre. Alors il existe une famille $(\lambda_i)_{i \in I(\bar{x})} \subset \mathbb{R}_+$ telle que:

$$\nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x}) = 0.$$

Remarque 1.3.6 1. Le théorème de Kuhn-Tucker s'applique pour des ensembles de contrainte de type inégalité. Si on a une contrainte de type égalité, on peut évidemment se ramener à deux contraintes de type inégalité en remarquant que

$$\mathbb{K} = \{x \in \mathbb{R}^N, g(x) = 0\} = \{x \in \mathbb{R}^N, g_1(x) \leq 0 \text{ et } g_2(x) \leq 0\}.$$

Cependant, si on pose ($g_1 = g$ et $g_2 = -g$), on remarque que la famille $\{\nabla g_1(\bar{x}), \nabla g_2(\bar{x})\} = \{\nabla g(\bar{x}), -\nabla g(\bar{x})\}$ n'est pas libre. On ne peut donc pas appliquer le théorème de Kuhn-Tucker sous la forme donnée précédemment dans ce cas (mais on peut il existe des versions du théorème de Kuhn-Tucker).

2. Dans la pratique, on a intérêt à écrire la conclusion du théorème de Kuhn-Tucker (i.e. l'existence de la famille $(\lambda_i)_{i \in I(\bar{x})}$ sous la forme du système de $N + p$ équations et $2p$ inéquations à résoudre suivant:

$$\left\{ \begin{array}{l} \nabla f(\bar{x}) + \sum_{i=1}^p \lambda_i \nabla g_i(\bar{x}) = 0 \\ \lambda_i g_i(\bar{x}) = 0, i = 1, \dots, p \\ g_i(\bar{x}) \leq 0, i = 1, \dots, p \\ \lambda_i \geq 0, i = 1, \dots, p \end{array} \right. .$$

Voir [4].

Chapitre 2

Programmation linéaire

On s'intéresse dans ce chapitre, aux problèmes de programmation linéaire.

On étudie l'existence et l'unicité de solutions de ces problèmes et on donne aussi quelques méthodes de résolution de ce type des problèmes.

2.1 Problème de programmation linéaire

Définition 2.1.1 *Un problème de programmation linéaire se présente sous la forme: trouver $\bar{x}$ tel que*

$$\begin{cases} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}^N, Cx \leq d\}, C \in A_{M,N}(\mathbb{R}), d \in \mathbb{R}^M \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = (a, x)_N \end{cases}.$$

Remarque 2.1.1 *Dans un problème de programmation linéaire, la fonction f est une fonctionnelle linéaire sur $\mathbb{R}^N$: $f(x) = \sum_{i=1}^N a_i x_i$, et l'ensemble $\mathbb{k}$ est encore de la forme*

$$\mathbb{k} = \left\{ x \in \mathbb{R}^N; \sum_{j=1}^N C_{ij} x_j \leq d_i, 1 \leq i \leq M \right\}.$$

2.2 Généralités sur la programmation linéaire

Un problème de programmation linéaire correspond à la minimisation d'une fonctionnelle linéaire:

$$f : x \in \mathbb{R}^N \longrightarrow f(x) = \sum_{i=1}^N a_i x_i = (a, x)_N.$$

où $a = (a_i)_{1 \leq i \leq N}$ est un vecteur donné de $\mathbb{R}^N$, lorsque le vecteur x décrit un ensemble de la forme

$$\mathbb{k} = \{x \in \mathbb{R}^N; \sum_{j=1}^N c_{ij} x_j \leq d_i, 1 \leq i \leq M\} = \{x \in \mathbb{R}^N; Cx \leq d\},$$

où $C = (C_{ij})_{\substack{1 \leq i \leq M \\ 1 \leq j \leq N}}$ est une matrice donnée de type (M, N) , et $d = (d_i)_{1 \leq i \leq M}$ est un vecteur de $\mathbb{R}^M$.

On peut en effet toujours se ramener à cette forme puisque toute égalité $\sum_{j=1}^N c'_{ij} x_j = d'_i$ apparaissant éventuellement dans la définition de l'ensemble $\mathbb{k}$ peut être remplacée par les deux inégalités:

$$\sum_{j=1}^N c'_{ij} x_j \leq d'_i \text{ et } \sum_{j=1}^N (-c'_{ij}) x_j \leq -d'_i.$$

Selon le type de question étudiée, nous serons amenés à distinguer trois formes canoniques d'un problème de programmation linéaire : trouver $\bar{x}$ tel que

$$\left\{ \begin{array}{l} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}^N; \sum_{j=1}^N C_{ij} x_j \leq d_i, 1 \leq i \leq M\} \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = \sum_{i=1}^N a_i x_i \end{array} \right., \quad (2.2.1)$$

(forme ci-dessus); ou bien: trouver $\bar{x}'$ tel que

$$\left\{ \begin{array}{l} \bar{x}' \in \mathbb{k}' = \{x' \in \mathbb{R}_+^{N'}; \sum_{j=1}^{N'} C'_{ij} x'_j \leq d'_i, 1 \leq i \leq M'\} \\ f'(\bar{x}') = \inf_{x' \in \mathbb{k}'} f'(x'), f'(x') = \sum_{i=1}^{N'} a'_i x'_i \end{array} \right., \quad (2.2.2)$$

(on rappelle que $\mathbb{R}_+^p = \{x = (x_i) \in \mathbb{R}^p; x_i \geq 0, 1 \leq i \leq p\}$ ou bien: trouver $\bar{x}''$ tel que

$$\left\{ \begin{array}{l} \bar{x}'' \in \mathbb{k}'' = \{x'' \in \mathbb{R}_+^{N''}; \sum_{j=1}^{N''} C''_{ij} x''_j \leq d''_i, 1 \leq i \leq M''\} \\ f''(\bar{x}'') = \inf_{x'' \in \mathbb{k}''} f''(x''), f''(x'') = \sum_{i=1}^{N''} a''_i x''_i \end{array} \right. . \quad (2.2.3)$$

Il est alors facile de voir que les trois formes canoniques sont équivalentes, au sens suivant: partant d'un problème posé sous l'une des formes, on peut toujours lui faire correspondre un problème posé sous l'une quelconque des deux autres formes, de telle façon que la connaissance de l'ensemble (qui est peut-être vide!) des solutions du problème "initial" entraîne celle de l'ensemble des solutions du "nouveau" problème, et inversement. La preuve est dans [1].

Remarque 2.2.1 Dans un souci évident de simplification de l'écriture, nous omettrons dorénavant l'indication des "primes" et "secondes" en exposants des variables. des fonctionnelles. etc.. apparaissant dans les problèmes considérés. Avant d'énoncer un résultat général d'existence, faisons deux remarques préliminaires très simples, qui illustrent bien le caractère "particulier" de la programmation linéaire, notamment vis-à-vis de la programmation quadratique.

Premièrement, il est clair qu'aucun point intérieur à l'ensemble $\mathbb{K}$ ne saurait être solution, sauf si la fonctionnelle f est nulle. Soit en effet $\bar{x}$ un point intérieur à l'ensemble $\mathbb{K}$, et $\varepsilon > 0$ le rayon d'une boule fermée de $\mathbb{R}^N$ centrée en $\bar{x}$ et tout entière contenue dans l'ensemble $\mathbb{K}$. La linéarité de la fonctionnelle $f : x \longrightarrow f(x) = (a, x)$ montre que $f(\bar{x} - \frac{\varepsilon}{\|a\|^2}a) = f(\bar{x}) - \varepsilon < f(\bar{x})$ si $a \neq 0$ et comme le point $\{\bar{x} - \frac{\varepsilon}{\|a\|^2}a\}$ appartient à l'ensemble $\mathbb{K}$, la conclusion en découle.

Deuxièmement, il se peut que le problème n'ait pas de solution lorsque l'ensemble $\mathbb{K}$ est non borné: considérer l'exemple de la fonction $f(x) = -x$ et de l'ensemble $\mathbb{K} = \mathbb{R}_+$. Il s'agit là d'une deuxième différence avec la programmation quadratique, qui provient naturellement de l'absence de coercivité d'une fonctionnelle linéaire. Voir [1].

Théorème 2.2.1 On considère le problème: trouver $\bar{x}$ tel que

$$\begin{cases} \bar{x} \in \mathbb{K} = \{x \in \mathbb{R}_+^N, Cx = d\}, C \in A_{M,N}(\mathbb{R}), \\ f(\bar{x}) = \inf_{x \in \mathbb{K}} f(x), f(x) = (a, x) \end{cases}.$$

On suppose l'ensemble $\mathbb{K}$ non vide. Alors on a l'alternative suivante:

- (i) ou bien $\inf_{x \in \mathbb{K}} f(x) = -\infty$.
- (ii) ou bien $\inf_{x \in \mathbb{K}} f(x) > -\infty$, et le problème admet au moins une solution.

Preuve. Voir [1]. ■

2.3 Exemples de problèmes de programmation linéaire

Une usine fabrique deux produits P_1 et P_2 à l'aide de matières premières q_1, q_2, q_3 . La fabrication d'une unité du produit P_1 nécessite 1 unité de q_1 , 2 unités de q_2 , 4 unités de q_3 ; la fabrication d'une unité du produit P_2 nécessite 6 unités de q_1 , 2 unités de q_2 , 1 unité de q_3 . L'usine dispose de 30 unités de q_1 , de 15 unités de q_2 , de 24 unités de q_3 .

Enfin, la vente d'une unité de P_1 rapporte un bénéfice de 2 écus, celle d'une unité de P_2 un bénéfice de 1 écu. L'objectif de l'usine étant la recherche d'un bénéfice maximal, comment doit-elle organiser sa production?

Appelant $\bar{x}_1$ et $\bar{x}_2$ les nombres "optimaux" d'unités des produits P_1 et P_2 à produire respectivement, ce problème d'organisation de production s'écrit: trouver $\bar{x} = (\bar{x}_1, \bar{x}_2)$ tel que:

$$\left\{ \begin{array}{l} \bar{x} \in \mathbb{k} = \{x = (x_1, x_2) \in \mathbb{R}_+^2; x_1 + 6x_2 \leq 30, 2x_1 + 2x_2 \leq 15, 4x_1 + x_2 \leq 24\}, \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = -2x_1 - x_2. \end{array} \right.$$

c'est donc bien un problème de programmation linéaire posé sous la forme (p₂) et qui se résout d'ailleurs "graphiquement" de façon très simple. En "déplaçant" les droites $\{f(x) = f(x_0)\}$, on s'aperçoit en effet que la solution est le sommet $\bar{x} = (\frac{11}{2}, 2)$ de l'ensemble $\mathbb{k}$, ici un pentagone (l'existence d'au moins une solution était de toute façon évidente, puisque l'ensemble $\mathbb{k}$ ci-dessus est non vide, fermé, borné).

C'est là le premier exemple d'une propriété générale, et fondamentale, des problèmes de programmation linéaire, que nous retiendrons pour l'instant sous la forme "vague" suivante: si le problème admet une solution, alors au moins $\bar{x}$ "sommet" de l'ensemble $\mathbb{k}$ (qui est un "polyèdre", borné ou non, de $\mathbb{R}^N$) est solution.

Il se peut d'ailleurs que plusieurs sommets soient solutions. C'eût été le cas dans le problème ci-dessus si, par exemple, la fonctionnelle f avait eu pour expression $f(x) = -4x_1 - x_2$. Alors le segment fermé joignant les deux points $(\frac{11}{2}, 2)$ et $(6, 0)$ eût été l'ensemble des solutions.

Remarque 2.3.1 Pour que la solution $\bar{x} = (\frac{11}{2}, 2)$ soit acceptable, encore faut-il que les unités de produit P_1 soient "divisibles". Sinon, on serait en présence d'un exemple de problème de programmation linéaire "en nombres entiers".

-Montrons au passage comment le problème ci-dessus peut être immédiatement converti en un problème de type (p₃): introduisant trois variables d'écart x_i , $i = 3, 4, 5$, on est ramené à trouver $\bar{x}$ tel que:

$$\begin{cases} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}_+^5; Cx = d\}, \\ f(x) = \inf_{x \in \mathbb{k}} f(x), f(x) = (a, x), \end{cases}$$

où

$$a = \begin{pmatrix} -2 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, C = \begin{pmatrix} 1 & 6 & 1 & & \\ 2 & 2 & & 1 & \\ 4 & 1 & & & 1 \end{pmatrix}, d = \begin{pmatrix} 30 \\ 15 \\ 24 \end{pmatrix}.$$

2.4 La méthode du simplexe

On considère un problème de programmation linéaire posé sous la forme (p₃): trouver $\bar{x}$ tel que:

$$\begin{cases} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}_+^N; \sum_{j=1}^N C_{ij}x_j = d_i, 1 \leq i \leq M\} \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = \sum_{i=1}^N a_i x_i \end{cases}.$$

Désignant par $C^j \in \mathbb{R}^M$, $1 \leq j \leq N$, les vecteurs colonnes de la matrice $C = (c_{ij}) \in A_{M,N}(\mathbb{R})$,

on remarque que les trois écritures suivantes sont équivalentes:

$$\sum_{j=1}^N C_{ij}x_j = d_i, 1 \leq i \leq M \iff Cx = d \iff \sum_{j=1}^N x_j C^j = d,$$

la dernière étant notamment d'un usage constant dans ce paragraphe.

En vue de préciser la notion de "sommet" de l'ensemble $\mathbb{k}$, commençons par quelques.

définitions générales: si $\mathbb{k}$ est une partie convexe d'un espace vectoriel, un point $\bar{x} \in \mathbb{k}$

est appelé point extrémal de l'ensemble $\mathbb{k}$ si l'implication:

$$\left. \begin{array}{l} x \in \mathbb{k}, y \in \mathbb{k}, 0 < \lambda < 1 \\ \bar{x} = \lambda x + (1 - \lambda)y \end{array} \right\} \implies \bar{x} = x = y,$$

est satisfaite. Un polyèdre de $\mathbb{R}^N$ est un ensemble de la forme:

$$\mathbb{k} = \{x \in \mathbb{R}^N; \sum_{j=1}^N C_{ij}x_j \leq d_i, 1 \leq i \leq p, \sum_{j=1}^N C_{ij}x_j = d_i, p+1 \leq i \leq M\}.$$

Enfin, un point extrémal d'un polyèdre est appelé sommet du polyèdre. Ces définitions généralisent de façon naturelle celles de polygone convexe (non nécessairement borné) du plan, et de sommet d'un tel polygone. Voir [1].

Exemple 2.4.1 Soit un polyèdre de $\mathbb{R}^3$, défini par $\mathbb{k} = \{x \in \mathbb{R}_+^3; x_1 + 2x_2 + 4x_3 = 4\}$. On vérifiera que les seuls sommets de l'ensemble $\mathbb{k}$ considéré sont les points: $(4, 0, 0), (0, 2, 0), (0, 0, 1)$. Commençons par caractériser les sommets d'un polyèdre de la forme

$$\mathbb{k} = \left\{x \in \mathbb{R}_+^N; \sum_{j=1}^N x_j C^j = d\right\}.$$

A un point quelconque x d'un tel ensemble, on associe l'ensemble d'indices

$$I^*(x) = \{1 \leq j \leq N; x_j > 0\}.$$

Si l'origine appartient à l'ensemble $\mathbb{k}$ (de façon équivalente si $d = 0$), auquel cas $I^*(0) = \emptyset$, on notera que l'origine est déjà un sommet, puisque

$$\left. \begin{array}{l} 0 = tx + (1 - t)y \\ x \geq 0, y \geq 0, 0 < t < 1 \end{array} \right\} \implies x = y = 0.$$

Examinons le cas des autres sommets.

Théorème 2.4.1 Un point $\bar{x} \in \mathbb{k}$ différent de l'origine est un sommet du polyèdre

$$\mathbb{k} = \{x \in \mathbb{R}_+^N; \sum_{j=1}^N x_j C^j = d\},$$

C^j : vecteurs de $\mathbb{R}^N$ si et seulement si les vecteurs $C^j, j \in I^*(\bar{x})$, où $I^*(\bar{x}) = \{1 \leq j \leq N; \bar{x}_j > 0\}$, sont linéairement indépendants.

Preuve. Voir [1]. ■

Théorème 2.4.2 *Si le problème: trouver $\bar{x}$ tel que*

$$\left\{ \begin{array}{l} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}_+^N; \sum_{j=1}^N x_j C^j = d\} \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), \quad f(x) = \sum_{j=1}^N a_j x_j, \end{array} \right. .$$

à une solution, alors (au moins) un sommet de l'ensemble $\mathbb{k}$ est aussi solution.

Preuve. Voir [1]. ■

Théorème 2.4.3 *Si le polyèdre:*

$$\mathbb{k} = \{x \in \mathbb{R}_+^N; Cx = d\}, C \in A_{M,N}(\mathbb{R}), d \in \mathbb{R}^M,$$

n'est pas vide, il possède au moins un sommet. Par ailleurs, les sommets sont en nombre fini. Voir [5].

La méthode du simplexe, due à G. B. Dantzig, repose sur le résultat ci-dessus: son principe est en effet d'évaluer la fonctionnelle f en certains sommets de l'ensemble $\mathbb{k}$, construits par récurrence suivant une stratégie particulière: partant d'un sommet $\bar{x}_0$, on construit une suite de sommets $\bar{x}_0, \bar{x}_1, \dots, \bar{x}_k, \bar{x}_{k+1}, \dots$ correspondant à des valeurs décroissantes de la fonctionnelle. Si le problème a une solution, et si l'on peut faire en sorte d'avoir uniquement des inégalités strictes $f(\bar{x}_k) > f(\bar{x}_{k+1})$ ce procédé conduit en un nombre fini d'itérations à un sommet qui est aussi solution, puisqu'il y a un nombre fini de sommets.

Remarque 2.4.1 (i) *Il arrive effectivement dans certains cas qu'on ne puisse obtenir l'inégalité stricte $f(\bar{x}_k) > f(\bar{x}_{k+1})$; on a seulement l'égalité $f(\bar{x}_k) = f(\bar{x}_{k+1})$. Cette circonstance peut conduire au phénomène dit de "cyclage", sur lequel nous reviendrons.*

(ii) *L'éventualité où $\inf_{x \in \mathbb{k}} f(x) = -\infty$ n'étant nullement exclue (reprendre l'exemple de la fonctionnelle $f(x) = -x$, et de l'ensemble $\mathbb{k} \in \mathbb{R}_+$, de sommet unique 0), il est remarquable que la méthode en question permette aussi de reconnaître cette éventualité, en l'absence toutefois du phénomène de "cyclage" signalé ci-dessus.*

(iii) On pourrait songer à évaluer systématiquement la fonctionnelle f en tous les sommets de l'ensemble $\mathbb{k}$, sans aucune stratégie particulière. Leur nombre étant déjà très grand pour des "petites" valeurs de M et N , ce procédé naïf est à proscrire complètement. Voir [1].

-Nous allons d'abord nous attacher à la description d'une itération de la méthode du simplexe, c'est-à-dire dans le cas le plus "courant", le passage d'un sommet à un autre sommet, d'autres éventualités étant en effet possibles. Ce n'est qu'ultérieurement que nous examinerons la question de la détermination d'un sommet initial $\bar{x}_0$ (l'"initialisation" de l'algorithme). Rappelons la définition du problème: il s'agit de trouver $\bar{x}$ tel que:

$$\left\{ \begin{array}{l} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}_+^N; \sum_{j=1}^N x_j C^j = d\} \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = \sum_{j=1}^N a_j x_j, \end{array} \right.,$$

où les vecteurs $C^j \in \mathbb{R}^M$, $1 \leq j \leq N$, sont les vecteurs colonnes d'une matrice $C = (c_{ij})$. Sans restreindre la généralité, on va supposer que: $\text{rang}(C) = M$.

(ce qui impose $M \leq N$). En effet, si $\text{rang}(C) = M' < M$, alors ou bien l'ensemble $\{x \in \mathbb{R}_+^N; \sum_{j=1}^N x_j C^j = d\}$ est vide, auquel cas le problème n'a pas de solution, ou bien on peut toujours extraire des équations:

$$\sum_{j=1}^N c_{ij} x_j = d_i, 1 \leq i \leq M,$$

un sous-ensemble de M' équations équivalentes dont la matrice est de rang M' .

Remarque 2.4.2 (1) On peut également supposer tous les vecteurs colonnes C^j de la matrice C non nuls. En effet si l'un d'eux est nul, disons C^N pour fixer les idées, alors ou bien $a_N > 0$, auquel cas toute solution du problème vérifie $\bar{x}_N = 0$ et on peut donc "éliminer la N -ème variable", ou bien $a_N < 0$ (si $a_N = 0$, la N -ème variable n'apparaît pas ...), auquel cas $\inf_{x \in \mathbb{k}} f(x) = -\infty$ et le problème n'a pas de solution.

(2) Dans le même ordre d'idées, il devrait être clair que le cas où $M = N$ est sans grand intérêt pour la suite....

-La description de la méthode du simplexe nous amène à une distinction préliminaire entre deux types de sommets. Rappelons (théorème 2.4.1) qu'un point $\bar{x} = (\bar{x}_i) \in \mathbb{R}^N$ est un sommet de l'ensemble $\mathbb{k}$ si et seulement si

$$\left\{ \begin{array}{l} \bar{x}_i \geq 0, 1 \leq i \leq N; \sum_{j=1}^N \bar{x}_j C^j = \sum_{j \in I^*(\bar{x})} \bar{x}_j C^j = d; \\ \text{les vecteurs } C^j \in \mathbb{R}^M, j \in I^*(\bar{x}), \text{ sont linéairement indépendants, où:} \end{array} \right.$$

$$I^*(\bar{x}) = \{1 \leq j \leq N, \bar{x}_j > 0\}.$$

On dit alors qu'un tel sommet est non dégénéré si $\text{card}(I^*(\bar{x})) = M$, et dégénéré si $\text{card}(I^*(\bar{x})) < M$; c'est ainsi que l'origine est un sommet dégénéré de l'ensemble $\mathbb{k}$ lorsqu'elle lui appartient, puisque $I^*(0) = \emptyset$.

Or, pour des raisons qui vont apparaître dans la description de la méthode, il est essentiel de pouvoir associer à tout sommet exactement M vecteurs colonnes C^j linéairement indépendants. C'est pourquoi dans le cas d'un sommet dégénéré $\bar{x}$, pour lequel $\text{card}(I^*(\bar{x})) = M' < M$, on complètera les vecteurs $C^j, j \in I^*(\bar{x})$, par $(M - M')$ vecteurs $C^j; j \notin I^*(\bar{x})$, de telle façon que les M vecteurs C^j ainsi obtenus soient linéairement indépendants. C'est toujours possible (puisque on a supposé $\text{rang}(C) = M$) mais en général de façon non unique.

On peut donc donner la définition équivalente: un point $\bar{x} = (\bar{x}_i) \in \mathbb{R}^N$ est un sommet de l'ensemble $\mathbb{k}$ si et seulement si il existe un ensemble:

$$I \subset \{1, 2, \dots, N\} \text{ avec } \text{card}(I) = M,$$

tel que:

$$\left\{ \begin{array}{l} \bar{x}_i \geq 0 \text{ si } i \in I, \bar{x}_i = 0 \text{ si } i \notin I; \sum_{i=1}^N \bar{x}_i C^i = \sum_{i \in I} \bar{x}_i C^i = d; \\ \text{les vecteurs } C_i, i \in I, \text{ sont linéairement indépendants,} \end{array} \right.$$

On dit alors que $(C^i)_{i \in I}$ est une base, associée au sommet $\bar{x}$. On retiendra que la base associée à un sommet non dégénéré est définie de façon unique, mais que la base associée à un sommet dégénéré n'est pas nécessairement définie de façon unique.

Remarque 2.4.3 Cette possibilité explique l'abus de langage parfois constaté (mais par ailleurs fort commode, par exemple dans la description de la méthode du simplexe) qui consiste à considérer un même sommet comme plusieurs sommets "différents" dans la mesure où plusieurs bases (réellement !) différentes peuvent lui être associées.

-Supposons donc connu un sommet $\bar{x} = (\bar{x}_i)_{i=1}^N$ de l'ensemble $\mathbb{k}$, de base associée $(C^i)_{i \in I}$. L'idée de la méthode est d'effectuer le passage au sommet suivant (quand c'est possible) en remplaçant l'un des vecteurs C^i , $i \in I$, par l'un des vecteurs C^j , $j \in I$, ce qui définira la base associée au sommet suivant. Soit j un indice n'appartenant pas à l'ensemble I . Le vecteur C^j s'écrit:

$$C^j = \sum_{i \in I} \gamma_i^j C^i,$$

les composantes γ_i^j étant définies de façon unique. Puisque:

$$\sum_{i \in I} (\bar{x}_i - \theta \gamma_i^j) C^i + \theta C^j = \sum_{i \in I} \bar{x}_i C^i = d, \text{ pour tout } \theta \in \mathbb{R},$$

on déduit que les points de composantes

$$\begin{cases} \bar{x}_i - \theta \gamma_i^j, & \text{si } i \in I, \\ \theta, & \text{si } i = j, \\ 0, & \text{si } i \notin I \cup \{j\}, \end{cases}$$

sont des point, de l'ensemble $\mathbb{k}$ pourvu que:

$$0 \leq \theta \leq \theta^j \stackrel{\text{déf}}{=} \min \left\{ \frac{\bar{x}_i}{\gamma_i^j}, i \in I \text{ et } \gamma_i^j > 0 \right\}.$$

Naturellement, on espère que le point correspondant à $\theta = \theta^j$ est un nouveau sommet lorsque $0 < \theta^j < +\infty$. C'est effectivement ce que l'on démontrera. Évaluée en ces points, la fonctionnelle f vaut:

$$\sum_{i \in I} a_i (\bar{x}_i - \theta \gamma_i^j) + \theta a_j = f(\bar{x}) + \theta (a_j - \sum_{i \in I} \gamma_i^j a_i),$$

de sorte que, pour pouvoir diminuer strictement la valeur de la fonctionnelle, il faut, d'une part, que le nombre θ^j soit > 0 et, d'autre part, que le nombre $\{a_j - \sum_{i \in I} \gamma_i^j a_i\}$ soit < 0 .

On est donc naturellement conduit à distinguer plusieurs éventualités, selon les signes des nombres:

$$(a_j - \sum_{i \in I} \gamma_i^j a_i), \max_{i \in I} \gamma_i^j, \theta^j = \min \left\{ \frac{\bar{x}_i}{\gamma_i^j}, i \in I \text{ et } \gamma_i^j > 0 \right\}, \text{ pour } j \notin I.$$

Si tous les nombres $(a_j - \sum_{i \in I} \gamma_i^j a_i)$, $j \notin I$, sont ≥ 0 (cas A), on a l'intuition que le sommet $\bar{x}$ est peut-être une solution (car on ne constate la non-décroissance de la fonctionnelle f que "dans la direction des sommets voisins"); c'est effectivement ce que l'on va montrer. Sinon (cas B), l'objectif est de se retrouver en un autre sommet, en annulant l'une au moins des "nouvelles" composantes $(\bar{x}_i - \theta \gamma_i^j)$, $i \in I$, pour une valeur finie de θ , qui ne peut être que $\theta = \theta^j$. C'est impossible si toutes les composantes γ_i^j sont ≤ 0 pour $i \in I$ (cas B1), auquel cas il est d'ailleurs clair que $\inf_{x \in \mathbb{k}} f(x) = -\infty$. Par contre, si l'une des composantes γ_i^j , $i \in I$, est < 0 , on montrera que le point $\bar{x}^+$ correspondant à $\theta = \theta^j$ (qui est alors un nombre fini) est effectivement un sommet, ce qui revient à montrer l'indépendance linéaire des vecteurs colonnes associés au point $\bar{x}^+$. Pour que celui-ci corresponde à l'inégalité stricte $f(\bar{x}^+) < f(\bar{x})$, encore faut-il que le nombre θ^j soit > 0 (cas B2): c'est le cas de progression "courante" de l'algorithme. Sinon (cas B3), on est peut-être en présence du phénomène de cyclage. Le résultat qui suit rassemble, en les précisant, ces diverses éventualités.

Théorème 2.4.4 *Soit $\bar{x}$ un sommet de l'ensemble $\mathbb{k}$, de base associée $(C^i)_{i \in I}$. Pour tout $j \notin I$, on pose $C^j = \sum_{i \in I} \gamma_i^j C^i$. Les seules éventualités possibles sont les suivantes:*

Cas A: $a_j - \sum_{i \in I} \gamma_i^j a_i \geq 0$, pour tout $j \notin I$. Alors le sommet $\bar{x}$ est solution du problème.

Cas B: Il existe (au moins) un indice $j \notin I$ pour lequel $\{a_j - \sum_{i \in I} \gamma_i^j a_i\} < 0$: cette éventualité se subdivise elle-même en trois cas:

Cas B1: Il existe (au moins) un indice $j \notin I$ pour lequel, on a simultanément:

$$a_j - \sum_{i \in I} \gamma_i^j a_i < 0 \text{ et } \gamma_i^j \leq 0 \text{ pour tout } i \in I.$$

Alors $\inf_{x \in \mathbb{k}} f(x) = -\infty$.

Cas B2: Il existe (au moins) un indice $j^+ \notin I$ pour lequel, on a simultanément:

$$\left\{ \begin{array}{l} a_{j^+} - \sum_{i \in I} \gamma_i^{j^+} a_i < 0; \\ \gamma_i^{j^+} > 0 \text{ pour au moins un indice } i \in I; \text{ pour } i \in I, \gamma_i^{j^+} > 0 \implies \bar{x}_i > 0. \end{array} \right.$$

Alors le point $\bar{x}^+$ de composantes

$$\bar{x}_i^+ = \bar{x}_i - \theta^{j^+} \gamma_i^{j^+} \text{ si } i \in I; \theta^{j^+} \text{ si } i = j^+; 0 \text{ si } i \notin I \cup \{j^+\},$$

où

$$\theta^{j^+} = \min \left\{ \frac{\bar{x}_i}{\gamma_i^{j^+}}; i \in I \text{ et } \gamma_i^{j^+} > 0 \right\} > 0,$$

est un sommet différent du sommet $\bar{x}$, pour lequel $f(\bar{x}) > f(\bar{x}_+)$.

Au sommet $\bar{x}^+$ est associée la base $(C^i)_{i \in I^+}$, où $I^+ = (I - \{j^-\}) \cup \{j^+\}$, $j^- \in I$, étant l'un quelconque des indices vérifiant:

$$j^- \in I \text{ et } \theta^{j^+} = \frac{\bar{x}_{j^-}}{\gamma_{j^-}^{j^+}}.$$

Cas B3: Il existe (au moins) un indice $j \notin I$ pour lequel $\{a_j - \sum_{i \in I} \gamma_i^j a_i\} < 0$, et pour chacun de ces indices $j \notin I$, on a simultanément:

$$\left\{ \begin{array}{l} \{a_j - \sum_{i \in I} \gamma_i^j a_i\} < 0; \\ \gamma_i^j > 0 \text{ et } \bar{x} = 0 \text{ pour au moins un indice } i \in I \text{ (le même)}. \end{array} \right. ;$$

Alors on peut associer au même sommet $\bar{x}$ une nouvelle base $(C^i)_{i \in I}$, où $(I - \{j^-\}) \cup \{j^+\}$, j^+ étant l'un quelconque des indices vérifiant:

$$j^+ \notin I \text{ et } \{a_{j^+} - \sum_{i \in I} \gamma_i^{j^+} a_i\} < 0,$$

et j^- étant l'un quelconque des indices vérifiant:

$$j^- \in I \text{ et } \gamma_{j^-}^{j^+} > 0.$$

Preuve. Voir [1]. ■

Remarque 2.4.4 Elle est commode d'interpréter les différents cas considérés en introduisant l'ensemble:

$$E \stackrel{\text{déf}}{=} \{j \notin I, a_j - \sum_{i \in I} \gamma_i^j a_i < 0\} = E_1 \cup E_2 \cup E_3,$$

avec

$$\begin{aligned} E_1 &= \{j \notin I, a_j - \sum_{i \in I} \gamma_i^j a_i < 0, \max_{i \in I} \gamma_i^j \leq 0\}, \\ E_2 &= \{j \notin I, a_j - \sum_{i \in I} \gamma_i^j a_i < 0, \max_{i \in I} \gamma_i^j > 0, \left\{ \begin{array}{l} \min_{i \in I} \left\{ \frac{\bar{x}_i}{\gamma_i^j} \right\} > 0, \\ \gamma_i^j > 0 \end{array} \right\}, \\ E_3 &= \{j \notin I, a_j - \sum_{i \in I} \gamma_i^j a_i < 0, \max_{i \in I} \gamma_i^j = 0, \left\{ \begin{array}{l} \min_{i \in I} \left\{ \frac{\bar{x}_i}{\gamma_i^j} \right\} > 0, \\ \gamma_i^j > 0 \end{array} \right\}. \end{aligned}$$

Nous avons en effet les correspondances suivantes:

$$\text{cas } A \iff E = \emptyset; \text{ cas } B \iff E \neq \emptyset;$$

$$\text{cas } B1 \iff E_1 \neq \emptyset; \text{ cas } B2 \iff E_2 \neq \emptyset; \text{ cas } B3 \iff E_3 = E,$$

qui montrent que les cas A et B, B1, B2 et B3 sont exclusifs, alors que les cas B1 et B2 ne le sont pas.

2.5 Méthodes de dualité

Supposons que les hypothèses suivantes sont vérifiées:

$$\left\{ \begin{array}{l} f \in C^1(\mathbb{R}^N, \mathbb{R}), \\ g_i \in C^1(\mathbb{R}^N, \mathbb{R}), \quad i = 1, \dots, p \\ \mathbb{k} = \{x \in \mathbb{R}^N, g_i(x) \leq 0, i = 1, \dots, p\} \end{array} \right. \quad (2.5.1)$$

et $\mathbb{k}$ est non vide.

On définit un problème "primal" comme étant le problème de minimisation d'origine, c'est-à-dire

$$\begin{cases} \bar{x} \in \mathbb{K} \\ f(\bar{x}) \leq f(x), \text{ pour tout } x \in \mathbb{K}. \end{cases} \quad (2.5.2)$$

On définit le "lagrangien" comme étant la fonction L définie de $\mathbb{R}^N \times \mathbb{R}^P$ dans $\mathbb{R}$ par:

$$L(x, \lambda) = f(x) + \lambda \cdot g(x) = f(x) + \sum_{i=1}^p \lambda_i g_i(x), \quad (2.5.3)$$

avec

$$g(x) = (g_1(x), \dots, g_p(x))^t \text{ et } \lambda = (\lambda_1(x), \dots, \lambda_p(x))^t,$$

On note C^+ l'ensemble défini par

$$C^+ = \{\lambda \in \mathbb{R}^p, \lambda = (\lambda_1, \dots, \lambda_p)^t, \lambda_i \geq 0 \text{ pour tout } i = 1, \dots, p\}.$$

Remarque 2.5.1 Le théorème de Kuhn-Tucker entraîne que si $\bar{x}$ est solution du problème primal (2.5.2) alors il existe $\lambda \in C^+$ tel que $D_1 L(\bar{x}, \lambda) = 0$ (c'est-à-dire $Df(x) + \lambda Dg(x) = 0$) et $\lambda g(\bar{x}) = 0$. On définit, alors l'application M de $\mathbb{R}^p$ dans $\mathbb{R}$ par:

$$M(\lambda) = \inf_{x \in \mathbb{R}^N} L(x, \lambda).$$

pour tout $\lambda \in \mathbb{R}^p$. On peut donc remarquer que $M(\lambda)$ réalise le minimum (en x) du problème sans contrainte, qui s'écrit, pour $\lambda \in \mathbb{R}^p$ fixé:

$$\begin{cases} x \in \mathbb{R}^N \\ L(x, \lambda) \leq L(y, \lambda) \text{ pour tout } y \in \mathbb{R}^N \end{cases}. \quad (2.5.4)$$

Voir [4].

Lemme 2.5.1 L'application M de $\mathbb{R}^p$ dans $\mathbb{R}$ définie par (2.5.3) est concave (ou encore l'application $-M$ est convexe), c'est-à-dire que pour tous $\lambda, \mu \in \mathbb{R}^p$ et pour tout $t \in]0, 1[$, on a:

$$M(t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(\mu).$$

Preuve. Soit $\lambda, \mu \in \mathbb{R}^p$ et $t \in]0, 1[$, on veut montrer que $M(t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(\mu)$. Soit $x \in \mathbb{R}^N$ alors: $L(x, t\lambda + (1-t)\mu) = f(x) + (t\lambda + (1-t)\mu)g(x) = tf(x) + (1-t)f(x) + (t\lambda + (1-t)\mu)g(x)$. On a donc $L(x, t\lambda + (1-t)\mu) = tL(x, \lambda) + (1-t)L(x, \mu)$. Par définition de M , on en déduit que pour tout $x \in \mathbb{R}^N$, $L(x, t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(\mu)$. Or, toujours par définition de M , $M(t\lambda + (1-t)\mu) = \inf_{x \in \mathbb{R}^N} L(x, t\lambda + (1-t)\mu) \geq tM(\lambda) + (1-t)M(\mu)$. Voir [4]. ■

On considère maintenant le problème d'optimisation dit "dual" suivant:

$$\begin{cases} \mu \in C^+, \\ M(\mu) \geq M(\lambda), \forall \lambda \in C^+. \end{cases} \quad (2.5.5)$$

Définition 2.5.1 Soit $L : \mathbb{R}^N \times \mathbb{R}^p \rightarrow \mathbb{R}$ et $(x, \mu) \in \mathbb{R}^N \times C^+$. On dit que (x, μ) est un point selle de L sur $\mathbb{R}^N \times C^+$ si $L(x, \lambda) \leq L(x, \mu) \leq L(y, \mu)$ pour tout $y \in \mathbb{R}^N$ et pour tout $\lambda \in C^+$.

Proposition 2.5.1 Sous les hypothèses (2.5.1), soit L définie par $L(x, \lambda) = f(x) + \lambda g(x)$ et $(x, \lambda) \in \mathbb{R}^N \times C^+$ un point selle de L sur $\mathbb{R}^N \times C^+$. Alors:

1. $\bar{x}$ est solution du problème (2.5.2),
2. μ est solution de (2.5.5),
3. $\bar{x}$ est solution du problème (2.5.4) avec $\lambda = \mu$.

On admettra cette proposition.

Réciproquement, on peut montrer que (sous des hypothèses convenables sur f et g), si μ est solution de (2.5.5), et si $\bar{x}$ solution de (2.5.4) avec $\lambda = \mu$, alors $(\bar{x}, \mu)$ est un point selle de L , et donc $\bar{x}$ est solution de (2.5.2). De ces résultats découle l'idée de base des méthodes de dualité: on cherche solution de (2.5.5).

On obtient ensuite une solution $\bar{x}$ du problème (2.5.2), en cherchant $\bar{x}$ comme solution du problème (2.5.4) avec $\lambda = \mu$ (qui est un problème de minimisation sans contraintes). La recherche de la solution μ du problème dual (2.5.5) peut se faire par exemple par l'algorithme très classique d'Uzawa, que nous décrivons maintenant.

Proposition 2.5.2 Sous les hypothèses (2.5.1), on suppose que pour tout $\lambda \in \mathbb{R}^N$, le problème (2.5.4) admet une solution unique, notée x et on suppose que l'application définie de

$\mathbb{R}^p$ dans $\mathbb{R}^N$ par $\lambda \rightarrow x_\lambda$ est différentiable. Alors $M(\lambda) = L(x_\lambda, \lambda)$, M est différentiable en λ pour tout λ , et $\nabla M(\lambda) = g(x_\lambda)$.

En conséquence, pour calculer $\nabla M(\lambda)$, on est ramené à chercher x_λ solution du problème de minimisation sans contrainte (2.5.4). On peut donc maintenant donner le détail de l'itération **générale de l'algorithme d'Uzawa**:

Itération de l'algorithme d'Uzawa. Soit $\mu_n \in C^+$ connu;

1. On cherche $x_n \in \mathbb{R}^N$ solution de

$$\begin{cases} x_n \in \mathbb{R}^N \\ L(x_n, \mu_n) \leq L(x, \mu_n), \forall x \in \mathbb{R}^N \end{cases}.$$

(On a donc $x_n = x_{\mu_n}$).

2. On calcule $\nabla M(\mu_n) = g(x_n)$,

3. $\bar{\mu}_{n+1} = \mu_n + \rho \nabla M(\mu_n) = \mu_n + \rho g(x_n) = ((\bar{\mu}_{n+1})_1, \dots, (\bar{\mu}_{n+1})_p)^t$,

4. $\mu_{n+1} = pC + (\bar{\mu}_{n+1})$, c'est-à-dire $\mu_{n+1} = ((\mu_{n+1})_1, \dots, (\mu_{n+1})_p)^t$ avec $(\mu_{n+1})_i = \max(0, (\bar{\mu}_{n+1})_i)$ pour tout $i = 1, \dots, p$. On a alors le résultat suivant de convergence de l'algorithme.

Proposition 2.5.3 (Convergence de l'algorithme d'Uzawa). Sous les hypothèses (2.5.1), on suppose de plus que: 1. il existe $\alpha > 0$ tel que $(\nabla f(x) - \nabla f(y))(x - y) \geq \alpha \|x - y\|^2$ pour tout $(x, y) \in (\mathbb{R}^N)^2$,

2. il existe $M_f > 0$, $\|\nabla f(x) - \nabla f(y)\| \leq M_f \|x - y\|$ pour tout $(x, y) \in (\mathbb{R}^N)^2$,

3. pour tout $\lambda \in C^+$, il existe un unique $x_\lambda \in \mathbb{R}^N$ tel que $L(x_\lambda, \lambda) \leq L(x, \lambda)$ pour tout $x \in \mathbb{R}$.

Alors si $0 < \rho < \frac{2\alpha}{M_f^2}$ la suite $((x_n, \mu_n))_n \in \mathbb{R}^N \times C^+$ donnée par l'algorithme d'Uzawa vérifie:

(i). $x_n \rightarrow \bar{x}$ quand $n \rightarrow +\infty$ où $\bar{x}$ est la solution du problème (2.5.2),

(ii). $(\mu_n)_{n \in \mathbb{N}}$ est bornée. Voir [4].

Remarque 2.5.2 (Sur l'algorithme d'Uzawa).

1. L'algorithme est très efficace si les contraintes sont affines: (i.e. si $g_i(x) = \alpha_i x + \beta_i$ pour tout $i = 1, \dots, p$, avec $\alpha_i \in \mathbb{R}^N$ et $\beta_i \in \mathbb{R}$).

2. Pour avoir l'hypothèse 3 du théorème, il suffit que les fonctions g_i soient convexes. (On a dans ce cas existence et unicité de la solution x_λ du problème (2.5.4) et existence et unicité de la solution $\bar{x}$ du problème (2.5.2). Voir [4].

2.6 Dualité et programmation linéaire

Étant donné la forme particulière des ensembles $\mathbb{k}$ qui y ont été considérés, ces résultats s'appliquent à des problèmes de programmation linéaire de la forme (P_1) ou (P_2) . Pour la première de ces formes: trouver $\bar{x}$ tel que

$$\begin{cases} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}^N, Cx \leq d\}, C \in A_{M,N}(\mathbb{R}), \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = (a, x)_N, \end{cases} \quad (2.6.1)$$

nous obtenons comme théorème immédiat des conditions nécessaires et suffisantes de minimum.

Théorème 2.6.1 (*Conditions nécessaires et suffisantes de minimum en programmation linéaire*). Soit $\bar{x}$ un point de l'ensemble $\mathbb{k} = \{x \in \mathbb{R}^N, Cx \leq d\}$, $C \in A_{M,N}(\mathbb{R})$, alors $f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x)$, où $f(x) = (a, x)_N$, si et seulement s'il existe un vecteur $\lambda \in \mathbb{R}^M$ tel que:

$$\begin{cases} \lambda \geq 0, C^T \lambda + a = 0, \\ (\lambda, C\bar{x} - d)_M = 0. \end{cases}$$

Preuve. Voir [1]. ■

Remarque 2.6.1 (i) Compte-tenu des relations $\lambda \geq 0$ et $Cx - d \leq 0$, on peut exprimer la dernière relation du théorème sous la forme équivalente $\lambda_i = 0$ si $(C_i, \bar{x})_M - d_i < 0$, $1 \leq i \leq M$, en notant C_i le i -ème vecteur colonne de la matrice transposée C^T (ce qui revient à noter C_i^T le i -ème vecteur ligne de la matrice C).

(ii) Alors que les vecteurs colonnes de la matrice C jouaient un rôle essentiel dans le paragraphe précédent, ce sont maintenant ses vecteurs lignes qui interviennent de façon naturelle. Cette observation sera mieux comprise lorsqu'on aura précisé les liens qui existent entre la méthode du simplexe et la dualité.

(iii) On vérifiera que, dans le cas particulier de la programmation linéaire considérée ici, les conditions nécessaires et suffisantes du théorème ci-dessus peuvent aussi être établies directement à partir du lemme de Farkas-Minkowski. Voir [1].

Théorème 2.6.2 (La dualité en programmation linéaire). Soit les deux problèmes : trouver $\bar{x}$ telle que :

$$\bar{x} \in \mathbb{K} = \{x \in \mathbb{R}_+^N, Cx \leq d\} \text{ et } f(\bar{x}) = \inf_{x \in \mathbb{K}} f(x), f(x) = (a, x)_N, \quad (2.6.2)$$

trouver A tel que :

$$\lambda \in A = \{y \in \mathbb{R}_+^M; C^T y \geq -a\} \text{ et } F(\lambda) = \sup_{y \in A} F(y), F(y) = -(d, y)_M, \quad (2.6.3)$$

(i) Si l'un des problèmes a une solution, alors l'autre problème a également une solution (en particulier, l'ensemble qui lui est associé est nécessairement non vide); de plus, $\sup_{y \in A} F(y) = F(\lambda) = f(\bar{x}) = \inf_{x \in \mathbb{K}} f(x)$ en notant respectivement $\bar{x}$ et λ des solutions quelconques de chacun des deux problèmes.

(ii) Une condition nécessaire et suffisante pour que l'un des deux problèmes (et donne l'autre aussi, d'après (i)) ait une solution est: $\mathbb{K} \neq \emptyset$ et $A \neq \emptyset$.

(iii) Un élément $\bar{x} \in \mathbb{R}_+^N$ est solution du problème (2.6.2) si et seulement s'il existe un élément $\lambda \in \mathbb{R}_+^M$, qui n'est autre qu'une solution quelconque du problème (2.6.3), tel que le couple $(\bar{x}, \lambda)$ soit un point-selle de la fonction :

$$L : (x, y) \in \mathbb{R}_+^N \times \mathbb{R}_+^M \rightarrow L : (x, y) = (a, x)_N + (y, Cx - d)_M,$$

c'est-à-dire que le couple $(\bar{x}, \lambda)$ vérifie :

$$(x, y) \in \mathbb{R}_+^N \times \mathbb{R}_+^M \text{ et } \sup_{y \in \mathbb{R}_+^M} L(\bar{x}, y) = L(\bar{x}, \lambda) = \inf_{x \in \mathbb{R}_+^N} L(x, \lambda).$$

Preuve. Voir [1]. ■

Remarque 2.6.2 De ce qui précède, on déduit qu'un élément $\bar{x} \in \mathbb{K}$ est solution du problème (2.6.2) si et seulement si il existe un élément $\lambda \in A$ tel que $F(\lambda) = f(\bar{x})$, la réciproque étant une conséquence de l'inégalité $\sup_{y \in A} F(y) \leq \inf_{x \in \mathbb{K}} f(x)$.

Théorème 2.6.3 (*La dualité et la méthode du simplexe*). On considère le problème: trouver $\tilde{\bar{x}}$ tel que:

$$\begin{cases} \tilde{\bar{x}} \in \tilde{\mathbb{k}} = \left\{ \tilde{x} \in \tilde{\mathbb{R}}_+^N, \tilde{C}\tilde{x} \leq d \right\}; \tilde{C} \in A_{M,\tilde{N}}(\mathbb{R}), d \in \mathbb{R}^M. \\ \tilde{f}(\tilde{\bar{x}}) = \inf_{\tilde{x} \in \tilde{\mathbb{k}}} \tilde{f}(\tilde{x}), \tilde{f}(\tilde{x}) = \tilde{a}^T \tilde{x} \text{ où } \tilde{a} \in \mathbb{R}^{\tilde{N}}. \end{cases} \quad (2.6.4)$$

et le problème équivalent: trouver $\bar{x}$ tel que

$$\begin{cases} \bar{x} \in \mathbb{k} = \{x \in \mathbb{R}_+^N, Cx = d\}. \\ f(\bar{x}) = \inf_{x \in \mathbb{k}} f(x), f(x) = a^T x. \end{cases} \quad (2.6.5)$$

où (I_N : matrice unité de $\mathbb{R}^N$):

$$C = \left(\begin{array}{c|c} \tilde{C} & I_N \end{array} \right) \in A_{M,N}(\mathbb{R}), N = \tilde{N} + M.$$

$$a^T = \left(\begin{array}{c|c} \tilde{a}^T & 0 \end{array} \right), a \in \mathbb{R}^N.$$

Lorsque la méthode du simplexe appliquée au problème (2.6.5) conduit à une solution, elle fournit du même coup une solution du problème dual du problème (2.6.4): trouver λ tel que

$$\begin{cases} \lambda \in A = \{y \in \mathbb{R}_+^M; \tilde{C}^T y + \tilde{a} \geq 0\}. \\ F(\lambda) = \sup_{y \in A} F(y), F(y) = -d^T y. \end{cases} \quad (2.6.6)$$

Preuve. Voir [1]. ■

Bibliographie

- [1] P. G. Ciarlet et J. L. Lions, Introduction a l'analyse numérique matricielle, et a l'optimisation, 1988.
- [2] Rafik Abdesselam, M1 IDS Cours-Optimisation appliquée, Université Lumière Lyon 2 Laboratoire ERIC, 2009-2010 .
- [3] Raphaële Herbin, Cours d'Analyse numérique, Université d'Aix-Marseille, 12 décembre 2011.
- [4] Raphaële Herbin, Cours d'Analyse numérique, Université d'Aix-Marseille Licence de mathématiques, 18 novembre 2013.
- [5] Raphaële Herbin, Université d'Aix-Marseille 1 Licence de mathématiques Cours d'Analyse numérique 2003-2004, 8 juillet 2004.

Key- words: Optimization-Linear Programming-Simplex-Duality.