

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche
Scientifique

Université A. Mira de Béjaïa
Faculté des Sciences Exactes
Département de Mathématique



MEMOIRE DE FIN DE CYCLE

*En vue de l'obtention d'un Master en
Mathématique*

Option : Statistique et Analyse décisionnelle

Thème

**Sur la programmation quadratique
convexe**

Présenté par:

M. SACI Faham

Soutenu devant le jury composé de :

Président	M ^{me}	M. Ourbih	Professeur	U.A. Mira, Béjaïa
Rapporteur	M ^{me}	N. Abassi	M.A.A	U.A. Mira, Béjaïa
Examineur	M ^{me}	S. Guebli	M.C.B	U.A. Mira, Béjaïa

Promotion 2014 /2015

Remerciements

Je remercie Dieu Tous Puissant de m' avoir accordé la force, la volonté et la patience afin accomplir ce modeste travail de recherche.

Je tiens à adresser mes sincères remerciements à ma promotrice, Mme Abassi Nacera, pour son aide et ses orientations toute au long de notre travail.

Je remercie également Mme Ourbih ,et Mme Guebli d'avoir accepter de juger mon travail.

Mes remerciements vont également à toute personne qui a contribué a la réalisation de ce modeste travail ;hicham, nassim,walid, bezza, lyes , kenza

merci tous!

Dédicace

Je dédie ce modeste travail à :

la mémoire de ma mère Allah yarhamha

mon père que dieu le garde pour nous.

mes soeurs et frères et leurs familles

Khalti que dieu la garde pour nous.

à mes oncles et leurs familles

Tous mes adorables amis.

À ma binome fatiha qui aurait du soutenir avec moi

Ainsi qu'à toute personne qui m'a soutenu de près où de loin.

Table des Matières

Introduction générale	1
1 Rappels mathématiques	3
1.1 Introduction	3
1.2 Rappel sur l'algèbre linéaire	3
1.2.1 Vecteurs et matrices	3
1.2.2 Matrices et vecteurs partitionnés	4
1.2.3 Espace vectoriel	4
1.2.4 Noyau , image et rang d'une matrice	5
1.2.5 Espace complémentaire orthogonal	6
1.2.6 Discussion générale sur l'existence et le nombre de solutions d'un système linéaire	6
1.3 Propriétés des formes quadratiques semi-définies positives	7
1.3.1 Gradient et Hessien d'une forme quadratique	7
1.3.2 Formes quadratiques définies et non définies	8
1.3.3 Propriétés des matrices définies positives et non négatives	8
1.4 Convexité	8
1.4.1 Ensemble convexe	9
1.4.2 Propriétés des ensembles convexes	9
1.4.3 Fonction convexe	9
1.4.4 Propriétés des fonctions convexe	9
1.5 Conclusion	10
2 Optimisation non linéaire	11
2.1 Introduction	11
2.1.1 Position su problème et définition	11
2.1.2 Conditions nécessaire de minimalité locale	12
2.2 Méthodes numériques pour la minimisation d'une fonction différentiable	12
2.2.1 Les méthodes du gradient	13
2.2.2 Méthode de Newton	14
2.3 Optimisation non linéaire avec contrainte	15
2.3.1 Conditions nécessaires de minimalité de Karush-Kuhn-Tucker	15
2.3.2 Condition d'optimalité pour le cas des contraintes linéaires de type égalité	17
2.3.3 Condition d'optimalité pour le cas des contraintes de type inégalités . .	18
2.4 Optimisation convexe	18
2.4.1 Problème quadratique convexe (PQC)	19
2.5 Dualité en programmation quadratique convexe	20
2.5.1 Dualité en programmation convexe	20
2.6 Conclusion	21

3	Méthodes de résolution en programmation quadratique convexe	22
3.1	Introduction	22
3.2	La méthode des points intérieurs	22
3.2.1	Position du problème	23
3.2.2	Technique de la méthode	23
3.2.3	La technique des chemins centraux	24
3.3	La méthode d'activation des contraintes	25
3.3.1	Définition	26
3.3.2	Critère d'optimalité	26
3.3.3	Itération de la méthode	26
3.3.4	Schéma de l'algorithme	28
3.4	La méthode du gradient projeté	31
3.4.1	Position de problème	31
3.4.2	Principe de la méthode	32
3.4.3	Algorithme de la méthode	32
3.5	Méthode adaptée de support	34
3.5.1	position du problème et définition	34
3.5.2	Formule d'accroissement de la fonction objectif	35
3.5.3	Critère d'optimalité	36
3.5.4	Critère de suboptimalité	36
3.6	Méthode de résolution	36
3.6.1	Construction d'une direction d'amélioration adaptée	37
3.6.2	Changement de plan	37
3.6.3	Estimation de la suboptimalité	38
3.6.4	Changement de support	39
3.7	Algorithme de Résolution	40
3.8	Exemple numérique	41
3.9	Récapitulatif	45
	Conclusion générale	46
	Bibliographie	47

Introduction générale

La programmation mathématique est une branche des mathématiques appliquées ayant pour objet l'étude théorique des problèmes d'optimisation, ainsi que la conception et la mise en œuvre des algorithmes de résolution. En effet, la programmation mathématique mono-critères résume à rechercher un n -uplet $x \in \mathbb{R}^n$ qui maximise (ou minimise) une fonction scalaire (fonction dite fonction objectif) sous des contraintes linéaires ou non linéaires et de type inégalités et (ou) égalités.

Suivant la nature de la fonction objectif, et celles des fonctions intervenant dans les contraintes, on tombe dans une catégorie particulière de programmation mathématique; par exemple : lorsque la fonction objectif est linéaire et les contraintes sont linéaires, on obtient un problème de programmation linéaire. Si la fonction objectif est quadratique, on a affaire à un problème de programmation quadratique; en dehors de ces cas, on parle de programmation non linéaire dont la programmation quadratique sujet de notre travail n'est qu'un cas particulier.

D'ailleurs la programmation quadratique est considérée comme une transition naturelle de la programmation linéaire vers la programmation non linéaire. En effet la majorité des méthodes développées pour le cas quadratique sont des extensions directes de celle de la programmation linéaire. De plus les algorithmes élaborés pour le cas de l'optimisation non linéaire reposent essentiellement sur l'approche quadratique.

L'optimisation quadratique trouve son application dans plusieurs domaines tels que l'économie, les sciences de l'ingénieur, la recherche opérationnelle et la commande optimale. Elle s'adapte mieux à la réalité en modélisant par exemple le risque et la distance. En effet le problème de gestion de portefeuille de Markowitz [26] est souvent utilisé par les économistes.

Historiquement, Barankin et Dorfman [2] furent les premiers à remarquer qu'en combinant les conditions d'optimalité de Lagrange avec celles du système original, la solution optimale d'un problème quadratique était une solution de base d'un système élargi ayant la propriété que seuls certains couples de variables figuraient dans la base. De son côté Markowitz montra qu'il est possible de modifier le système élargi et d'engendrer paramétriquement une classe de solutions de base ayant la propriété particulière ci-dessus et convergeant vers l'optimum en un nombre fini d'itérations. Enfin, Wolf [23] montra qu'en modifiant légèrement la méthode de simplexe d'une façon à ne pas autoriser l'introduction d'une variable dans la base si sa variable complémentaire s'y trouvait déjà, on parvenait aisément à l'optimum recherché.

D'autres méthodes ont été développées pour la résolution de ce type de problème, parmi lesquelles on peut citer :

- *Méthode d'activation des contraintes* : [27] Elle est utilisée dans le cas des contraintes d'inégalité, et où on considèrera certaines contraintes d'inégalité comme des contraintes d'égalité pendant un certain nombre d'itérations. La première méthode est mise au point

par Fletcher en 1971 [12] ; par la suite d'autres auteurs ont fait des raffinements numériques à cette dernière tel que Gil et Al en 1978 ainsi que Gould en 1991. Goldfarb et Idnani [2] en 1983 ont développé une méthode duale pour les cas des programmes quadratiques strictement convexes.

- *Méthode des points intérieurs* : ces méthodes, ont été développées en 1984 par Karmarkar [19] dans le cadre de problèmes linéaires, avant d'être généralisées à d'autres problèmes plus généraux notamment la programmation quadratique et ou convexe. Dans cette méthode, les itérés approchent la solution par l'intérieur de domaine et nécessitent un nombre d'itérations qui croît de façon polynomiale avec le nombre de variables.
- *La Méthode adaptée de support* : [21, 22, 28] Cette méthode est développée par R. Gabasov et F.M. Killirova [28] pour résoudre initialement le problème d'un contrôle optimal, puis étendue à la résolution des problèmes de la programmation linéaire et quadratique convexe sous forme générale.

Cette dernière, qui constitue une variante de la méthode directe de support qui est une généralisation de la méthode du simplexe, est développée par les mêmes auteurs dans les années 70, dont le principe est partant d'une solution initiale et chaque itération consiste à trouver une direction d'amélioration et un pas maximal le long de cette direction en changeant tous les indices non optimaux.

Le problème de cette méthode est pratiquement le même que celui de la méthode directe de support, au lieu d'utiliser une direction standard, on utilisera une direction dite adaptée ; i.e : ces composantes non basiques sont calculées en fonction des composantes de la solution courante dans le but d'assurer un accroissement maximal.

L'avantage de cette méthode réside dans le fait qu'elle permet de calculer la solution optimale à ϵ -près où $\epsilon \geq 0$ choisi à l'avance.

Ce travail commence par une introduction qui s'articule autour des trois chapitres. Dans le premier chapitre nous faisons quelques rappels algébriques, certains résultats classiques concernant les fonctions quadratiques, ainsi que la notion de la convexité.

Le deuxième chapitre, nous avons séparé le cas d'optimisation non linéaire sans contraintes et avec contraintes et nous avons présenté deux méthodes de minimisation de fonction différentiables à savoir la méthode du gradient à pas optimal et la méthode de Newton.

Dans le dernier chapitre nous présenterons en détails les méthodes existantes pour la résolution des problèmes quadratiques convexes cités ci-dessus. Et en fin nous terminons notre travail avec une conclusion générale et quelques perspectives.

Chapitre 1

Rappels mathématiques

1.1 Introduction

Dans ce chapitre nous allons donner quelques notions sur l'algèbre linéaire et les propriétés des formes quadratique , ainsi que la notion des ensembles et fonctions convexes .

1.2 Rappel sur l'algèbre linéaire

1.2.1 Vecteurs et matrices

Définition 1.2.1 Soient $n, m \in \mathbb{N}^*$. une matrice d'ordre $m \times n$ à coefficients dans \mathbb{R} est un tableau a deux dimensions, ayant m lignes et n colonnes, représenté sous la forme suivante :

$$A = A(I, J) = (a_{ij}, i \in I, j \in J) = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

où $I = \{1, 2, \dots, m\}$ et $J = \{1, 2, \dots, n\}$ représentent respectivement l'ensemble des indices des lignes et des colonnes de A. Pour des calculs pratiques, la matrice A se note aussi :

$$A = (a_1, a_2, \dots, a_n) = \begin{pmatrix} A'_1 \\ A'_2 \\ \vdots \\ A'_i \\ \vdots \\ A'_m \end{pmatrix}$$

où $a_j = A(I, j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ est un vecteur colonne de dimension m .

$A'_i = A(i, J) = (a_{i1}, a_{i2}, \dots, a_{in})$ et un vecteur ligne de dimension n .
Le symbole $(\cdot)'$ est celui de la transposition.

1.2.2 Matrices et vecteurs partitionnés

On peut faire le produit d'une matrice A et d'un vecteur x Après les avoir partitionnés judicieusement. On dit alors qu'on a effectué un produit en blocs. En effet si l'on a

$$A = (A_1 \mid A_2), x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Alors on peut écrire

$$Ax = (A_1 \mid A_2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A_1x_1 + A_2x_2$$

De même pour

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

l'équation $Ax = b$ peut alors s'écrire

$$\begin{cases} A_{11}x_1 + A_{12}x_2 = b_1 \\ A_{21}x_1 + A_{22}x_2 = b_2 \end{cases}$$

On peut partitionner une matrice d'une manière arbitraire. Par exemple, si $A = A(I, J)$ est une matrice d'ordre $(m \times n)$ et que J_B et J_N sont deux sous-ensembles quelconques de J , tels que $\det(J_B) = m$, $J_B \cup J_N = J$, $J_B \cap J_N = \emptyset$ alors on peut partitionner A de la façon suivante :

$$A = (a_1, a_2, \dots, a_j, \dots, a_n) = [A_B \mid A_N],$$

avec $A_B = A(I, J_B)$, $A_N = A(I, J_N)$. Si $x = x(J) = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$, $x_B = x(J_B)$, $x_N = x(J_N)$, alors on peut écrire

$$\begin{aligned} Ax &= \sum_{j=1}^n a_j x_j = \sum_{j \in J_B} a_j x_j + \sum_{j \in J_N} a_j x_j \\ &= A(I, J_B)x(J_B) + A(I, J_N)x(J_N) \\ &= A_B x_B + A_N x_N \end{aligned}$$

1.2.3 Espace vectoriel

Un vecteur x de \mathbb{R}^n est une collection ordonnée $x = (x_1, x_2, \dots, x_n)'$ de n réels x_j , $j = 1, 2, \dots, n$ appelés composantes de x . Le nombre n est appelé dimension du n vecteur. L'espace \mathbb{R}^n est l'ensemble de toutes les collections de ce type. Il est muni des deux opérations linéaires de base :

- **Addition** : La somme des deux vecteurs $x, y \in \mathbb{R}^n$ est un vecteur de \mathbb{R}^n , défini par :

$$x + y = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)' \tag{1.1}$$

- **Multiplication par des réels** : Soit $x = (x_1, x_2, \dots, x_n)'$ un vecteur de \mathbb{R}^n . La multiplication du vecteur x par le réel λ est aussi un vecteur de \mathbb{R}^n , défini comme suit :

$$\lambda x = (\lambda x_1, \lambda x_2, \dots, \lambda x_n)' \tag{1.2}$$

La structure que nous obtenons (l'ensemble de tous les vecteurs n -dimensionnels avec les deux opérations qu'on vient de définir) s'appelle l'espace vectoriel réel \mathbb{R}^n n -dimensionnel.

Sous espace linéaire

Un sous-ensemble \mathbf{S} non vide de \mathbb{R}^n est appelé sous-espace linéaire de \mathbb{R}^n si les deux opérations (1.1) et (1.2) sont stable dans \mathbf{S} , autrement dit :

$$\forall x, y \in \mathbf{S}, \quad \forall \alpha, \beta \in \mathbb{R} \Rightarrow \alpha x + \beta y \in \mathbf{S}$$

.

Sous-espace affine

Un sous-ensemble A de E est sous-espace affine si :

$$\forall x, y \in A \forall \alpha \in \mathbb{R}, \quad \alpha x + (1 - \alpha)y \in A$$

Autrement dit, un sous-espace affine contient toujours la droite passant par deux de ces points x et y .

Combinaison linéaire

Étant donnés k vecteurs a_1, a_2, \dots, a_k , et un ensemble de k scalaires $\lambda_1, \lambda_2, \dots, \lambda_k$. Le vecteur b est dit combinaison linéaire des vecteurs $\{a_i\}_{i=1\dots k}$ s'il s'écrit sous la forme :

$$b = \lambda_1 a_1 + \lambda_2 a_2 + \dots + \lambda_k a_k. \quad (1.3)$$

L'expression (1.3) peut être écrite sous forme du produit d'une matrice et d'un vecteur comme suit :

$$b = (a_1, a_2, \dots, a_k) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{pmatrix} = A\lambda, \quad (1.4)$$

où a_i est la $i^{\text{ème}}$ colonne de la matrice A , et λ_i est le $i^{\text{ème}}$ élément du vecteur colonne λ .

Une combinaison linéaire où tous les coefficients sont nuls est appelée combinaison linéaire triviale; dans le cas contraire elle est appelée *combinaison non triviale*.

1.2.4 Noyau , image et rang d'une matrice

Définition 1.2.2 Soit A une matrice d'ordre $m \times n$. Le noyau de A est l'ensemble

$$N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$$

L'image ou l'espace image de A est l'ensemble :

$$R(A) = \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n, Ax = y\}$$

Définition 1.2.3 On appelle rang d'une matrice, la dimension de l'image de la matrice A :

$$\text{rang}(A) = \dim R(A)$$

1.2.5 Espace complémentaire orthogonal

Le produit scalaire (produit euclidien) de deux n -vecteurs x et y est le scalaire défini par :

$$x'y = \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

Le produit scalaire possède les propriétés suivantes :

- Commutativité
- Distributivité par rapport à l'addition des vecteurs.
- Positivité

Si $x'y = 0$, les vecteurs x et y sont dit orthogonaux et on note $x \perp y$.

Deux parties A et B de \mathbb{R}^n sont dites orthogonales si :

$$\forall a \in A, \forall b \in B, a'b = 0$$

1.2.6 Discussion générale sur l'existence et le nombre de solutions d'un système linéaire

Soient m et n deux nombres entiers. Un système de m équations à n inconnues x_1, x_2, \dots, x_n s'écrit comme suit :

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \quad \dots \quad \dots \quad \dots = \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m, \end{cases} \quad (1.5)$$

où les coefficients a_{ij} sont des réels. Les nombres b_1, b_2, \dots, b_m sont appelées les membres libres du système 1.5 ou les seconds membres. En posant

$$A = (a_{ij}, 1 \leq i \leq m, 1 \leq j \leq n), x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

le système (1.5) peut s'écrire sous la forme matricielle suivante :

$$Ax = b. \quad (1.6)$$

Tout vecteur x vérifiant les équations (1.5) s'appelle solution du système. Le système (1.5) est dit compatible s'il possède une ou plusieurs solutions. Dans le cas contraire, il est dit incompatible ou impossible. D'une manière générale, le système (1.6) possède une solution si le vecteur $b \in R(A)$, i.e, au sous-espace engendré par les vecteurs colonnes de la matrice A . Lorsque le vecteur b est nul, le système (1.6) est dit homogène. Tout système homogène possède la solution triviale $x = 0$.

Définition 1.2.4 Le système linéaire (1.6) est dit de rang complet en lignes si $\text{rang}(A) = m$, avec $m \leq n$, et de rang complet en colonnes si $\text{rang}(A) = n$, avec $m \geq n$.

Lemme 1.2.1 Soit $m \leq n$ et $\text{rang}A = m$, Alors le système $Ax = b$ admet toujours des solutions, quel que soit le second membre b :

- (a) une solution unique si $m = n$,
- (b) une infinité de solutions si $m < n$.

1.3 Propriétés des formes quadratiques semi-définies positives

Définition 1.3.1 Une fonction réelle de la forme suivante :

$$F(x) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j \quad (1.7)$$

est dite forme quadratique de n variables x_1, x_2, \dots, x_n .

En posant $x' = (x_1, x_2, \dots, x_n)$, et $A = (a_{ij}, 1 \leq i, j \leq n)$, la formule (1.7) s'écrit sous la forme suivante :

$$F(x) = x'Ax$$

1.3.1 Gradient et Hessian d'une forme quadratique

Définition 1.3.2 Soit une fonction de classe C^1 $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Le gradient de la fonction f est défini par :

$$g(x) = \nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = 2Dx + c$$

où $\frac{\partial f}{\partial x_i}$ est les dérivées partielles de $f(x)$ par rapport à x_i

Définition 1.3.3 Soit une fonction de classe C^2 $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Le hessian de la fonction f est définie par :

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Définition 1.3.4 Soit une fonction de classe C^1 $f : \mathbb{R}^n$. La dérivée directionnelle de f dans la direction d au point x est :

$$\begin{aligned} \frac{\partial f}{\partial d}(x) &= \lim_{h \rightarrow 0} \frac{f(x + hd) - f(x)}{h} \\ &= \frac{d}{dh} f(x + hd) \Big|_{h=0} \\ &= \frac{\partial f}{\partial x_1}(x + hd) \Big|_{h=0} d_1 + \dots + \frac{\partial f}{\partial x_n}(x + hd) \Big|_{h=0} d_n \\ &= \nabla' f(x) d \end{aligned}$$

Si $\|d\| = 1$ alors la dérivée directionnelle est le taux d'accroissement de f dans la direction d au point x . Le taux d'accroissement est maximal dans la direction du gradient.

1.3.2 Formes quadratiques définies et non définies

Soit la forme quadratique $F(x) = x'Dx$

Définition 1.3.5 $F(x)$ est dite définie positive si $x'Dx > 0, \forall x \in \mathbb{R}^n$ et $x \neq 0$. Elle est dite semi-définie positive ou définie non négative si $x'Dx \geq 0, \forall x \in \mathbb{R}^n$ et $x \neq 0$

Définition 1.3.6 $F(x)$ est dite définie négative si $x'Dx < 0, \forall x \in \mathbb{R}^n$ et $x \neq 0$. Elle est définie non positive si $x'Dx \leq 0 \forall x \in \mathbb{R}^n$ et $x \neq 0$

Définition 1.3.7 Une matrice symétrique D dite matrice définie positive (non négative) et se note $D > 0, (D \geq 0)$, si elle associée à une forme quadratique définie positive (non négative).

Définition 1.3.8 Une forme quadratique $F(x)$ est dite non définie si $F(x)$ est positive pour certaines valeurs de x et négative pour d'autres.

1.3.3 Propriétés des matrices définies positives et non négatives

Les matrices symétriques définies positives ont des propriétés très intéressantes. En voici quelques unes :

Propriété 1.3.1 Soit une matrice symétrique $D = (d_{ij}, 1 \leq i, j \leq n)$ si D est définie positive (non négative), alors on a :
 $d_{ij} > 0 (d_{ij} \geq 0), \forall i = 1, 2, 3 \dots, n$.

Propriété 1.3.2 Soit la matrice D partitionnée de la manière suivante :

$$D = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}$$

Si $D > 0 (D \geq 0)$, alors les sous matrices principales D_{11} et D_{22} sont aussi définies positives (non négatives). d'une manière générale, toute sous matrice principale d'une matrice définie positive (non négative) est définie positive (non négative)

Propriété 1.3.3 Un élément diagonal d'une matrice symétrique D définie non négative ne peut s'annuler que si les autres éléments de la même ligne et colonne s'annulent aussi.

Propriété 1.3.4 Soit D une matrice symétrique définie non négative. Si x est un point quelconque mais fixé de \mathbb{R}^n tel que $x'Dx = 0$, on a alors $Dx = 0$

1.4 Convexité

La convexité joue un rôle très important dans la théorie classique de l'optimisation. Elle est un outil indispensable pour la recherche des conditions à la fois nécessaires et suffisantes d'optimalité.

1.4.1 Ensemble convexe

Définition 1.4.1 Un ensemble \mathbf{C} dans \mathbb{R}^n est dit convexe si $\forall x_1, x_2 \in \mathbf{C}, \lambda \in [0, 1]$, le vecteur $x = \lambda x_1 + (1 - \lambda)x_2$ appartient à \mathbf{C} .

1.4.2 Propriétés des ensembles convexes

Si \mathbf{C}_1 et \mathbf{C}_2 sont deux ensembles convexes de \mathbb{R}^n , alors $K = \mathbf{C}_1 \cap \mathbf{C}_2$ est convexe et $K = \{x \mid x = x_1 + x_2, x_1 \in \mathbf{C}_1 \text{ et } x_2 \in \mathbf{C}_2\}$ est convexe.

1.4.3 Fonction convexe

une fonction réelle f est définie sur un ensemble convexe \mathbf{C} de \mathbb{R}^n , est dite convexe, si pour tous les points x, y de \mathbf{C} et pour tout nombre réel positif ou nul λ tel que $0 \leq \lambda \leq 1$, l'inégalité suivante est vérifiée :

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \quad (1.8)$$

Définition 1.4.2 Une fonction convexe $f(x)$, $x \in \mathbf{C}$ est dite strictement convexe si l'inégalité (1.2) est stricte pour tous les points x_1, x_2 de \mathbf{C} , avec $x_1 \neq x_2$ et $\lambda \in]0, 1[$.

1.4.4 Propriétés des fonctions convexe

Propriété 1.4.1 Soit f une fonction réelle définie sur un ensemble convexe $\mathbf{C} \subset \mathbb{R}^n$. Alors f est convexe si et seulement si son épigraphe

$$\text{epi}(f) = \{(x, r) : x \in \mathbf{C}, r \geq f(x)\} \subset \mathbb{R}^n \times \mathbb{R}$$

est un ensemble convexe.

Propriété 1.4.2 Soit f une fonction réelle définie sur un ensemble convexe $\mathbf{C} \subset \mathbb{R}^n$. Alors f est convexe si et seulement si :

$$f\left(\sum_{k=1}^n \lambda_k s_k\right) \leq \sum_{k=1}^n \lambda_k f(s_k) \quad (1.9)$$

ou $s_i \in \mathbf{C}, i = 1, 2, \dots, p, \lambda_i \geq 0, \sum_{i=1}^p \lambda_i = 1$

Propriété 1.4.3 Soit une fonction réelle de classe C^1 , définie sur un ensemble convexe $\mathbf{C} \subset \mathbb{R}^n$. Alors f est convexe si et seulement si :

$$f(y) - f(x) \geq (y - x)' \nabla f(x), \forall x, y \in \mathbf{C} \quad (1.10)$$

Propriété 1.4.4 Soit une fonction réelle de classe C^2 , définie sur un ensemble convexe $\mathbf{C} \subset \mathbb{R}^n$. Alors f est convexe si et seulement si :

$$(y - x)' H(x) (y - x) \geq 0, \forall x, y \in \mathbf{C} \quad (1.11)$$

Propriété 1.4.5 Si $\mathbf{C} \subset \mathbb{R}^n$ est un ouvert convexe, alors f est convexe si et seulement si

$$H(x) \geq 0, \forall x \in \mathbf{C}$$

Nous remarquons qu'une fonction quadratique semi-définie positive est une fonction convexe.

Définition 1.4.3 Le vecteur $d \in \mathbb{R}^n$, $d \neq 0$ est appelé direction admissible en point $x \in \mathbf{S}$ s'il existe un réel $\alpha > 0$ tel que $x + \theta d \in \mathbf{S}$, $\forall \theta \in [0, \alpha]$.

Si x est un point intérieur, alors toutes les directions sont admissibles.

1.5 Conclusion

Dans ce chapitre nous avons fait un petit rappels sur l'algèbre linéaire , les propriétés des formes quadratiques et enfin nous avons vu la notion de la convexité et les propriétés des fonctions convexe

Chapitre 2

Optimisation non linéaire

2.1 Introduction

Ce chapitre, est consacré à l'optimisation non linéaire, nous allons reprendre des résultats fondamentaux de l'optimisation non linéaire sans contraintes et avec contraintes, tous ce qui est en relation avec les conditions de minimalité (1 er ordre et 2 ème ordre), ainsi que les condition de Karush-Kuhn-Tucker (KKT)

2.1.1 Position su problème et définition

Soit f une fonction non linéaire, définie de \mathbb{R}^n dans \mathbb{R} et de classe C^1 . Le problème de programmation non linéaire consiste à trouver $x^* \in \mathbf{S} \subset \mathbb{R}^n$ tel que

$$f(x^*) = \min f(x), x \in \mathbf{S} \quad (2.1)$$

L'ensemble des contraintes \mathbf{S} est donné en général par des équations et des inéquations linéaires ou non. Il peut être représenté de la manière suivante :

$$\mathbf{S} = \{x \in \mathbb{R}^n / g(x) \leq 0\} \quad (2.2)$$

où la fonction g est une fonction vectorielle définie de \mathbb{R}^n dans \mathbb{R}^m et de classe C^1 , avec

$$g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix} \quad \nabla g(x) = (\nabla g_1(x), \nabla g_2(x), \dots, \nabla g_m(x))$$
$$\nabla g(x) = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_2}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_1} \\ \frac{\partial g_1}{\partial x_2} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_m}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_1}{\partial x_n} & \frac{\partial g_2}{\partial x_n} & \dots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}$$

Les résultats fondamentaux de l'optimisation non linéaire sont obtenus dans le cas ou la fonction objectif f est non linéaire et où les contraintes qui définissent l'ensemble \mathbf{S} sont linéaires. Le problème que l'on étudie ici dans ce travail est celui de la recherche du minimum d'une fonction quadratique convexe.

Définition 2.1.1 Soit f une fonction réelle définie sur un ensemble ouvert \mathbf{S} de \mathbb{R}^n . La fonction f admet un minimum local (strict) en $x^* \in \mathbf{S}$ si $\exists \mathbf{B}(x^*, \epsilon) = \{x / \|x - x^*\| < \epsilon\} \subset \mathbf{S}$ tel que :

$$f(x) \geq f(x^*), \forall x \in \mathbf{B}(x^*, \epsilon) \quad (f(x) > f(x^*)) \quad (2.3)$$

Définition 2.1.2 Soit f une fonction réelle, définie sur un ensemble de \mathbf{S} de \mathbb{R}^n . La fonction f admet un minimum global en $x^* \in \mathbf{S}$ si :

$$f(x) \geq f(x^*), \forall x \in \mathbf{S} \quad (2.4)$$

2.1.2 Conditions nécessaire de minimalité locale

Définition 2.1.3 Le vecteur $d \in \mathbb{R}^n$, $d \neq 0$ est appelé direction admissible en point $x \in \mathbf{S}$ s'il existe un réel $\alpha > 0$ tel que $x + \theta d \in \mathbf{S}$, $\forall \theta \in [0, \alpha]$.

Si x est un point intérieur, alors toutes les directions sont admissible.

Lemme 2.1.1 Soit f une fonction réelle de classe \mathbb{C}^1 définie sur l'ensemble S de \mathbb{R}^1 , Si x^* est un minimum local de f sur S , alors pour toute direction d admissible en x^* on a :

$$d' \nabla f(x^*) \geq 0$$

Théorème 2.1.1 Si x^* est un minimum local de f sur \mathbb{R}^n et si f est différentiable en x^* , alors :

$$\nabla f(x^*) = 0 \quad (2.5)$$

le point x^* vérifiant la condition (2.5) est appelé un point stationnaire.

Théorème 2.1.2 Si x^* est un minimum local de f sur \mathbb{R}^n et si f est deux fois différentiable en x^* , alors

- (i) $\nabla f(x^*) = 0$ (stationnarité)
- (ii) $H(x)$ est semi-définie positive.

2.2 Méthodes numériques pour la minimisation d'une fonction différentiable

Les méthodes numériques de minimisation d'une fonction sur \mathbb{R}^n consistent toutes à rechercher un point stationnaire \bar{x} à partir d'une approximation initiale donnée x^0 . L'algorithme général se construit sur la forme suivante :

$$x^{k+1} = x^k + \theta_k l^k, \quad k = 0, 1, \dots \quad (2.6)$$

où l^k est K^{ime} direction d'amélioration ou de descente et θ_k le K^{ime} pas le long de cette direction. ces méthodes numériques se différencient seulement par le choix de l^k et θ_k .

Les méthodes du gradient sont les plus couramment utilisées parmi les algorithmes de minimisations des fonctions différentiables.

2.2.1 Les méthodes du gradient

Toutes les méthodes du gradient sont basées sur le schéma suivant :

$$x^{k+1} = x^k + \theta_k l^k, \quad k = 0, 1, \dots$$

les différents algorithmes existants ne se différencient que par le choix de θ_k le long de la direction de descente qui est l'anti-gradient.

Parmi ces dernières on distingue notamment :

- L'algorithme du gradient à pas fixe ($\theta_k = \theta = \text{constant}$)
- L'algorithme du gradient à pas optimal (où l'algorithme de la descente la plus rapide)
- L'algorithme du gradient à pas prédéterminé (où méthode de la série divergente)

Ces algorithmes en général ne sont pas finis ; on doit alors définir un test d'arrêt . Ce dernier est basé sur l'un des critères suivants :

- i) $\|\nabla f(x^k)\| < \epsilon$, $\epsilon > 0$ donné $\epsilon = 10^{-6}\|\nabla f(x^0)\|$
- ii) $\text{Max}|\frac{\partial f}{\partial x_i}| < \epsilon, i = 1, 2, \dots, n$, ϵ donné.
- iii) $|f(x^k) - f(x^{k-1})| < \epsilon$, ϵ donné.

Une fois que le point stationnaire $\bar{x} \simeq x^k$ est trouvé , il restera alors à vérifier que le vecteur x^k correspond bien à un minimum, car ça peut être un point selle.

Dans ce travail nous allons citer que l'algorithme du gradient a pas optimal.

Algorithme du gradient à pas optimal

Pour cette méthode, les itérations sont construites selon la formule itérative :

$$x^{k+1} = x^k - \theta_k \nabla f(x^k), \quad k = 0, 1, 2, \dots \quad (2.7)$$

où le pas θ_k vérifie la condition

$$f(x^k - \theta_k \nabla f(x^k)) = \min_{\theta \geq 0} f(x^k - \theta \nabla f(x^k))$$

Cette condition signifie que le mouvement le long de la direction de l'anti-gradient se poursuit tant que la fonction $f(x)$ décroît.

Si $\lim_{\|x\| \rightarrow \infty} f(x) = \infty$; alors la suite $\{x^k\}$ construite selon la forme itérative (2.7) vérifie la relation suivante :

$$\lim_{k \rightarrow \infty} \nabla f(x^k) = 0$$

Propriété 2.2.1 L'algorithme du gradient à pas optimal possède la propriété suivante : deux directions de déplacement successives sont orthogonales .

En effet, le pas θ_h est calculé de façon à minimiser la fonction [10]

$$\varphi(\theta) = f(x^k + \theta l^k); \quad l^k = -\nabla f(x^k).$$

On doit avoir

$$\varphi'(\theta_k) = \nabla f'(x^k + \theta_k l^k) \cdot l^k = \nabla f'(x^{k+1}) l^k = 0$$

D'où l'on déduit $\langle l^{k+1}, l^k \rangle = 0$

Cette propriété des directions successives orthogonales a des effets néfastes pour des fonctions dual conditionnées. En effet, le nombre d'itération peut être considérable pour minimiser ce genre de fonction.

2.2.2 Méthode de Newton

Lors de la minimisation d'une fonction $f(x)$ sur \mathbb{R}^n , le travail consiste à chercher un point \bar{x} tel que

$$\nabla f(x^*) = g(x^*) = 0, \quad g : \mathbb{R}^n \rightarrow \mathbb{R}$$

la méthode de newton appliquée à l'équation $g(x^*) = \nabla f(x^*) = 0$ construit alors les approximations successives de la manière suivante :

$$x^{k+1} = x^k - \left[\frac{\partial g}{\partial x}(x^k) \right]^{-1} g(x^k) \quad , \quad k = 0, 1, 2, \dots \quad (2.8)$$

où

$$g = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_i \\ \vdots \\ g_n \end{pmatrix}, \quad \frac{\partial g}{\partial x}(x^k) = \left(\frac{\partial g_i}{\partial x_j}(x^k), \quad 1 \leq i, j \leq l \right)$$

En tenant compte du fait que $\frac{\partial g}{\partial x}(x^k) = \nabla^2 f(x^k) = H(x^k)$ La formule itérative (3.27) s'écrit alors :

$$x^{k+1} = x^k - H^{-1}(x^k) \nabla f(x^k), \quad k = 0, 1, 2, \dots \quad (2.9)$$

Le point x^{k+1} peut être interprété comme la racine de l'approximation linéaire de la fonction $g(x) = \nabla f(x)$ au voisinage de x^k .

une autre interprétation de la méthode de newton est la suivante : elle consiste à minimiser l'approximation quadratique $q(x)$ de $f(x)$ au voisinage de x^k :

$$f(x) \simeq q(x) = f(x^k) + \nabla f'(x^k)(x - x^k) + \frac{1}{2}(x - x^k)' \nabla^2 f(x^k)(x - x^k)$$

Si le hessien $H(x^k) = \nabla^2 f(x^k)$ est défini positif, le point x^{k+1} de la méthode de newton est alors pris comme étant le minimum de la fonction $q(x)$:

$$\nabla q(x^{k+1}) = \nabla f(x^k) + \nabla^2 f(x^k)(x^{k+1} - x^k) = 0$$

D'où la formule itérative :

$$x^{k+1} = x^k - [\nabla^2 f(x^k)]^{-1} \nabla f(x^k) \quad (2.10)$$

On remarque que dans la méthode de Newton, la direction ainsi que le pas sont fixés ;

$$l^k = -[\nabla^2 f(x^k)]^{-1} \nabla f(x^k), \quad \theta_k = 1$$

La méthode de newton possède une vitesse de convergence quadratique, mais la convergence ne peut être assurée que localement. En effet, si l'approximation initiale x^0 est trop éloignée de \bar{x} la méthode peut ne pas converger.

Remarque 2.2.1 Lorsque le hessien $\nabla^2 f(x^k)$ n'est pas défini positif, la direction de déplacement l^k dans la méthode de newton peut ne pas être une direction de descente.

Dans ce cas, la méthode de newton peut fournir aussi des maxima locaux ou des points -selles.

Remarque 2.2.2 Dans le cas d'une fonction quadratique convexe, la méthode de newton donne la solution \bar{x} en une seule itération :

$$x^1 = x^0 - D^{-1}(Dx^0 + c) = x^0 - D^{-1}c = -D^{-1}c = x^*$$

Cette unique itération, n'est que la solution du système linéaire $Dx - c$.

Modification de la méthode de newton

Les inconvénient de la méthode de newton sont liés à la nécessité de calculer à chaque itération le hessien et son inverse, ainsi qu'au fait que le domaine de convergence est petit. Pour pallier à la première difficulté on réécrit l'itération générale (3.28) comme suit

$$x^{k+1} = x^k - H^{-1}(x^k) \nabla f(x^k) \quad (2.11)$$

ou le calcul du hessien et de son inverse ce fait seulement on x^0 qui l'approximation initiale. pour pallier à la deuxième difficulté on modifie la formule (2.10) de la manière suivante

$$x^{k+1} = x^k - \theta_k H^{-1}(x^k) \nabla f(x^k), \quad 0 < \theta \leq 1 \quad (2.12)$$

En effet, le domaine de convergence s'agrandit dans la méthode (3.30) lorsqu'on fait diminuer θ_k .

En revanche, la vitesse de convergence dans ce cas ne sera pas aussi rapide que pour la méthode de newton proprement dite.

2.3 Optimisation non linéaire avec contrainte

Un problème d'optimisation non linéaire avec contraintes se formule de la manière suivante :

$$\min F(x), (x \in \mathbf{S}) \quad (2.13)$$

où $\mathbf{S} = \{x \in \mathbb{R}^n / g_i(x) = 0, i = 1, 2, \dots, k, g_i(x) \leq 0, i = k + 1, \dots, m\}$ désigne l'ensemble des solutions réalisable. On suppose que les fonctions F et $g_i (i = 1, \dots, m)$ sont de classe C^1 ; c'est à dire continument différentiables sur \mathbb{R}^n

Définition 2.3.1

- Un vecteur $x \in \mathbb{R}^n$ est appelé solution réalisable ou plan de problème (2.13) s'il vérifie toutes les contraintes de problème i.e $x \in \mathbf{S}$.
- Une solution réalisable x^* est appelé solution optimale du problème (2.13) si

$$F(x^*) \leq F(x), \forall x \in \mathbf{S}$$

et on note $\min_{x \in \mathbf{S}} F(x) = F(x^*)$

- $x^* \in \mathbf{S}$ est appelé minimum local de problème (2.13) s'il existe un réel $\epsilon > 0$, tel que :

$$F(x^*) \leq F(x), \quad \forall x \in \mathbf{S} \cap \mathbb{B}(x^*, \epsilon)$$

, où $\mathbb{B}(x^*, \epsilon)$ est la boule de centre x^* et de rayon ϵ

2.3.1 Conditions nécessaires de minimalité de Karush-Kuhn-Tucker

Soit f une fonction non linéaire, définie de \mathbb{R}^n dans \mathbb{R} et de classe C^1 . Le problème de minimisation de la fonction f avec des contraintes linéaires de type inégalités se présente ainsi :

$$f(x) \rightarrow \min \quad (2.14)$$

sous les contraintes suivantes :

$$\mathbf{S} = \{x \in \mathbb{R}^n / g_i(x) = A'_i x - b_i \leq 0, i \in I = \{1, 2, 3, \dots, m\}\} \quad (2.15)$$

où b est un vecteur de \mathbb{R}^m et A est une matrice d'ordre $m \times n$, formée des vecteurs colonnes et des vecteurs lignes suivants :

$$A = (a_1, a_2, \dots, a_n); \quad A = \begin{pmatrix} A'_1 \\ A'_2 \\ \vdots \\ A'_m \end{pmatrix}; \quad \text{et} \quad g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix}$$

est une fonction définie de \mathbb{R}^n dans \mathbb{R}^m .

Tout vecteur x vérifiant les inégalités (2.15) est appelé solution admissible (réalisable) ou plan du problème (2.1).

Pour que l'ensemble des solutions réalisables \mathbf{S} ne soit pas vide ou ne soit pas réduit à un point isolé, on considérera que $\text{rang } A = m < n$.

Définition 2.3.2 L'ensemble des indices actifs au point x^* est formé des indices i qui vérifient $g_i(x) = 0$. Il est noté comme suit :

$$I_a = I_a(x^*) = \{i \in I / g_i(x^*) = 0\} \tag{2.16}$$

Définition 2.3.3 La fonction $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$ est appelée fonction de Lagrange associée au problème de minimisation de f sur \mathbf{S} , où le vecteur $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \in \mathbb{R}^m, \lambda \geq 0$, formé des multiplications de Lagrange λ_i , est unique .

Le théorème de KKT est fondamental et donne sous certaines conditions qu'on appelle " qualification des contraintes" , une condition nécessaire de minimalité locale pour le problème (2.1).

faire l'hypothèse de qualification des contraintes en x^* qu'on note $(QC)_{x^*}$, revient à chercher l'existence d'une direction d pointant à l'intérieur du cône convexe défini par :

$$K = \{d \in \mathbb{R}^n / \nabla g_i(x^*)'d \leq 0, \forall i \in I_a(x^*)\}. \tag{2.17}$$

C'est à dire que :

$$(QC)_{x^*} \Leftrightarrow \exists d / \nabla g_i(x^*)'d < 0, i \in I_a(x^*) \tag{2.18}$$

Évidemment la vérification directe de (QC) peut être difficile en pratique, c'est pourquoi on a recherché des conditions suffisantes pour que (QC) soit réalisé. Les résultats les plus importants sont rassemblés dans le lemme ci-dessous :

Lemme 2.3.1 Pour que (QC) soit vérifiée en tout point $x \in \mathbf{S}$ il suffit que l'une des conditions (a) et (b) soit réalisée :

- (a) toutes les fonctions g_i sont linéaires (Karlin 1959);
- (b) toutes les fonctions g_i sont convexes et \mathbf{S} a un intérieur non vide (Slater 1950) pour que (QC) soit vérifiée en un point $x^* \in \mathbf{S}$, il suffit que l'on ait :
- (c) les gradients $\nabla g_i(x^*)$, des contraintes saturées en x^* sont linéairement indépendants (Facco et McCormick 1968).

Théorème 2.3.1 Théorème 2.3.2 (théorème de Karush-Kuhn-Tucker 1951)

Si x^* est un minimum local de f sur \mathbf{S} , et si $(QC)_{x^*}$ a lieu, alors il existe $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ tel que :

- (i) $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$
- (ii) $\lambda_i^* g_i(x^*) = 0, i \in I$

ce qu'exprime (i) n'est autre que $\nabla L(x^*, \lambda^*) = 0$, ce qui n'est d'après Lagrange que l'étendu de la condition de minimalité $\nabla f(x^*) = 0$, établie pour un problème sans contraintes, et ce, moyennant l'intervention des auxiliaires qui sont les multiplicateurs λ^* .

2.3.2 Condition d'optimalité pour le cas des contraintes linéaires de type égalité

Le problème se formule sous la forme suivante :

$$\begin{cases} \min F(x), \\ g_i(x) = A'_i x - b_i = 0, \quad \forall i \in I = \{1, 2, \dots, m\} \end{cases} \quad (2.19)$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$ est une fonction continument différentiable, b est un vecteur de \mathbb{R}^m et A une matrice d'ordre $m \times n$, formée des vecteurs colonnes et ligne suivant :

$$A = (a_1, a_2, \dots, a_n); \quad A = \begin{pmatrix} A'_1 \\ A'_2 \\ \vdots \\ A'_m \end{pmatrix}; \quad g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \vdots \\ g_m(x) \end{pmatrix},$$

est une fonction vectorielle définie de $\mathbb{R}^n \rightarrow \mathbb{R}^m$ pour que l'ensemble des solutions réalisables S ne soit pas vide ou ne soit pas réduit à un point isolé, on considérera que $\text{rang} A = m < n$.

Définition 2.3.4 La fonction $L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$ est appelée fonction de Lagrange associée au problème de (2.19) où le vecteur $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m) \in \mathbb{R}^m, \lambda \geq 0$, formé des multiplications de Lagrange λ_i , est unique grâce à la remarque (2.3.2).

Théorème 2.3.3 Soit x^* un minimum du problème (2.19), Alors, il existe nécessairement un vecteur $\lambda \in \mathbb{R}^m$ vérifiant :

$$\nabla F(x^*) + A' \lambda = 0 \quad (2.20)$$

Si de plus A est de rang complet en ligne, alors λ est unique.

La condition (2.20) peut être donné autrement, en utilisant la fonction de Lagrange

$$\nabla_x L(x, \lambda) = \nabla F(x) + A' \lambda = 0 \quad (2.21)$$

De plus, un minimum local est tout d'abord un point réalisable, qui vérifie

$$Ax = b \Rightarrow \nabla_\lambda(x, \lambda) = Ax - b = 0 \quad (2.22)$$

En combinant les relations (2.21) et (2.22), on obtient alors la condition nécessaire d'optimalité de premier ordre pour le problème (2.19)

Théorème 2.3.4 (théorème de Lagrange)

Soit x^* un minimum local (ou global) pour le problème (2.19), Alors il existe un vecteur multiplicateur de Lagrange $\lambda^* \in \mathbb{R}^m$, tel que :

$$\nabla L(x^*, \lambda^*) = 0 \Leftrightarrow \begin{cases} \nabla_{\lambda^*} L(x^*, \lambda^*) = 0 \\ \nabla_{x^*} L(x^*, \lambda^*) = 0 \end{cases} \quad (2.23)$$

Le couple (x^*, λ^*) est appelé point stationnaire de la fonction de Lagrange. La condition nécessaire du second ordre pour le problème (2.19) est la suivante

Théorème 2.3.5 Soit x^* un minimum local pour le problème (2.19) et λ^* un vecteur multiplicateur de Lagrange vérifiant (2.23), alors la matrice $\nabla^2 F(x^*)$ est semi-définie positive sur l'ensemble des points de la variété linéaire $Ay = 0$. Autrement dit :

$$y' \nabla^2 F(x^*) u \geq 0, \quad \forall y \in \nu = \{y \in \mathbb{R}^n \quad : \quad Ay = 0\} \quad (2.24)$$

La condition suffisante de second ordre est la suivante :

Théorème 2.3.6 Soit (x^*, λ^*) un couple de vecteurs vérifiant la condition nécessaire d'optimalité de premier ordre du problème (2.19) i.e ;

$$\nabla L(x^*, \lambda^*) = 0$$

pour que x^0 soit un minimum local du problème (2.19), il est suffisant que la matrice $\nabla^2 F(x^*)$ soit définie positive sur le sous espace vectoriel ν .

2.3.3 Condition d'optimalité pour le cas des contraintes de type inégalités

Considérons maintenant un problème non linéaire avec contraintes linéaires de type inégalité

$$\begin{cases} \min F(x), \\ g_i(x) = A'_i x - b_i \leq 0, \quad \forall i \in I = \{1, 2, \dots, m\} \end{cases} \quad (2.25)$$

Définition 2.3.5 Soit x une solution réalisable du problème(2.25) . l'ensemble des contraintes actives (saturées) au point x est l'ensemble d'indices suivant :

$$I_0 = I_0(x) = \{i \in I \quad : \quad A'_i x = b_i\}$$

Théorème 2.3.7 (théorème de Karush-Kuhn-Tucker 1951)

Si x^0 est un minimum local de f sur \mathbf{S} ,et si $(QC)_{x^*}$ a lieu, alors il existe $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$ tel que :

- (i) $\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0$
- (ii) $\lambda_i^* g_i(x^*) = 0, i \in I$

ce qu'exprime (i) n'est autre que $\nabla L(x^*, \lambda^*) = 0$, ce qui n'est d'après Lagrange que l'étendu de la condition de minimalité $\nabla f(x^*) = 0$, établie pour un problème sans contraintes, et ce, moyennant l'intervention des auxiliaires qui sont les multiplicateurs λ^* .

2.4 Optimisation convexe

L'hypothèse de la convexité apporte élégance et simplicité à la théorie de l'optimisation. En particulier, les condition nécessaires d'optimalité deviennent également suffisantes , et tout le résultat acquiert un caractère global.

Définition 2.4.1 On dit qu'un problème de programmation mathématique est convexe (res. strictement convexe) s'il consiste à minimiser une fonction convexe (res. strictement convexe) sur un domaine convexe.

L'étude des problèmes convexes et des algorithmes de résolution correspondants est l'objet de la programmation convexe. L'hypothèse de convexité est cruciale en optimisation. Notons que

- Les problèmes convexes sont synonymes de minimisation
- Les problèmes convexes sont les \prec bons \succ problèmes de la théorie : ceux pour les quels il existe des algorithmes de résolutions efficaces.
- l'hypothèse de convexité ne garantit cependant ni l'existence ni l'unicité d'une éventuelle solution.

Voici quelques propriétés pour tout problème de programmation convexe :

Propriété 2.4.1 Soit f une fonction définie sur un convexe \mathbf{R}^n . Alors l'ensemble \mathbf{M} des point où f atteint son minimum est convexe.

Propriété 2.4.2 Tout problème strictement convexe admet au plus une solution.

Propriété 2.4.3 Soit $f : \mathbf{C} \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Alors tout minimum local est un minimum global.

Propriété 2.4.4 Si la fonction f est strictement convexe, alors son minimum global est atteint en un seul point x^* .

Définition 2.4.2 Considérons le problème défini dans (2.14) et (2.15) où f est une fonction convexe \mathbf{C}^1 .

Soit (x^*, λ) un couple de vecteurs vérifiant les conditions de KKT :

- (i) $\nabla_{l_x}(x^*, \lambda^*) = 0, \quad \lambda_i^* \geq 0, \quad i \in I;$
- (ii) $\lambda_i^* g_i(x^*) = 0, \quad i \in I.$

Alors le vecteur x^* constitue un minimum global de f .

2.4.1 Problème quadratique convexe (PQC)

Il s'agit d'une classe de problème d'optimisation où la fonction objectif est quadratique, s'écrivant sous la forme $F(x) = \frac{1}{2}x'Dx + C'x$, avec D symétrique, que l'on minimise sur un polyèdre convexe fermé. Ce genre de programmes est convexe dès que la matrice D est semi-définie positive.

' On remarquera qu'un problème linéaire n'est autre qu'un problème quadratique dégénéré ($D = 0$), et c'est toujours un problème convexe.

L'étude des problèmes quadratiques convexes constitue un domaine propre de la théorie de la programmation quadratique; le résultat le plus remarquable est le suivant :

Théorème 2.4.1 *Tout problème quadratique convexe dont la valeur est finie admet au moins une solution.*

2.5 Dualité en programmation quadratique convexe

Le concept de dualité joue un rôle très important en programmation mathématique. Le but est de trouver une formulation alternative équivalente du problème de programmation mathématique, qui convient le plus au calcul ou qui a une signification théorique importante. Le problème original est appelé problème primal et le problème transformé est le problème dual. Souvent, les variables du dual peuvent être interprétées comme des multiplicateurs de Lagrange pour le cas linéaire et prennent la valeur λ^* comme solution duale, quand λ^* est le multiplicateur associé à la solution optimale primale x^* . Cependant, dans le cas non linéaire il existe toujours une fonction objectif, souvent reliée à la fonction de Lagrange, qui doit être optimisée. Ici, on traitera de la dualité associée à un problème de programmation convexe comme problème primal. Il est important de remarquer que si le problème primal n'est pas convexe, alors le problème dual peut bien ne pas avoir de solution à partir de laquelle la solution primale peut être déduite. Donc ce n'est pas toujours possible d'appliquer la dualité comme technique générale dans le but de chercher une solution.

2.5.1 Dualité en programmation convexe

Définition 2.5.1 (*Problème primal*) On définit le problème primal comme un problème de minimisation, qui consiste à trouver un vecteur x^* , s'il existe, tel que

$$(PCP) \begin{cases} F(x^*) = \min F(x), \\ x^* \in S = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i = 1, 2, \dots, m\} \end{cases} \quad (2.26)$$

où la fonction F est convexe sur l'ensemble S défini par les fonctions $g_i(x)$, $i = 1 \dots m$ qui sont aussi convexes sur \mathbb{R}^n .

Définition 2.5.2 On définit le problème dual du (PCP) comme un problème de maximisation, qui consiste à trouver deux vecteurs $k^0 \in \mathbb{R}^n$ et $y^0 \in \mathbb{R}^m$, s'ils existent tel que :

$$(PCP) \begin{cases} L(k^0, y^0) = \max(L(k, y)), \\ (k^0, y^0) \in V = \{(k, y) \in \mathbb{R}^n \times \mathbb{R}^m / \nabla_k L(k, y) = 0, y \geq 0\} \end{cases} \quad (2.27)$$

ou $L(k, y) = F(k) + y'g(k)$ est la fonction de Lagrange associée au problème (PCP). Les relations existantes entre le programme primal et de son dual sont données par les deux théorèmes suivants :

Théorème 2.5.1 (*Théorème faible de la dualité, Wolfe 1961*)

Soit $F(x)$ et $L(k, y)$ les fonctions objectifs respectivement du problème primal et son dual.

On a alors

$$L(k, y) \leq F(x), \quad \forall (k, y) \in V, \forall x \in S.$$

Théorème 2.5.2 (*Théorème fort de la dualité, Wolfe 1961*)

Soit x^* une solution optimale du problème primal(PCP). Alors il existe $y^* \in \mathbb{R}^m, y^* \geq 0$, tel que le couple (x^*, y^*) et une solution optimale du dual (PCD) et on a :

$$F(x^*) = L(x^*, y^*).$$

2.6 Conclusion

Dans ce chapitre nous avons présenté les résultats fondamentaux pour l'optimisation non linéaire, les différentes méthode de minimisation numérique essentiellement les méthodes de gradient et la méthode de newton et enfin nous avons vu des notion sur la dualité.

Chapitre 3

Méthodes de résolution en programmation quadratique convexe

3.1 Introduction

Dans ce chapitre , nous introduisons les différentes approches développées pour la résolution des problèmes de programmation quadratique convexe avec contraintes d'inégalités linéaires. Dans la littérature, plusieurs méthodes ont été développées pour la réalisations des problèmes de programmation quadratique. Elles sont généralement des extensions de celles développées pour la programmation linéaire (PL) et peuvent être regroupées en trois classes.

La première classe est constituée des méthodes d'activation des contraintes qui sont une généralisation de la méthode du simplexe pour la (PL). Cette approche est dédiée pour la résolution des problèmes quadratiques convexes et non convexes de taille moyenne ; pour le cas non convexe, on se contente généralement de la recherche d'un minimum local. On peut aussi distinguer des méthodes primales qui cherchent à satisfaire la faisabilité duale en maintenant la faisabilité primale et la condition de complémentarité . par contre, les méthodes duales sont seulement conçues pour le cas convexe et cherche à satisfaire la faisabilité primale en maintenant la faisabilité duale et la condition de complémentarité vérifiée.

La deuxième approche est formée des méthodes de points intérieurs , réputées pour leurs bonne complexité polynomiale et qui sont efficaces pour la résolution des problèmes quadratiques de grande taille. Elles tentent de satisfaire la condition de complémentarité tout en maintenant la faisabilité du primal et du dual.

La dernière classe est constituée des méthodes adaptées de support qui sont intermédiaires entre les deux premières approches. Elles traitent les problèmes quadratiques tels qu'ils se présentent et utilisent un critère d'arrêt ϵ -optimal.

3.2 La méthode des points intérieurs

En 1984, Karmarkar a montré qu'une méthode de points intérieurs (MPI) peut résoudre des programmes linéaires en un temps polynomial. Les deux décennies suivantes ont vu d'énormes efforts fournis par les spécialistes de l'optimisation pour étudier les propriétés théoriques et pratiques des différentes MPIs. Une des premières découvertes est que ces méthodes peuvent être vue comme une modification de la méthode de Newton, qui est capable de traiter des contraintes d'inégalités . Nesterov et Nemirovski ont montré que cette approche est applicable à une large classe de problèmes que celles des problèmes linéaires. Par exemple les problèmes de programmation quadratique peuvent être résolus polynomialement comme d'autres programmes convexe en utilisant des MPIs.

Les MPIs ont des caractéristiques communes qui se distinguent de celle de la méthode du simplexe. Chaque itération est couteuse à calculer mais peut faire un progrès significatif vers la solution. Par contre, le simplexe demande généralement un plus grand nombre d'itération non couteuse. La méthode du simplexe chemine autour de la bordure du polytope réalisable, testant une séquence de sommets (points extrêmes) jusqu'à ce qu'il trouve un point optimal. Les MPIs approximent la solution Soit de l'intérieur ou de l'extérieur du domaine réalisable et ne s'approchent de la frontière qu'en limite.

3.2.1 Position du problème

On considère, pour simplifier, le problème avec contraintes d'inégalités seulement, celles pour lesquelles les algorithmes de points intérieurs ont été conçus. On l'écrit sous la forme :

$$\begin{cases} \min f(x) \\ a_i x \leq b_i \quad (i \in I) \\ a_i x = b_i \quad (i \in J) \end{cases}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^1 , I, J , finis.

3.2.2 Technique de la méthode

En introduisant le vecteur des variables d'écart $s = b - Ax$, on peut réécrire les conditions d'optimalité qui sont à la fois nécessaires et suffisantes pour l'optimalité de couple $(x, \lambda) = (x^*, \lambda^*)$ comme suit :

$$\begin{cases} Dx + A'\lambda + c = 0, \\ Ax + s - b = 0, \\ s_i \lambda_i = 0, \quad \forall i \in \mathcal{I} \\ (\lambda, s) \geq 0. \end{cases} \quad (3.1)$$

Les méthodes de points intérieurs de type primal-dual calculent les solutions (x, λ, s) du système en appliquant une variante de la méthode de Newton aux trois premières équation du système (3.1) en modifiant les directions de recherche et le pas de telle sorte que les inégalité de la quatrième condition du système (3.1) restent satisfaites à chaque itération.

Définissons maintenant l'application non linéaire suivante $F : \mathbb{R}^{n+2m} \rightarrow \mathbb{R}^{n+2m}$:

$$F(x, \lambda, s) = \begin{bmatrix} Dx + A'\lambda + c \\ Ax + s - b \\ S\Lambda e \end{bmatrix}, \quad (s, \lambda) \geq 0 \quad (3.2)$$

où $S = \text{diag}\{s_1, s_2, \dots, s_m\}$ et $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_m\}$ sont des matrices carrées diagonales d'ordre m , ainsi que $e = (1, 1, 1, \dots, 1)'$ est un m -vecteur dont chaque composante égale à 1. Notons que la fonction F est linéaire en ses deux premiers termes et non linéaire en son dernier terme.

Les méthodes de points intérieurs du type primal-dual des itérés (x^k, λ^k, s^k) qui vérifient les bornes strictes, c'est-à-dire que $x^k > 0$ et $s^k > 0$. cette propriété est à l'origine du terme point intérieur. En appliquant la méthode de Newton à $F(x, \lambda, s) = 0$ la solution obtenue peut ne pas vérifier les condition $(x, \lambda) \geq 0$.

La méthode de Newton forme un modèle linéaire de F autour du point courant et obtient la direction de recherche $(\Delta x, \Delta \lambda, \Delta s)$ en résolvant le système d'équations suivant :

$$\begin{bmatrix} D & A' & 0 \\ A & 0 & I \\ 0 & \Lambda S & \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -\Lambda S e \end{bmatrix} \quad (3.3)$$

Notons que le déplacement maximal le long de cette direction est généralement non permis puisque il pourrait violer les inégalités $(\lambda, s) \geq 0$. Afin d'éviter cette difficulté, on calcule la direction d'amélioration de telle sorte que le nouveau point est :

$$(x^+, \lambda^+, s^+) = (x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s)$$

pour un certain paramètre $\alpha \in [0, 1]$. Malheureusement, on peut souvent prendre seulement un petit pas le long de la direction ($\alpha \ll 1$) avant la violation de la condition $(\lambda, x) > 0$.

3.2.3 La technique des chemins centraux

Ci-dessous, on présente la technique des chemins centraux qui modifie la procédure de la base de Newton.

Un chemin central \mathbb{C} est un arc de points réalisables stricts qui joue un rôle vital dans les algorithmes de type primal-dual. Il est paramétré par un scalaire $\tau > 0$, et chaque point $(x_\tau, \lambda_\tau, s_\tau) \in \mathbb{C}$ résout le système :

$$F(x_\tau, \lambda_\tau, s_\tau) = \begin{bmatrix} 0 \\ 0 \\ \tau e \end{bmatrix}, (\lambda_\tau, s_\tau) > 0 \quad (3.4)$$

. Cette équation n'est autre que l'équation de Newton pour laquelle la condition de complémentarité est remplacée par $\lambda_i s_i = \tau$ ($i = 1, \dots, m$). elle approxime (3.1) de plus en plus lorsque τ tend vers zéro. Si \mathbb{C} converge lorsque τ tend vers zéro, il doit converger vers la solution primale-duale du système (3.1). Donc, le chemin central nous guide vers la solution optimale le long d'une route en gardant les produits $\lambda_i s_i$ strictement positifs et il tend vers zéro avec un même tau.

choisissons un paramètre de centrage $\sigma \in [0, 1]$ et introduisant la mesure de la dualité de μ définie par :

$$\mu = \frac{1}{n} \sum_{i=1}^n \lambda_i s_i = \frac{\lambda' s}{n} \quad (3.5)$$

qui mesure la moyenne des produits $\lambda_i s_i$. En posant $\tau = \sigma \mu$ et en appliquant la méthode de Newton au système (3.4); on obtient :

$$\begin{bmatrix} D & A' & 0 \\ A & 0 & I \\ 0 & \Lambda & S \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ (-\Lambda S + \sigma \mu) e \end{bmatrix} \quad (3.6)$$

Le déplacement $(\Delta x, \Delta \lambda, \Delta s)$ est une direction de newton vers le point $(x_\tau, \lambda_\tau, s_\tau) \in \mathbb{C}$ pour lesquels les produits $\lambda_i s_i$ sont égale à $\tau = \sigma \mu$. par contre, le déplacement (3.3) pointe directement du point pour lesquelles les conditions de KKT (3.4) sont satisfaites. Si $\sigma = 1$, les équations (3.6) définissent une direction centrée et si $\sigma = 0$, (3.6) définissent une direction standard de Newton.

pour la majorité des problèmes, un point de départ strictement réalisable (x^0, λ^0, s^0) est difficile à calculer. Les méthodes de points intérieurs irréalisables remédient à ça en imposant uniquement que les composantes λ^0 et s^0 soient strictement positives. Par conséquent on change légèrement le système (3.6) en définissant les résidus des deux équations de la manière suivante :

$$r_d = Dx + A'\lambda + c; \quad r_p = Ax - b + s$$

Alors comme (3.6) pour le calcul de la direction d'amélioration, on résout le système modifié suivant :

$$\begin{bmatrix} D & A' & 0 \\ A & 0 & I \\ 0 & \Lambda & S \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \lambda \\ \Delta s \end{bmatrix} = \begin{bmatrix} r_d \\ r_p \\ (-\Lambda S + \sigma \mu)e \end{bmatrix} \quad (3.7)$$

La prochaine itération est obtenue en posant

$$x^+, \lambda^+, s^+ = (x, \lambda, s) + \alpha(\Delta x, \Delta \lambda, \Delta s),$$

ou α est choisi de manière à retenir les composantes de λ_i et s_i strictement positives :

$$\begin{aligned} \lambda^+ &= \lambda + \alpha \Delta \lambda > 0, \quad i = 1, 2, \dots, n, \\ s^+ &= s + \alpha \Delta s > 0, \quad i = 1, 2, \dots, n, \end{aligned}$$

Il n'est pas difficile de voir que la valeur maximale α_{max} qui maintient ces conditions est donnée par la formule suivante :

$$\alpha_{max} = \min\left\{\min_{i:\Delta\lambda_i < 0} -\frac{\lambda_i}{\Delta\lambda_i}, \min_{i:\Delta s_i < 0} -\frac{s_i}{\Delta s_i}\right\} \quad (3.8)$$

Nous pouvons nous éloigner de cette valeur maximale et empêcher ainsi que chaque composante λ_i et s_i d'être aussi proche de zéro en définissant le pas α , comme suit :

$$\alpha = \min\{1, \eta \alpha_{max}\} \quad (3.9)$$

ou η est un coefficient proche de 1 et généralement on le prend égal à 0,999.

dans cette partie nous avons présenté la méthode des point intérieurs du type primale-duale. Les algorithmes de points intérieurs ont la propriété remarquable de converger en quelques dizaine d'itérations sur la plupart des exemples réels.

3.3 La méthode d'activation des contraintes

Introduction

La méthode activation des contraintes (ASM : Active set method) [21] est une méthode classique développée au début des années soixante-dix pour la résolution des problèmes linéaires et quadratiques ; elle s'applique pour des problèmes d'optimisation avec contraintes linéaires de types inégalités ou mixtes.

cette méthode se présente sous forme de trois classe primale,duale, et primale-duale, dans cette partie on ne présente que la méthode primale, cette dernière est itérative et s'exécute en deux phase[5] :

- La première phase consiste à trouver une solution réalisable de départ, qui est obtenue en exécutant la première phase d'un solveur de programmation linéaire.
- La deuxième phase consiste à appliquer la méthode ASM proprement dite. son principe général est de résoudre à chaque itération un programme quadratique avec contraintes d'égalités, correspondants aux indices actifs. Par la suite, elle ajuste cet ensemble pour identifier les contraintes active à l'optimum.

3.3.1 Définition

Soit le problème suivant :

$$\begin{cases} F(x) = \frac{1}{2}x'Dx + c'x \longrightarrow \min, \\ A'_i x = b_i \quad i \in \varepsilon \\ A'_i x \leq b_i, \quad i \in \mathcal{I} \end{cases} \quad (3.10)$$

ou D une matrice carrée symétrique d'ordre n et semi-définie positive ; $I = \varepsilon \cup \mathcal{I} = \{1, 2, 3, \dots, m\}$ est un ensemble fini d'indices des contraintes du problème $A_i, i \in I, c, x$ sont des n -vecteurs et $b = (b_i, i \in I)'$ est un m -vecteur.

- Un point x vérifiant les contraintes du problème (3.10) est appelé solution réalisable.
- une contrainte $i \in \mathcal{I}$ est dite active ou (saturée) au point x si

$$A'_i x = b_i$$

L'ensemble des contraintes actives au point x , noté I_0 est l'ensemble d'indices défini comme suit :

$$I_0(x) = \{i \in I : A'_i x = b_i\}, \text{ avec } I = \varepsilon \cup \mathcal{I}$$

3.3.2 Critère d'optimalité

En associant aux contraintes du problème (3.10) le m -vecteur multiplicateur λ , la fonction de Lagrange s'écrit :

$$L(x, \lambda) = \frac{1}{2}x'Dx + c'x + \sum_{i \in \varepsilon \cup \mathcal{I}} \lambda(A'_i x - b_i). \quad (3.11)$$

En plus, on définit l'ensemble actif $I_0(x^0)$ à la solution optimale x^0 comme étant les indices des contraintes pour lesquels l'égalité est vérifiée :

$$I_0(x) = \{i \in I : A'_i x = b_i\}$$

La condition nécessaire d'optimalité de KKT du premier ordre est la suivante :

$$\begin{aligned} Dx^0 + c + \sum_{i \in I} \lambda_i^0 (A'_i x^0 - b_i) &= 0 \\ A'_i x^0 &= b_i, \text{ pour tout } i \in I_0(x^0), \\ A'_i x^0 &\leq b_i, \text{ pour tout } i \in I \setminus I_0(x^0), \\ \lambda_i^0 &\in \mathbb{R}, i \in \varepsilon, \lambda_i^0 \geq 0, \text{ pour tout } i \in \mathcal{I}; \\ \lambda_i^0 (A'_i x^0 - b_i) &= 0, \text{ pour } i \in \mathcal{I}. \end{aligned}$$

3.3.3 Itération de la méthode

À une itération donnée de l'algorithme, soit x la solution courante de I_0 l'ensemble d'indices actifs correspondant. une itération de l'algorithme consiste à exécuter les opérations suivantes :

- (1) on résout le sous-problème quadratique correspondant à l'ensemble de travail suivant :

$$\begin{cases} F(x) = \frac{1}{2}x'Dx + c'x \longrightarrow \min, \\ A'_i x \leq b_i, \quad i \in I_0 \end{cases} \quad (3.12)$$

La solution optimale pour (3.12) est de la forme $x + p$, ou p est une certaine correction (direction de descente) de la solution courante x . Alors le n -vecteur p est calculé de telle sorte que la fonction objectif diminue. Par conséquent, le programme quadratique à résoudre est le suivant :

$$\begin{cases} F(x) = \frac{1}{2}p'Dp + g'x \longrightarrow \min, \\ A'_i p = 0, \quad i \in I_0 \end{cases} \quad (3.13)$$

ou $g = Dx + c$ est le gradient de F au point x .

On envisage deux cas possibles :

- ★ Si $p = 0$, aller à l'étape **(2)**
- ★ sinon ($p \neq 0$), posons :

$$\bar{x} = x + \theta p \quad \text{avec} \quad 0 < \theta \leq 1 \quad (3.14)$$

Le pas θ doit être choisi de telle sorte que la nouvelle solution soit réalisable. pour $i \in I_0$ n'importe quelle valeur de θ va maintenir la faisabilité de \bar{x} . Par contre toutes contraintes non saturées doivent vérifier :

$$A'_i(x + \theta p) \leq b_i, \quad \forall i \in I/I_0.$$

Alors, la valeur optimal θ^0 est la suivante :

$$\theta^0 = \min\{1, \theta_{i_0}\} \quad (3.15)$$

ou

$$\theta_{i_0} = \min\{\theta_i, i \in I/I_0\}, \quad \text{avec} \quad \theta_i = \begin{cases} \frac{b_i - A'_i x}{A'_i p}, & \text{si } A'_i > 0, \\ \infty, & \text{si } A'_i \leq 0, i \in I/I_0 \end{cases}$$

Si $\theta^0 = 1$, alors dans ce cas l'ensemble de travail ne change pas, et on aura :

$$\bar{x} = x + p, \quad \bar{I}_0 = I_0$$

Sinon, la contrainte i_0 est ajoutée à l'ensemble de travail. Par conséquent, le nouveau ensemble de travail est donc $\bar{I}_0 = I_0 \cup \{i_0\}$.

(2) cette étape consiste à vérifier si la nouvelle solution obtenue \bar{x} est optimale ou non dans le problème(3.10). Ceci se fait en calculant les multiplicateurs de Lagrange associés au problème (3.12). Rappelons que pour $i \notin I_0$, $\lambda_i = 0$. Les autres composantes du vecteur λ ont été calculés lors de la résolution du problème (3.12).

(a) Si $\lambda_i \geq 0, \forall i \in I \cap I_0$, alors la condition d'optimalité de KKT est remplie pour le problème (3.10). Donc, la solution optimale est trouvées, et on arrête l'algorithme.

(b) Sinon, il existe une composante de vecteur λ vérifiant $\lambda_{i_1} < 0$, avec $i_1 \in I_0 \cap I/I_0$. Alors on peut améliorer la valeur de la fonction objectif en supprimant la contrainte i_1 de l'ensemble de travail. Si plusieurs contraintes ne vérifient pas le critère d'optimalité, Alors la contrainte à supprimer doit vérifier la relation suivante :

$$\lambda_{i_1} = \min\{\lambda_i : \lambda_i < 0, i \in I \cap I_0\}. \quad (3.16)$$

Donc le nouveau ensemble de travail est $\bar{I}_0 = I_0 \cup \{i_1\}$. ce processus de résolution est répété jusqu'à ce que l'optimum soit trouvé.

3.3.4 Schéma de l'algorithme

Soit x une solution réalisable du problème (3.10) et I_0 l'ensemble du travail correspondant. Le schéma de l'algorithme présente les étapes suivantes :

(1) Calculer p en résolvant le programme quadratique (3.13).

- Si $p = 0$, alors aller à (3).

- Sinon, on pose :

$$\bar{x} = x + \theta^0 p;$$

ou θ^0 est définie par la relation (3.15). Aller à (2).

(2) Changement de l'ensemble actif.

- si $\theta^0 = \theta_{i_1} < 1$ alors :

$$\bar{I}_0 = I_0 \cup \{i_0\}$$

- Sinon ($\theta^0 = 1$), on pose :

$$\bar{I}_0 = I_0$$

Aller à (3)

(3) calculer les multiplicateur de Lagrange de (3.12).

- Si $\lambda_i \geq 0, \forall i \in \mathbb{I} \cap I_0$, alors l'algorithme s'arrête, avec \bar{x} solution optimale.

- Sinon , déterminer :

$$\lambda_{i_1} = \min\{\lambda_i : \lambda_i < 0, i \in \mathbb{I} \cap I_0\}.$$

posons $\bar{I}_0 = I_0 \cup \{i_1\}$. Aller à (1).

Exemple numérique

Considérons le programme quadratique suivant :

$$\begin{cases} F(x) = 2x^2 + x_1x_2 + x_2^2 - 12x_1 - 10x_2 \longrightarrow \min, \\ x_1 + x_2 \leq 4, \\ -x_1 \leq 0, \\ -x_2 \leq 0. \end{cases} \quad (3.17)$$

Les données de ce problème sont les suivantes :

$$D = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}, c = \begin{pmatrix} -12 \\ -10 \end{pmatrix}, A = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}$$

La méthode qu'on utilise dans cet exemple pour la résolution du problème (3.13) est la méthode d'élimination simple des variables.

Itération 1 : Soit $x^1 = (0, 0)'$ la solution réalisable de départ pour ce problème. L'ensemble d'indices actifs en ce point est $I_0^1 = \{2, 3\}$. Alors pour le problème (3.12), on a :

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}; b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; Z = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Il est clair que $p^1 = (0, 0)'$ est optimal pour (3.13), puisque en fait aucun autre point n'est réalisable pour cet ensemble actif. Le vecteur des multiplicateur de Lagrange associé est :

$$\lambda^1 = -A^{-1}g(x^1) = -\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -12 \\ -10 \end{pmatrix} = \begin{pmatrix} -12 \\ 10 \end{pmatrix}.$$

Aucun indice actif n'est optimal et par conséquent la deuxième contrainte sera supprimée de l'ensemble de travail. donc, on aura :

$$x^2 = x^1, I_0^2 = I_0^1 \setminus 2 = \{3\}$$

Itération 2 : Résolution du problème (3.12) avec les paramètres suivants :

$$A = (0, 1), \quad b = 0, \quad Z = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad Z'DZ = 4$$

Le déplacement optimal est le suivant :

$$p^2 = -Z(Z'DZ)^{-1}Z'g = -\frac{1}{4} \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1 \ 0) \begin{pmatrix} -12 \\ -10 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Le nouveau point obtenue est

$$x^3 = x^2 + p^2 = \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

qui est réalisable. Le multiplicateur de Lagrange associé à la contrainte saturée est alors :

$$\lambda_3 = -[A_B^{-1}]' [g_B + D(J_B, J)p] = -(-1) \left[-10 + (1 \ 2) \begin{pmatrix} 3 \\ 0 \end{pmatrix} \right] = -7 < 0.$$

Par conséquent, la troisième contrainte est supprimée de l'ensemble des indices actifs; donc, on aura :

$$I_0^3 = I_0^2 \setminus 3 = \emptyset$$

Itération 3 : A cette itération aucune contrainte n'est active au point x^3 . Par conséquent, la solution du problème(3.17) sans contrainte nous donne :

$$x^* = -D^{-1}c = -\begin{pmatrix} \frac{2}{7} & \frac{-1}{7} \\ \frac{-1}{7} & \frac{4}{7} \end{pmatrix} \begin{pmatrix} -12 \\ -10 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix},$$

qui n'est pas réalisable. Le déplacement du point x^3 à x^* est le vecteur :

$$p^4 = x^* - x^3 = \begin{pmatrix} -1 \\ 4 \end{pmatrix}.$$

Calcul du θ^0 :

$$A'_1 p^4 = (1 \ 1) \begin{pmatrix} -1 \\ 4 \end{pmatrix} = 4 > 0 \Rightarrow \theta_1 = \frac{b_1 - A'_1 x^3}{A'_1 p^4} = \frac{4 - (1 \ 1) \begin{pmatrix} 3 \\ 0 \end{pmatrix}}{3} = \frac{1}{3};$$

$$A'_2 p^4 = (-1 \ 0) \begin{pmatrix} -1 \\ 4 \end{pmatrix} = 1 > 0 \Rightarrow \theta_2 = \frac{b_2 - A'_2 x^3}{A'_2 p^4} = \frac{0 - (-1 \ 0) \begin{pmatrix} 3 \\ 0 \end{pmatrix}}{1} = 3;$$

$$A'_3 p^4 = (0 \quad -1) \begin{pmatrix} -1 \\ 4 \end{pmatrix} = -4 \leq 0 \Rightarrow \theta_3 = \infty.$$

$$\theta_{i_0} = \min\{\theta_1, \theta_2, \theta_3\} = \theta_1 = \frac{1}{3} \Rightarrow i_0 = 1 \quad \text{alors} \quad \theta^0 = \min\{1, \theta_{i_0}\} = \theta_{i_0} = \frac{1}{3}.$$

La nouvelle solution obtenue, ainsi que l'ensemble de travail correspondant sont les suivants :

$$x^4 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} -1 \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{8}{3} \\ \frac{4}{3} \end{pmatrix}; \quad I_0^4 = I_0^3 \cup 1 = \{1\}.$$

Itération 4 : Données correspondant à x^4 et I_0^4 :

$$A = (1 \quad 1), b = 4, Z = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad Z'DZ = 4, \quad g = g(x^4) = \begin{pmatrix} 0 \\ -\frac{14}{3} \end{pmatrix}$$

Le déplacement optimal est le suivant :

$$p^4 = -Z(Z'DZ)^{-1}Z'g = -\frac{1}{4} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1 \quad -1) \begin{pmatrix} 0 \\ -\frac{14}{3} \end{pmatrix} = \begin{pmatrix} -\frac{7}{6} \\ \frac{7}{6} \end{pmatrix}$$

Calcul θ^0 :

$$A'_2 p^4 = (-1 \quad 0) \begin{pmatrix} -\frac{7}{6} \\ \frac{7}{6} \end{pmatrix} = \frac{7}{6} > 0 \Rightarrow \theta_2 = \frac{b_2 - A'_2 x^4}{A'_2 p^4} = \frac{0 - (-1 \quad 0) \begin{pmatrix} \frac{8}{3} \\ \frac{4}{3} \end{pmatrix}}{\frac{7}{6}} = \frac{16}{7}$$

$$A'_3 p^4 = (0 \quad -1) \begin{pmatrix} -\frac{7}{6} \\ \frac{7}{6} \end{pmatrix} = -\frac{7}{6} \leq 0 \Rightarrow \theta_3 = \infty$$

$\theta^0 = \min\{1, \theta_2, \theta_3\} = 1$ alors on aura :

$$x^5 = x^4 + p^4 = \begin{pmatrix} \frac{8}{3} \\ \frac{4}{3} \end{pmatrix} + \begin{pmatrix} -\frac{7}{6} \\ \frac{7}{6} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{15}{6} \end{pmatrix}$$

La valeur du multiplicateur de KKT associé à la première contrainte est :

$$\lambda_3 = -[A_B^{-1}]'(g_B + D(J_B, J)p) = -\left[-\frac{14}{3} + (2, 1) \begin{pmatrix} -\frac{7}{6} \\ \frac{7}{6} \end{pmatrix}\right] = \frac{7}{2} \geq 0$$

Donc x^5 est la solution optimale pour le problème (3.17).

La méthode d'activation des contraintes a l'avantage de traiter les problèmes d'optimisation linéaire et quadratique tels qu'ils se présentent, sans chercher à modifier leurs contraintes; cela permet d'avoir un gain en espace mémoire et en temps d'exécution sur la machine. L'inconvénient de cette méthode est que l'ensemble actif change lentement contrairement à la méthode du gradient projeté qui permet un changement rapide de l'ensemble actif.

3.4 La méthode du gradient projeté

Introduction

La méthode du gradient projeté s'inspire des méthodes usuelles de gradient qu'on a vu dans la deuxième section dans le deuxième chapitre. Supposons, d'une façon générale, que l'on souhaite minimiser une fonctionnelle $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sur un ensemble de contraintes C . Si l'on construit une suite d'itérés de la forme $x^{k+1} = x^k + \theta_k d^k$, où d^k est une direction de descente, La question qui se pose est ce que si x^k appartient à C , alors x^{k+1} appartiendra encore à C . Il faut donc ramener x^{k+1} dans C , ce que l'on fait en utilisant une projection.

3.4.1 Position de problème

On considère le problème :

$$\begin{cases} \min f(x) \\ a_i x \leq b_i \quad (i \in I) \\ a_i x = b_i \quad (i \in J) \end{cases}$$

où $f : \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^1 ; I c'est l'ensemble des indices des contraintes d'inégalités et J l'ensemble des contraintes d'égalité sont deux ensemble finis.

Pour tout x solution réalisable $I(x)$ est l'ensemble des indices des contraintes d'inégalité saturées par x . On fait l'hypothèse que x vérifie les critère de qualification de l'indépendance linéaire, c'est à dire que les vecteurs a_i , $i \in I(x) \cup J$ sont linéairement indépendants.

3.4.2 Principe de la méthode

la méthode du gradient projeté traite les problème avec contraintes d'égalité, Cependant, pour les problèmes avec contraintes d'inégalité linéaires, on sera amené à considérer une succession de problèmes avec contraintes d'égalité, et la matrice A pourra évoluer à chaque itération, par ajout ou suppression d'une ligne. On part d'un x solution réalisable, on détermine la direction d définie par la projection de $-\nabla f(x)$ sur l'espace

$$L = \{y : a_i y = 0, \quad i \in I(x) \cup J\}$$

et

$$d = P_L(-\nabla f(x))$$

où P_L est l'opérateur de projection orthogonale sur L .

$$-\nabla f(x) = d + \sum_{i \in I(x) \cup J} \lambda_i a_i$$

Par construction, le déplacement de la direction d ne viole pas les contraintes saturées.

Si $d \neq 0$ le déplacement se fait jusqu'à atteindre le minimum de f dans la direction d mais en respectant les contraintes non saturées, on obtient un nouveau point et on réitère.

Si $d = 0$, alors

$$-\nabla f(x) = \sum_{i \in I(x) \cup J} \lambda_i a_i$$

. Si $\forall i \in I(x)$, $\lambda_i \geq 0$, alors les conditions de KKT sont vérifiées, et le critère d'optimalité est satisfait; on arrête.

Si $\exists j \in I(x)$, $\lambda_j < 0$ alors cela signifie que la contrainte j est inactive sur x . On la retire de $I(x)$ et on réitère.

3.4.3 Algorithme de la méthode

Projection sur un convexe fermé

Soit K un convexe fermé et soit $\mu \in \mathbb{R}^n$. On appelle projeté de μ sur K , l'unique élément $\mu \in K$ solution de :

$$\|u - \mu\| = \min_{v \in K} \|v - \mu\|.$$

Le projeté de μ est noté $u = P_K(\mu)$, et est caractérisé par :

$$(u - \mu, u - v) \leq 0 \quad \forall v \in K$$

Soit K un convexe fermé de \mathbb{R}^n . Soit J une fonctionnelle convexe sur K . Considérons le problème :

$$(P) \quad \min_{v \in K} J(v)$$

Théorème Pour tout $\rho > 0$, on a :

$$\|u \text{ est minimum} \Leftrightarrow u = P_K(u - \rho \nabla J(u)).$$

Preuve Soit : $\rho > 0$

$$u \text{ est min} \Leftrightarrow (\nabla J(u), u - v) \leq 0 \quad \forall v \in K$$

$$\Leftrightarrow (\rho \nabla J(u), u - v) \leq 0 \quad \forall v \in K,$$

$$\Leftrightarrow (u - \underbrace{(u - \rho \nabla J(u))}_{\text{point fixe}}, u - v) \leq 0 \quad \forall v \in K,$$

$$\Leftrightarrow u = P_K(u - \rho \nabla J(u))$$

Algorithme du gradient projeté

1 On choisit : une condition initiale $u_0 \in K$, un pas $\rho > 0$ et une précision $\eta > 0$. On calcule $u_1 = P_K(u_0 - \rho J'(u_0))$.

2 Tant que $\|u_k - u_{k-1}\| > \eta$, on définit u_{k+1} par

$$u_{k+1} = P_K(u_k - \rho J'(u_k))$$

Remarques

- On ne peut plus prendre comme test d'arrêt $\|\nabla J(u_k)\| \sim 0$ car pour le problème avec contraintes le minimum u ne satisfait pas forcément " $\nabla J(u) = 0$ "

- On a $\|u_k - u_{k+1}\| < \eta \Leftrightarrow u_{k-1} \simeq P_K(u_{k-1} - \rho \nabla J(u_{k-1}))$, donc le test d'arrêt signifie bien que u_{k-1} est proche du point fixe $u = P_K(u - \rho \nabla J(u))$

Théorème 3.4.1 Soit $K \neq \emptyset$ un convexe fermé de \mathbb{R}^n , et soit

$$J : v \rightarrow \frac{1}{2}(Av, v) - (b, v),$$

où $A \in \mathbb{R}^{n \times n}$ est symétrique, définie positive, et $b \in \mathbb{R}^n$.

Si $0 < \rho < \frac{2}{\lambda_{\max}(A)}$, alors quel que soit $u_0 \in K$, la suite (u_k) définie par (6) converge vers la solution u du problème (P).

Preuve Notons d'abord que $J'(v) = Av - b \forall v \in K$.

$$\|u_{k+1} - u\| = \|P_K(u_k - \rho(Au_k - b)) - P_K(u - \rho(Au - b))\|$$

$$\leq \|(I - \rho A)(u_k - u)\|$$

$$\leq \gamma_\rho \|u_k - u\|$$

$$\text{où } \gamma_\rho |\lambda_{\max}(I - \rho A)| < 1 \text{ pour tout } 0 < \rho < \frac{2}{\lambda_{\max}(A)}$$

Donc

$$\|u_n - u\| \leq (\gamma_\rho)^n \|u_0 - u\| \rightarrow 0.$$

- Le théorème précédent peut être généralisé à une classe plus large de problèmes d'optimisation convexe avec contraintes.
- D'un point de vue pratique, la projection d'un élément $v \in \mathbb{R}^n$ sur un convexe quelconque peut être très difficile à déterminer.
- **Un cas facile** Si :

$$K = \prod_{i=1}^n [a_i, b_i], \text{ alors}$$

si on prend $y = (y_1, \dots, y_n)^t$, la projection $x = P_K(y)$ a pour composantes :

$$x_i = \min(\max(\max(a_i, y_i), b_i)) \text{ pour } 1 \leq i \leq n$$

3.5 Méthode adaptée de support

Introduction

Le principe est de démarrer d'un plan du support initial et chaque itération consiste à trouver une direction d'amélioration et un pas maximal le long de cette direction de façon à améliorer la valeur de la fonction objectif, tout en assurant de ne pas sortir du domaine admissible déterminé par les contraintes du problème, en utilisant la métrique dite adaptée qui permet le changement de tous les indices non optimaux en fonction desquels on construit une direction d'amélioration et le pas le long de cette direction.

3.5.1 position du problème et définition

Le problème de programmation quadratique convexe se présente comme suit :

$$F(x) = \frac{1}{2}x'Dx + c'x \longrightarrow \min, \quad (3.18)$$

$$Ax = b \quad (3.19)$$

$$d^- \leq x \leq d^+ \quad (3.20)$$

où $D' = D \geq 0$; c, d^-, d^+ et x sont des n -vecteurs; b un m -vecteur; $\text{rang}A = m < n$.

Soit $I = \{1, 2, \dots, m\}$: l'ensemble d'indices des lignes de A ;

$J = \{1, 2, \dots, n\}$: l'ensemble d'indices des colonnes de A ;

$J = \{1, 2, \dots, n\} = J_B \cup J_N$, avec $J_B \cap J_N = \emptyset, |J_B| = m$

On peut alors fractionner les vecteurs et la matrice de la manière suivante :

$d^- = d^-(J) = (d_j^-, j \in J)$; $d^+ = d^+(J)$; $x = x(J) = (x_j, j \in J_N)$;

$x = \begin{pmatrix} x_B \\ - \\ x_N \end{pmatrix}$, $x_B = x(J_B) = (x_j; j \in J_B)$, $x_N = x(J_N) = (x_j, j \in J_N)$;

$x = \begin{pmatrix} c_B \\ - \\ c_N \end{pmatrix}$, $c_B = c(J_B)$, $c_N = c(J_N)$,

$A = A(I, J) = (a_{ij}, 1 \leq i \leq m, 1 \leq j \leq n)$; $A = (a_j, j \in J), a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ $A = (A_B | A_N)$, $A_B = A(I, J_B)$, $A_N = (I, J_N)$.

Définition 3.5.1 Un vecteur x vérifiant les contraintes (3.19),(3.20) est appelé *plan* ou *solution réalisable* du problème (3.19),(3.20).

Définition 3.5.2 Un plan x^0 est dit *optimal* si

$$F(x^0) = \frac{1}{2}x^{0'}Dx^0 + c'x^0 = \min \frac{1}{2}x'Dx + c'x$$

Où x est pris parmi tous les vecteurs vérifiant les contraintes (3.19),(3.20).

Définition 3.5.3 Un plan x^ϵ est appelé ϵ -*optimal* ou *suboptimal* si :

$$F(x^\epsilon) - F(x^0) = \frac{1}{2}(x^\epsilon)'Dx^\epsilon + c'x^\epsilon - \frac{1}{2}x^{0'}Dx^0 + c'x^0 \leq \epsilon$$

Où ϵ est un nombre arbitraire positif ou nul et x^0 est une solution optimale du problème(3.18),(3.20).

Définition 3.5.4 L'ensemble $J_B \subset J$, $|J_B| = m$, est appelé *support* des contraintes (3.19),(3.20) si

$$\det A_B \neq 0$$

Définition 3.5.5 le couple $\{x, J_B\}$, formé du plan x et du support J_B est appelé plan de support des contraintes.

Définition 3.5.6 Le plan de support est dit non dégénéré si

$$d_j^- < x_j < d_j^+, \quad j \in J_B.$$

3.5.2 Formule d'accroissement de la fonction objectif

Soit un plan de support des contraintes du problème (3.18),(3.20)et considérons un autre plan quelconque $\bar{x} = x + \Delta x$. L'accroissement de la fonction objectif s'écrit alors :

$$F(\bar{x}) - F(x) = g'(x)\Delta x + \frac{1}{2}(\Delta x)'D(\Delta x) \quad (3.21)$$

où $g(x) = Dx + c$ est le gradient de la fonction (3.18), avec $g(x) = g(J) = (g_j, j \in J)$; $g = \begin{pmatrix} g_B \\ -- \\ g_N \end{pmatrix}$, $g_B = g(J_B)$; $g_N = g(J_N)$

Par ailleurs on a $\begin{cases} Ax = b \\ A\bar{x} = b \end{cases} \Leftrightarrow A\bar{x} = A(\Delta x + x) = A\Delta x + Ax = b \Rightarrow A\Delta x = 0$

En posant $\Delta x = \begin{pmatrix} \Delta x_N \\ -- \\ \Delta x_B \end{pmatrix}$ $\Delta x_B = \Delta x(J_B)$, $\Delta x_N = \Delta x(J_N)$, l'égalité $A\Delta x = 0$ peut s'écrire $A_B\Delta x_B + A_N\Delta x_N = 0$ d'où

$$\Delta x_B = -A^{-1}A_N\Delta x_N \quad (3.22)$$

On définit le vecteur des potentiels u et le vecteur des estimation E tels que :

$$u' = g'_B A_B^{-1}, E' = E'(J) = g' - u'A = (E'_B, E'_N)$$

où

$$E'_B = 0, \quad E'_N = g'_N - u'A_N$$

La formule (3.21), devient alors :

$$F(\bar{x}) - F(x) = g'_B\Delta x_B + g'_N\Delta x_N + \frac{1}{2}(\Delta x_B, \Delta x_N)'D(\Delta x_B, \Delta x_N).$$

En vertu de (3.22), on obtient donc :

$$\begin{aligned} F(\bar{x}) - F(x) &= g'_B(-A_b^{-1}A_N\Delta x_N) + g'_N\Delta x_N + \frac{1}{2}(-A_B^{-1}A_N\Delta x_N, \Delta x_N)'D - A_B^{-1}A_N\Delta x_N, \Delta x_N) \\ &= [-g'_B A_b^{-1}A_N + g'_N]\Delta x_N + \frac{1}{2}\Delta'x_N \begin{pmatrix} -A_B^{-1}A_N \\ I_N \end{pmatrix}' D \begin{pmatrix} -A_B^{-1}A_N \\ I_N \end{pmatrix} \Delta x_N, \end{aligned}$$

où $I_N = I(J_N, J_N)$ est la matrice identité d'ordre $(n - m)$.

posons :

$$Z = Z(I, J_N) = \begin{pmatrix} -A_B^{-1} A_N \\ I_N \end{pmatrix}; M = M(J_N, J_N) = Z' D Z \quad (3.23)$$

Finalement (3.21) aura la forme suivante :

$$F(\bar{x}) - F(x) = E'_N \Delta x_N + \frac{1}{2} \Delta' x_N M \Delta x_N \quad (3.24)$$

3.5.3 Critère d'optimalité

Théorème 3.5.1 Soit $\{x, J_B\}$ un plan de support des contraintes (3.1)-(3.3). Alors les relations :

$$\begin{cases} E_j \geq 0, & \text{si } x_j = d_j^-, \\ E_j \leq 0, & \text{si } x_j = d_j^+ \\ E_j = 0, & \text{si } d_j^- < x_j < d_j^+ \text{ pour } \forall j \in J_N, \end{cases} \quad (3.25)$$

sont suffisantes pour l'optimalité du plan x .

ces mêmes relations sont aussi nécessaires, si le plan de support des contraintes est non dégénéré.

3.5.4 Critère de suboptimalité

Les relations (3.25) n'étant pas vérifiées, calculons alors le nombre $\beta(x, J_B)$ appelé *estimation de suboptimalité* qui estime l'écart entre la valeur optimale $F(x^0)$ est une autre valeur $F(x)$ d'un plan de support des contraintes quelconque $\{x, J_B\}$.

Théorème 3.5.2 Soit $\{x, J_B\}$ un plan de support des contraintes du problème (3.1)-(3.3), et ϵ , un nombre positif ou nul arbitraire. Si le nombre :

$$\beta(x, J_B) = \sum_{\substack{j \in J_N \\ E_j > 0}} E_j (x_j - d_j^-) + \sum_{\substack{j \in J_N \\ E_j < 0}} E_j (x_j - d_j^+) \leq \epsilon, \quad (3.26)$$

alors $\{x, J_B\}$ est ϵ -optimal.

3.6 Méthode de résolution

Introduisons les définitions suivantes :

Définition 3.6.1 On appelle support de la fonction objectif (3.18) l'ensemble des indices $J_S \subset J_N$ tel que $\det M_S = \det M(J_S, J_S) \neq 0$, où M est la matrice (3.23). On posera $J_{NN} = J_N \setminus J_S$.

Définition 3.6.2 On appelle support du problème (3.18),(3.20) l'ensemble $J_P = \{J_B, J_S\}$ formé du support des contraintes J_B et celui de la fonction objectif J_S .

Définition 3.6.3 On appelle plan de support du problème (3.18),(3.20) la paire $\{x, J_P\}$ formée du plan x et du support J_P ; il est dit accordé si $E(J_S) = 0$.

3.6.1 Construction d'une direction d'amélioration adaptée

Les plans de support admettent plusieurs métriques pour les direction d'amélioration. Considérons la métrique suivante pour les composantes non basiques de la direction admissible l :

$$d_j^- - x_j \leq l_j \leq d_j^+ - x_j, \quad j \in J_{NN} = J_N \setminus J_S. \quad (3.27)$$

Cette métrique dépend du plan courant x , et de ce fait , elle est dite adaptée. Afin de calculer les composantes de la direction admissible d'amélioration l , considérons l'accroissement

$$\Delta F = F(x + l) - F(x) = \sum_{\substack{j \in J_N \\ E_j \geq 0}} E_j l_j + \sum_{\substack{j \in J_N \\ E_j \leq 0}} E_j l_j + \frac{1}{2} l_N^T M l_N.$$

En tenant compte de la métrique (3.27), la partie linéaire de ΔF atteint son minimum pour les valeurs des composantes de $l(J_{NN})$ suivantes :

$$l_j = \begin{cases} d_j^- - x_j & \text{si } E_j > 0, \\ d_j^+ - x_j & \text{si } E_j < 0, \\ 0 & \text{si } E_j = 0, \text{ pour } j \in J_{NN} \end{cases} \quad (3.28)$$

quant aux composantes de $l(J_S)$, elles se déduisent à partir de $E_j(x + l) = 0$, $j \in J_S$. Donc

$$M(J_S, J_S)l(J_S) + M(J_S, J_{NN})l(J_{NN}) = 0$$

D'où

$$l(J_S) = -M_S^{-1}M(J_S, J_{NN})l(J_{NN}) \quad (3.29)$$

les composantes $l(J_B)$ seront déduites de $Al = 0$, d'où

$$l(J_B) = -A_B^{-1}(A_S l_S + A_{NN} l_{NN}). \quad (3.30)$$

3.6.2 Changement de plan

On construit alors le nouveau plan \bar{x} sous la forme

$$\bar{x} = x + \theta^0 l;$$

où l est la direction d'amélioration définie par les formules (3.28),(3.30) et le nombre θ^0 est le pas le long de cette direction , avec $\theta^0 = \min \{1, \theta_{j1}, \theta_{js}, \theta_F\}$. Les nombres $1, \theta_{j1}, \theta_{js}$, se calculent de façon à ce que les contraintes directes sur le vecteur \bar{x} soient vérifiées :

$$d_j^- - x_j \leq \theta^0 l_j \leq d_j^+ - x_j, \quad \text{si } j \in J_{NN}, \quad (3.31)$$

$$d_j^- - x_j \leq \theta^0 l_j \leq d_j^+ - x_j, \quad \text{si } j \in J_B, \quad (3.32)$$

$$d_j^- - x_j \leq \theta^0 l_j \leq d_j^+ - x_j, \quad \text{si } j \in J_S, \quad (3.33)$$

Donc $\theta_{j1} = \min_{j \in J_B} \theta_j$; $\theta_{js} = \min_{j \in J_S} \theta_j$, où

$$\theta_j = \begin{cases} \frac{d_j^+ - x_j}{l_j}, & \text{si } l_j > 0 \\ \frac{d_j^- - x_j}{l_j}, & \text{si } l_j < 0 \\ \infty & \text{si } l_j = 0 \end{cases}$$

Le nombre $\theta^0 = 1$ représente le pas correspondant aux indices de J_{NN} .

Quant à θ_F , il se calcule de façon que le passage de x à \bar{x} assure une relaxation maximale de la fonction objectif tout en gardant le même signe pour E_j et \bar{E}_j ; où :

$$E'_N = g'(x)Z, \quad \bar{E}_N(\bar{x}) = \bar{E}_N(x + \theta^0 l)$$

$$\bar{E}_N(\bar{x}) = E_N(x) + \theta^0 M l_N = E_N(x) + \theta^0 \delta_N$$

On posera donc $\theta_F = \sigma_{j^*} = \min \sigma_j$, $j \in J_{NN}$, avec

$$\sigma_j = \begin{cases} -\frac{E_j}{\delta_j}, & \text{si } E_j \delta_j < 0, \\ \forall & \text{si } E_j \delta_j \geq 0 \end{cases}$$

$$\sigma_j = M(j, J_N) l_N$$

3.6.3 Estimation de la suboptimalité

Calculons la nouvelle estimation de suboptimalité $\beta(\bar{x}, J_B)$ en fonction de $\beta(x, J_B)$:

$$\begin{aligned} \beta(\bar{x}, J_B) &= \sum_{\substack{j \in J_N \\ \bar{E}_j > 0}} \bar{E}_j(\bar{x}_j - d_j^-) + \sum_{\substack{j \in J_N \\ \bar{E}_j < 0}} \bar{E}_j(\bar{x}_j - d_j^+) \\ &= \sum_{\substack{j \in J_N \\ \bar{E}_j > 0}} \bar{E}_j(x_j + \theta^0 j_j - d_j^-) + \sum_{\substack{j \in J_N \\ \bar{E}_j < 0}} \bar{E}_j(x_j + \theta^0 l_j - d_j^+) \\ &= \sum_{\substack{j \in J_N \\ E_j > 0}} (E_j + \theta^0 l_j - d_j^-) + \sum_{\substack{j \in J_N \\ E_j < 0}} (E_j + \theta^0 \delta_j)(x_j + \theta^0 l_j - d_j^+) \\ &= \sum_{\substack{j \in J_N \\ E_j > 0}} E_j(x_j - d_j^-) + \sum_{\substack{j \in J_N \\ E_j > 0}} E_j \theta^0 l_j + \sum_{\substack{j \in J_N \\ E_j > 0}} \theta^0 \delta_j(x_j - d_j^-) + \sum_{\substack{j \in J_N \\ E_j > 0}} \theta^0 \delta_j \theta^0 l_j + \\ &\quad \sum_{\substack{j \in J_N \\ E_j < 0}} E_j(x_j - d_j^+) + \sum_{\substack{j \in J_N \\ E_j < 0}} \theta^0 E_j l_j + \sum_{\substack{j \in J_N \\ E_j < 0}} \theta^0 \delta_j(x_j - d_j^+) + \sum_{\substack{j \in J_N \\ E_j < 0}} \theta^0 \delta_j \theta^0 l_j. \end{aligned}$$

En vertu des relations (3.30), on aura alors

$$\begin{aligned} \beta(\bar{x}, J_B) &= \beta(x, J_B) = \theta^0 \sum_{\substack{j \in J_N \\ E_j > 0}} E_j(d_j^- - x_j) + \theta^0 \sum_{\substack{j \in J_N \\ E_j < 0}} E_j(d_j^+ - x_j) + \theta^0 \sum_{\substack{j \in J_N \\ E_j > 0}} \delta_j(-l_j) + \\ &\quad \theta^0 \sum_{\substack{j \in J_N \\ E_j < 0}} \delta_j(-l_j) + \theta^{02} \sum_{\substack{j \in J_N \\ E_j > 0}} \delta_j l_j + \theta^{02} \sum_{\substack{j \in J_N \\ E_j < 0}} \delta_j l_j \\ &= \beta(x, J_B) - \theta^0 \beta(x, J_B) - \theta^0 \delta'_N l_N + \theta^{02} \delta'_N l_N. \end{aligned}$$

Donc, on obtient la formule suivante :

$$\beta(\bar{x}, J_B) = (1 - \theta^0) \beta(x, J_B) - \theta^0 (1 - \theta^0) l' D l. \quad (3.34)$$

3.6.4 Changement de support

- Pour $\theta^0 = 1$, alors le vecteur $x^0 = \bar{x} = x + l$ est une solution optimale du problème (3.18),(3.20).

Le changement de support s'effectue lorsque $\theta^0 < 1$ et $\beta(\bar{x}, J_B) > \epsilon$. Trois cas peuvent alors se présenter :

- $\theta^0 = \theta_{j_1} < 1$: contrairement à la méthode du simplexe, le choix de l'indice j_0 n'est pas unique, ce qui fait la particularité de cette méthode. Lorsque ce cas se réalise pour un indice $j_1 \in J_B$, on a alors

$$l_{j_1} = - \sum_{j \in J_N} e'_{j_1} A_B^{-1} a_j l_j = - \sum_{j \in J_N} x_{j_1 j} l_j \neq 0,$$

où e est un vecteur unitaire de dimension m . Il existe alors $j_0 \in J_N$ tel que $x_{j_1 j_0} \neq 0$. Cette dernière condition nous assure par conséquent que $\bar{J}_B = (J_B \setminus j_1) \cup j_0$ est bel et bien un support.

Si on peut avoir $x_{j_1 j_0} \neq 0$, avec $j_0 \in J_S$, on posera donc

$$\bar{J}_B = (J_B \setminus j_1) \cup j_0, \bar{J}_S = (J_S \setminus j_0)$$

Sinon, on choisira un indice $j_0 \in J_{NN}$ tel que $l_{j_0} \neq 0$ et on posera

$$\bar{J}_B = (J_B \setminus j_i) \cup j_0 \text{ et } \bar{J}_S = J_S$$

- $\theta^0 = \theta_{j_s}$: on pose $\bar{J}_B = J_B$, $\bar{J}_S = J_S \setminus j_s$.
- $\theta^0 = \theta_F$: dans ce cas on aura $\bar{J}_B = J_B$, $\bar{J}_S = J_S \cup j_*$.

Pour ces trois cas sus-cités, on recommencera une nouvelle itération avec le plan de support $\{\bar{x}, \bar{J}_P\}$ si $\beta(\bar{x}, \bar{J}_B) > \epsilon$.

3.7 Algorithme de Résolution

Soit un nombre réel positif ou nul quelconque ϵ et un plan de support initial $\{x, J_P\}$ tel que $J_P = \{J_B, J_S\}$, avec $J_S = \emptyset$ pour plus de facilité.

Algorithme

Debut

1) Soit un plan de support initial $\{x, J_P\}$ du problème (3.1)-(3.3);

2) Calculer le vecteur des estimations $E'_N = E'(J_N) = g'_N - u'A_N$;

Si $\beta(x, J_B) \leq \epsilon$ Alors x est ϵ -optimal; Aller à Fin.

Sinon Aller en 3;

3) Changement de plan

Calculer la direction d'amélioration l ;

Calculer le pas $\theta^0 = \min\{1, \theta_{j_1}, \theta_{j_s}, \theta_F\}$;

Calculer $\bar{x} = x + \theta^0 l$;

Calculer l'estimation de suboptimalité $\beta(\bar{x}) = (1 - \theta^0)(\beta - \theta^0 l' D l)$;

Si $\beta(\bar{x}) \leq \epsilon$ Alors \bar{x} est ϵ -optimal; Aller à Fin;

Sinon Aller en 4;

4) Changement de support

Si $\theta^0 = 1$ Alors $\bar{x} = x + 1$; $\beta(\bar{x}) = 0$; \bar{x} est optimal; Aller à Fin;

Sinon Si $\theta^0 = \theta_{j_i}$, Alors

Si $j_0 \in J_S$, Alors $\bar{J}_B = (J_B \setminus j_i) \cup j_0$ et $\bar{J}_S = J_S \setminus j_0$.

Sinon $j_0 \in J_{NN}$ Alors $\bar{J}_B = (J_B \setminus j_i) \cup j_0$ et $\bar{J}_S = J_S$

Sinon Si $\theta^0 = \theta_{j_s}$ Alors $\bar{J}_B = J_B$, $\bar{J}_S = J_S \setminus j_s$;

Sinon Si $\theta^0 = \theta_F$ Alors $\bar{J}_B = J_B$, $\bar{J}_S = J_S \cup j_*$;

Si $\beta(\bar{x}, \bar{J}_B) > \epsilon$; alors aller en 2 en partant de $\{\bar{x}, \bar{J}_P\}$.

Fin

3.8 Exemple numérique

Soit le problème de programmation quadratique suivant :

$$F(x) = 4x_1^2 - 4x_1x_2 + 2x_2^2 + 2x_1 + x_2 + 3x_3 - x_4 \rightarrow \min x_1 - 4x_2 + x_3 = 3;$$

$$2x_1 + x_2 + x_4 = 4;$$

$$-2 \leq x_1 \leq 2, 0 \leq x_2 \leq 4, 2 \leq x_3 \leq 5, -3 \leq x_4 \leq 6,$$

avec

$$D = \begin{pmatrix} 8 & -4 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, c = \begin{pmatrix} 2 \\ 1 \\ -3 \\ -1 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, d^- = \begin{pmatrix} -2 \\ 0 \\ 2 \\ -3 \end{pmatrix}, d^+ = \begin{pmatrix} 2 \\ 4 \\ 5 \\ 6 \end{pmatrix}.$$

Soit $x = (0, 0, 3, 4)$ un plan initial du problème. Posons $J_B = \{3, 4\}$, $J_N = \{1, 2\}$. Nous avons alors $A_B = (a_3, a_4) = I_2$; $A_N = (a_1, a_2) = \begin{pmatrix} 1 & -4 \\ 2 & 1 \end{pmatrix}$.

Déterminons la matrice $M = Z'DZ$:

$$Z = Z(J, J_N) = \begin{pmatrix} -A_B^{-1}A_N \\ I_N \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 4 \\ -2 & -1 \end{pmatrix}$$

D'où

$$M = Z'DZ = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 4 & -1 \end{pmatrix} \begin{pmatrix} 8 & -4 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 4 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} 8 & -4 \\ -4 & 4 \end{pmatrix}$$

Calculons le vecteur gradient $g(x) = (g_B, g_N)$

$$g(x) = Dx + c = \begin{pmatrix} 8 & -4 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 3 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ -3 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ -3 \\ -1 \end{pmatrix},$$

$$g_B = \begin{pmatrix} -3 \\ -1 \end{pmatrix} \quad \text{et} \quad g_N = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Calculons le vecteur des estimations $E_N = g_N - (g_B' A_B^{-1} A_N)'$

$$E_N = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \left((-3, -1) \begin{pmatrix} 1 & -4 \\ 2 & 1 \end{pmatrix} \right)' = \begin{pmatrix} 7 \\ -10 \end{pmatrix}.$$

Posons $J_S = \emptyset, J_{NN} = J_N \setminus J_S = \{1, 2\}$. La paire $\{x, J_P\}$, avec $J_P = \{J_B, J_S\}$ est alors le plan de support du problème considéré. Calculons l'estimation de suboptimalité :

$$\begin{aligned}\beta(x, J_B) &= \sum_{E_j > 0} E_j(x_j - d_j^-) + \sum_{E_j < 0} E_j(x_j - d_j^+) \\ &= E_1(x_1 - d_1^-) + E_2(x_2 - d_2^+) \\ &= 7(0 + 2) + (-10)(0 - 4) \\ &= 54 > \varepsilon.\end{aligned}$$

Itération 1 :

$$l_1 = d_1^- - x_1 \quad \text{car } E_1 > 0, l_2 = d_2^+ - x_2 \quad \text{car } E_2 < 0,$$

on alors :

$$l(J_{NN}) = \begin{pmatrix} -2 \\ 4 \end{pmatrix}.$$

$$l(J_S) = (l_j, j \in J_S) = 0, \text{ car } J_S = \emptyset$$

$$l(J_B) = (l_j, j \in J_B) = -A_B^{-1} A_{NN} l_{NN} = \begin{pmatrix} -1 & 4 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 18 \\ 0 \end{pmatrix}$$

D'où

$$\delta_N = \delta(J_N) = \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = M l_N = \begin{pmatrix} 8 & -4 \\ -4 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ 4 \end{pmatrix} = \begin{pmatrix} -32 \\ 24 \end{pmatrix}$$

Calculons le pas θ le long de cette direction :

Pour $j \in J_{NN}$, nous avons : $\theta^0 = 1$.

Pour $j \in J_B$, nous avons :

$$\theta_{j_1} = \min_{j \in J_B} \theta_j = \min\{\theta_3, \theta_4\}, \text{ avec } \begin{cases} \theta_3 = 1/9 \\ \theta_4 = \infty, \text{ car } l_4 = 0. \end{cases}$$

$$\theta_{j_1} = \theta_3 = 1/9 \Rightarrow j_1 = 3.$$

Pour $j \in J_S$ nous avons : $\theta_{J_S} = \infty, \text{ car } J_S = \emptyset$.

Pour θ_F , nous avons $\sigma_{j^*} = \min\{\sigma_1, \sigma_2\}$, tel que : $\sigma_1 = \frac{-E_1}{\delta_1} = 7/32; \sigma_2 = \frac{-E_2}{\delta_2} = 5/12$

D'où $\theta^0 = \min\{1, 1/9, \infty, 7/32\} = \theta_3 = 1/9$.

On a alors

$$\bar{x} = x + \theta^0 l = \begin{pmatrix} -2/9 \\ 4/9 \\ 5 \\ 4 \end{pmatrix}$$

et

$$\beta(\bar{x}, J_B) = (1 - \theta^0)(\beta(x, J_B) - \theta^0 l' D l) = \frac{2608}{81} \simeq 32.19.$$

Soit $X = X(J_B) = A_B^{-1} a_{j_0} = \begin{pmatrix} x_{x_1 j_0} \\ x_{x_2 j_0} \end{pmatrix}$, avec $j_0 \in J_N$ non optimal.

$$X = X(J_B) = \begin{pmatrix} x_{31} \\ x_{41} \end{pmatrix} = A_B^{-1} a_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \Rightarrow x_{j_1 j_0} = 1 \neq 0, \text{ avec } j_1 = 3, j_0 = 1.$$

D'où

$$\bar{J}_B = \{3, 4\} \setminus \{3\} \cup \{1\}$$

avec $\bar{A}_B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$ dont sa matrice inverse est $\bar{A}_B^{-1} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$. Donc

$$\bar{J}_B = \{1, 4\}; \bar{J}_S = \emptyset; \bar{J}_{NN} = \{2, 3\}; \bar{J}_P = \{\bar{J}_B, \bar{J}_S\}.$$

On commence la deuxième itération par le plan de support $\{\bar{x}, \bar{J}_p\}$:

Calculons alors le nouveau vecteur des estimations :

$$g(\bar{x}) = D\bar{x} + c = \begin{pmatrix} 8 & -4 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -2/9 \\ 4/9 \\ 5 \\ 4 \end{pmatrix} + \begin{pmatrix} 2 \\ 4 \\ -3 \\ -1 \end{pmatrix} = \begin{pmatrix} -14/9 \\ 33/9 \\ -3 \\ -1 \end{pmatrix}, \bar{E}_N = \begin{pmatrix} 58/9 \\ -31/9 \end{pmatrix}$$

ainsi que la nouvelle estimation de suboptimalité :

$$\beta(\bar{x}, \bar{J}_B) = \bar{E}_2(\bar{x}_2 - d_2^-) + \bar{E}_3(\bar{x}_3 - d_3^+) = 232/81 \simeq 2.86.$$

Déterminons la matrice M :

$$M = Z'DZ = \begin{pmatrix} 4 & 1 & 0 & -9 \\ -1 & 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 8 & -4 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 4 & -1 \\ 1 & 0 \\ 0 & 1 \\ -9 & 2 \end{pmatrix} = \begin{pmatrix} 100 & -28 \\ -28 & 8 \end{pmatrix}$$

Calculons la direction l :

$$l(\bar{J}_{NN}) = \begin{pmatrix} l_2 \\ l_3 \end{pmatrix} = \begin{pmatrix} -4/9 \\ 0 \end{pmatrix}.$$

$$l(\bar{J}_B) = (l_j, j \in \bar{J}_B) = -A_B^{-1} A_{NN} l_{NN} = \begin{pmatrix} 4 & -1 \\ -9 & 2 \end{pmatrix} \begin{pmatrix} -4/9 \\ 0 \end{pmatrix} = \begin{pmatrix} -16/9 \\ 4 \end{pmatrix}$$

D'où

$$\delta_N = \begin{pmatrix} \delta_2 \\ \delta_3 \end{pmatrix} = M l_N = \begin{pmatrix} 100 & -28 \\ -28 & 8 \end{pmatrix} \begin{pmatrix} -4/9 \\ 0 \end{pmatrix} = \begin{pmatrix} -400/9 \\ 112/9 \end{pmatrix}$$

Calculons le pas le long de cette nouvelle direction :

Pour $j \in \bar{J}_{NN}$, nous avons : $\theta^0 = 1$.

Pour $j \in \bar{J}_S$, nous avons :

$$\theta_{ji} = \min_{j \in \bar{J}_B} \theta_j = \min\{\theta, \theta\}, \text{ avec } \begin{cases} \theta = \\ \theta_4 = 1/2. \end{cases}$$

$$\theta_{j1} = \theta_4 = 1/2.$$

Pour $j \in \bar{J}_S$ nous avons : $\theta_{j*} = \infty$, car $\bar{J}_S = \emptyset$,

et finalement

$$\sigma_{j*} = \min_{j \in \bar{J}_{NN}} \sigma_j = \min\{\sigma_2, \sigma_3\}, \text{ avec } \begin{cases} \sigma_2 = 29/200 \\ \sigma_3 = 31/112. \end{cases}$$

$$\sigma_{j*} = \sigma_2 = 29/200.$$

Ainsi $\theta^0 = \sigma_{j^*} = \sigma_2 = 29/200$.

Nous avons alors :

$$\bar{\bar{x}} = \bar{x} + \theta^0 l = \begin{pmatrix} -24/50 \\ 19/50 \\ 5 \\ 229/50 \end{pmatrix}.$$

et

$$\bar{\bar{J}}_B = \bar{J}_B = \{1, 4\}; \bar{\bar{J}}_{NN} = \{3\}; \bar{\bar{J}}_P = \{\bar{\bar{J}}_B, \bar{\bar{J}}_S\}.$$

Donc le nouveau vecteur des estimations vaut :

$$g(\bar{\bar{x}}) = D\bar{\bar{x}} + c = \begin{pmatrix} 8 & -4 & 0 & 0 \\ -4 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -24/50 \\ 19/50 \\ 5 \\ 229/50 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ -3 \\ -1 \end{pmatrix}, \bar{\bar{E}}_N = \left(0 \quad \frac{-41}{25}\right),$$

et

$$\beta(\bar{\bar{x}}, \bar{J}_B) = (1 - \theta^0)(\beta(\bar{x}, \bar{J}_B) - \theta^0 l' D l) = 0.$$

La valeur de $\beta(\bar{\bar{x}}, \bar{J}_B)$ étant nulle, le critère d'optimalité est alors vérifié.

D'où $x^0 = (-24/50, 19/50, 5, 229/50)$ est plan optimal, avec $F(x^0) = \frac{-911}{50} = -18.22$.

la méthode adaptée est une extension de la méthode directe du support, sa particularité est de manipuler les variables de décisions tel que se présentent initialement sans modification préliminaires. de plus, cet algorithme est doté d'un critère d'arrêt qui peut donner une solution approchée avec une précision choisie à l'avance.

3.9 Récapitulatif

Méthodes	Année	Auteurs	Caractéristiques
Points intérieurs	1984	<ul style="list-style-type: none"> - N.K.Karmakar - Nesterov - Nemirovski 	<ul style="list-style-type: none"> - Problème convexe et non convexe de taille moyenne. - Chaque itération peut faire un progrès vers la solution - Approxime la solution soit de l'intérieur ou de l'extérieur du domaine réalisable
Activation des contraintes	<ul style="list-style-type: none"> - 1971 - 1978 - 1983 - 1997 	<ul style="list-style-type: none"> - R Flechter - Gil et Al - D. Goldfarb et A.idnani - Boland 	<ul style="list-style-type: none"> - Résolution des problèmes de taille raisonnable. - Traiter les problèmes tel qu'ils se présentent sans modifier leur contraintes. - L'ensemble actif change lentement.
Gradient projeté	1974	P.E Gil et W. Murray	<ul style="list-style-type: none"> - Coût élevé pour le calcul de projection. - Meilleure convergence. - changement rapide de l'ensemble actif. -
Méthode adaptée de support	- 1983	<ul style="list-style-type: none"> - R. Gabassov - F.M Kirilova 	<ul style="list-style-type: none"> - Fait changer tous les indices non optimaux. - Traiter les problèmes tel qu'ils se présentent sans modifier leur contraintes.. - peut avoir une solution ϵ-optimale. - Converge en un nombre fini d'itérations

Conclusion générale

Ce travail est une synthèse des quatre méthodes de la résolution des problèmes de programmation quadratique convexe (activation des contraintes, les points intérieurs, le gradient projeté et la méthode adaptée); pour cela nous avons d'abord rappelé les notions fondamentales sur l'algèbre linéaire, la programmation quadratique et la convexité; ensuite nous avons présenté les résultats classiques sur l'optimisation non linéaire, sans contrainte où on a fait un petit rappel sur les méthodes de minimisation (méthode de gradient, méthode de Newton) et avec contraintes, les conditions d'optimalité de Karush-Kuhn-Tucker (KKT), ainsi que les théorèmes de dualité en programmation quadratique convexe.

perspectives

- *Faire expérimentation numérique pour comparer les quatre méthodes.*
- *Appliquer ces méthodes pour résoudre des cas pratiques comme :*
 - *Gestion de production*
 - *Gestion de portefeuille*
 - *Les moindres carrées*

Bibliographie

- [1] J. Abadie. *On the Kuhn-Tucker theorem*. In J. Abadie, editor, *Nonlinear Programming North Holland*, Amsterdam, 1967.
- [2] E. W. Barankin and R. Dorfman. *On quadratic programming, volume II*. Publications in statistics, University of California, 1958.
- [3] E. M. L. Beale. On quadratic programming. *Naval Research Logistics Quarterly*, , 1959.
- [4] M. Bentoubache. Nouvelle méthode pour la résolution des problèmes de programmation linéaire sous forme canonique et à variables bornées., Mémoire de Magister, Université de Béjaïa, 2005.
- [5] B. Brahemi. méthodes primales duales pour la résolution des problèmes de programmation quadratique convexe., Mémoire de Magister, Université de Béjaïa, 2006.
- [6] B. Brahmi. méthodes primales duales pour la résolution des problèmes de programmation quadratique : Extension et applications, *thèse de doctorat, Université de Béjaïa*,
- [7] A. B. Berkelaar. Sensitivity analysis in degenerate quadratic programming. Technical Report 30, Econometric Institute, Erasmus University, Rotterdam, The Netherlands, 1997.
- [8] M. O. Bibi. Optimization of linear dynamic system with double constraint on the trajectories. In *Collected abstracts of the 16th conference on systems modeling and optimization*, Compiègne (France), 1993.
- [9] M. O. Bibi Support method for solving a linear quadratic problem with polyhedral constraints on control, 1996.
- [10] M. O. Bibi. Méthodes adaptées de la programmation linéaire. Cours de Post-graduation en Recherche Opérationnelle, Université de Béjaïa, 2004.
- [11] Laouar Abdelhak *Méthodes du support pour la résolution d'un problème de programmation quadratique convexe à variable mixtes*, Mémoire de magister en Mathématiques appliquées, Université de Béjaïa 2010.
- [12] FLETCHER R, A general quadratic programming algorithm. *Journal of institute of mathematics and its applications*. , 1996.
- [13] Kranich E. Interior point methods for mathematical programmings, A bibliography, 1991.
- [14] M. O. Bibi and A. Faradji. Méthode duale pour la minimisation d'une fonctionnelle quadratique convexe avec contraintes simples. Résumés des communications de la 4^{me} rencontre de Recherche Opérationnelle, U.S.T.H.B, Alger, 1996.
- [15] M. O. Bibi and N. Ikhenche. Optimisation par la méthode adaptée d'un problème linéaire-quadratique convexe à variables bornées. Actes du Colloque international MSS'4, U.S.T.H.B, Alger, 2004.
- [16] G. B. Dantzig. programming in linear structure. *Econometrica*, 1949
- [17] A. Faradji. Algorithmes de minimisation d'une fonctionnelle quadratique. Mémoire de magister, Université de Tizi-Ouzou, 1998.

- [18] R. Gabassov. Adaptive method of linear programming. Preprints of the University of Karlsruhe, Institute of Statistics and Mathematics, Karlsruhe, Germany, 1993.
- [19] N.K.Karmarkar. A new polynomial-time algorithm for linear programming, *combinatorica* 1984
- [20] N. Ikhenche. Méthode de support pour la minimisation d'une fonctionnelle quadratique convexe. Mémoire de magister, Université de Béjaia, 2004
- [21] R.Gabassov, F.M.Kirillova. Méthode de programmation linéaire. Université de Minsk, 1977
- [22] R.Gabassov, F.M.Kirillova. Méthode d'optimisation. Université de Minsk, 1981
- [23] S. Radjef. Sur la programmation linéaire multiobjectifs. Mémoire de magister, Université de Béjaia, 2001
- [24] S. Radjef. Application de la méthode adaptée aux problèmes multicritères. Thèse de doctorat, Université de Béjaia, 2011
- [25] P.WOLF. The simplex method for quadratic programming. *Econometrica*, 1959
- [26] H.M. Markowitz. The optimization of quadratic function subject to linear constraint. *Naval research logistic quarterly*, 1956
- [27] R.Fletcher. Practical methods of optimization. Constrained optimization. John Wiley and sons edition, 1981
- [28] R.Gabassov, F.M. Kirillova, méthode d'optimisation. édition de l'université, Minsk 1981