الجمهوريـة الجزائـريـة الديمقراطيـة الشعبيـة
République Algérienne Démocratique et Populaire
وزارة التعليـم العالـي والبحـث العلمـي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Nº Réf :……………

Centre Universitaire
Abd elhafid boussouf Mila

Institut des sciences et de la technologie

Département de Mathématiques et Informatique

Mémoire préparé En vue de l'obtention du diplôme de Licence

en: - Filière mathématiques

## Thème

# Programmation quadratique

Préparé par :

- Belegherib Nassima
- Belkessour Sara
- Guessas Ilham
- Bounaas Chaib Noorelhouda

Encadrer par : Baiche Kanzia
Grade : M.A.B

Année universitaire :2014/2015

# Remerciements

Nous remercions **«DIEU»** qui nous a donné du courage et de la volonté d'avoir réussit dans notre étude.

Nous tenons remercier notre encadreur **Mm BAICHE KANZIA** nous à proposé ce sujet de recherche, et nous à encadré et surtout par ses conseils, sa compréhension, sa gentillesse, ses encouragements et sa bibliographie.

Nous tenons remercier a mes *Parents, source inépuisable d'amour, d'affection et de soutier, a tous famille chaque individu par votre nom, à tous mes amis et à mes camarades de promotion 2014.*

Nous voudrions remercier aussi tout personnes qui a contribué de loin ou de prés à la réalisation de ce mémoire.

# Table des matières

**Introduction Générale**

L'optimisation joue un role très important en recherche opérationnelle (en économie et micro-économie), en mathématiques appliques, en analyse numérique et en statistiques.

Il existe deux grandes classes d'optimisation, celles sans contraintes et avec contraintes.

Le problème d'optimisation consiste alors à déterminer les variables de décison conduisant aux meilleurs conditions de fonctionnement du système (ce qui revient à minimiser ou maximiser).

Les méthodes de programmation quadratique successive sont considérées aujourd'hui comme les plus efficaces par de nombreux auteurs [S et Al. (1984)] , [B-H-V. (1984)] , [A-K. (1985)], [M-F. (1987)] , [C et Al. (1987)] . On sait qu'elles possèdent une vitesse de convergence superlinéaire (et meme quadratique dans le cas des méthodes de Newton) , ce qui, de ce point vue, les rend supérieures aux autres méthodes. On trouvera des éléments de comparaison avec les méthodes de gradient réduit dans [B-H-W. (1982)] et avec les méthodes de lagrangiens augmentes dans [A-K. (1985)] et [B-H-V. (1984) ] .

Ce mémoire est composé d'une introduction générale, trois chapitres et une conclusion générale.

Dans le premier chapitre nous présentons des généralités et des notions fondamentales qui sont necéssaires pour la suite de ce travail.

Dans le deuxième chapitre, on s'interesse à l'optimisation avec et sans contrainte.

Dans le dernier chapitre, nous donnons une présentation d'un problème de programmation quadratique, nous étudions les méthodes du gradient avec projection et les méthodes de Newton.

**Notation 0.0.1**

$C$ : *corps des nombres rationnelles.*

$\mathbb{R}$ : *corps des nombre réelles.*

$epi\,(F)$ : *epigraphe de* $f$.

$\|\,.\,\|$ : *la norme.*

$E$ : *espace vectoriel.*

$\lambda$ : *conjugué.*

$|\,.\,|$ : *module.*

$\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

$M^T = M$ : *la matrice symetrique.*

$B = \{e_1, e_2, ....., e_n\}$ : *une base de* $E$.

$e_i =$ *une vecteur de :* $\{1, ..2, ..., i\}$.

$C, k$ : *les ensembles des contraintes.*

$V$ : *un voisinage.*

$C^0$ : *continue.*

$\frac{\partial f}{\partial x}$ : *dérivée partielle de* $f$ *par rapport à* $x$.

$C^1 = \frac{\partial f}{\partial x}$*(existe et continue).*

$C^2 = \frac{\partial^2 f}{\partial x_i \partial x_j}$*(existe et continue).*

$E$ : *espace vectoriel.*

$\mathbb{R}^* = \mathbb{R} / \{0\}$.

$\mathbb{R}_+ = [0, +\infty[$.

$\nabla f$ : *vecteur gradient de la fonction* $f$.

$\nabla^2 f\,(x^*) = \frac{\partial^2 f}{\partial x_i \partial x_j} \forall i = 1, n \quad , j = 1, m$ : *la matrice hessienne.*

$K^0$ : *l'intérieur de l'ensemble* $K$.

$f'$ : *dérivée de la fonction* $f$.

$f''$ : *deuxième dérivées de fonction* $f$.

*dom : domaine.*

$x^T$ : *transposé de* $x$.

$B\,(x^*, \alpha)$ : *boule ouvert de centre* $x^*$*et de rayon* $\alpha$.

$L\,(E)$ : *l'ensemble des fonctions linéaire.*

$LM$ : *l'ensemble des minimiseurs locaux de* $f$.

$< . >$ : *produit scalaire.*

$\lambda^T$ : *vecteur multiplicateur de lagrange.*

# Chapitre 1

# Ensemble convexe et fonction convexe :

Dans ce chapitre, nous allons rappler les plus importants résultats d'analyse convexe ( les ensembles convexes et les fonctions convexes ), des fonctions affines, des fonctions quadratiques pour le dévellopement de ce travail.

## 1.1 Ensemble convexe :

**Définition 1.1.1**
*Soit $E$ un espace vectoriel et $A$ un partie de $E$, on dit que $A$ est convexe si :*

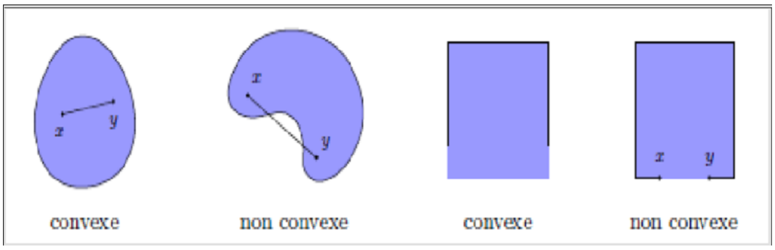
$$\forall x,\ y \in A,\ \forall \lambda \in [0,1] \qquad \lambda x + (1-\lambda)\, y \in A.$$

**Définition 1.1.2**
*Un sous ensemble $C$ de $\mathbb{R}$ est un convexe si pour chacun des couples $(x, y)$ de $C^2$ le segement $[x, y]$ est inclus dans $C$. Les convexes de $\mathbb{R}$ sont les intervalles de $\mathbb{R}$.*

### 1.1.1 Exemple et propriétes :

On a $C$, $C_1$, $C_2$ sont des ensembles convexes.
Ci dessous $E$, $E_1$, $E_2$ et $F$ sont des espaces vectoriels :
1. La somme $C_1 + C_2 = \{x_1 + x_2\ :\ x_1 \in C_1,\ x_2 \in C_2\}$ est convexe.
2. Le produit scalaire $\alpha C = \{\alpha x\ :\ x \in C\}$ est convexe.

3. Si $\{C_i\}_{i\in I}$ est une famille quelconque de convexes de $E$, alors leur intersection $\cap_{i\in I} C_i$ est un convexe.




Figure(1.1)- Ensembles convexes et non convexes.

4. Soit $A : E \to F$ une application lineare. L'image directe $A\,(C)$ est convexe et l'image réciproque $A^{-1}\,(C)$ est convexe.

5. Soient $C_1 \subset E_1$ et $C_2 \subset E_2$ alors $C_1 * C_2$ est convexe dans $E_1 * E_2$ si et seulment si $C_1$ et $C_2$ sont convexes.

**Exemple 1.1.3**

$S = \{(x_1, x_2) = (2,3)\,,\ (x_1, x_2) = (1,0)\}\,.$

*$S$ n'est pas convexe* $\Rightarrow \exists\ x,\ y \in S,\ \exists\ \lambda \in [0,1] \quad / \quad \lambda x + (1-\lambda)\,y \notin S.$

*On pose* $X = (2,3)\,,\ Y = (1,0)\,,\ \lambda = \frac{3}{4}.$

$Z = \lambda X + (1-\lambda)\,Y = \frac{3}{4}\,(2,3) + \frac{1}{4}\,(1,0) = \left(\frac{7}{4}, \frac{9}{4}\right) \notin S.$

*Donc $S$ n'est pas convexe.*

## 1.2 Fonction Convexe :

**Définition 1.2.1**

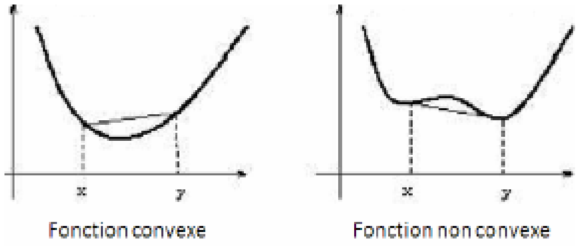
*Soit $[a,b]$ un intervalle de $\mathbb{R}$. Soit $f$ définie sur $[a,b]$ à valeurs dans $\mathbb{R}$, on dit que $f$ est convexe si pour tous les couples $(x,y)$ des points de cet intervalle $[a,b]$ et pour tous $\lambda \in [0,1]$, l'inégalité suivante, dit <<inégalité de convexité >> est vérifiée :*

$$f\,(\lambda x + (1-\lambda)\,y) \leq \lambda f\,(x) + (1-\lambda)\,f\,(y)\,.$$

*La fonction $f$ est dite strictement convexe si l'inégalité de convexité est stricte lorsque le couple $(x,y)$ vérifiée : $x \neq y$ avec $\lambda \in \,]0,1[\,.$*

*Autrement dit si :*

$$\begin{aligned} \forall\,[x,y] &\subset [a,b]\,, x \neq y, \forall \lambda \in \,]0,1[\,. \\ f\,(\lambda x + (1-\lambda)\,y) &< \lambda f\,(x) + (1-\lambda)\,f\,(y)\,. \end{aligned}$$




Figure(1.2)- Fonction convexe et non convexe.

**Définition 1.2.2**
*L'épigraphe de la fonction $F$ définie par :*

$$epi\,(F) = \left\{(x, \lambda) \in \mathbb{R}^n * \mathbb{R} \quad / \quad F(x) \leq \lambda \in \mathbb{R}^{n+1}\right\}.$$

**Définition 1.2.3**
*Un fonction $F : \mathbb{R}^n \to \bar{\mathbb{R}} = [-\infty, +\infty]$ est dite convexe si son épigraphe est un ensemble convexe de $\mathbb{R}^{n+1}$.*
*ie : $F : \mathbb{R}^n \to \mathbb{R}$ convexe $\Leftrightarrow$ $epi\,(F)$ convexe dans $\mathbb{R}^{n+1}$.*

**Exemple 1.2.4**
*Soit $f(x) = \|x\|$, $x \in \mathbb{R}^n$.*
*On montre que pour $x$, $y \in \mathbb{R}^n$, $\lambda \in [0, 1]$.*

$$\begin{aligned}
\|\lambda x + (1-\lambda) y\| &\leq \|\lambda x\| + \|(1-\lambda) y\| \\
&\leq |\lambda| \, \|x\| + |(1-\lambda)| \, \|y\| \\
&\leq \lambda \|x\| + (1-\lambda) \|y\| \\
&\leq \lambda f(x) + (1-\lambda) f(y).
\end{aligned}$$

**Exemple 1.2.5**
*Soit $f(x) = x^2$.*
*On à $epi\,(f)$ convexe dans $\mathbb{R}^2 \Rightarrow f$ est fonction convexe.*

## 1.3 Fonctions affines :

### 1.3.1 Rappel sur la fonction affine :

**Définition 1.3.1**
*On appelle fonction affine une fonction du type $x \to ax + b$, où $a$ et $b$ sont des nombres réelle.*

**Exemple 1.3.2**
$f : x \to -2x + 3$.
*$f$ est une fonction affine. L'image de $2$ est $-1$ $((-2) * 2 + 3 = -1)$. L'antécédent de $7$ est $(-2)$ (résoudre l'équation $-2x + 3 = 7$).*

### 1.3.2 Représentation graphique d'une fonction affine :

**Propriété :**

Dans un repère, la représentation graphique d'une fonction affine est une droite.

On dit que $y = ax + b$ est une équation de cette droite. Le nombre $a$ est appelé coefficient directeur de la droite et $b$ est l'ordonnée à l'origine.

**Propriété :**

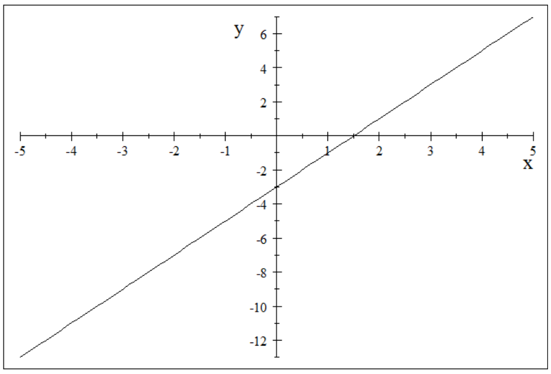
Appelons $(d)$ la droite d'équation $y = ax + b$. Appelons $M$ un point des coordonnées $(x_M, y_M)$.

Si $M \in (d)$, alors ses coordonnées vérifient l'égalité :

$$y_M = ax_M + b.$$

Réciproquement, si les coordonnées de $M$ vérifient l'égalité $y_M = ax_M + b$, alors $M \in (d)$.

**Exemple 1.3.3**
*Représenter graphiquement la fonction affine $x \to 2x - 3$. D'après ce qui précède, on sait qu'il s'agit d'une droite. Pour tracer cette droite, il faut deux points. $y = 2x - 3$ est l'équation de la droite à tracer.*
*a) Si $x = 0$, alors $y = -3$ donc le point de coordonnées $(0, -3)$ appartient à la droite.*
*b) Si $x = 2$, alors $y = 1$ donc le point de coordonnées $(2, 1)$ appartient à la droite.*




### 1.3.3 Détermination d'une fonction affine :

**Propriété :**

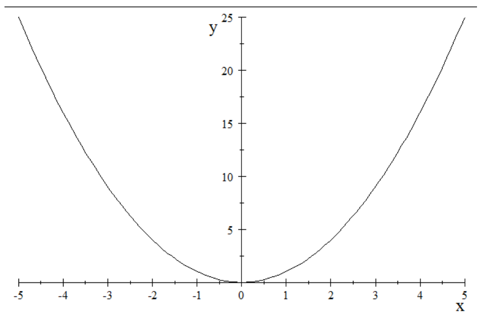
Si $f$ est une fonction affine, alors les accroissements de $f(x)$ et de $x$ sont proportionnels, le coefficient de proportionnalité étant $a$. Autrement dit, si $f(x) = ax + b$, alors quels que soient les nombres $x_1$ et $x_2$ : $f(x_2) - f(x_1) = a(x_2 - x_1)$.

**Exemple 1.3.4**
*Déterminer la fonction affine telle que :*
*a)* $2$ *est une image de* $(-3)$.
*b)* $(-3)$ *est une image de* $7$.
*Une fonction affine est de la forme* $x \to ax + b$.
*D'après la propriété précédente, on a :*
$f(2) - f(-3) = a(2 - (-3))$.
$-3 - 7 = 5a$.
$5a = -10$.
$a = -\frac{10}{5} = -2$.
*On sait maintenant que la fonction affine est de la forme* $x \to -2x + b$. *L'image de* $2$ *est* $(-4) + b$ *donc* $(-4) + b = -3$. *On déduit :* $b = (-3) + 4 = 1$. *La fonction affine cherchée est* $x \to -2x + 1$.

## 1.4 Les fonctions quadratiques :

### 1.4.1 La fonction quadratique de base :

$f(x) = x^2$



### 1.4.2 La fonction quadratique transformée :

Elle s'écrit sous deux formes : - Générale : $f(x) = ax^2 + bx + c$.
- Canonique : $f(x) = a(x-h)^2 + k$.

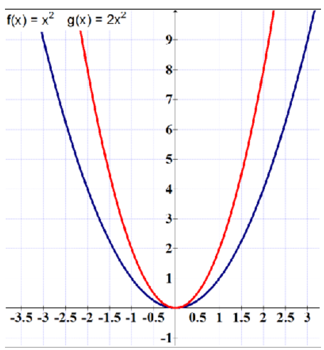
**Introduction des paramètres :**

– Le paramètre « $a$ » : $f(x) = x^2 \longrightarrow g(x) = ax^2$.

Ce paramètre entraîne un changement d'échelle verticale en multipliant par « $a$ » les ordonnées des couples de la fonction de base.

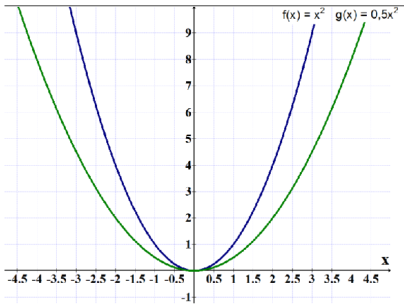
Si $a > 1$ :

$f(x) = x^2$ et $g(x) = 2x^2$



Si $0 < a < 1$ :

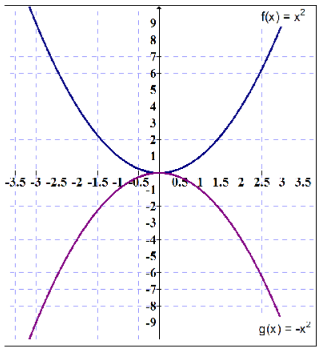
$f(x) = x^2$ et $g(x) = 0.5x^2$




Si $a < 0$ :

Il y a une réflexion de la courbe par rapport à l'axe des « $x$ ».

$f(x) = x^2$ et $g(x) = -x^2$

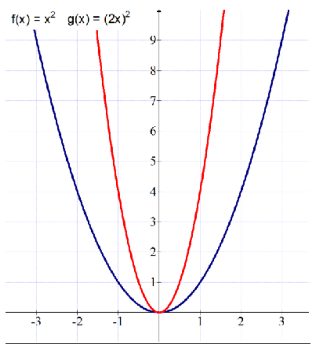



– Le paramètre « $b$ » : $f(x) = x^2 \longrightarrow g(x) = (bx)^2$.

Ce paramètre entraîne un changement d'échelle horizontal en divisant par « $b$ » les abscisses des couples de la fonction de base.

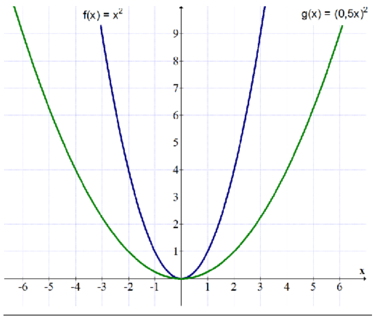
Si $b > 1$ :

$f(x) = x^2$ et $g(x) = (2x)^2$



Si $0 < b < 1$ :

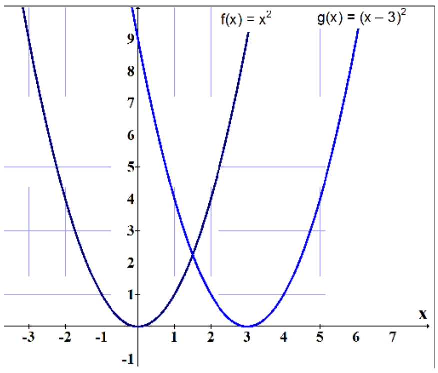
– $f\left(x\right) = x^2$   et  $g\left(x\right) = \left(0.5x\right)^2$




– Le paramètre « $h$ » :  $f(x) = x^2 \longrightarrow g(x) = (x - h)^2$.

Ce paramètre entraîne une translation horizontale de « $h$ » unités vers la droite si « $h$ » est positif ou « $h$ » unités vers la gauche si « $h$ » est négatif.

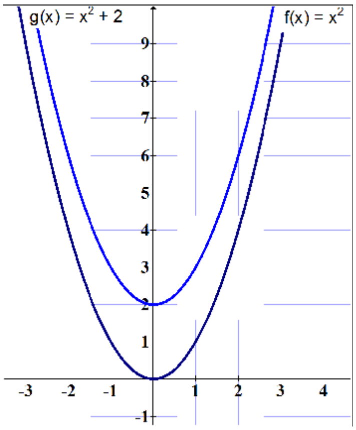
$f\left(x\right) = x^2$   et  $g\left(x\right) = \left(x - 3\right)^2$




– Le paramètre « $k$ » :  $f(x) = x^2 \longrightarrow g(x) = x^2 + k$.

Ce paramètre entraîne une translation verticale de « $k$ » unités vers le haut si « $k$ » est positif ou « $k$ » unités vers le bas si « $k$ » est négatif.

$f(x) = x^2 \quad$ et $\quad g(x) = x^2 + 2$




En mettant ensemble les trois paramètres on obtient la forme canonique :
$f(x) = a(x - h)^2 + k$.
– Le paramètre « $b$ » est compris dans la valeur de « $a$ » :

$$\begin{aligned} f(x) &= a(b(x-h))^2 + k \\ &= a(b^2(x-h)^2) + k \\ &= ab^2(x-h)^2 + k \\ &= a(x-h)^2 + k. \end{aligned}$$

**Exemple 1.4.1**
$f(x) = 2(3x - 6)^2 + 1$.
$f(x) = 2(3(x - 2))^2 + 1$.
$f(x) = 2 * 3^2(x - 2)^2 + 1 \longrightarrow f(x) = 18(x - 2)^2 + 1$.
*En transformant algébriquement la forme canonique, on arrive à la deuxième forme de la fonction quadratique :*
$f(x) = ax^2 + bx + c$.

**Exemple 1.4.2**

$$\begin{aligned} f(x) &= 2(x-1)^2 + 3 \\ &= 2(x^2 - 2x + 1) + 3 \\ &= 2x^2 - 4x + 2 + 3 \\ &= 2x^2 - 4x + 5. \end{aligned}$$

*De façon générale :*

$$\begin{aligned} f(x) &= a(x-h)^2+k \\ &= a(x^2-2xh+h^2)+k \\ &= ax^2-2axh+ah^2+k. \end{aligned}$$

*On a que :* $b=-2ah \rightarrow h=\frac{-b}{2a}$.
*On a aussi :* $c=ah^2+k$.
$k=c-ah^2 \rightarrow k=c-a\left(\frac{-b}{2a}\right)^2 \rightarrow k=c-a\left(\frac{b^2}{4a^2}\right) \rightarrow k=c-\frac{b^2}{4a} \rightarrow k=\frac{4ac-b^2}{4a}$.
*Ainsi, à partir de la forme générale, on peut trouver les coordonnées* $(h,k)$ *du sommet de la parabole.*

$$\textbf{\textit{Sommet}}\ (h,k)=\left(\frac{-b}{2a},\frac{4ac-b^2}{4a}\right).$$

*On peut aussi passer de la forme générale à la forme canonique.*
$h(x)=2x^2-2x+4$.
$h=\frac{-b}{2a}=\frac{-(-2)}{4}=\frac{1}{2}$ . *et* $k=\frac{4(2)(4)-(-2)^2}{4(2)}=\frac{28}{8}=\frac{7}{2}$.
*Forme canonique :*

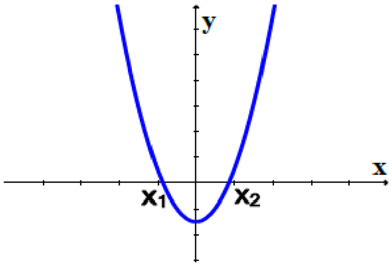
$$h\left(x\right)=2\left(x-\frac{1}{2}\right)^2+\frac{7}{2}.$$

**Zéros de la fonction quadratique :**

À partir de la forme canonique : $f(x)=a(x-h)^2+k$.
$f(x)=a(x-h)^2+k$.
$0=a(x-h)^2+k$.
$-k=a(x-h)^2$.
$\frac{-k}{a}=(x-h)^2 \rightarrow \sqrt{\frac{-k}{a}}=x-h$ ou $-\sqrt{\frac{-k}{a}}=x-h$.
Donc $x_1=\sqrt{\frac{-k}{a}}+h$ et $x_2=-\sqrt{\frac{-k}{a}}+h$.
1. Si $\frac{-k}{a}>0$, on a 2 zéros réels.
2. Si $\frac{-k}{a}=0$, on a 1 zéro réel.
3. Si $\frac{-k}{a}<0$, on a aucun zéro réel.
À partir de la forme générale : $f(x)=ax^2+bx+c$.
$x_1=\frac{-b-\sqrt{b^2-4ac}}{2a}$ et $x_2=\frac{-b+\sqrt{b^2-4ac}}{2a}$.
4. Si $b^2-4ac>0$, on a 2 zéros réels.
5. Si $b^2-4ac=0$, on a 1 zéro réel.
6. Si $b^2-4ac<0$, on a aucun zéro réel.

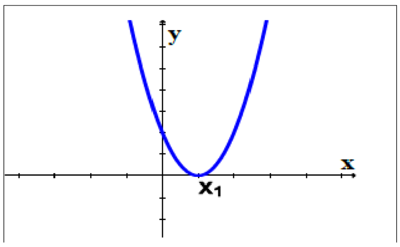
**Signes et variations de la fonction quadratique :**

- $a > 0$ et 2 zéros :



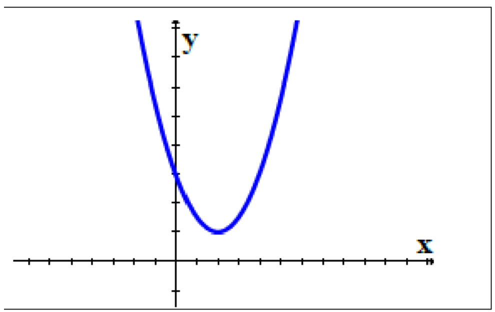

Positive sur $]-\infty, x_1]U[x_2, +\infty[$.
Négative sur $[x_1, x_2]$.
Décroissante sur $]-\infty, h]$.
Croissante sur $[h, +\infty[$.

- $a > 0$ et 1 zéro :



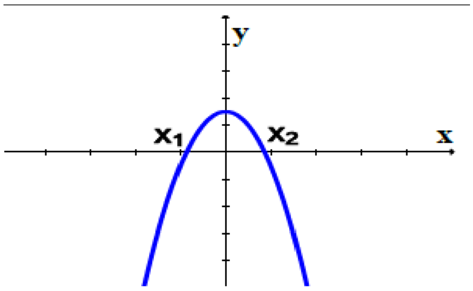

Positive sur $\mathbb{R}$.
Décroissante sur $]-\infty, h]$.
Croissante sur $[h, +\infty[$.

- $a > 0$ et aucun zéro :

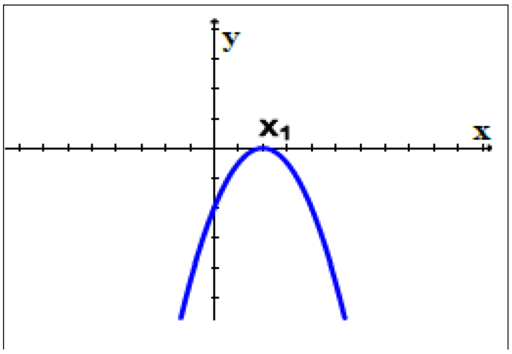



Positive sur $\mathbb{R}$.
Décroissante sur $]-\infty, h]$.
Croissante sur $[h, +\infty[$.

– $a < 0$ et 2 zéros :

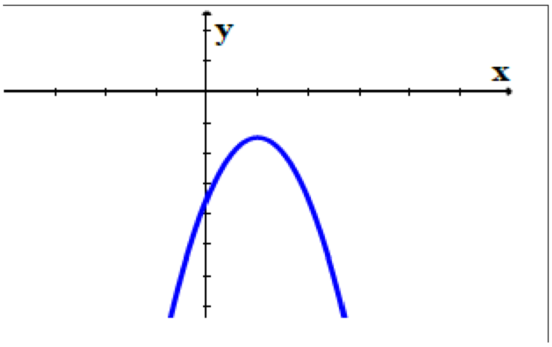



Positive sur $[x_1, x_2]$.
Négative sur $]-\infty, x_1] \, U \, [x_2, +\infty[$ .
Décroissante sur $[h, +\infty[$ .
Croissante sur $]-\infty, h]$ .
– $a < 0$ et 1 zéro :




Négative sur $\mathbb{R}$.
Décroissante sur $[h, +\infty[$.
Croissante sur $] - \infty, h]$.

– $a < 0$ et aucun zéro :




Négative sur $\mathbb{R}$.
Décroissante sur $[h, +\infty[$.
Croissante sur $]-\infty, h]$.

## 1.5 Forme quadratique à une matrice symétrique par rapport à une base :

**Proposition 1.5.1**
*Si* $M(a) = a_{ij}$ *est une matrice symétrique* $(t_M = M)$ *alors :*
$\left(x = \sum_{i=1}^{n} x_i e_i\right)$.
$q(x) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1<i\leq j\leq n} a_{ij} x_i x_j$.

**Exemple 1.5.2**
$M = \begin{pmatrix} 3 & 1 & 4 \\ 1 & -1 & 2 \\ 4 & 2 & 5 \end{pmatrix}$.
*La forme quadratique associée à* $M$ *est :*
$q : \mathbb{R}^3 \to \mathbb{R}$.
$(x_1, x_2, x_3) \to q(x_1, x_2, x_3)$.
$q(x_1, x_2, x_3) = \sum_{i=1}^{3} a_{ii} x_i^2 + 2 \sum_{1\leq i<j\leq 3} a_{ij} x_i x_j = 3x_1^2 - x_2^2 + 5x_3^2 + 2(x_1x_2 + 4x_1x_3 + 2x_2x_3)$.
$= 3x_1^2 - x_2^2 + 5x_3^2 + 2x_1x_2 + 8x_1x_3 + 4x_2x_3$.

# Chapitre 2

# Optimisation sans contrainte et avec contrainte :

L'optimisation est une branche de mathématique cherchant à modéliser, à analyser et à résourdre des problèmes de meilleur élément (décision) d'un ensemble afin de maximiser ou de minimiser un critère qualitatif qui mesure la qualité de la décision. Le mot optimisation vient du latin "optimum" qui signifie "le meilleur".

## 2.1 Problème d'optimisation :

**Définition 2.1.1 (problème d'optimisation)**
*C'est la modélisation d'un problème. Un modèle est une constuction mathématique utilisée pour représenter certains aspects significatifs du problème réel. Il y a deux types de modèle par l'optimisation qui s'écrit sous la forme générale suivante :*

$$(p)\begin{cases} min\ f(x) \\ x \end{cases}.$$

*telle que $f$ est une fonction à plusieurs variables, à valeurs réelles et elle est appelée la fonction objectif ou la fonction cout du problème $(p)$.*
*1) Si $\alpha = \min_{x\in\mathbb{R}_{x\in\mathbb{R}^n}} f(x)$ s'appelle modélisation sans contrainte.*
*2)Si $\alpha = \min_{x\in C} f(x)$ s'appelle modélisation par optimisation avec contrainte ou $C$ est un sous-ensemble de $\mathbb{R}^n$.*

## 2.2 Solutions optimales et réalisables :

### 2.2.1 Solution réalisable :

**Définition 2.2.1**
*Un point $x^* \in C$ (c'est-à-dire vérifiant les contraintes de $(p)$) est appellé solution réalisable de $(p)$.*

### 2.2.2 Solution optimale global :

**Définition 2.2.2**
*Une solution réalisable qui minimise $f$ sur $C$ est appellée solution optimale global de $(p)$, nous la noterons $x^*$. L'ensemble des solutions globales est noté :* $\arg\min_C f(x)$.

### 2.2.3 Solution optimale local :

**Définition 2.2.3**
*Un point $x^* \in C$ est une solution optimale local de $(p)$ s'il existe un voisinage $V$ de $x^*$ tel que :*

$$f(x^*) \leq f(x) ; \forall x \in V.$$

*On note par :*
$\log\min_C f(x)$ *(l'ensemble des solutions optimales local de $(p)$ ).*

## 2.3 Optimisation sans contrainte :

### 2.3.1 Présentation du problème :

La forme d'un problème d'optimisation sans contrainte :

$$(p) \begin{cases} \min f(x) \\ x \in \mathbb{R}^n \end{cases} .$$

Soit $f \in C^0(E, R)$ et $E$ un espace vectoriel normé. On cherche soit un minimum global de $f$, c'est-à-dire : $x^* \in E$ tel que :

$$f(x^*) \leq f(x), \ \forall x \in E. \quad (1)$$

Ou un minimum local, c'est-à-dire : $x^* \in E$ tel que : $\exists\ \alpha > 0$.

$$f(x^*) \leq f(x)\,,\ \forall x \in B(x^*, \alpha)\,. \tag{2}$$

tel que : $B(x^*, \alpha) = \{x \in E,\ \ \| x - x^*\| < \alpha\}\,.$

### 2.3.2 Existence et unicité :

**Existence :**

**Théorème 2.3.1**
*Le problème (p) admet une solution optimale $x^* \in X$, si $f : X \subset \mathbb{R}^n \to \mathbb{R}$ est continue et si de plus $X$ est un ensemble compact (fermé bornée) qui vérifie :*

$$f(x^*) \leq f(x)\,, \forall x \in X.$$

**Théorème 2.3.2**
*Soit $f : \mathbb{R}^* \to \mathbb{R}$ une fonction tell que :*
*(i) $f$ est continue.*
*(ii) $f(x) \to +\infty$ quand $\|x\| \to +\infty$.*
*Alors il existe $x^* \in \mathbb{R}^n$ tel que $f(x^*) \leq f(x)\,,\ \forall x \in \mathbb{R}^n$.*

**Unicité :**

**Théorème 2.3.3**
*Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction strictement convexe, alors le problème admet au plus un $x^* \in \mathbb{R}^n$ tel que :*

$$f(x^*) \leq f(x)\,,\ \forall x \in \mathbb{R}^n.$$

### 2.3.3 Conditions d'optimalités :

Dans le but d'analyser ou de résoudre de manière efficace un problème d'optimisation, il est très important de pouvoir disposer des conditions d'optimalité.

**Conditions nécessaires d'optimalité :**

**Théorème 2.3.4 (condition nécessaire du premier ordre)**
*Soient $f : R^n \to R$ continuellement différentiable $C^1$ et $x^* \in \mathbb{R}^n$. Si $x^*$ est un minimum local de $f$. Alors $\nabla f(x^*) = 0$.*

**Remarque 2.3.5**
*Si $f$ est convexe, la condition nécessaire du premier ordre est également suffisante pour que $x^*$ soit un minimum global.*
*Dans le cas ou $f$ deux fois continuellement différentiable $C^2$, une autre condition nécessaire est donnée par la théorème, elle appellée condition nécessaire du second ordre car elle fait intervenir la matrice hessienne de $f$ $\left(\nabla^2 f(x^*) = \frac{\partial^2 f}{\partial x_i \partial x_j}\right)$.*

**Définition 2.3.6 (condition nécessaire du second ordre)**
*Soient $f : \mathbb{R}^n \to \mathbb{R}$ deux fois continuellement différentiable et $x^* \in \mathbb{R}^n$. Si $x^*$ est un minimum local de $f$ alors :*
*$\nabla f(x^*) = 0$ et $\nabla^2 f(x^*)$ est semi-définie positive.*

$$\nabla^2 f(x^*) \geq 0 \Rightarrow \forall x \in \mathbb{R}^n, x^T \nabla^2 f(x^*) \ x \geq 0.$$

**Exemple 2.3.7**

$$\begin{cases} \min f(x) = 1 - \exp^{-x^2} \\ \qquad x \in \mathbb{R}^n \end{cases}.$$

$\nabla f(x) = f'(x) = 2x \exp^{-x^2}$.
$f'(x) = 0 \Rightarrow x = 0$.
$\nabla^2 f(x) = f''(x) = 2\exp^{-x^2} - 4x^2 \exp^{-x^2} = \exp^{-x^2}(2 - 4x^2)$.
$\nabla^2 f(0) = 2 > 0 \Rightarrow x^* = 0$ *est un point* min *local.*

**Exemple 2.3.8**

$$\begin{cases} \min f(x_1, x_2) = x_1^2 + x_2^2 - 2x_1x_2 + x_1 \\ \qquad x_1, \ x_2 \in \mathbb{R}^n. \end{cases}.$$

$\nabla f(x) = \begin{pmatrix} 2x_1 - 2x_2 + 1 \\ 2x_2 - 2x_1 \end{pmatrix}$.

$\nabla f(x) = 0 \Rightarrow \begin{cases} 2x_2 - 2x_1 = 1 \\ 2x_2 - 2x_1 = 0 \end{cases}$.

*Donc le problème n'admet pas de point min local.*

## 2.4 Optimisation avec contrainte :

### 2.4.1 Présentation du problème :

Soit $E = \mathbb{R}^n$, soit $f \in C^0(E, \mathbb{R})$, et soit $K$ un sous ensemble de $E$. On s'intéresse à la recherche de $x^* \in K$ tel que :

$$\begin{cases} f(x^*) = \min_K f \\ x^* \in K \end{cases} . \qquad (3) .$$

Ce problème est un problème de minimisation avec contrainte au sens où l'on cherche $x$ qui minimise $f$ en astreignant $x$ a être dans $K$.

Ces contraints sont données à l'aide de fonction au moins continues plusieurs formes :

**Contraintes égalités :**

On pose :

$$K = \{x \in E,\ g_i(x) = 0,\ i = 1, ...., p\} .$$

On verra plus loin que le problème de minimisation de $f$ peut alors être résolu grace au théorème des multiplicateurs de lagrange.

**Contraintes inégalités :**

On pose :

$$K = \{x \in E,\ g_i(x) \leq 0,\ i = 1, ......, p\} .$$

On verra plus loin que le problème de minimisation de $f$ peut alors être résolu grace au théorème de Kuhn-Tucker.

**Contraintes d'égalités et d'inégalités :**

L'ensemble des contraintes est donnée par :

$$K = \{x \in E,\ g_i(x) = 0,\ \forall i = 1, ....., p \text{ et } g_i(x) \leq 0,\ \forall i = 1, ....., p\} .$$

### 2.4.2 Existence et unicité de la solution optimale :

**Théorème 2.4.1 (Existence )**
*Soit* $E = \mathbb{R}^n$, *et* $f \in C^0(E, \mathbb{R})$ .

*(i) Si $K$ est un sous-ensemble fermé et borné de $E$, alors il existe $x^* \in K$ tel que :*

$$f(x^*) = \min_K f.$$

*(ii) Si $K$ est un sous-ensemble fermé de $E$, et si $f$ est croissante à l'infini c'est-à-dire que : $f(x) \to +\infty$ quand $\|x\| \to +\infty$, alors $\exists\ x^* \in K$ tel que :*

$$f(x^*) = \min_K f.$$

**Théorème 2.4.2 (Unicité)**
*Soit $E = \mathbb{R}^n$ et $f \in C^0(E, \mathbb{R})$. On suppose que $f$ est strictement convexe et que $K$ est convexe. Alors il existe au plus un élément $x^*$ de $K$ tel que :*

$$f(x^*) = \min_K f.$$

**Théorème 2.4.3 (Existence et unicité )**
*Soit $E = \mathbb{R}^n$, $f \in C^0(E, \mathbb{R}^n)$ une fonction stricrement convexe et $K$ un sous-ensemble convexe fermé de $E$. Si $K$ est borné ou si $f$ est croissante à l'infini, c'est-à-dire si $f(x) \to +\infty$ quand $\|x\| \to +\infty$, alors il existe un unique élément $x^*$ de $K$ solution du problème de minimisation (3), tel que :*

$$f(x^*) = \min_K f.$$

### 2.4.3 Conditions d'optimalités :

**Condition simple d'optimalité :**

**Proposition 2.4.4**
*Soient $E = \mathbb{R}^n$, $f \in C^0(E, \mathbb{R})$, et $x^* \in K$ tel que :*

$$f(x^*) = \min_K f.$$

**Théorème 2.4.5**
*On suppose que $f$ est différentiable en $x^*$ :*
*(i) Si $x^* \in K^0$ alors $\nabla f(x^*) = 0$ ($K^0$ l'intérieur de l'ensemble $K$).*
*(ii) Si $K$ est convexe, alors $\nabla f(x^*)(x - x^*) \geq 0$ pour tout $x \in K$.*

**Contraintes égalités :**

Dans tout paragraphe, on considérera les hypothèses et notation suivantes :
$f \in C^0(\mathbb{R}^n, \mathbb{R})$, $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, $i = 1......p$.
$K = \{x \in \mathbb{R}^n,\ g_i(x) = 0,\ \forall i = 1.......p\}$.
$g = (g_1, ....., g_p)^t \in C^1(\mathbb{R}^n, \mathbb{R}^p)$.

**Remarque 2.4.6**
*Comme* $g \in C^1(\mathbb{R}^n, \mathbb{R}^p)$, *si* $x \in \mathbb{R}^n$, *alors* $Dg(x) \in L(\mathbb{R}^n, \mathbb{R}^p)$, *ce qui revient à dire. En confondant l'application linéaire* $Dg(x)$ *avec sa matrice, que* $Dg(x) \in M_{p,n}(\mathbb{R})$. *Par définition,* $\mathrm{Im}(Dg(x)) = \{Dg(x)z,\ z \in \mathbb{R}^n\} \subset \mathbb{R}^p$.

**Théorème 2.4.7 (Multiplicateurs de lagrange)**
*Soit* $x^* \in K$ *tel que* $f(x^*) = \min_K f$. *On suppose que* $f$ *est différentiable en* $x$ *et* $\dim(\mathrm{Im}(Dg(x^*))) = p$ *(ou rang* $(Dg(x^*)) = p$*), alors : il existe* $(\lambda_{1,.....,\lambda p})^t \in \mathbb{R}^p$ *tel que :*

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

**Contraintes inégalités :**

Soit $f \in C^0(\mathbb{R}^n, \mathbb{R})$ et $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, $i = 1.....p$. On considère maintenant un ensemble $K$ de la forme :

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

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

$$\begin{cases} \dfrac{x^* \in K}{f(x^*) \leq f(x),\ \forall x \in K} \end{cases}.$$

**Théorème 2.4.8 (Kuhn-Tucker)**
*Soit* $f \in C^0(\mathbb{R}^n, \mathbb{R})$, *soit* $g_i \in C^1(\mathbb{R}^n, \mathbb{R})$, *pour* $i = 1....p$, *et soit :* $K = \{x \in \mathbb{R}^n,\ g_i(x) \leq 0,\ \forall i = 1, ......$
*On suppose qu'il existe* $x^*$ *solution de* $(3)$, *et on pose* $I(x^*) = \{i \in \{1.....p\}\ /\ \ g_i(x^*) = 0\}$.
*On suppose que* $f$ *est diffétentaible en* $x^*$ *et que la famille de* $\mathbb{R}^n$ $\{\nabla g_i(x^*),\ i \in I(x^*)\}$ *est libre. Alors il existe une famille* $(\lambda_i)_{i \in I(x^*)} \subset \mathbb{R}_+$.

$$\nabla f(x^*) + \Sigma_{i \in I(x^*)} \lambda_i \nabla g_i(x^*) = 0.$$

# Chapitre 3

# Programmation quadratique :

## 3.1 Introduction :

Le problème de programmation quadratique se résolvent à l'aide de méthodes du gradient avec projection et les méthodes de Newton.

## 3.2 Présentation d'un problème de Programmation quadratique :

Un problème de programmation quadratique est un problème qui peut se mettre sous la forme suivante :

$$\left\{ \begin{array}{c} \min_{x\in\mathbb{R}^{n,m}} \sum_{i=1}^{n} \left( \sum_{j=1}^{n} a_{ij}x_ix_j + b_ix_i \right) \\ i \in [1,n] \sum_{i=1}^{n} c_{i,j}x_i \leq 0 \ \forall j \in 1.m \end{array} \right. \text{..........} (1) .$$

Ou encore, en utilisant le formalisme matriciel :

$$\left\{ \begin{array}{c} \min X^T AX + BX \\ C_jX \leq 0 \ \forall j \in 1..m \end{array} \right. .$$

## 3.3 La méthode de gradient et la méthode de Newton :

### 3.3.1 La méthode du gradient avec projection :

**La méthode de descente :**

**Définition 3.3.1**

*Soient* $f \in C(E, R)$ *et* $E = R^n$.

*1. Soit* $x \in E$*, on dit que* $d \in E \,/\, \{0\}$ *est une direction de descente en* $x$ *s'il existe* $\rho_0 > 0$ *tel que :*

$$f(x + \rho d) \leq f(x), \ \forall \rho \in [0, \rho_0].$$

*2. Soit* $x \in E$*, on dit que* $d \in E \,/\, \{0\}$ *est une direction de descente stricte en* $x$ *s'il existe* $\rho_0 > 0$ *tel que :*

$$f(x + \rho d) \leq f(x), \ \forall \rho \in ]0, \rho_0].$$

*3. Une "méthode de descente" pour la recherche de* $x^*$ *tel que :* $f(x) = \inf_E f$ *consiste à construire une suite* $(x_n)_n$ *de la manière suivante :*

*a) Inisialisation* $x_0 \in E$;

*b) Itérationn : on suppose* $x_0......x_n$ *connus*$(n \geq 0)$;

*i) On cherche* $d_n$ *direction de descente stricte de* $x_n$.

*ii) On prend* $x_{n+1} = x_n + d_n$ *avec "bien choisi".*

*Initialisation de* $(x_n \rightarrow x_0 \in E)$ .

*Tantque (critère d'arrêt) faire*

*Choix d'une direction (de descente)* $d_n$*,*

*Choix d'un pas de descente* $\rho_n > 0$ *telque :* $x_n + \rho_n d_n \in E$

$x_{n+1} \longleftarrow x_n + \rho_n d_n$

*Fin*

*Retourner* $x_n$......................................................................*(Algorithme I).*

**Proposition 3.3.2**

*Soient* $E = R^n$, $f \in C^1(E, R)$, $x \in E$ *et* $d \in E / \{0\}$ , *alors :*

*1) Si* $d$ *direction de descente en* $x$ *alors* $d\ \nabla f(x) \leq 0$.

*2) Si* $\nabla f(x) \neq 0$ *alors* $d = -\nabla f(x)$ *est une direction de descente stricte en* $x$.

**Projecton sur un convexe fermé :**

**Proposition 3.3.3**
*Soit $E$ un espace de Hilbert, muni d'une norme $\|.\|$ induite par un produit scalair $(.,.)$ et soit $K$ un convexe fermé non vide de $E$. Alors, tout $x \in K$ il existe un unique $x_0 \in K$ tel que :*
$\|x - x_0\| \leq \|x - y\|$ *pour tout* $y \in K$,
*On not* $x_0 = P_k(x)$ *: La projection orthogonale de* $x$ *sur* $K$*. On a également :*
$x_0 = p_k(x)$ *si et seulement si* $(x-\ x_0,\ x_0 - y) \geq 0,\ \forall y \in K$.
*Dans le cadre des algorithmes de minimisation avec contraintes que nous allons développer maintenant, nous considèrons* $E = R^n$, $f \in C^1(R^n, R)$ *une fonction convexe, et* $K$ *fermé convexe non vide. On cherche à calculer une solution approchée de* $x$*, solution du problème* (3) .

**Algorithme du gradient avec projection :**

On considère le problème :

$$\begin{cases} minf(x) \\ x \in K \end{cases} \quad .............. (p)\,.$$

avec : $\phi \neq K \subset R^n$convexe fermé et $f \in C^1$ sur $K$. Soit $\rho > 0$ fixé. On a la condition : $x^* = proj_k(x^* - \rho\nabla f(x^*))$.

est une condition nécessaire pour que $x^*$ soit une solution optimale de $(p)$ elle est suffisante.

si $f$ est convexe. (Si la condition est vérifiée, on dit que $x^*$ est un point critique).

**Algorithme du gradient à pas fixe avec projection sur K (GPFK) :** La méthode est comme suit :

1. Initialisation : on part de $x^* \in C$, on fait $k = 0$.
2. On calcule $y_k = proj_k(x_k - \rho\nabla f(x_k))$.
3. Test d'arrêt : $y_k = x_k$.
4. Sinon on prend pour direction de descente $d_k = y_k - x_k$.
5. On fait une recherche linéaire à partire de $x_k$ dans la direction $d_k$. On obtient $t_k$.
6. On fait $x_{k+1} = x_k + \rho t_k d_k$, $k = k + 1$ et on retourne en 2.

Si $d_k = 0$ alors $x_k$ est un point critique et solution optimale dans le cas où $f$ est convexe.

Vérifions dans le cas contraire que $d_k$ est une direction de descente. Par définition de la projection on a :

$$< y_k - x_k + \alpha \nabla f\,(x)\,,\; x - y_k > \;\; \geq 0, \forall x \in k.$$

Prenons $x = x_k$, alors :

$$0 > -\|d_k\|^2 \geq \alpha\,(f\,(x_k)\,, d_k)\,.$$

D'autre part la convexité de $K$ implique :

$$x_k + t(y_k - x_k) \in K \text{ pour tout } t \in ]0, 1[.$$

**Théorème 3.3.4 (convergence de l'agorithme GPFK)**
*Soit $f \in C^1\;(R^n, R)$, et $K$ convexe fermé non vide. On suppose que :*
*1) Il existe $\alpha > 0$, tel que : $< \nabla f(x) - \nabla f(y),\; x - y > \; \geq \alpha \mid x - y \mid^2$ pour tout $(x, y) \in \mathbb{R}^n * \mathbb{R}^n$.*
*2) Il existe $M > 0$ tel que :$\mid \nabla f(x) - \nabla f(y) \mid \leq M \mid x - y \mid$ pour tout $(x, y) \in R^n * R^n$.*
*alors :*
*a) Il existe un unique élément $x \in K$ solution du probléme (3).*
*b) Si $0 < \rho < \frac{2\alpha}{M^2}$, la suite $(x_n)$ définie par l'algorithme $(PGFK)$ converge vers $x^*$ lorsque $n \to +\infty$.*

**Algorithme du gradient à pas optimal avec projection sur K (GPOK) :** L'algoritme du gradient à pas optimal avec projection sur $K$ s'écrit :

Initialisation : $x_0 \in K$.

Itération : $x_n$ connu.

$d_n = -\nabla f(x_n)$,calculer $\rho_n$ optimal dans la direction $d_n$ tel que : variable à chaque itération : $x_{n+1} = P_k(x_n + \rho_n d_n)$.

La démonstration de convergence de cet algorithme se déduit de celle de l'algorithme à pas fixe.

### 3.3.2 Les méthodes de Newton :

**La méthode de Newton à pas optimal et fixe :**

La méthode de Newton correspond à une méthode de descente pour laquelle le choix de la direction est donnée par le gradient .

$$d = -\nabla_{\nabla^2 f} f(x) = -\left(\nabla^2 f\right)^{-1} \nabla f\left(x\right).$$

Ce choix, permis lorsque la Hessienne est inversible, définit un algorithme robuste (Un algorithme $I$ est dit robuste si la direction $d$ calculée à chaque itération est une direction de descente). Lorsque $\nabla^2 f$ est une matrice symétrique définie positive.

Ce choix de pas utilise l'information de second ordre donnée par la Hessienne : en utilisant un développement limité de $f$ au second ordre, on obtient :

$$f(x+v) \approx f(x) + < \nabla f(x), v > + \frac{1}{2} < v, \nabla^2 f\left(x\right) v > .$$

Une idée naturelle est de choisir comme direction de descente celle qui minimise cette approximation.

En différenciant par rapport à $v$, on obtient :

$$\nabla f(x) + \nabla^2 f(x) v = 0.$$

et donc :

$$v = -\left(\nabla^2 f\right)^{-1} \nabla f\left(x\right) \ldots\ldots\ldots\ldots. (a).$$

En fait, cette quantité contient non seulement la direction de descente optimale mais aussi le pas optimal à choisir pour minimiser le développement limité au second ordre.

**La méthode de Newton pour la résolution numérique d'équations :**

La méthode de Newton pour résoudre des équations du type $F(x) = 0$ s'applique donc au cas particulier où $F = \nabla f$.

La méthode de Newton pour la résolution d'équations utilise une information à l'ordre 1 sur $F$. En supposant par exemple que $F$ est $C^1(R^n, R^n)$, on utilise le développement de Taylor :

$$F(x+v) = F(x) + DF(x)(v) + o(\|v\|).$$

L'idée est de résoudre l'équation "approximante" $F(x) + DF(x)(v) = 0$ qui a une unique solution si $DF(x)$ est inversible donnée par $v = -DF(x)^{-1}(F(x))$.

Évidemment, on retrouve bien la formule $(a)$ lorsqu'on remplace $F$ par $\nabla f$ et $DF$ par $\nabla^2 f$. On peut donc proposer l'algorithme suivant :

Initialisation de $x \leftarrow x_0 \in R^n$.

/ est le seuil de tolérance sur la valeur de la fonction/.

$Tantque(\|F(x)\| \geq \eta)\ faire$

$x \leftarrow x - DF(x)^{-1}(F(x))$.

$Fin$

$Retourner\ x$.............. (algorithme 1) .

**Convergence locale quadratique de la méthode de Newton :**

**Définition 3.3.5**
*Une fonction* $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. *Est dite Lipschitizienne sur* $\mathbb{R}^n$ *s'il existe* $K \in \mathbb{R}_+$ *tq :*
$\forall x,\ y \in \mathbb{R}^n$

$$\|f(x) - f(y)\| \leq K\|x - y\|.$$

Soit $F : R^n \to R^n$ une application $C^1(R^n, R^n)$ telle que $DF$ soit localement Lipschitz.

On suppose qu'il existe $x^* \in R^n$ tel que $F(x^*) = 0$ et $dF(x^*)$ inversible. Alors, il existe

$\varepsilon > 0$ tel que l'algorithme 1 initialisé dans la boule $B(x^*, \varepsilon)$ converge et on a dans ce cas :

$$\|x_{k+1} - x^*\| \leq C\|x_k - x^*\|^2.$$

pour une certaine constante $C$.

**La méthodes de quasi-Newton :**

L'idée centrale des méthodes de quasi-Newton est de mettre à jour une approximation du Hessien de $f$ noté $B_k$ à chaque itération de la méthode de descente. Le modèle de second-ordre est donc :

$$f(y) \approx f(x_k) + < \nabla f(x_k), y - x_k > + \frac{1}{2} < y - x_k, B_k(y - x_k) > .$$

où $B_k$ est symétrique définie positive et mise à jour à chaque itération afin de vérifier la condition de quasi-Newton :

$$B_k(x_{k-1} - x_k) = \nabla f(x_{k-1}) - \nabla f(x_k) \text{........} (1.1) .$$

Cette condition est l'extension en dimension supérieure de la méthode de la sécante sauf que la matrice $B_k$ peut (doit) garder une certaine mémoire des mises à jour précédentes. On remarque qu'es deux conditions précédentes ($B_k$ symétrique définie positive et l'équation (1.1) dite équation de la sécante) ne peuvent être satisfaites simultanément que si :

$$< x_{k-1} - x_k, \nabla f(x_{k-1}) - \nabla f(x_k) > \ = \ < x_{k-1} - x_k, B_k(x_{k-1} - x_k) > \ > 0.$$

**La méthode de Newton globale :**

Le but de cette partie est de modifier la méthode de Newton pour qu'elle converge vers une solution quelle que soit la position initiale des points. En effet, on a vu dans la première partie que la méthode de Newton pouvait diverger dans certains cas. Aussi, devons nous essayer de corriger ce défaut.

Pour simplifier le problème, nous nous placerons dans le cas d'un problème sans contrainte d'inégalité. Le problème à résoudre est donc :

$$\begin{cases} \min f(x) \\ c(x) = 0 \\ X \in \mathbb{R}^n \end{cases} .$$

**Fonction de pénalisation :**

**Définition 3.3.6**

*La pénalisation est un concept qui permet de transformer un problème d'optimisation avec contrainte et sans contrainte. Le but est d'introduire une nouvelle fonction dont la valeur est pénalisée lorsque l'on ne respecte pas les contraintes.*

*On remplace le problème :*

$$\begin{cases} \min f(x) \\ x \in X \end{cases} .$$

*par le problème :*

$$\begin{cases} \min \theta_\delta(x) \\ x \in \mathbb{R}^n \end{cases} .$$

*avec* $\theta_\delta(x) = f(x) + \delta\Pi(x)$ *.*

$r$ *est appellé le facteur de pénalisation.*

*Il y a différents choix possibles pour la fonction* $\Pi$*. C'est elle qui force à respecter les contraintes d'égalités, nous allons pour notre part choisir la pénalisation de Han.*

*Dans ce cas, on pose :* $\Pi(x) = \|c(x)^*\|_2$ *on a alors :*

$$\theta'_\delta(x_k, d_k) = -d_k^t L_k d_k + (\lambda_k^{pq})^t c_k - \delta\|c_k\|_2 ....... (*)$$

*$d_k$ est une direction de descente si $\theta'_\delta(x_k, d_k) < 0$, c'est à dire qu'il faut :*
*$d_k^t L_k d_k \geq 0$ et $\delta > \|\lambda_k^{pq}\|_2$.*
*Soit $\delta_k$ la valeur de $\delta$ à l'itération $k$, le but est de modifier l'ancien $\delta$ de manière à avoir :*

$$\delta_k > \|\lambda_k^{pq}\|_2 + \bar{\delta}.$$

*avec $\bar{\delta}$ une petite constante pour cela on procédera de la manière suivante :*

*-lors de la première itération, on prend*

$$\bar{\delta} = \max\left(\sqrt{epsmachine}, \|\lambda_1\|_D \ / \ 100\right) \ \textit{et} \ \delta_1 = \|\lambda_1\|_D + \bar{\delta}.$$

*-lors des itérations suivantes, il faut réaliser la condition $(*)$ sans toute fois avoir un $\delta_k$ trop grand.*
*Si*

$$\delta_{k-1} < \|\lambda_k^{pq}\|_D + \bar{\delta}, \delta_k = \max\left(1.56_{k-1}\|\lambda_k^{pq}\|_D + \bar{\delta}\right).$$

*Sinon si*

$$\delta_{k-1} > 1.1\left(\|\lambda_k^{pq}\|_D + \bar{\delta}\right), \delta_k = \left(\delta_{k-1} + \|\lambda_k^{pq}\|_D + \bar{\delta}\right)/2.$$

*Sinon*

$$\delta_k = \delta_{k-1}.$$

**Définition 3.3.7**
*-On calcule une direction de descente $d_k$ vérifiant les conditions d'optimalité du premier ordre.*
*-On calcule la valeur de $\delta_k$.*
*-On calcule $\alpha_k$ de manière à faire décroitre suffisamment la fonction de pénalisation de Han.*
*-On prend $x_{k-1} = x_k + \alpha_k d_k$ et on recommence une nouvelle itération jusqu'à vérifier une des conditions d'arrêt.*

**Convergence de la méthode de Newton à pas optimal :**

Soit $f \in C^2(\Omega, \mathbb{R})$ et $x_0 \in \Omega$ tels que :

$$0 < CId \leq \nabla^2 f(x) \leq KId.$$

pour $x \in S_0 = \{x \in \Omega \ / \ f(x) \leq f(x_0)\}$.

Que l'on suppose fermé dans $R^n$ alors on a :
- la suite des valeurs $f(x_k)$ est strictement décroissante ou la suite $x_k$ est stationnaire.
- si $inf\ f > -\infty$ alors la suite $\nabla f(x_k)$ converge vers 0, $S_0$ contient au moins un minimiseur local de $f$ et $lim_{k \to +\infty}\ d(x_k, LM) = 0$ où $LM$ est l'ensemble des minimiseurs locaux de $f$.
- Si $S_0$ est convexe alors $x_k$ converge vers l'unique minimiseur de $f$ sur $\Omega$.

**La méthode de Newton modifiée :**

**Théorème 3.3.8**
*Soit $\alpha$, $\beta$ deux réels tels que : $0 < \alpha < \beta < 1$.et $f$ une fonction de classe $C^1(\mathbb{R})$. On suppose que $f'(0) < 0$ et qu'il existe $t_0 > 0$ telque :*

$$f(t_0) \geq f(0) + \alpha t_0 f'(0).$$

*alors il existe un intervalle non vide $I \subseteq ]0, t_0[$ telque tous $t \in I$ définissant la règle de Wolfe.*

**Exemple 3.3.9 (d'algorithme pour satisfaire la règle de Wolfe)**
*Initialisation de $\varepsilon = 1, a = 0, b = \infty$.*
$Tantque$ *(Règle de Wolfe non satisfaite)* $faire$
$Si\ (f(x + \varepsilon d) > f(x) + \alpha\varepsilon < \nabla f(x), d >)\ Alors$
$b = \varepsilon$
$\varepsilon \longleftarrow \frac{a+b}{2}$
$Sinon$
$Si(< \nabla f(x + \varepsilon d), d > \beta < \nabla f(x), d >)Alors$
$a = \varepsilon$
$\varepsilon \longleftarrow \frac{a+b}{2}$ *si* $b < 1$ *et* $\varepsilon \longleftarrow 2a$ *sinon.*
$Sinon$
$Retourner\ \varepsilon$

$FinSi$
$FinSi$
$Fin.$

**Définition 3.3.10**
*Initialisation de* $x \longleftarrow x_0 \in \Omega$, $H \longleftarrow \nabla^2 f(x)$ *et* $v > 0$. (*$\eta$ est seuil de tolérance sur le gradient*).

$Tantque\ (\|\nabla f(x)\| \geq \eta)\ faire$

*Si* ($H$ *définie positive*) *Alors*

$Si\ \left(\frac{|<H^{-1}\nabla f(x),\ \nabla f(x)>|}{\|\nabla f(x)\|\|H^{-1}(\nabla f(x))\|} \geq V\right)$

$d \longleftarrow -H^{-1}\nabla f(x)$
$FinSi$
$Sinon$
$d \longleftarrow -\nabla f(x)$
$FinSi$
$\varepsilon =$ *choix du pas par la règle de Wolfe*
$x \longleftarrow x + \varepsilon d$
$H \longleftarrow \nabla^2 f(x)$
$Fin$
*Retourner* $x$ ..........................(*Algorithme* 2 : *un algorithme de Newton modifiée*).

## Conclusion Générale

En conclusion, on peut dire que la programmation quadratique est la méthode d'optimisation la plus performante pour résourdre les problèmes d'optimisation avec contraintes d'égalité et d'inégalité par des méthodes itératifs qui sont les méthodes de gradient avec projection et les méthodes de Newton.

# Bibliographie


[1] J-B. Hiriart-Urruty, optimisation et analyse convexe. PUF, (1998).

[2] M. AZI, cours optimisation, centre universitaire de Mila, 2014-2015.

[3] H. Yacine, cours d'algèbre, centre universitaire de Mila, 2013-2014.

[4] J-P Corriou, méthode numérique et optimisation théorie et pratique pour l'ingénieur, septembre 2010.

[5] J-P. Guillement, résumé d'optimisation, 2010-2012.

[6] Mechel Bielaire, introduction à l'optimisation défférentiable.

[7] Grégoire Allaire, analyse numérique et optimisation. France, 2008.

[8] J-C. Culioli, introduction à l'optimisation. France, 2012.

[9] Jean-Baptiste Delome, programmation quadratique et semidéfinie, Mod 2005.

[10] B-Rousselet, introduction à l'optimisation, avril 1998.

[11] Francois-Xavier. Vialard, optimisation numérique, mai 2012.

[12] Raphaèle-Herbin, cours d'analyse numérique, octobre 2008.

[13] Fabrice Rossi, optimisation sous contraintes, décembre 2009/janvier 2010.