# III- Optimisation non linéaire avec contraintes :

## I/- Généralités :

On s'interesse au problème de minimisation d'une fonction $f$ définie de $\mathbb{R}^n$ dans $\mathbb{R}$ comportant des contraintes.

Le problème est de la forme :

$$(P)\begin{cases} \min f(x) \\ x \in C,\ C \subset \mathbb{R}^n \\ C \neq \emptyset,\ \text{fermé} \end{cases}$$

Ou bien :

$$(P)\begin{cases} \min f(x) \\ \text{sous-contraintes:} \\ h_i(x) = 0,\ \forall\, i = \overline{1,p}. \\ g_j(x) \leq 0,\ \forall\, j = \overline{1,q}. \\ x \in \mathbb{R}. \end{cases}$$

. Ainsi $C = \left\{ x \in \mathbb{R}^n,\ h_i(x) = 0,\ \forall\, i = \overline{1,p} \text{ et } g_j(x) \leq 0,\ \forall\, j = \overline{1,q} \right\}$.

. Les fonctions $f$, $h_i$ $(i = \overline{1,p})$ et $g_j$ $(j = \overline{1,q})$ définis de $\mathbb{R}^n$ dans $\mathbb{R}$ sont supposées de classe $\mathcal{C}^1$.

. La fonction $f$ est appelée fonction objectif.

. $h_i$ $(i = \overline{1,p})$ et $g_j$ $(j = \overline{1,q})$ sont appelés contraintes de problème.

. $C$ domaine des solutions réalisables ou admissible.

**Définition 1-1 :** - Si pour $x \in C$ et pour $j \in \{1, \dots, q\}$, On a : $g_j(x) = 0$, On dit que la contrainte $g_j$ est saturée (ou active) au point $x$.

Une contrainte qui n'est pas active en $x$ est dite contrainte inactive ou non saturée au points $x$ (c-à-d $g_j(x) < 0$)

$I(x) = \left\{ j \in \{1, 2, \dots, q\},\ g_j(x) = 0 \right\}$ est appelé ensemble des indices des contraintes actives en $x$.

**Définition 1-2 :** - On dit qu'une direction "$d$" est admissible en $x^\circ \in C$ si :

. pour $i = \overline{1,p}$ $\quad d^t \cdot \nabla h_i(x^\circ) = \langle \nabla h_i(x^\circ), d \rangle = 0$.

• Pour $j = \overline{1,q}$ $\quad g_j(x^0) = 0$ $\quad$ alors : $d^t \nabla g_j(x^0) = \langle \nabla g_j(x^0), d \rangle \leqslant 0$.
suive une direction admissible à partir d'un $x^0$ de $C$ permet de rester dan $C$, On le quitte tangentiellement. (  )


Exemple :

$$\min f(x)$$

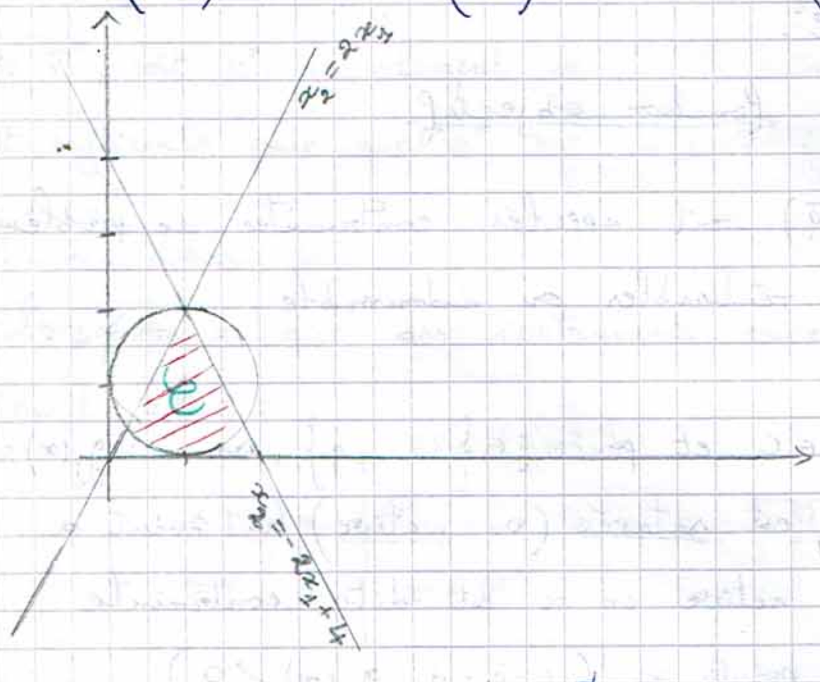
$$g_1(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \leq 0$$

$$g_2(x_1, x_2) = x_2 - 2x_1 \leqslant 0.$$

$$g_3(x_1, x_2) = x_2 + 2x_1 - 4 \leqslant 0.$$

$\forall (x_1, x_2) \in \mathbb{R}^2$.

$d_0 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$, $\quad d_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\quad x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$




On détermine les contraintes active en $x = (1, 2)^t$

$$\begin{cases} g_1(1,2) = 0 & \text{(active)} \\ g_2(1,2) = 0 & \text{(active)}. \\ g_3(1,2) = 0 & \text{(active)} \end{cases}$$

1 $\quad \nabla g_i$ ?

$$\nabla g_1(x) = \begin{pmatrix} 2(x_1 - 1) \\ 2(x_2 - 1) \end{pmatrix} \quad ; \quad \nabla g_1(1,2) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

$$\nabla g_2(x) = g_2(x_1, x_2) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

$$\nabla g_3(x) = \nabla g_3(1,2) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

# $d_0$ est-elle une direction admissible au point $x$ pour $f$.

$$\langle \nabla g_1(1,2), d_0 \rangle = (0, -1)\begin{pmatrix} 0 \\ 2 \end{pmatrix} = -2 \leqslant 0$$

$$\langle \nabla g_2(1,2), d_0 \rangle = (0, -1)\begin{pmatrix} -2 \\ 1 \end{pmatrix} = -1 \leqslant 0$$

$$\langle \nabla g_3(1,2), d_0 \rangle = (0, -1)\begin{pmatrix} 2 \\ 1 \end{pmatrix} = -1 \leq 0$$

$d_0$ est une direction admissible au point $x = (1,2)^t$.

# $d_1$ est-elle une direction admissible au point $x$ pour $f$

$$\langle \nabla g_1(1,2), d_1 \rangle = (0, 1)\begin{pmatrix} 0 \\ 2 \end{pmatrix} = +2 \geqslant 0$$

alors $d_1$ n'est pas une direction admissible au point $x = (1,2)$.

## 2_ Résultat d'existence et d'unicité:

### Théorème 1-1: (Existence)

Supposons que $f$ est propre, continue, que $C$ est fermé et non vide et que l'une des conditions suivantes est vérifiée :

1°/ Soit $C$ est borné

2°/ Soit $f$ est coercive. (quadratique)

Alors le problème (P) admet au moins une solution.

### Théorème 1-2: (Existence et unicité)

$f$ fonction propre, continue et strictement convexe. $C$ est un sous ensemble fermé et convexe de $\mathbb{R}^n$.

Si $C$ est borné ou $f$ est convexe alors il existe un unique $x^* \in C$ qui est solution du problème.

## 3- Condition d'optimalité :

### 3-1- Condition d'optimalité du $1^{er}$ ordre :

Théorème 3-1 : (Condition nécessaire du $1^{er}$ ordre)

Si $f$ est gâteau différentiable et si $C$ est un convexe fermé de $\mathbb{R}^n$, alors toute solution $x^*$ de $(P)$ vérifie la condition suivante :

$$\forall x \in C \quad \langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad \ldots (1-1)$$

Théorème 3-2 : (Condition nécessaire et suffisante du $1^{er}$ ordre dans le cas convexe)

Si $f$ convexe, gâteau différentiable et si $C$ est convexe fermé de $\mathbb{R}^n$, soit $x^*$ un élément de $C$, la condition $(1-1)$ est nécessaire et suffisante pour que $x^*$ soit solution $(P)$.

### 3-2- Contrainte en égalité :

Lorsqu'on a que des contraintes en égalité ce problème $(P)$ s'écrit, alors :

$$(P_C) = \begin{cases} \min f(x) \\ h_i(x) = 0 , \quad i = \overline{1, p} \\ x \in \mathbb{R}^n \end{cases}$$

Théorème 1-5 : (Condition $1^{er}$ ordre contrainte en égalité) :

On suppose : • $f$ et $h_i$, $i = \overline{1,p}$ sont de classe $C^1$.

• Les "$p$" vecteurs de $\mathbb{R}^n$ : $\nabla h_1(x^*), \ldots, \nabla h_p(x^*)$ sont linéairement indépendant $(p \leq n)$ alors il existe "$p$" réels $\lambda_1^*, \ldots, \lambda_p^*$ tel que : $\nabla f(x^*) + \sum \lambda_i \nabla h_i(x) = 0 \quad \ldots (1-2)$

Définition 1-5 :

Les réels $\lambda_1^*, \dots, \lambda_p^*$ du théorème 1-5 sont appelés multiplicateurs de lagrange associé à une contraintes $h_1, \dots, h_p$ au points $x^*$ respectif.

Théorème 1-6 : (Condition d'optimalité non qualifié).

On suppose que $f$, $h_i$ pour $i = \overline{1,p}$ et $g_j$ pour $j = \overline{1,q}$ sont de classe $C^1$, soit $x^*$ une solution du problème alors il existe $\lambda^* = (\lambda_1^*, \dots, \lambda_p^*) \in \mathbb{R}^p$ et $\mu^* = (\mu_1^*, \dots, \mu_q^*) \in \mathbb{R}_+^q$ et $\mu_0^* \in \mathbb{R}_+$.
tels que :

- $\forall j \in \{0, 1, \dots, q\} \quad \mu_j^* \geqslant 0 \qquad (1\text{-}3\text{-}\ )$
- $h_i(x^*) = 0 \ , \ \forall i = \overline{1,p}$ et $g_j(x^*) \leqslant 0 \ , \ \forall j = \overline{1,q}$. $(1\text{-}3\text{-}b)$
- $\forall j \in \{1, 2, \dots, q\} \quad \mu_j^* . g_j(x^*) = 0 \qquad (1\text{-}3\text{-}c)$
- $\mu_0^* \nabla f(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) + \sum_{i=1}^{p} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{q} \mu_j . \nabla g_j(x^*) = 0$

$(1\text{-}3\text{-}d)$.

Remarque :

1/ $\lambda_1^*, \dots, \lambda_p^*, \mu_1^*, \dots, \mu_q^*$ sont appelés multiplicateurs de lagrange.

2/- La relation $(1\text{-}3\text{-}b)$ est appelé relation de réalisabilité.

3/- La relation $(1\text{-}3\text{-}c)$ est une relation de complémentarité.

4/- Les conditions du théorème (1-6) sont dite non qualifié car le scalaire $\mu_0^*$ peut être nul et on n'a aucune information sur $f$ puisqu'elle n'appartiennent nulle part dans les conditions d'optimalité. Il faut donc donner des conditions qui assurent que $\mu_j^* \neq 0$.
De telles conditions, sont appelés condition de qualification ou de régularité lorsqu'elles sont vérifiées, le problème est dit qualifié.

Définition 1-4 : (point régulier) (C.Q-1)

On dit qu'un élément $x^*$ de $\mathbb{R}^n$ est régulier pour les contraintes $h_i$, $\forall i=\overline{1,p}$ et $g_j$, $\forall j=\overline{1,q}$.

# s'il est réalisable.

$h_i(x^*)=0$, $\forall i=\overline{1,p}$ ; $g_j(x^*)\leqslant 0$, $\forall j=\overline{1,q}$.

# Les vecteurs $\nabla h_1(x^*),\dots,\nabla h_p(x^*)$ sont linéairement ~~indépendant~~ indépendant.

# et si on peut trouver une direction $d$ de $\mathbb{R}^n-\{0\}$ tel que :

$$\langle \nabla h_i(x^*), d\rangle = 0 \quad , \forall i=\overline{1,p}$$

$$\langle \nabla g_j(x^*), d\rangle < 0 \quad , \forall j\in I(x^*)$$ (pour les contraintes actives).

On dit aussi que $x^*$ vérifie les conditions de qualification de Mangasarian - Formowitz notées (C.Q.1).

Définition 1-5 :
- On dit qu'un élément $x^*\in\mathbb{R}^n$ vérifie les conditions : (C.Q.2) pour les contraintes $h_i$ $(i=\overline{1,p})$ et $g_j$ $(j=\overline{1,q})$

# s'il est réalisable.

# et les vecteurs $\nabla h_i(x^*)$, $\nabla g_j(x^*)$, $\forall i=\overline{1,p}$ et $\forall j=\overline{1,q}$ sont linéairement indépendant.

Théorème 1-7 : (Condition de Karush - Kuhen - Tucker)

On suppose que $f$, $h_i$ $(i=\overline{1,p})$ et $g_j$ $(j=\overline{1,q})$ sont de classe $C^1$. soit $x^*$ une solution du problème (P). On suppose que $x^*$ est un point régulier pour les contraintes $h_i$ $(i=\overline{1,p})$ et $g_j$ $(j=\overline{1,q})$ alors il existe $\lambda^*=(\lambda_1^*,\dots,\lambda_p^*)\in\mathbb{R}^p$ et $\mu^*=(\mu_1^*,\dots,\mu_q^*)\in\mathbb{R}^q$ tel que :

# $\forall j=\overline{1,q}$ $\quad \mu_j^*\geqslant 0 \quad$ (1-9-a)

# $h_i(x^*)=0$, $\forall i=\overline{1,p}$ et $g_j(x^*)\leqslant 0$, $\forall j=\overline{1,q}$.

\# $\forall j = \overline{1,q}$ , $\mu_j^* \, g_j(x^*) = 0$ (1-9-c)

\# $\nabla f(x^*) + \sum \lambda_i^* \nabla h_i(x^*) + \sum \mu_j^* \nabla g_j(x^*) = 0$.

Remarque :

Les conditions (1-9) (càd : (1-9-a), (1-9-b), (1-9-c) et (1-9-d) sont appelées condition de Karush-Kuhn-Tucker.

Définition 1-1 : "Lagrangien"

On appelle lagrangien du problème (P) la fonction définie de $\mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \longrightarrow \mathbb{R}$ par :

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \sum_{i=1}^{p} \lambda_i h_i(x) + \sum_{j=1}^{q} \mu_j g_j(x)$$

Remarque : La relation (1-9-d) s'écrit :

$\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0$

Où $\nabla_x$ désigne le gradient du lagrangien par rapport à la variable $x$.

Dans le cas convexe : Nous avons le résultat suivant :

Théorème 1-8 : (Condition nécessaire suffisante dans le cas convexe)

On suppose que $f$, $h_i$ $(i = \overline{1,p})$ et $g_j (j = \overline{1,q})$ sont de $C^1$, que $f$ et $g_j$ sont des fonctions convexes et $h_i$ sont affines.
On suppose aussi que $x^*$ est un point régulier pour les contraintes $h_i$ et $g_j$, alors :

$$x^* \text{ solution de } (P) \underset{\Longleftarrow}{\overset{(\text{thm } 1\text{-}7) \atop \Longrightarrow}{\Longleftrightarrow}} \text{les conditions (1-9) sont satisfaites.}$$

Exemple : On considère la fonction suivante :

$$f(x,y) = \frac{1}{2}(x,y)\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + xy + 2y^2.$$

On considère le problème $(P_1)$ suivant :

$$(P_1)\begin{cases} \min f(x,y) \\ x+y \geqslant 1 \\ y \geqslant 0 \\ (x,y)\in \mathbb{R}^2 \end{cases} \Longrightarrow (P_1)\begin{cases} \min f(x,y) \\ g_1(x,y) = -x-y+1 \leqslant 0 \\ g_2(x,y) = -y \leqslant 0 \\ (x,y)\in \mathbb{R}^2 \end{cases}$$

1/- On montre que le problème admet une solution unique :

- $f$ propre.
- $f$ continue.
- $f$ est coercive. (car : c'est une fonction quadratique dont le hessien est fermé $*$).
- Le domaine $f$ est non vide et fermé alors $(P_1)$ admet au moins un minimum.
- $f$ est strictement convexe (car $*$) Donc : $(P_1)$ admet un unique minimum.

2) $X = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

a). Quelles sont les contraintes actives en $X$ ?

$g_1(X) = g_1(1,0) = 0$ (active)

$g_2(X) = g_2(1,0) = 0$ (active)

$$I(X) = I(1,0) = \{1,2\}.$$

b)- $X$ est-il un point régulier ?

$$\left.\begin{matrix} g_1(X) = 0 \leq 0 \\ g_2(X) = 0 \leq 0 \end{matrix}\right\} X \text{ est realisable.}$$

- Il existe un unique $d = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$, $d \in \mathbb{R}^n$.

$\langle \nabla g_j(x), d \rangle < 0$, $\forall j \in I(x)$

$$\langle \nabla g_1(x), d \rangle = (d_1, d_2)\begin{pmatrix} -1 \\ -1 \end{pmatrix} = -d_1 - d_2 < 0$$

$$\langle \nabla g_2(x), d \rangle = (d_1, d_2)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -d_2 < 0$$

$$\begin{cases} -d_2 < 0 \\ -d_1 - d_2 < 0 \end{cases} \Rightarrow \begin{cases} d_2 > 0 \\ d_1 > -d_2 \end{cases}$$

c) $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ vérifie t-il les conditions de Karush-Kuhn-Tucker ?

$$\nabla f(x) = \begin{pmatrix} 2x + 4 \\ x + 4y \end{pmatrix} \Leftrightarrow \nabla f(1,0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$\nabla g_1(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad ; \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \mu_1 \begin{pmatrix} -1 \\ -1 \end{pmatrix} + \mu_2 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{cases} -\mu_1 = -2 \\ -\mu_1 - \mu_2 = -1 \end{cases} \Leftrightarrow \begin{cases} \mu_1 = 2 > 0 \\ \mu_2 = -1 < 0 \end{cases}$$

Conclusion : Le point $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ n'est pas la solution du problème $(P_1)$, car : $\mu_2 = -1 < 0$.

## 4 - Condition d'optimalité du second ordre :

**Théorème 1-10 :**

On suppose que $f$, $h_i$ $(i = \overline{1,p})$ et $g_j$ $(j = \overline{1,q})$ sont de $C^2$.
On suppose aussi que $x^*$ est un minimum du problème et que $x^*$ est un point régulier, alors il existe $\lambda^* = (\lambda_1^*, \dots, \lambda_p^*) \in \mathbb{R}^p$ et $\mu^* = (\mu_1^*, \dots, \mu_q^*)$ tel que :

Les conditions (1-9) sont satisfaites, pour toute direction $d \in \mathbb{R}^n - \{0\}$ vérifiant :

$$\begin{cases} \langle \nabla h_i(x^*), d \rangle = 0 , & \forall i = \overline{1,p} \\ \langle \nabla g_j(x^*), d \rangle = 0 , & \forall j \in I^+(x^*) \\ \langle \nabla g_j(x^*), d \rangle \leq 0 , & \forall j \in I(x^*) \setminus I^+(x^*) \end{cases} \quad (1-10)$$

Où $I^+(x^*) = \{ j \in \{1, \dots, q\} \}$

$$g_j(x^*) = 0 \quad \text{et} \quad \mu_j^* > 0.$$

On a : $\langle \nabla^2_{xx} (\mathcal{L}(x^*, \lambda^*, \mu^*) d, d \rangle \geq 0 \quad \dots (1-11)$

avec $\nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*)$ désigne la seconde dérivées de $\mathcal{L}$ au point $(x^*, \lambda^*, \mu^*)$.

Définition : . L'ensemble $I^*(x^*)$ est l'ensemble des indices des contraintes fortement actives.

Lorsque $I^*(x^*) = I(x^*)$ alors $g_j(x^*) = 0 \iff \mu_j^* > 0$.

On dit qu'il y a une stricte complimentarité.

Théorème : On suppose que $f$, $h_i$ $(i = \overline{1,p})$ et $g_j$ $(j = \overline{1,q})$ sont $C^2$, soit $x^* \in \mathbb{R}^n$ vérifiant les conditions $(1-9)$ avec les multiplicateurs de lagrange $(\lambda^*, \mu^*)$.

La matrice hessienne du lagrange au points $(x^*, \lambda^*, \mu^*)$

$$H(x^*) = \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*)$$

$$= \nabla^2 f(x) + \sum \lambda_i^* \nabla h_i(x^*) + \sum \mu_j^* \nabla^2 g_j(x^*).$$

est définie positive sur sous-espace.

$$\mathcal{C} = \{ d \in \mathbb{R}^n : \langle \nabla h_i(x^*), d \rangle = 0, \forall i = \overline{1,p} \text{ et } \langle \nabla g_j(x^*), d \rangle = 0, \forall j \in I^+(x^*) \}.$$

(Pour montrer que la matrice définie positive $d^t H(x), d > 0$

Alors $x^*$ est un minimum strict de $f$ sur $C$.

# II - Les algorithmes :

## Introduction :

- Les méthodes d'optimisation avec contraintes se divisent en deux familles :
  - Méthode directes ou primales.
  - Méthode utilisant la notion du duabilité.

Les méthodes primales opèrent directement sur le problème donné. Elles engendrent une suite minimisante de solution, réalisables, leur avantage, c'est qu'elles présentent une solution réalisable approchée si les intersections sont interrompues. Leur inconvénient, c'est au niveau de la convergence vers le minimum global.

Les méthodes duables sont plus robustes. Leur inconvénient, c'est qu'elle ne fournissent pas une solution approchée

## 1 - Méthode du gradient projeté

Cette méthode s'inspire de la méthode du gradient dans le cas sans contraintes.

Néanmoins, lorsqu'on minimise sur un ensemble $C$ et qu'à l'itération $k$, $x^k \in C$, On n'est pas sur que l'itéré suivant

$$x^{k+1} = x^k - \nabla f(x^k) \text{ soit dans } C.$$

Il faut donc le ramener dans $C$ grâce à une projection sur $C$.

Cette méthode est résumé dans l'algorithme suivant :

### Algorithme : "Méthode du gradient projeté" :

1 - Initialisation :

$k = 0$ Choix de $x^0 \in C$ ; $\rho_0 > 0$, $\varepsilon > 0$

2 - Itération $k$ :

$$x^{k+1} = \Pi_C \big( \underbrace{x^k - \rho_k \nabla f(x^k)}_{} \big)$$

si $x \in C$, on obtient : $x^k - \rho_k \nabla f(x^k)$

3- Critère d'arrêt :

Si $\| x^{k+1} - x^k \| < \varepsilon$ stop.

Sinon $k = k + 1$ et on retourne à l'étape 2.

Définition : . Le point $x^*$ est le projeté de $x$ sur $C$.

. L'application $\Pi_C : \mathbb{R}^n \longrightarrow C$

$$x \longrightarrow \Pi_C(x) = x^*.$$

qui à $x$ associé son projeté $x^*$ s'appelle la projection sur $C$.

On définit la distance d'un point $x$ à l'ensemble $C$ par :

$d(x, C) = \inf \| x - y \| \ ; \ y \in C$.

Avantage : - La méthode du gradient projeté est méthode simple.

Inconvénient : - Difficile a appliquée puisqu'à chaque itération il faut calculer le projeté de $x^{k+1}$ sur $C$.

2- Méthode de Lagrange - Newton pour la contrainte en égalité :

a)- Cas quadratique :

- Considérons le problème quadratique suivant avec des contraintes d'égalité :

$$(P_Q) \begin{cases} \min f(x) = \frac{1}{2} x^t Q - C^t x . \\ \min f(x) = \frac{1}{2} < x Q , x > - < C , x > \\ \text{Sous-contraintes : } \quad Ax = b \Longleftrightarrow Ax - b = 0 \end{cases}$$

Où . $Q$ : matrice carré d'ordre $n$.

. $x$ et $C$ : vecteur colonne de $\mathbb{R}^n$.

. $A$ : matrice de type $(p \times n)$.

. $b$ : vecteur colonne de $\mathbb{R}^p$

- Les conditions de Karush-Kuhn-Tucker pour $(P_Q)$ :

- $\exists \lambda = (\lambda_1, \dots, \lambda_p) \in \mathbb{R}^p$.
- $Ax - b = 0$
- $\nabla_x \left( \frac{1}{2} x^t Q x - C^t x \right) + \nabla_x \left( \lambda^t . (Ax - b) \right) = 0$

Ceci implique :

$$\begin{cases} Q x + A^t . \lambda = C \\ A x = b \end{cases}$$

$$\left[ \begin{array}{c|c} Q & A \\ \hline A & 0 \end{array} \right] \left[ \begin{array}{c} x \\ \hline \lambda \end{array} \right] = \left[ \begin{array}{c} C \\ \hline b \end{array} \right] \quad \dots \quad (S)$$

Ainsi, la solution optimal $\begin{pmatrix} x^* \\ \lambda^* \end{pmatrix}$ du problème $(P_Q)$ est solution du système linéaire (S)

Si $M = \left[ \begin{array}{c|c} Q & A^t \\ \hline A & 0 \end{array} \right]$ est inversible

Alors (S) est un système de Cramer et le problème $(P_Q)$ admet une solution unique.

Exemple :

$$\begin{cases} \min f(x,y) = x^2 + xy + 2y^2 \\ h_1(x,y) = x + y - 1 = 0 \\ h_2(x,y) = -2x + y = 0 \\ (x,y) \in \mathbb{R}^2. \end{cases}$$

On a : $f(x,y) = \frac{1}{2}(2x^2 + 2xy + 4y^2) + \frac{1}{2}\begin{pmatrix} x \\ y \end{pmatrix} \underbrace{\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}}_{Q} (x,y)$

$+ \underbrace{(0,0)}_{C} \begin{pmatrix} x \\ y \end{pmatrix}$.

$$\begin{cases} h_1(x,y) = 0 \\ h_2(x,y) = 0 \end{cases} \iff \begin{cases} x + y - 1 = 0 \\ -2x + y = 0 \end{cases}$$

$$\iff \underbrace{\begin{pmatrix} 1 & 1 \\ -2 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{b}$$

Soit $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \in \mathbb{R}^2$ multiplicateur de lagrange associés à $\lambda_1$ et $\lambda_2$.

$$\left[\begin{array}{cc|cc} 2 & 1 & 1 & 2 \\ 1 & 4 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 \end{array}\right] \left[\begin{array}{c} x \\ y \\ \hline \lambda_1 \\ \lambda_2 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ \hline 1 \\ 0 \end{array}\right] \quad \ldots (S')$$

$$(S') \Rightarrow \begin{cases} 2x + y - \lambda_1 - 2\lambda_2 = 0 & \ldots (1) \\ x + 4y + \lambda_1 + \lambda_2 = 0 & \ldots (2) \\ x + y = 1 & \ldots (3) \\ -2x + y = 0 & \ldots (4) \end{cases}$$

(3) et (4) $\Longrightarrow$ $x = \frac{1}{3}$ , $y = \frac{2}{3}$.

On remplace dans (1) et (2) pour avoir :

$$\begin{cases} \lambda_1 - 2\lambda_2 = \frac{-4}{3} \\ \lambda_1 + \lambda_2 = -3 \end{cases}$$

$\lambda_1 = \frac{-22}{9} \in \mathbb{R}$ ; $\lambda_2 = \frac{-5}{9} \in \mathbb{R}$.

## b)- Cas générale :

Considérons le cas général suivant

$$(P_e)\begin{cases} \min f(x) \\ h(x)=0 \quad \text{avec} : h=(h_1,h_2,\dots,h_p) \end{cases}$$

Condition Karush - Kuhn - Tucker pour $(P_e)$ donnent :

- $\exists \lambda \in \mathbb{R}^p$
- $h(x)=0 \iff h_i(x)=0 \ , \forall\, i=\overline{1,p}$.
- $\nabla_x \mathcal{L}(x,\lambda)=0$.

avec :

$$\mathcal{L}(x,\lambda)= f(x)+\sum_{i=1}^{p} \lambda_i h_i(x)$$

$$= f(x)+\langle \lambda , h(x)\rangle$$

$$= f(x)+\lambda^t h(x)$$

On obtient alors :

$$\begin{cases} \nabla f(x)+\sum \lambda_i \nabla h_i(x)=0 & \dots (3-2) \\ h_i(x) \quad ( h_i(x)=0 \ , \forall\, i=\overline{1,p} & \dots (3-3) \end{cases}$$

On linéairise (3-2) et (3-3), et partir d'un point $(x^k,\lambda^k)$ au voisinage de $x^k$ pour $(x^{k+1},\lambda^{k+1})$ la solution du système linéairise.

Linéairisation de la fonction $f$ au voisinage d'un point $(x_0,y_0)$

$$F(x,y)= f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)\cdot(x-x_0)+\frac{\partial f}{\partial y}(y,y_0)(y-y_0)$$

y est le $\lambda$

$$\nabla f(x^k)+\sum_{i=1}^{p}\lambda_i \nabla h_i(x^k)+\left[\nabla^2 f(x^k)+\sum_{i=1}^{p}\lambda_i^k \cdot \nabla^2 h_i(x^k)\right]$$

$$(x^{k+1}-x^k)+\sum_{i=1}^{p}\nabla h_i(x^k)(\lambda_i^{k+1}-\lambda_i^k)=0$$

$$h(x^k)+\langle \nabla h(x^k), x^{k+1}-x^k\rangle=0 \ , \ \forall\, i=\overline{1,p}.$$

Donc :

$$\left[\nabla^2 f(x^k) + \sum \lambda_i^k \cdot \nabla h_i^k(x^k)\right](x^{k+1} - x) + \sum_{i=1}^{p} \nabla h_i(x^k)(\lambda_i^{k+1} - \lambda_i)$$
$$= -\nabla f(x^k) - \sum \lambda_i^k \nabla h_i(x^k).$$

$$\langle \nabla h_i(x^k), x^{k+1} - x^k \rangle = -h(x^k), \quad i = \overline{1,p}.$$

On obtient alors le système linéaire suivant :

$$\left[\begin{array}{c|c} \nabla_x^2 \mathcal{L}(x^k, \lambda^k) & \nabla h_1(x^k) \dots \nabla h_p(x^k) \\ \hline \nabla h_1(x^k) & \\ \vdots & 0 \\ \nabla h_p(x^k) & \end{array}\right] \left[\begin{array}{c} x^{k+1} - x^k \\ \hline \lambda^{k+1} - \lambda^k \end{array}\right] = \left[\begin{array}{c} -\nabla \mathcal{L}(x^k, \lambda^k) \\ \hline -\lambda_1(x^k) \\ \vdots \\ -\lambda_p(x^k) \end{array}\right]$$

Donc :

$$\left[\begin{array}{c|c} \nabla_x^2 \mathcal{L}(x^k, \lambda^k) & (\nabla h(x^k))^t \\ \hline \nabla h(x^k) & 0 \end{array}\right] \left[\begin{array}{c} x^{k+1} - x^k \\ \hline \lambda^{k+1} - \lambda^k \end{array}\right] = \left[\begin{array}{c} -\nabla \mathcal{L}(x^k, \lambda^k) \\ \hline -h_i(x^k) \end{array}\right]$$

**Algorithme : Lagrange Newton :**

1_ Initialisation :

$k = 0$ choix de $x^0 \in \mathbb{R}^n$ et de $\lambda^0 \in \mathbb{R}^p$ et $\varepsilon > 0$.

2_ Itération $k$ :

On connait $(x^k, \lambda^k)$

Résoudre :

$$\left[\begin{array}{c|c} \nabla_{xx}^2 \mathcal{L}(x^k, \lambda^k) & \nabla h(x^k)^t \\ \hline \nabla h(x^k) & 0 \end{array}\right] \left[\begin{array}{c} d_k \\ \hline y_k \end{array}\right] = -\left[\begin{array}{c} \nabla_x \mathcal{L}(x^k, \lambda^k) \\ \hline h(x^k) \end{array}\right]$$

$x^{k+1} = d_k + x^k$ et $\lambda^{k+1} = y_k + \lambda^k$.

3_ Critère d'arrêt : si $\| x^{k+1} - x^k \| < \varepsilon$ stop

Sinon $k = k+1$ et retourner à l'étape (2).

**Remarque :**

- Inconvénient de cette méthode $(3.2)$ ; $(3.3)$ sont vérifiées seulement pour le minimum mais aussi pour le maximum avec contraintes.
- La suite $(x^k, \lambda^k)$ peut ne pas converger si $(x^0, \lambda^0)$ et choisit assez loin de $(x^*, \lambda^*)$ qu'est la solution de problème.

**Exemple :**

$$\begin{cases} \min f(x,y) = \frac{1}{3}x^3 + xy + \frac{1}{3}y^3 \\ h_1(x,y) = (x-1)^2 + y^2 + 1 \\ h_2(x,y) = x^2 + (y-1)^2 - 1 = 0 \end{cases}$$

$$X = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{et} \quad \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

$$\nabla f(X) = \begin{pmatrix} x^2 + y \\ x + y^2 \end{pmatrix} \; ; \quad \nabla^2 f(X) = \begin{pmatrix} 2x & 1 \\ 1 & 2y \end{pmatrix}$$

$$\nabla h_1(X) = \begin{pmatrix} 2(x-1) \\ 2y \end{pmatrix} \; ; \quad \nabla^2 h_1(X) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\nabla h_2(X) = \begin{pmatrix} 2x \\ 2(y-1) \end{pmatrix} \; ; \quad \nabla^2 h_2(X) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

$$\nabla_x \mathcal{L}(X, \lambda) = \nabla f(X) + \lambda_1 \nabla h_1(X) + \lambda_2 \nabla h_2(X)$$

$$= \begin{pmatrix} x^2 + y + 2\lambda_1 x - 2\lambda_1 + 2\lambda_2 x \\ x + y^2 + 2\lambda_1 y + 2\lambda_2 y - 2\lambda_2 \end{pmatrix}$$

$$\nabla^2_{xx} \mathcal{L}(x, \lambda) = \nabla^2 f(x) + \lambda_1 \nabla h_1(x) + \lambda_2 \nabla h_2(x)$$

$$= \begin{pmatrix} 2x + 2\lambda_1 + 2\lambda_2 & 1 \\ 1 & 2y + 2\lambda_1 + 2\lambda_2 \end{pmatrix}$$

$$\left[\begin{array}{cc:cc} 2x^k + 2\lambda_1^k + 2\lambda_2^k & 1 & 2x^k - 2 & 2x^k \\ 1 & 2y^k + 2\lambda_1^k + 2\lambda_2^k & 2y^k & 2y^k - 2 \\ \hdashline 2x^k - 2 & 2y^k & 0 & 0 \\ 2x^k & 2y^k - 2 & 0 & 0 \end{array}\right] \left[\begin{array}{c} x^{k+1} - x^k \\ y^{k+1} - y^k \\ \hdashline \lambda_1^{k+1} - \lambda_1^k \\ \lambda_2^{k+1} - \lambda_2^k \end{array}\right]$$

$$= - \left[\begin{array}{c} (x^k)^2 + y^k + 2\lambda_1 x^k - 2\lambda_1^k + 2\lambda_2^k x^k \\ x^k + (y^k)^2 + 2\lambda_1^k y^k + 2\lambda_2^k y^k - 2\lambda_2^k \\ \hdashline (x^k - 1)^2 + (y^k)^2 - 1 \\ (x^k)^2 + (y^k - 1)^2 - 1 \end{array}\right]$$