*Chapitre IV*

## OPTIMISATION SANS CONTRAINTES

L'optimisation sans contraintes est une sous-discipline de l'optimisation mathématique, dans laquelle l'ensemble admissible est l'espace tout entier. Ces problèmes sont plus simples à analyser et à résoudre que les problèmes d'optimisation avec contraintes.

### IV.1. Généralités

L'optimisation a un vocabulaire particulier, pour cela nous allons introduire quelques notations et définitions classiques. Tout d'abord, nous donnons la formulation générale d'un problème d'optimisation, donc on a besoin d':

-Une fonction objective ou fonction de coût ou critère, à minimiser, noté $f : R^n \to R$

-Un ensemble $S \subseteq R^n$ où l'on va chercher la solution, on dit que S est l'ensemble des éléments admissibles du problème ou bien l'ensemble des contraintes.

On cherche à minimiser *f* sur S c'est à dire on cherche *x* de S tel que :

$$f(x) = \min_{y \in S} f(y) \Leftrightarrow f(x) \leq f(y), \quad \forall y \in S$$

**Remarque :**

Lorsque l'on utilise la notation « inf » pour un problème d'optimisation c'est à-dire "inf *f(x)*" cela indique que l'on ne sait pas, si la valeur du minimum est atteinte. On utilise de préférence la notation "minf(x)" mais il ne s'agit pas d'une convention universelle. Pour les problèmes de maximation, les notations "sup et max" remplace "inf et min" respectivement.

L'optimisation se scinde en deux type de problèmes: sans et avec contraintes, dans les deux cas, le but est de trouver les valeurs qui maximisent ou minimisent une fonction.

**a-Optimisation sans contrainte**

Nous allons étudier le problème d'optimisation sans contraintes où on effectue la minimisation de la fonction $f : R^n \to R$ sur tout l'espace $R^n$. Nous considérons donc le problème formulé de la façon suivante :

$$(P) \quad \{\min f(y); \quad y \in R^n\}$$

qui peut réécrire sous la forme

$$\{trouver\ x \in R^n \ \ tel\ que\ \ f(x) \leq f(y), \quad \forall y \in R^n\}$$

**b-Optimisation sous contrainte**

Le problème d'optimisation avec contraintes s'écrit sous la forme :

$$(P)\ \{\min\ f(y);\ \ y \in S\}$$

$$\{trouver\ x \in S\ tel\ que\ \ f(x) \le f(y),\ \ \forall y \in S\}$$

S est l'ensemble des contraintes sous forme d'égalité et d'inégalité telle que pour *(p,q) de N*

$$S = \left\{ \begin{matrix} x \in R^n,\ \ g_i(x) = 0, i \in \{1,....,p\} \\ h_j(x) \le 0;\ j \in \{1,....,q\} \end{matrix} \right\},$$

tel que:

-La fonction $g(x) = \left(g_1(x); g_2(x),....., g_p(x)\right)$ est une fonction de plusieurs variables *x de* $R^n$ à valeurs dans R, qui représente les contraintes en égalités.

-La fonction $h(x) = \left(h_1(x); h_2(x),....., h_q(x)\right)$ est une fonction de plusieurs variables *x* de $R^n$ à valeurs dans R, qui représente les contraintes en inégalités.

Nous noterons ces types de problèmes ainsi :

**-(PCE)** Problème avec contraintes d'égalité.

**-(PCI)** Problème avec contraintes d'inégalité.

**Exemple :**

1-Quelle est la valeur maximale de la fonction :

$$f(A,B,C) = \sin A \sin B \sin C,$$

sous la contrainte : $A + B + C = \pi$.

C'est un problème d'optimisation avec contraintes d'égalité.

2-Maximiser la fonction objective

$$f(x_1, x_2) = 1200\ x_1 + 1000\ x_2$$

sous les contraintes : $3\ x_1 + 4\ x_2 \le 160$, $6\ x_1 + 3\ x_2 \le 180$, $x_1 \ge 0$ ; $x_2 \ge 0$.

C'est un problème d'optimisation avec contraintes d'inégalité.

***Remarque :***

1- Plus généralement, on peut remplacer l'espace $R^n$ par un espace vectoriel topologique sur R de dimension fini.

2- Pour la maximisation, faire la minimisation de la fonction *(-f):*

$$\max\ f(x) = -\min\ f(-x).$$

### IV.*2. Quelques exemples d'optimisation*

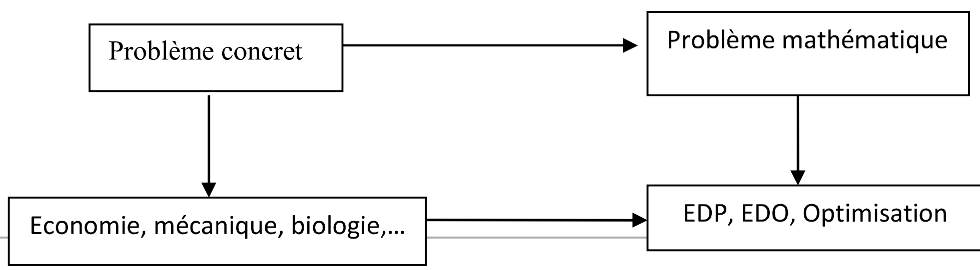
### IV.2.a : Exemple 1

On veut acheter les aliments au plus bas prix possible qui fourniront les besoins journaliers de calories ( au moins 2000) protéines ( 30 unités) et fibre ( 20 unités), les aliments disponibles sont la viande ( un kilo à un prix de 500 dinars fournit 800 calories, 10 unités  de protéines et 2 de fibre) des fruits ( 200 cal, 6 et 4 unités à prix de 200 dinars) et du riz (100 cal, 5 et 4 unités à un prix de 150 dinars).

Pour répondre à la question, on va donner un modèle mathématique à ce problème.

***Modélisation :***

Un **modèle mathématique** est une traduction des données du problème  dans le but de lui appliquer les outils, les techniques et les théories mathématiques.




Dans la modélisation mathématique de problème d'optimisation, on distingue trois étapes :

1. Identification des variables de décisions ; ce sont les paramètres sur lesquels l'utilisateur peut agir pour faire évoluer le système considéré.
2. Définition d'une fonction coût ou fonction permettant d'évaluer l'état du système (ex : rendement, performance,…).
3. Description des contraintes imposées aux variables de décision.

Le problème d'optimisation consiste alors à déterminer les variables de décision conduisant aux meilleures conditions de fonctionnement du système (ce qui revient à minimiser ou maximiser la fonction coût), tout en respectant les contraintes d'utilisation définies à l'étape 3

Donc ; pour donner le modèle mathématique du problème précédent il faut passer par les trois étapes suivantes :

-Les variables

- La fonction économique (objective)

-Les contraintes

a)Les variables :

Notons par :

***x :*** la quantité de la viande

***y*** : la quantité de fruit

***z :*** la quantité du riz

***b***)La fonction objective :

Le prix total est : $f(x,y,z)=(500x+200y+150z)$

c)Les contraintes :

$$\begin{cases} 800x+200y+100z \geq 2000 \\ 10x+6y+5z \geq 30 \\ 2x+4y+4z \geq 20, \\ x \geq 0, \quad y \geq 0, \quad z \geq 0 \end{cases},$$

Donc, notre problème est modélisé en un problème mathématique de la façon suivante :

$$(P) \quad \{\min \quad (500x+200y+150z); (x,y,z) \in S\}$$

tel que $S=\{(x,y,z) \in R_+^{\,3} \,/\, 800x+200y+100z \geq 2000,\ 10x+6y+5z \geq 30,\ 2x+4y+4z \geq 20\}$

**IV .2.b. Exemple 2 : Un problème classique du sac à dos**

Soit $n$ objets de poids respectifs $p_1,\ldots,p_n \in \mathbb{R}$, et d'utilités respectives $u_1,\ldots,u_n \in \mathbb{R}$, et $P \in \mathbb{R}$ un poids maximal que l'on  est disposé à porter. On pose $x_i=1$ si on met le i-ème objet dans le sac-à-dos, et $x_i=0$ sinon. On veut maximiser l'utilité du sac-à-dos sous contrainte de poids. Donc, notre problème est modélisé en un problème mathématique de la façon suivante :

$$\max_{x_i \in \{0,1\}} \sum_{1 \leq i \leq n} x_i u_i \quad \text{sous les contraintes} \quad \sum_{1 \leq i \leq n} x_i p_i \leq P$$

**I .2.c. Exemple 3 :**

On a besoin d'au moins 60 litres de peinture pour peindre les corridors d'un édifice. Pour effectuer ce travail, on utilise de la peinture blanche et de la peinture bleue. Selon le contremaître, on doit utiliser au plus 2 fois plus de peinture bleue que de peinture blanche. On évalue la surface à peindre à au plus 240 m$^2$ . Selon le fournisseur de peinture, un litre de peinture blanche couvre 2 m$^2$ et coûte 10 d, tandis qu'un litre de peinture bleue couvre 3 m$^2$ et coûte 12 d.

Combien de litres de chaque couleur le contremaître doit-il utiliser pour minimiser ses dépenses ?

***IV. 3. Solution Optimale***

Soit $f : R^n \to R$, On appelle problème d'optimisation sans contraintes, le problème *(P)* suivant :

$$(P) : \{\min f(x),\ \ x \in R^n\}$$

C'est un problème d'optimisation sans conditions sur les variables. Les minima locaux et globaux de *f* sur $R^n$ sont définis de la manière suivante :

***Définition :***

1-On dit que la fonction *f* du problème ($P$) possède un minimum global en $\hat{x} \in R^n$ ssi :

$$\forall x \in R^n,\ \ f(x) \geq f(\hat{x}).$$

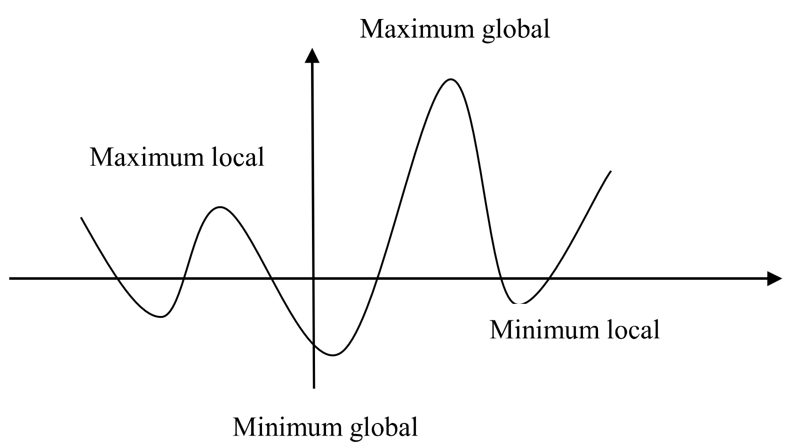
2-On dit que le point $\hat{x} \in R^n$ est un minimum locale du *(P)* ssi il existe un voisinage $V_\varepsilon(\hat{x})$ de $\hat{x}$ tel que :

$$\forall x \in V_\varepsilon(\hat{x}),\ \ f(x) \geq f(\hat{x}).$$

3-On dit que $\hat{x} \in R^n$ est un minimum local strict de *(P)* ssi : il existe un voisinage $V_\varepsilon(\hat{x})$ de $\hat{x}$ tel que :

$$\forall x \in V_\varepsilon(\hat{x}),\ \ x \neq \hat{x},\ \ f(x) > f(\hat{x}).$$

4-On dit que $\hat{x} \in R^n$ est un minimum local isolée de (P) s' il existe un voisinage $V_\varepsilon(\hat{x})$ de $\hat{x}$ tel que $\hat{x}$ est la seule solution optimal de *(P)*.



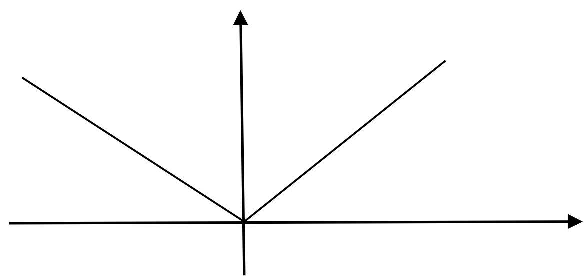

***Remarque :***

- Un minimum global est clairement un minimum local
- Si on dit simplement minimum on comprend minimum global.
- Toute solution optimale isolée est une solution optimale locale stricte. La réciproque n'est pas toujours vraie.

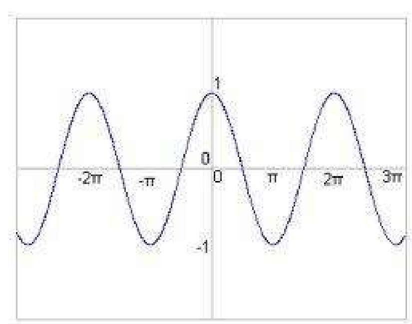
***Exemples :***

1- Soit la fonction : $f(x) = Cte, \quad \forall x \in R$ . Tout point $x$ de $R$ est un minimum local et global mais il n'existe aucune solution stricte ou isolée.

2- Soit la fonction : $f(x) = |x|, \quad \forall x \in R$. Le point $x=0$ est un minimum locale, globale, stricte et aussi isolée.



3- Pour la fonction $f(x) = \cos(x)$, il existe une infinité de minima et maxima globaux.




4- Pour la fonction $f(x) = x\cos x$, il existe une infinité de minima et maxima locaux mais aucun minimum ou maximum global.

**Remarque :** Les questions qui nous intéressent dans les problèmes d'optimisation sont les suivantes :
1) Existence et unicité de la solution du problème *(P).*
2) Caractérisation des solutions (les points réalisant le minimum)
3) Calcul, expression explicite du minimum et algorithmes permettant de l'approcher.

On utilise la plupart du temps des outils de nature différente pour répondre a chacune de ces questions, par exemple le problème de l'existence se résout en général en exploitant les propriétés t*opologiques* de l'ensemble S ( dans le cas sans contraintes S=$R^n$) et de la fonction *f.*

### *IV.4. Résultats d'existence et d'unicité*

Avant d'étudier les propriétés de la solution du problème *(P),* il faut assurer de leur existence. Le théorème le bien connu sur l'existence des minima et des maxima est

***Théorème d'existence 1 : (Théorème de Weierstrass)***

Soient S un ensemble compact (fermé et borné) non vide de $R^n$ et $f : S \to R$ une fonction continue sur S, alors *f* admet au moins un min $\hat{x}$ sur S. Autrement dit, il existe un point $\hat{x}$ de S minimum global de *f* sur S i.e. :

$$\forall x \in S, \quad f(x) \geq f(\hat{x}).$$

***Preuve :*** Soit $(x_n)$ une suite minimisante de *f* sur S c'est-à-dire d'éléments de S telle que

$$x_n \in S, \ \forall n \in N \ et \ \lim_{n \to +\infty} f(x_n) = \inf_{x \in S} \ f(x)$$

Comme S est borné, la suite minimisante soit bornée, donc on peut extraite une sous-suite notée $(x_n)$ qui converge vers un élément $\hat{x}$ de S car S est un ensemble fermé. Cette suite extraite vérifie

$$\lim_{n \to +\infty} f(x_n) = f(\hat{x})$$

car f est continue,
d'où par unicité de la limite

$$f(\hat{x}) = \inf_{x \in S} \ f(x)$$

Donc *f* réalise son minimum dans S.

***Remarque*** :

1- De la même façon, il existe un point de maximum global de *f* sur S.
2- Si la fonction *f* n'est pas continue alors elle n'admet pas un minimum.
3- Une fonction continue sur un fermé borné n'atteint pas toujours son minimum dans un espace de dimension infinie.
4- L'existence d'une suite minimisante provient de la définition de l'inf.
5- Rappelons qu'un compact S de $R^n$ est caractérisé par la propriété que toute suite de points de l'ensemble S admet une valeur d'adhérence dans l'ensemble.
   Intuitivement, on comprend qu'un fermé borné vérifie une telle propriété : une suite de points de l'ensemble S ne peut pas partir "à l' infini" puisque l'ensemble est borné, elle doit donc s'accumuler quelque part dans $R^n$ et un point d'accumulation de la suite est nécessairement dans l'ensemble puisque l'ensemble S est fermé.

Dans le cas des problèmes d'optimisation sans contraintes *($S=R^n$),* on va utiliser le théorème suivant

***Théorème d'existence 2 :***

Soit $f : R^n \to R$ une fonction continue et coercive ( i.e. infini à l'infini : $\lim_{\|x\|\to+\infty} f(x) = +\infty$ ) alors $f$ admet au moins un minimum sur $R^n$. Autrement dit, il existe un point $\hat{x}$ de $R^n$ minimum global de $f$ sur $R^n$

$$\forall x \in R^n, \quad f(x) \geq f(\hat{x}).$$

***Preuve :***

Soit $(x_n)$ une suite minimisante dans $R^n$ c'est à dire

$$\lim_{n\to+\infty} f(x_n) = \inf_{x\in R^n} f(x)$$

Comme $f$ est coercive, la suite $(x_n)$ ne peut pas partir a l'infini donc elle est bornée et il existe une sous-suite extraite notée $(x_n)$ de $R^n$ qui converge vers un $\hat{x}$ de $R^n$ : $x_n \to \hat{x} \in R^n$.

Or $f$ est continue, donc

$$\lim_{n\to+\infty} f(x_n) = f(\hat{x})$$

d'où par unicité de la limite $\qquad f(\hat{x}) = \inf_{x\in R^n} f(x)$.

Donc ; $f$ réalise son minimum dans $R^n$.

***Théorème d'unicité :***

Si de plus $f : R^n \to R$ strictement convexe, alors il existe un unique minimum $\hat{x} \in R^n$ de $f$ tel que :

$$\forall x \in R^n, \quad f(x) \geq f(\hat{x}).$$

***Preuve :***

Soit $f$ strictement convexe, supposons qu'il existe $\hat{x}$, $\bar{x} \in R^n$ deux min de la fonction $f$ sur $R^n$ tels que

$$\bar{x} \neq \hat{x} \ \ et \ \ f(\hat{x}) = f(\bar{x}) = \inf_{x\in R^n} f(x).$$

Soit $\tilde{x} = \lambda\bar{x} + (1-\lambda)\hat{x} \in R^n$ avec $\lambda \in (0,1)$ et comme $f$ est strictement convexe

$$f(\tilde{x}) < \lambda f(\hat{x}) + (1-\lambda) f(\bar{x}) = f(\hat{x}) = \min_{x\in R^n} f(x).$$

Ceci fournit une contradiction, ce qui est impossible; donc $\hat{x} = \bar{x}$.

Nous pouvons maintenant énoncer un deuxième résultat d'existence et d'unicité

***Théorème 2 :***

Soit $f$ une fonction $C^1(R^n,R)$, on suppose que $f$ est $\alpha - elliptique$ c'est dire qu'il existe $\alpha > 0$ (appelée la constante d'ellipticité) tel que :

$$\forall (x,y) \in R^n \times R^n, \ (\nabla f(x) - \nabla f(y), x - y) \geq \alpha \|x-y\|^2.$$

Alors, $f$ est strictement convexe et coercive. En particulier le problème *(P)* admet une solution unique.

***Preuve :*** Exercice

Il faut maintenant donner des conditions d'optimalités pour pouvoir calculer la ( ou les) solutions, mais avant ça, on a besoin de quelques définitions.

## IV. 5.Direction de descente

Une direction de descente est une direction le long de laquelle la fonction à minimiser a une dérivée directionnelle strictement négative. Ces directions sont utilisées par les méthodes à directions de descente.

***Définitions3 :***

Soit $f$ : $R^n \to R$, $\hat{x} \in R^n$ et « d » un vecteur de $R^n/\{0\}$. Le vecteur « d » est dit de descente au point $\hat{x} \in R^n$ si et seulement si

$$\exists \delta > 0, \quad \forall \lambda \in \left]0, \delta\right[, \quad f(\hat{x} + \lambda d) < f(\hat{x}).$$

***Théorème :***

Soit $f$ : $R^n \to \mathbb{R}$ une fonction différentiable au point $\hat{x} \in R^n$ et $d \in R^n$ tel que

$$f'(\hat{x}, d) = (\nabla f(\hat{x}), d) < 0$$

Alors, 'd' est une direction de descente de $f$ en $\hat{x} \in R^n$.

***Preuve*** : Soit $f$ une fonction différentiable au point $\hat{x} \in R^n$, donc

$$f(\hat{x} + \lambda d) = f(\hat{x}) + \lambda.\nabla f(\hat{x})^t.d + \lambda.\|d\|.\alpha(\hat{x}, \lambda d), \quad \forall x \in S, \quad où \ \alpha(\hat{x}, \lambda d) \to 0 \ \ quand \ \ \lambda \to 0.$$

Donc
$$\frac{f(\hat{x} + \lambda d) - f(\hat{x})}{\lambda} = \nabla f(\hat{x})^t.d + \|d\|.\alpha(\hat{x}, \lambda d)$$

et

$$\lim_{\lambda \to 0} \frac{f(\hat{x} + \lambda d) - f(\hat{x})}{\lambda} = \nabla f(\hat{x})^t.d \Rightarrow f'(\hat{x}, d) = (\nabla f(\hat{x}), d)$$

et comme

$$(\nabla f(\hat{x}), d) < 0 \Rightarrow f'(\hat{x}, d) < 0,$$

donc :
$$\exists \delta > 0, \ \ tq \ \ \frac{f(\hat{x} + \lambda d) - f(\hat{x})}{\lambda} < 0, \forall \lambda \in \left]0, \delta\right[$$
$$\exists \delta > 0, \quad \forall \lambda \in \left]0, \delta\right[, \quad f(\hat{x} + \lambda d) < f(\hat{x}).$$

Donc le vecteur « d » est une direction de descente.

***Remarques :***

1- Par définition du gradient, il revient au même de dire que 'd' fait avec l'oppose du gradient ($-\nabla f(x)$) un angle θ, appelé ***angle de descente***, qui est strictement plus petit que 90◦ :

$$\theta = \arccos \frac{(-\nabla f(x), d)}{\|\nabla f(x)\|.\|d\|} \in \left]0, \frac{\pi}{2}\right[$$

La notion d'angle définie ci-dessus dépend du produit scalaire et n'est pas invariante par rotation des vecteurs. L'ensemble des directions de descente de $f$ en x,

$\{d \in R^n, (\nabla f(x), d) < 0\}$ forme un demi-espace ouvert de $R^n$.

2- Par définition de la dérivée, on voit que si *d* est une direction de descente

$$f(\hat{x}+\lambda d) < f(\hat{x}) \text{ pour tout } \lambda > 0 \text{ suffisamment petit.}$$

et donc que *f* décroit strictement dans la direction « d ». De telles directions sont intéressantes en optimisation car, pour faire décroitre *f*, il suffit de faire un déplacement le long de « d ». Les méthodes a directions de descente utilisent cette idée pour minimiser la fonction *f*.

## ***IV.6. Conditions d'optimalité :CO***

Dans cette section, nous allons chercher à obtenir des conditions nécessaires et parfois suffisantes de minimalité. L'objectif est d'une certaine manière beaucoup plus pratique puisque ces conditions d'optimalité seront le plus souvent utilisées pour tenter de calculer un minimum ou le maximum. Ces conditions vont donc s'exprimer à l'aide de la dérivée première ou seconde.

### ***IV.6.1.Condition d'optimalité nécessaires du premier ordre :CON1***

Les conditions que nous allons donné sont des conditions différentiables qui portent sur la dérivée de la fonction à minimiser.

***Théorème :***

Soit $f:\mathbb{R}^n \to \mathbb{R}$ une fonction différentiel au point $\hat{x}$ de $R^n$. Si $\hat{x}$ est un optimum local du problème *(P)*, alors $\nabla f(\hat{x}) = 0$

***Preuve :***

Supposons le contraire c'est-à-dire $\nabla f(\hat{x}) \neq 0$ donc on un vecteur $d = -\nabla f(\hat{x})$ qui est une direction de descente i.e $\exists \delta > 0,\ \forall \lambda \in ]0,\delta[,\ f(\hat{x}+\lambda d) < f(\hat{x})$.

et $\hat{x} \in R^n$ est un optimum local du *f* donc :

$$\exists V_\varepsilon(\hat{x}),\ \forall x \in V_\varepsilon(\hat{x}),\ f(x) \geq f(\hat{x}).$$

Soit $V_\varepsilon(\hat{x})$ un voisinage quelconque de $\hat{x}$ :

$$V_\varepsilon(\hat{x}) \cap [\hat{x}, \hat{x}+\lambda\nabla f(\hat{x})] \neq \phi,\ \ \lambda \in ]0,\delta[ \Rightarrow\ \exists\ x_\lambda = \hat{x}+\lambda d \in V_\varepsilon(\hat{x}),\ \ f(x_\lambda) < f(\hat{x}).$$

Contradiction avec le fait que $\hat{x} \in R^n$ est une solution optimale locale du (P).

On conclut que $\nabla f(\hat{x}) = 0$

***Remarque :***

1- La réciproque du théorème précédent n'est pas vraie.

2- Tout point $\hat{x} \in R^n$ vérifiant $\nabla f(\hat{x}) = 0$ est appelé « point critique » ou « point stationnaire ».

3- La relation $\nabla f(\hat{x}) = 0$ est aussi appelée « équation d'Euler ».

4- Ce théorème n'a pas de sens si la fonction *f* n'est pas différentiable.

5- Il faut bien noter que la condition du premier ordre ne fournit qu'une condition nécessaire d'optimalité. Autrement dit, l'application de ce critère fournira en général trop de points parmi lesquels il faudra déterminer la ou les solutions du problème *(P)* usuellement en comparant les valeurs de la fonction objective *f* en ces points.

6- Si *f* est convexe, la condition nécessaire du premier ordre est également suffisante pour que $\hat{x} \in R^n$ soit un minimum de $f$.

***Théorème :***

Soit $f$: $R^n \rightarrow R$ une fonction convexe et différentiel au point $\hat{x}$ de $R^n$. Un point $\hat{x}$ réalise un optimum de $f$ sur $R^n$ si et seulement si

$$\nabla f(\hat{x}) = 0$$

**Preuve :** On a vu que la condition est toujours nécessaire, montrons qu'elle est suffisante.

Soit $\hat{x} \in R^n$ tel que $\nabla f(\hat{x}) = 0$ et comme *f* est convexe donc

$$\forall x \in R^n, f(x) \geq f(\hat{x}) + (\nabla f(\hat{x}))^t (x - \hat{x}) \Rightarrow f(x) \geq f(\hat{x}),$$

On a donc immédiatement le fait que $\hat{x}$ réalise un minimum de *f* sur $R^n$:

Nous donnons maintenant une condition nécessaire permettant de préciser encore les éventuels minima. Cette condition va faire intervenir la dérivée seconde de *f*.

## ***IV.6.2.Condition d'optimalité nécessaires du deuxième ordre CON2***

Dans le cas où *f* est deux fois différentiable, une autre condition nécessaire est donnée par le théorème suivant :

***Théorème :***

Soit $f$ : $R^n \rightarrow R$ une fonction deux fois différentiable au point $\hat{x} \in R^n$ . Si *f* possède un minimum local en $\hat{x}$ alors :

1- $\nabla f(\hat{x}) = 0$ et

2- La matrice Hessienne $H(\hat{x})$ soit semi définie positive

***Preuve :***

Soit *f* une fonction deux fois différentiable au point $\hat{x} \in R^n$ et *d* un vecteur de $R^n$, donc pour $\lambda$ assez petit, on a

$$f(\hat{x}+\lambda d) = f(\hat{x}) + \lambda(\nabla f(\hat{x}), d) + \frac{1}{2}\lambda^2 .d^t H(\hat{x}).d + \lambda^2 .\|d\|.^2 \alpha(\hat{x}, \lambda d), \quad \forall x \in S, \quad où\ \alpha(\hat{x}, \lambda d) \to 0 \ \ quand \ \ \lambda \to 0.$$

le point $\hat{x} \in R^n$ est le minimum de $f$ c'est-à-dire $\nabla f(\hat{x}) = 0$

donc
$$\frac{f(\hat{x}+\lambda d) - f(\hat{x})}{\lambda^2} \geq 0, \forall \lambda \in ]0, \delta[$$
$$\Rightarrow \frac{1}{2} d^t H(\hat{x}).d + \|d\|.^2 \alpha(\hat{x}, \lambda d) \geq 0 \ , \forall \lambda \in ]0, \delta[$$

Par passage à la limite lorsque $\lambda \to 0 \Rightarrow \frac{1}{2} d^t H(\hat{x}).d \geq 0, \ \ \forall d \in R^n$

Donc, la matrice hessienne $H(\hat{x})$ est semi définie positive.

***Remarque :***

La réciproque du théorème précédent n'est pas vraie (pour la fonction $f(x)=x^3$, les conditions d'optimalité du premier et de deuxième ordre sont vérifiées pour *x=0 mais* « 0 » n'est pas un minimum local.

## ***IV.6.3.Condition d'optimalité suffisante du deuxième ordre COS2***

Les conditions données précédemment sont nécessaires, c'est-à-dire qu'elles doivent être satisfaites pour tout minimum local, cependant, tout vecteur vérifiant ces conditions n'est pas nécessairement un minimum local. Le théorème suivant établit une condition suffisante pour qu'un vecteur soit un minimum local, si f est deux fois continuellement différentiable.

***Théorème :***

Soit $f : R^n \to R$ une fonction deux fois différentiable au point $\hat{x} \in R^n$ . Si $\nabla f(\hat{x}) = 0$ et la matrice hessienne $H(\hat{x})$ est définie positive, alors *f* possède un minimum local en $\hat{x} \in R^n$.

***Preuve :***

Soit *f* une fonction deux fois différentiable au point $\hat{x} \in R^n$ , donc *f* s'écrit sous la forme

$$f(x) = f(\hat{x}) + (\nabla f(\hat{x}))^t (x-\hat{x}) + \frac{1}{2}(x-\hat{x})^t H(\hat{x})(x-\hat{x}) + \|x-\hat{x}\|^2 \alpha(\hat{x}, x-\hat{x}), \quad \forall x \in R^n,$$
$$où \ \ \alpha(\hat{x}, x-\hat{x}) \to 0 \ \ quand \ \ x \to \hat{x}.$$

Supposons que $\hat{x} \in R^n$ n'est pas un min local stricte c'est-à-dire $\forall V_\varepsilon(\hat{x})$ , $\exists x_k \in V_\varepsilon(\hat{x}), \ \ f(x_k) \leq f(\hat{x}).$

Notons $$d_k = \frac{x_k - \hat{x}}{\|x_k - \hat{x}\|^2},\quad donc : \|d_k\|^2 = 1\ ,$$

d'après le théorème de Bolzano Weierstrass

$$\exists N_1 \in N,\ \ d_k \underset{k\to+\infty}{\to} \tilde{d},\ \ k \in N_1$$

donc

$$f(x_k) - f(\hat{x}) = \frac{1}{2}(x_k - \hat{x})^t H(\hat{x})(x_k - \hat{x}) + \|x_k - \hat{x}\|^2 \alpha(\hat{x}, x_k - \hat{x}),\ \ ,$$

$$\frac{f(x_k) - f(\hat{x})}{\|x_k - \hat{x}\|^2} = \frac{1}{2} d_k{}^t H(\hat{x}).d_k + \alpha(\hat{x}, \lambda d)$$

Or $$f(x_k) \leq f(\hat{x}) \Rightarrow f(x_k) - f(\hat{x}) \leq 0,$$

donc $$\forall k,\ \ \frac{1}{2} d_k{}^t H(\hat{x}).d_k + \alpha(\hat{x}, x_k - \hat{x}) \leq 0,$$

passant à la limite $$k \to +\infty \Rightarrow \tilde{d}^t H(\hat{x}).\tilde{d} \leq 0,\ \ \tilde{d} \neq 0\ \ car\ d_k \to \hat{d}\ \ et\ \ \|d_k\| = 1$$

Donc, la matrice hessienne $H(\hat{x})$ n'est pas définie positive.

***Exemple :***

Considérons la fonction $f : R^2 \to$ R

$$f(x)=f(x_1,x_2)=6x_1^2+x_2^3+6x_1x_2+3x_2^2$$

Nous allons identifier tous les points critique de cette fonction, et allons utiliser la première condition d'optimalité afin de déterminer leur nature (min, max, ou ni min ni max). La fonction *f* est différentiable, alors

***Etape 1* : Identification des points critiques :**

La fonction *f* est différentiable et son gradient est: $\nabla f(x)=( 12x_1+6x_2, 3x_2^2+6x_1+6x_2)$

La condition nécessaire de premier ordre *:* $\nabla f(x_1,x_2)=(0,0)$ nous permet d'identifier deux points critique : $x=(0,0)^t$ *et* $x=(1/2,-1)^t$

***Etape 2* : Nature des points critiques**

La matrice hessienne de cette fonction est

$$H(x,y) = \begin{pmatrix} 12 & 6 \\ 6 & 6x_2 + 6 \end{pmatrix}$$

Les conditions de deuxième ordre nous permettent de classer ces points critiques

- Pour $x=(0,0)^t$ on a $det\ H(0,0)=36>0$ et ses valeurs propres sont positives.

  La matrice hesienne est alors définie positive, et donc $x=(0,0)^t$ est un minimum local.

- pour $x=((1/2),-1)^t$ don $det\ H((1/2),-1)=-36<0.$

  La matrice hesienne n'est pas semi-définie positive, et donc $x=((1/2),-1)^t$ n'est pas un minimum local.

**Remarques :**

Si $\nabla f(x) = 0$ et la matrice hessienne H (x) est indéfinie (n'est ni semi-définie positive, ni définie positive, ni semi-définie négative, ni définie négative). Alors, le point $x$ est appelé **point selle.**

***Optimisation classique sans contrainte***

On considère ici des fonctions $f$ définies dans ℝ (une fonction d'une variable réelle). Si la fonction $f$ et sa dérivée $f'$ sont continues en un point $x$ ou la croissante de la fonction devient décroissante, alors elle possède un maximum. Le raisonnement contraire est valable pour un minimum.

Il existe deux critère pour déterminer si un point candidat (le point qui annuler la première dérivée de $f$) est un point extrémum (minimum ou maximum).

**1-**Premier critère est:

-Si le signe de la dérivée est positif puis devient négatif quand $x$ croit, alors le point candidat est un maximum de la fonction.

-Si le signe de la dérivée est négatif puis devient positive quand x croit, alors le point candidat est un minimum de la fonction.

**2-**Deuxième critère :

fait appel à la dérivée seconde de la fonction $f(x)$. Soit $P(x_0;f(x_0))$ le point en lequel $f'(x_0)=0$. Alors si en ce point:

1- $f''(x_0)<0$, il s'agit d'un maximum.

2- $f''(x_0)>0$, il s'agit d'un minimum.

***Exemple :*** Soit la fonction $y=f(x)=x^3-3x^2+5$

$f$ est dérivable sur R, Calculons sa première et sa deuxième dérivée

$$f'(x)=3x^2-6x=3x(x-2),\ f''(x)==6x-6=6(x-1).$$

On obtient les points candidats de la fonction $f$ en résolvant l'équation

$$f'(x)=0 \text{ donc } 3x(x-2)=0 \Rightarrow x_1=0 \text{ et } x_2=2,$$

et les valeurs correspondantes de $y$ sont $y_1=5$ et $y_2=1$.

Déterminons la nature des points candidats en utilisant le résultat précédent

$$f''(0)=-6<0 \quad \text{et} \quad f''(2)=6>0.$$

La fonction a donc un maximum en $x_1=0$ et un minimum en $x_2=2$. Les coordonnées du maximum sont (0;5) et celles du minimum (2;1).

Par conséquent, dans un voisinage convenablement choisi du point $x_1=0$, la valeur de la fonction est plus petite que $f(x)$. On dit que la fonction possède un maximum local ou relatif au point $x_1=0$. Au point $x_2=2$, on dit que la fonction possède un minimum local ou relatif. Ces deux valeurs, maximum et minimum, sont appelées des extremums locaux ou relatifs.

**Exercices du chapitre 2**

**Exercice 1 :** ---------------------------------------------------------------------------------------------------

Déterminer les minima et les maxima, s'ils existent, de la fonction $f$ définie sur $IR^2$ par

1- $f(x_1, x_2) = ax_1^2 + bx_2^2$ , $(a, b \in R)$,  2- $f(x_1,x_2) = x_1^2 - 3x_1\, x_2 + x_1^2$

3- $f(x, y) = x/(1 + x^2 + y^2)$ ,  4- $f(x, y = x((\ln x)^2 + y^2)$

**Exercice 2 :**---------------------------------------------------------------------------------------------------

Dans le théorème d'existence du minimum (sans contraintes) d'une fonction *f,* montrer que l'on peut remplacer la propriété « infinie à l'infini » par la condition plus faible

$$\inf_{x \in R^n} f(x) \leq \lim_{r \to +\infty} (\inf_{\|x\| \geq r} f(x)),\ \ r > 0$$

**Exercice 3 :**---------------------------------------------------------------------------------------------------

Soit $f$ une fonction $C^1$ de $\mathbf{R^n}$ dans **R,** et (*P*) est le problème $\{ \min f(x), x \in R^n \}$, on suppose qu'il existe $\alpha > 0$ tel que : $\forall (x, y) \in R^n \times R^n,\ (\nabla f(x) - \nabla f(y), x - y) \geq \alpha \|x - y\|^2$ (on dit que $f$ est $\alpha$ -elliptique)

1- Montrer que : $f(y) - f(x) \geq (\nabla f(x), y - x) + \frac{\alpha}{2} \|x - y\|^2,\ \forall (x, y) \in R^n \times R^N$
2- Montrer que $f$ est strictement convexe et coercive.
3- Déduire que le problème *(P)* admet une solution unique.
4- On suppose que $f$ est deux fois différentiable en tous points de $R^n$, montrer que $f$ est $\alpha$ -elliptique ssi : $\forall (x, y) \in R^n \times R^n,\ (D^2 f(x).y, y) \geq \alpha \|y\|^2$

**Exercice 4 :** ---------------------------------------------------------------------------------------------------

Soient *u* et *v* deux vecteurs non de $R^2$. On considère la fonction *g* de *R* dans *R*, telle que

$$\forall (x, y) \in R^2,\ \ g(x, y) = \|xu + yv\|$$

Montrer que g est minorée par un réel strictement positif *m* sur l'ensemble $C = \{(x, y) \in R^2,\ \ x^2 + y^2 = 1\}$

En déduire que $\forall (x, y) \in R^2,\ \ g(x, y) \geq m\sqrt{x^2 + y^2}$

2. Soit $(p_1,_{\dots}\, p_k)$ une suite de *k* points de $R^2$ avec k > 2. On pose $p_i = (\alpha_i, \beta_j)$ et on suppose que la suite $(\alpha_i)$ est strictement croissante. On considère alors la fonction $f$ *de* $R^2$ telle que

$$\forall (x, y) \in R^2,\ \ f(x, y) = \sum_{i=1}^{k} (\beta_i - x\alpha_i - y)^2$$

-Montrer que $f$ est coercive. En déduire qu'il existe $(x_0; y_0)$ dans $R^2$ tel que : $f(x_0, y_0) = \min_{(x,y) \in R^2} f(x, y)$

3. Calculer un tel $(x_0; y_0)$ et montrer qu'il est unique.

**Exercice 5 :** ------------------------------------------------------------------------------------------------

1- Soient $p_1 = 52,\ \ p_2 = 44$ les prix respectifs de deux produits. Soient $q_1,\ \ q_2$ les quantités respectives de ces produits. La fonction coût est $C = q_1^2 + q_1 q_2 + q_2^2$ .
Trouver les quantités $q_1,\ q_2$ maximisant le bénéfice.

2- Un fabricant de postes de télévision produit $q$ postes par semaine à un coût total :
$C=6q^2+80q+5000$. C'est un monopoleur et son prix s'exprime par la relation $P=1080-4q$
Montrons que le bénéfice net maximum est atteint lorsque la production est de 50 postes par semaine.

**Exercice 6 :** -----------------------------------------------------------------------------------------------------

On considère la fonction $f$ définie sur $R^3$ par : $f(x,y,z) = 3x^2 + 3y^2 + 3z^2 - 2xy - 2x - 10y$

1. Montrer que $f$ est coercive sur $R^3$. Que peut-on en déduire ?

2. Calculer les points où $f$ est minimale sur $R^3$.

**Exercice 7:** -----------------------------------------------------------------------------------------------------

On considère le problème suivant : Trouver $u \in R^n$ tel que $\|Bu - C\| = \inf\limits_{v \in R^n} \|Bv - C\|$ *(1)*

où $B$ est une matrice réelle de taille $m * n$, $c \in R^n$ et $\|.\|$ désigne la norme euclidienne sur $R^n$.

Dans la suite on note aussi par (., .) le produit scalaire sur $R^n$.

1- Montrer que *(1)* admet toujours au moins une solution.
2- Montrer que cette solution est unique si et seulement si $B$ est injective. Quelle est la Condition nécessaire sur $m$ et $n$ pour que $B$ soit injective ?

3- Montrer que le problème *(1)* est équivalent au problème *(2)* ci-dessous :
Trouver $u \in R^n$ tel que $f(u) = \inf\limits_{v \in R^n} f(v)$ *(2)*

où on a posé $f(v) = 1/2\ (B^t Bv, v) - (B^t c, v)$.

4- La fonction $f$ est-elle différentiable ?
5- Si cela est possible, calculer le gradient et la hessienne de $f$.
6- Donner la condition d'optimalité au premier ordre pour le problème (2).
Cette condition est-elle suffisante ici pour obtenir un minimum?

7- Dans le cas où $B$ est de rang $n$, montrer que $f$ est strictement convexe. Que peut-on alors conclure ?

**Exercice 8:** ------------------------------------------------------------------------------------------------

On considère la fonction $f$ de $R^n$ dans R définie par la relation :

$$\forall x \in R^n,\ f(x) = \sin(\|x\|^2) = \sin\left(\sum_{i=1}^{n} x^2{}_i\right)$$

1. Montrer que la fonction $f$ est différentiable sur $R^n$ et calculer son gradient.
2. Montrer que la fonction $f$ est $C^2$ et calculer sa matrice Hessienne.
3. Déterminer tous les extrema (minima et maxima) de $f$ sur $R^n$. Représenter graphiquement ces extrema dans le cas où n = 2.

**Exercice 09 :**-----------------------------------------------------------------------------------------------------

Soit $f : R^3 \to R, f(x,y,z) = x^2 + y^2 + z^2 + xy + yz + xz - 3x - 4y - z + 4$

1- Mettre $f$ sous la forme d'une fonction quadratique $f(x) = \frac{1}{2}(Ax,x) - (b,x) + c$ avec $A$ est une matrice symétrique, $b \in R^3,\ c \in R.$

2- Soient $\lambda_{\min}$ *et* $\lambda_{\max}$ la plus petite et la plus grande valeur propre de $A$. Montrer que

$$\forall x \in R^3,\ \ \lambda_{\min} \|x\|_2^2 \le (Ax,x) \le \lambda_{\max} \|x\|_2^2,$$

3- Montrer que si $A$ est définie positive alors $f$ « infinie à l'infini » et admet un minimum unique sur $R^3$

4- En déduire que si $A$ est non positive alors $f$ n'admet pas de minimum sur $R^3$

**Exercice 10:** -----------------------------------------------------------------------------------------------------

On se propose d'approcher un nuage de points donnés par les couples de réels $(t_i,x_i)$, $i=1,\dots,N$ par une parabole d'équation $x(t)=at^2 +bt+c$ ou (a,b,c) sont trois réels à déterminer.

1- Exprimer le problème ci-dessus sous forme de problème de minimisation.
2- Ce problème de minimisation a-t-il une solution ? est-elle unique ?
3- Ecrire la condition d'optimalité permettant de trouver le minimum

**Exercice 11:** -----------------------------------------------------------------------------------------------------

A- On considère la fonction $g$ de $R^2$ dans R définie par la relation :

$$\forall (x,y) \in R^2 ;\ g(x; y) = x^2 + y^2 + xy$$

1. Montrer que la fonction g est coercive et strictement convexe sur $R^2$.
2. En déduire que $g$ possède un unique minimum sur $R^2$.
3. Donner la condition d'optimalité.

**Exercice 12:** -----------------------------------------------------------------------------------------------------

I- Soit $f$ la fonction numérique à variable réelle définie par :

$$f(x) = \sqrt{x^4 + 1} - x + 3$$

1. Montrer que $f$ admet un minimum sur R et Déterminer les points critiques de $f$.

3. Parmi les points critiques, lesquels correspondent à des minima ?

II- Soient $p$ et $q$ deux réels strictement positifs tels que : $\frac{1}{p} + \frac{1}{q} = 1$. En utilisant la concavité de la fonction logarithme, montrer que:

$$\forall x, y \in R_+^*,\ \ xy \le \frac{x^p}{p} + \frac{x^q}{q}.$$