V EXTREMUM

▸ 1. Les définitions de base

★ Il convient d'avoir une parfaite connaissance des définitions qui suivent

Déf. 37 f est une application d'une partie D de $\mathbb{R}^n$ dans $\mathbb{R}$.

1. f possède un **maximum sur D** s'il existe un élément A de D tel que : $\forall X \in D,\ f(X) \leqslant f(A)$.

Dans ce cas $\boxed{\text{le}}$ maximum de f sur D est $f(A)$, c'est le plus grand élément de $f(D)$ et on le note : $\underset{X\in D}{\text{Max}}\, f(X)$. On parle encore de **maximum global** ou de maximum absolu de f sur D.

2. f possède un **minimum sur D** s'il existe un élément A de D tel que : $\forall X \in D,\ f(X) \geqslant f(A)$.

Dans ce cas $\boxed{\text{le}}$ minimum de f sur D est $f(A)$, c'est le plus petit élément de $f(D)$ et on le note : $\underset{X\in D}{\text{Min}}\, f(X)$. On parle encore de **minimum global** ou de minimum absolu de f sur D.

★★★ La situation est celle de la définition précédente. Si f possède un maximum (resp. minimum) sur D celui-ci est **UNIQUE**. Néanmoins il peut exister plusieurs points de D (et même une infinité) qui "réalisent" ce maximum (resp. minimum).

S'il n'existe qu'un point de D qui réalise le maximum (resp. minimum) de f sur D on parle de maximum (resp. minimum) strict.

★ On peut étendre ces définitions au cas f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ et D une partie de son domaine de définition.

Déf. 38 f est une application d'une partie D de $\mathbb{R}^n$ dans $\mathbb{R}$ et A est un point de D.

1. f possède un **maximum local en A** s'il existe un réel strictement positif r tel que :

$$\forall X \in B(A,r)\cap D,\ f(X) \leqslant f(A) \quad \text{ou} \quad \forall H \in B(O,r),\ A+H \in D \Rightarrow\ f(A+H) \leqslant f(A).$$

2. f possède un **minimum local en A** s'il existe un réel strictement positif r tel que :

$$\forall X \in B(A,r)\cap D,\ f(A) \leqslant f(X) \quad \text{ou} \quad \forall H \in B(O,r),\ A+H \in D \Rightarrow\ f(A) \leqslant f(A+H).$$

3. f possède un **extremum local** en A si f possède un maximum ou un minimum local en A.

★★ S'il est important de savoir écrire que f possède un maximum ou un minimum local en A il est tout aussi important de savoir écrire que f ne possède pas d'extremum local en A. D'où la proposition suivante.

Prop. 17 **PP** f est une application d'une partie D de $\mathbb{R}^n$ dans $\mathbb{R}$ et A est un point de D. Les assertions suivantes sont équivalentes.

i) f ne possède pas d'extremum local en A ;

ii) **Pour toute boule de centre A** il existe deux éléments X_1 et X_2 appartenant à cette boule tels que $f(X_1) > f(A)$ et $f(X_2) < f(A)$;

ii') $\forall r \in \mathbb{R}^{+*},\ \exists X_1 \in B(A,r),\ \exists X_2 \in B(A,r),\ f(X_1) > f(A)$ et $f(X_2) < f(A)$;

iii) $\forall r \in \mathbb{R}^{+*},\ \exists H_1 \in B(O,r),\ \exists H_2 \in B(O,r),\ f(A+H_1) > f(A)$ et $f(A+H_2) < f(A)$.

P Soit f une application d'un ouvert Ω, de $\mathbb{R}^n$, dans $\mathbb{R}$.

- f possède un maximum local en A s'il existe un réel strictement positif r tel que : $\forall X \in B(A,r),\ f(X) \leqslant f(A)$.
- f possède un minimum local en A s'il existe un réel strictement positif r tel que : $\forall X \in B(A,r),\ f(A) \leqslant f(X)$.

Déf. 39 f est une application d'une partie D de $\mathbb{R}^n$ dans $\mathbb{R}$ et A est un point de D.

1. f possède un **maximum local strict en A** s'il existe un réel strictement positif r tel que :

$$\forall X \in B(A,r) \cap D - \{A\},\ f(X) < f(A) \quad \text{ou} \quad \forall H \in B(O,r)\{O\},\ A+H \in D \Rightarrow f(A+H) < f(A).$$

2. f possède un **minimum local strict en A** s'il existe un réel strictement positif r tel que :

$$\forall X \in B(A,r) \cap D - \{A\},\ f(A) < f(X) \quad \text{ou} \quad \forall H \in B(O,r) - \{O\},\ A+H \in D \Rightarrow f(A) < f(A+H).$$

3. f possède un **extremum local strict** en A si f possède un maximum ou un minimum local strict en A.

▶ 2. Rappels : borne supérieure et inférieure, fonction majorée, minorée...

Déf. 40 Soit P une partie de $\mathbb{R}$.

La **borne supérieure** de P est, le plus petit élément, s'il existe, de l'ensemble des majorants de P ; notation $\text{Sup}\, P$.

La **borne inférieure** de P est le plus grand élément, s'il existe, de l'ensemble des minorants de P ; notation $\text{Inf}\, P$.

Th. 53
- Toute partie non vide et majorée de $\mathbb{R}$ possède une borne supérieure.
- Toute partie non vide et minorée de $\mathbb{R}$ possède une borne inférieure.

Déf. 41 f une application d'une partie de $\mathbb{R}^n$ dans $\mathbb{R}$ et D est une partie du domaine de définition de f.

f est majorée sur D si $f(D)$ est une partie majorée de $\mathbb{R}$; autrement dit s'il existe un élément M de $\mathbb{R}$ tel que : $\forall X \in D,\ f(X) \leqslant M$.

f est minorée sur D si $f(D)$ est une partie minorée de $\mathbb{R}$; autrement dit s'il existe un élément m de $\mathbb{R}$ tel que : $\forall X \in D,\ m \leqslant f(X)$.

f est bornée sur D si f est majorée et minorée sur D ou si $|f|$ est majorée sur D.

Déf. 42 f une application d'une partie de $\mathbb{R}^n$ dans $\mathbb{R}$ et D est une partie non vide du domaine de définition de f.

Si f est majorée sur D, la borne supérieure de $f(D)$ s'appelle **la borne supérieure de f sur D** et se note $\underset{X \in D}{\text{Sup}}\, f(X)$.

Si f est minorée sur D, la borne inférieure de $f(D)$ s'appelle **la borne inférieure de f sur D** et se note $\underset{X \in D}{\text{Inf}}\, f(X)$.

★ Les hypothèses sont celles de la définition précédente.

Si f possède un maximum (resp. minimum) sur D : $\underset{X \in D}{\text{Sup}}\, f(X) = \underset{X \in D}{\text{Max}}\, f(X)$ (resp. $\underset{X \in D}{\text{Inf}}\, f(X) = \underset{X \in D}{\text{Min}}\, f(X)$).

▶ 3. Fonction continue sur un fermé borné.

Déf. 43 Une partie D de $\mathbb{R}^n$ est **bornée** s'il existe un réel M tel que $\forall X \in D,\ \|X\| \leqslant M$.

Th. 54 $\boxed{\textbf{1.}}$ $[a,b]$ est un $\boxed{\text{segment}}$ de $\mathbb{R}$ et f est une fonction numérique de la variable réelle $\boxed{\text{continue}}$ sur [a,b].
f possède un maximum et un minimum sur $[a,b]$. Autrement dit il existe deux éléments c et e dans $[a,b]$ tels que :

$$\forall t \in [a,b],\ f(e) \leqslant f(t) \leqslant f(c)$$

Notons que f est bornée sur $[a,b]$ et atteint ses bornes ;

$\boxed{\textbf{2.}}$ D est une partie non vide $\boxed{\text{fermée et bornée}}$ de $\mathbb{R}^n$ et f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ $\boxed{\text{continue}}$ sur D.
f possède un maximum et un minimum sur D. Autrement dit il existe deux éléments C et E de D tels que :

$$\forall X \in D,\ f(E) \leqslant f(X) \leqslant f(C)$$

Notons que f est bornée sur D et atteint ses bornes ;

$\boxed{\textit{Remarques}}$ 1. Ce théorème est fondamental et contient un résultat d'**existence** déterminant dans beaucoup de questions.

2. Pour nous dans la pratique il nous donnera l'existence d'un maximum et d'un minimum. Il restera à les déterminer et à savoir là où ils sont atteints.

▶ 4. Extremum : une condition nécessaire

Th. 55 f est une application d'un $\boxed{\textbf{OUVERT}}$ Ω de $\mathbb{R}^n$ dans $\mathbb{R}$ et A un point de Ω.

On suppose que f $\boxed{\text{admet des dérivées partielles premières en } A}$ par rapport à toutes les variables.

$\boxed{\text{Si}}$ f possède un extremum local en A $\boxed{\text{alors}}$:

$$\forall i \in [\![1,n]\!],\ \frac{\partial f}{\partial x_i}(A) = 0 \qquad \text{ou} \qquad \nabla f(A) = 0_{\mathbb{R}^n}.$$

★★ Toutes les hypothèses de ce théorème sont importantes. Il convient donc de les vérifier proprement et rapidement avant d'en utiliser la conclusion.

Il est clair que cette condition est nécessaire mais pas suffisante. Prendre $A = (0,0)$ et $f : (x,y) \to xy$ pour s'en convaincre.

Déf. 44 f est une application d'un ouvert Ω de $\mathbb{R}^n$ dans $\mathbb{R}$.

On appelle **point critique de f** tout point A de Ω tel que toutes les dérivées partielles premières de f en A existent et sont nulles ; autrement dit un point A de Ω tel que $\nabla f(A)$ existe et vaut $0_{\mathbb{R}^n}$.

Déf. 45 f est une application d'un ouvert Ω de $\mathbb{R}^n$ dans $\mathbb{R}$.

On appelle **point selle ou point colle de f** tout point critique de f où f ne possède pas d'extremum local.

▸ 5. Extremum : une condition suffisante

Th. 56 Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$ et A un point de cet ouvert.

q_A est la forme quadratique de $\mathbb{R}^n$ associée à la Hessienne de f en A.

On suppose que $\boxed{A \text{ est un point critique de } f}$.

1) Si : $\forall H \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\},\ q_A(H) > 0$, alors f admet un minimum local (strict) en A.

2) Si : $\forall H \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\},\ q_A(H) < 0$, alors f admet un maximum local strict en A.

3) Si : $\exists (H_1, H_2) \in (\mathbb{R}^n)^2,\ q_A(H_1) > 0$ et $q_A(H_2) < 0$ alors f n'admet pas d'extremum local en A.

4) Si : $\forall H \in \mathbb{R}^n, q_A(H) \geqslant 0$ sans plus... (resp. $\leqslant 0$), la conclusion n'est pas immédiate...

★ Attention le programme n'exige que les deux premiers points.

★ Pour étudier le signe de q_A on peut :

• décomposer cette forme quadratique comme combinaison linéaire de carrés de formes linéaires indépendantes ;

• utiliser le signe des valeurs propres de la Hessienne de f en A.

Les résultats du complément 1 précisent le second point et il faut savoir les démontrer.

Th. 57 **le cas $\mathbf{n = 2}$**. Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^2$ et A un point de cet ouvert.

On suppose que A est un point critique de f et on pose : $r = \dfrac{\partial^2 f}{\partial x^2}(A) \quad s = \dfrac{\partial^2 f}{\partial x \partial y}(A) \quad t = \dfrac{\partial^2 f}{\partial y^2}(A)$.

1. Si $rt - s^2 > 0$ et $r > 0$, f admet un minimum local strict en A.
2. Si $rt - s^2 > 0$ et $r < 0$, f admet un maximum local strict en A.
3. Si $rt - s^2 < 0$, f n'admet pas d'extremum local en A.
4. Si $rt - s^2 = 0$ la conclusion n'est pas immédiate...

★ Attention le programme exige les quatre points !

Dans quelques livres on raisonne sur $s^2 - rt$ ce qui change légèrement (!) le théorème.

▸ 6. Position du graphe par rapport à l'hyperplan tangent

Th. 58 Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$ et A un point de cet ouvert.

q_A est la forme quadratique de $\mathbb{R}^n$ associée à la Hessienne de f en A.

1) Si : $\forall H \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\},\ q_A(H) > 0$, alors au voisinage de A le graphe de f est au-dessus de son hyperplan tangent en A.

2) Si : $\forall H \in \mathbb{R}^n - \{0_{\mathbb{R}^n}\},\ q_A(H) < 0$, alors au voisinage de A le graphe de f est au-dessous de son hyperplan tangent en A.

▶ 7. "Rappel" : signe d'une forme quadratique.

Th. 59 $\boxed{\text{SD}}$ Soit q la forme quadratique sur $\mathbb{R}^n$ associée à une matrice symétrique S de $\mathcal{M}_n(\mathbb{R})$.

1. Si toutes les valeurs propres de S sont positives ou nulles : $\forall X \in \mathbb{R}^n,\ q(X) \geqslant 0$.

2. Si toutes les valeurs propres de S sont négatives ou nulles : $\forall X \in \mathbb{R}^n,\ q(X) \leqslant 0$.

3. Si toutes les valeurs propres de S sont strictement positives : $\forall X \in \mathbb{R}^n,\ X \neq 0_{\mathbb{R}^n} \Rightarrow q(X) > 0$.

4. Si toutes les valeurs propres de S sont strictement négatives : $\forall X \in \mathbb{R}^n,\ X \neq 0_{\mathbb{R}^n} \Rightarrow q(X) < 0$.

5. Si S possède au moins une valeur propre strictement positive et au moins une valeur propre strictement négative, il existe deux éléments (non nuls) X_1 et X_2 de $\mathbb{R}^n$ tels que $q(X_1) > 0$ et $q(X_2) < 0$.

Prop. 18 $\boxed{\text{SD}}$ Soit q la forme quadratique sur $\mathbb{R}^n$ associée à une matrice symétrique S de $\mathcal{M}_n(\mathbb{R})$.

Les assertions suivantes sont équivalentes.

i) $\forall X \in \mathbb{R}^n,\ q(X) \geqslant 0$ (resp. $\leqslant 0$).

ii) Il existe une matrice M de $\mathcal{M}_n(\mathbb{R})$ telle que $S = {}^tMM$ (resp. $S = -\,{}^tMM$).

Prop. 19 $\boxed{\text{SD}}$ Soit q la forme quadratique sur $\mathbb{R}^n$ associée à une matrice symétrique S de $\mathcal{M}_n(\mathbb{R})$.

Les assertions suivantes sont équivalentes.

i) $\forall X \in \mathbb{R}^n,\ X \neq 0_{\mathbb{R}^n} \Rightarrow q(X) > 0$ (resp. < 0)

ii) Il existe une matrice inversible M de $\mathcal{M}_n(\mathbb{R})$ telle que $S = {}^tMM$ (resp. $S = -\,{}^tMM$).

▶ 8. Extremum : la condition suffisante again

Th. 60 Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$ et A un point de cet ouvert.

q_A est la forme quadratique de $\mathbb{R}^n$ associée à la Hessienne de f en A.

On suppose que $\boxed{A \text{ est un point critique de } f}$.

1) Si toutes les valeurs propres de $\nabla^2 f(A)$ sont strictement positives, alors f admet un minimum local (strict) en A.

2) Si toutes les valeurs propres de $\nabla^2 f(A)$ sont strictement négatives, alors f admet un maximum local (strict) en A.

3) Si une des valeurs propres de $\nabla^2 f(A)$ est strictement positive et si une est strictement négative alors f n'admet pas d'extremum local en A.

4) Si toutes les valeurs propres de $\nabla^2 f(A)$ sont positives (ou nulles) ou si toutes les valeurs propres de $\nabla^2 f(A)$ sont négatives (ou nulles), la conclusion n'est pas immédiate...

▶ 9. Complément : une condition suffisante intéressante (mais de mammouth).

Th. 61 $\boxed{\text{SD}}$ Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert Ω de $\mathbb{R}^n$.

On suppose que $\forall B \in \Omega,\ \forall X \in \mathbb{R}^n,\ q_B(X) \geqslant 0$ (resp. $\leqslant 0$).

Si A est un point critique de f alors, f possède un minimum (resp. maximum) local en A.

★ Ceci est une conséquence simple de l'égalité de Taylor-Lagrange à l'ordre 1.

Th. 62 $\boxed{\textbf{SD}}$ Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert $\boxed{\textbf{CONVEXE}}$ Ω de $\mathbb{R}^n$.

On suppose que $\forall B \in \Omega,\ \forall X \in \mathbb{R}^n,\ q_B(X) \geqslant 0$ (resp. $\leqslant 0$).

Si A est un point critique de f alors, f possède un minimum (resp. maximum) global en A.

▸ 10. Complément : le cas des fonctions convexes ou concaves.

Th. 63 Soit f une fonction de classe $\mathcal{C}^1$ sur un ouvert convexe Ω de $\mathbb{R}^n$. Soit A un point critique de f.

Si f est convexe sur Ω, f possède un minimum global en A.

Si f est concave sur Ω, f possède un maximum global en A.

Th. 64 Soit f une fonction de classe $\mathcal{C}^2$ sur un ouvert convexe Ω de $\mathbb{R}^n$. Soit A un point critique de f.

f est convexe sur Ω si et seulement si $\forall A \in \Omega,\ \forall X \in \mathbb{R}^n,\ q_A(X) \geqslant 0$. f est concave sur Ω si et seulement si $\forall A \in \Omega,\ \forall X \in \mathbb{R}^n,\ q_A(X) \leqslant 0$.

▸ 11. Pour finir.

$\boxed{\textbf{PP}}$ Soit Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^2$.

Pour étudier les extremums de f il convient dans une première étape de déterminer l'ensemble des points critiques de f.

Ceci étant fait pour chaque point A de cet ensemble on examine si f possède en ce point un extremum.

Deux possibilités (au moins) nous sont offertes :

- Utiliser les résultats du paragraphe précédent (c'est à dire les conditions d'ordre 2).

- Etudier à la main le signe de $f(X) - f(A)$ au voisinage de A ou le signe de $f(A+H) - f(A)$ au voisinage de O.

VI EXTREMUM SOUS CONTRAINTES

▶ 1. Introduction

Considérons un gérant de portefeuilles. Un client s'adresse à lui pour se constituer un portefeuille à risque minimum.

Le gérant a le choix entre n actions A_1, A_2, ..., A_n

Pour tout élément k de $[\![1, n]\!]$, on considère que la variation de l'action A_k pendant une période T est une variable aléatoire X_k d'espérance m_k et d'écart-type σ_k. On suppose que les variables aléatoires X_1, X_2, ..., X_n sont indépendantes.

Le gérant doit donc constituer un portefeuille dont la variance est minimum.

Son problème est donc de déterminer, pour tout k dans $[\![1, n]\!]$, la proportion x_k d'action A_k à mettre dans le portefeuille.

Il doit donc trouver x_1, x_2, ..., x_n dans $]0, +\infty[$ tel que $V(x_1\,X_1 + x_2\,X_2 + \cdots + x_n\,X_n)$ soit minimum et tel que $x_1 + x_2 + \cdots + x_n = 1$.

Or $V(x_1\,X_1 + x_2\,X_2 + \cdots + x_n\,X_n) = x_1^2\,\sigma_1^2 + x_2^2\,\sigma_2^2 + \cdots + x_n^2\,\sigma_n^2$ (indépendance des variables X_1, X_2, ..., X_n).

Il doit donc minimiser $x_1^2\,\sigma_1^2 + x_2^2\,\sigma_2^2 + \cdots + x_n^2\,\sigma_n^2$ sous la contrainte $x_1 + x_2 + \cdots + x_n = 1$.

Supposons en plus que le client espère tirer un rendement égal à M pendant une période T donnée.

Notons que $E(x_1\,X_1 + x_2\,X_2 + \cdots + x_n\,X_n) = x_1\,m_1 + x_2\,m_2 + \cdots + x_n\,m_n$.

Le gérant doit alors minimiser $x_1^2\,\sigma_1^2 + x_2^2\,\sigma_2^2 + \cdots + x_n^2\,\sigma_n^2$ sous les contraintes $\begin{cases} x_1 + x_2 + \cdots + x_n = 1 \\ x_1\,m_1 + x_2\,m_2 + \cdots + x_n\,m_n = M \end{cases}$.

Dans les deux cas il ne s'agit pas simplement d'étudier les extremums d'une fonction numérique f de n variables mais de trouver les extremums de la restriction de cette fonction à une partie $\mathcal{C}$ de $\mathbb{R}^n$.

Ici la fonction f est définie sur l'ouvert $\Omega = (]0, +\infty[)^n$ par :

$$\forall X = (x_1, x_2, \ldots, x_n) \in \Omega,\ f(X) = x_1^2\,\sigma_1^2 + x_2^2\,\sigma_2^2 + \cdots + x_n^2\,\sigma_n^2.$$

Dans le premier cas $\mathcal{C} = \{X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid x_1 + x_2 + \cdots + x_n = 1\}$.

Dans le second : $\mathcal{C} = \{X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \mid x_1 + x_2 + \cdots + x_n = 1 \text{ et } m_1\,x_1 + m_2\,x_2 + \cdots + m_n\,x_n = M\}$.

Posons : $\forall X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n,\ g_1(X) = x_1 + x_2 + \cdots + x_n$.

Posons encore : $\forall X = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n,\ g_2(X) = m_1\,x_1 + m_2\,x_2 + \cdots + m_n\,x_n$.

Dans le premier cas $\mathcal{C} = \{X \in \mathbb{R}^n \mid g_1(X) = 1\}$ dans le second $\mathcal{C} = \{X \in \mathbb{R}^n \mid g_1(X) = 1 \text{ et } g_2(X) = m\}$.

Observons que g_1 et g_2 sont deux formes linéaires sur $\mathbb{R}^n$.

On peut déjà ébaucher une méthode de résolution du problème reposant sur la substitution.

Dans le premier cas la contrainte permet d'exprimer x_n en fonction de x_1, x_2, ..., x_{n-1}. En réinjectant dans f on se ramène à l'optimisation d'une fonction de $n - 1$ variables.

Dans le second cas la contrainte doit permettre d'exprimer deux variables en fonction des autres. En réinjectant dans f on se ramène à l'optimisation d'une fonction de $n - 2$ variables.

Les quelques lignes qui suivent ont pour but de donner un cadre général au problème précédent et de dégager, ici encore, une condition nécessaire pour qu'un point soit solution du problème.

▶ 2. La notion d'extremum sous contrainte d'égalités linéaires.

Déf. 46 Ω est un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$.

p est un élément de $\mathbb{N}^*$, g_1, g_2, ..., g_p sont p formes linéaires sur $\mathbb{R}^n$ et b_1, b_2, ..., b_p sont p réels.

$\mathcal{C}$ est l'ensemble des éléments X de $\mathbb{R}^n$ tels que $\begin{cases} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \dots\dots\dots \\ g_p(X) = b_p \end{cases}$

A est un point de $\Omega \cap \mathcal{C}$. On dit que f **possède un maximum (resp. minimum) local sous la contrainte $\mathcal{C}$ en A** si la restriction de f à $\mathcal{C}$ possède un maximum (resp. minimum) local en A ; autrement dit s'il existe un réel r strictement positif tel que :

$$\forall X \in \mathcal{C} \cap B(A,r),\ f(X) \leqslant f(A) \quad (\text{resp. } f(A) \leqslant f(X)).$$

★ On peut de toute évidence définir les notions d'extremum local, de maximum local strict, de minimum local strict, d'extremum local strict, de maximum global, de minimum global,... sous la contrainte $\mathcal{C}$.

▶ 3. La condition nécessaire d'existence d'extremum sous contrainte d'égalités linéaires.

Th. 65 Ω est un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$ sur Ω.

p est un élément de $\mathbb{N}^*$, g_1, g_2, ..., g_p sont p formes linéaires sur $\mathbb{R}^n$ et b_1, b_2, ..., b_p sont p réels.

$\mathcal{C}$ est l'ensemble des éléments $X = (x_1, x_2, \dots, x_n)$ de $\mathbb{R}^n$ tels que $\begin{cases} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \dots\dots\dots \\ g_p(X) = b_p \end{cases}$.

On pose : $\mathcal{H} = \{X \in \mathbb{R}^n \mid g_1(X) = g_2(X) = \cdots = g_p(X) = 0\} = \operatorname{Ker} g_1 \cap \operatorname{Ker} g_2 \cap \ldots \cap \operatorname{Ker} g_p$.

1. $\mathcal{H}$ est un sous-espace vectoriel de $\mathbb{R}^n$.

2. Si $\mathcal{C}$ n'est pas vide alors $\mathcal{C}$ est un sous-espace affine de direction $\mathcal{H}$.

3. **PP** A est un point de $\Omega \cap \mathcal{C}$. Si f possède un extremum local sous la contrainte $\mathcal{C}$ en A alors le gradient $\nabla f(A)$ de f en A est orthogonal à $\mathcal{H}$.

Prop. 20 Les hypothèses sont celles du théorème précédent

1. Pour tout i dans $[\![1, p]\!]$, le gradient de g_i en un point de $\mathbb{R}^n$ ne dépend pas de ce point ($X \to \nabla g_i(X)$ est constante sur $\mathbb{R}^n$).

2. L'orthogonal de $\mathcal{H}$ est le sous-espace vectoriel engendré par la famille $\big(\nabla g_1(X), \nabla g_2(X), \dots, \nabla g_p(X)\big)$ où X est un élément quelconque de $\mathbb{R}^n$.

3. A est un point de $\Omega \cap \mathcal{C}$. Si f possède un extremum local sous la contrainte $\mathcal{C}$ en A alors, pour tout élément H (non nul) de $\mathcal{H}$, la dérivée de f en A dans la direction de H est nulle ($f'_H(A) = 0$).

Déf. 47 Les hypothèses sont celles du théorème précédent. Un point A de $\Omega \cap \mathcal{C}$ tel que $\nabla f(A)$ soit orthogonal à $\mathcal{H}$ est appelé **point critique de f dans l'optimisation sous la contrainte $\mathcal{C}$.**

★ **P** **Pratiquement** Les hypothèses sont celles des résultats précédents. Pour résoudre le problème il convient d'abord de chercher les points A de $\mathcal{C}$ tel que $\nabla f(A)$ soit orthogonal $\mathcal{H}$, c'est à dire les points critiques de f dans l'optimisation sous la contarinte $\mathcal{C}$.

Dans une deuxième étape on étudie si ces points sont solutions du problème. Pour ce faire, si A est un tel point, on regarde si le signe de $f(X) - f(A)$ est constant pour X dans $\mathcal{C}$ et dans un voisinage de A.

Expression matricielle de la contrainte

p est un élément de $\mathbb{N}^*$, g_1, g_2, ..., g_p sont p formes linéaires sur $\mathbb{R}^n$ et b_1, b_2, ..., b_p sont p réels.

$\mathcal{C}$ est l'ensemble des éléments X de $\mathbb{R}^n$ tels que $\begin{cases} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \dots\dots\dots\dots \\ g_p(X) = b_p \end{cases}$

Pour tout élément i de $[\![1,p]\!]$, notons $(a_{i1}\, a_{i2}\, \dots a_{in})$ la matrice de g_i relativement aux bases canoniques de $\mathbb{R}^n$ et de $\mathbb{R}$.

$$\boxed{\forall i \in [\![1,p]\!],\ \forall X \in (x_1, x_2, \dots, x_n),\ g_i(X) = a_{i1}\, x_1 + a_{i2}\, x_2 + \cdots a_{in}\, x_n}$$

$$\boxed{\forall X \in (x_1, x_2, \dots, x_n),\ X \in \mathcal{C} \Longleftrightarrow \begin{cases} a_{11}\, x_1 + a_{12}\, x_2 + \cdots a_{1n}\, x_n = b_1 \\ a_{21}\, x_1 + a_{22}\, x_2 + \cdots a_{2n}\, x_n = b_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_{p1}\, x_1 + a_{p2}\, x_2 + \cdots a_{pn}\, x_n = b_p \end{cases}}$$

Considérons alors la matrice $C = (a_{ij})$ de $\mathcal{M}_{p,n}(\mathbb{R})$ et la matrice colonne $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$. On a alors :

$$\boxed{\forall X \in (x_1, x_2, \dots, x_n),\ X \in \mathcal{C} \Longleftrightarrow C \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = B}$$

Notons donc que $\mathcal{C}$ est à l'ensemble des solutions d'un système linéaire de p équations à n inconnues.

▶ 4. Une modeste condition suffisante.

Th. 66 Ω est un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^2$ sur Ω.

p est un élément de $\mathbb{N}^*$, g_1, g_2, ..., g_p sont p formes linéaires sur $\mathbb{R}^n$ et b_1, b_2, ..., b_p sont p réels.

$\mathcal{C}$ est l'ensemble des éléments X de $\mathbb{R}^n$ tels que $\begin{cases} g_1(X) = b_1 \\ g_2(X) = b_2 \\ \dots\dots\dots\dots \\ g_p(X) = b_p \end{cases}$.

On pose : $\mathcal{H} = \{X \in \mathbb{R}^n \mid g_1(X) = g_2(X) = \cdots = g_p(X) = 0\} = \operatorname{Ker} g_1 \cap \operatorname{Ker} g_2 \cap \ldots \cap \operatorname{Ker} g_p$.

On suppose que A est un point de $\mathcal{C} \cap \Omega$ tel que $\nabla f(A)$ soit orthogonal à $\mathcal{H}$.

On suppose encore que $\forall B \in \mathcal{C} \cap \Omega,\ \forall H \in \mathcal{H},\ q_B(H) \geqslant 0$ (resp. $\leqslant 0$).

Alors f possède en A un minimum (resp. maximum) local sous la contrainte $\mathcal{C}$.

▶ 5. Vers le Lagrangien (hors programme).

De toute évidence ce qui précède nous laisse un peu sur notre faim et appelle à quelques généralisations.

On a bien envie de passer de "sous contrainte d'égalités linéaires" à "sous contrainte d'égalités" (voir à sous contrainte d'égalités et d'inégalités).

En clair il serait intéressant de remplacer les $g_i(X) = b_i$ ou $g_i(X) - b_i = 0$, avec g_i forme linéaire sur $\mathbb{R}^n$, par $h_i(X) = 0$ avec h_i fonction "quelconque" de $\mathbb{R}^n$ dans $\mathbb{R}$.

Reprenons encore les hypothèses des résultats précédents. Posons $\forall i \in [\![1,p]\!],\ h_i = g_i - b_i$.

Alors $\boxed{\mathcal{C} = \{X \in \mathbb{R}^n \mid h_1(X) = h_2(X) = \cdots = h_p(X) = 0\}}$. De plus $\forall i \in [\![1, p]\!],\ \forall X \in \mathbb{R}^n,\ \nabla h_i(X) = \nabla g_i(X)$.

Supposons que f possède un extremum local sous la contrainte $\mathcal{C}$ en un point A de Ω.

Alors A est élément de $\mathcal{C}$ et le gradient de f en A est orthogonal à $\mathcal{H}$.

L'orthogonal de $\mathcal{H}$ est Vect $\Big(\nabla g_1(X), \nabla g_2(X), \cdots, \nabla g_p(X)\Big)$ où X est un élément quelconque de $\mathbb{R}^n$.

Notons que $\forall X \in \mathbb{R}^n$, Vect $\Big(\nabla g_1(X), \nabla g_2(X), \cdots, \nabla g_p(X)\Big) = \text{Vect}\Big(\nabla h_1(X), \nabla h_2(X), \cdots, \nabla h_p(X)\Big)$.

Ainsi $h_1(A) = h_2(A) = \cdots = h_p(A) = 0$ et il existe un élément $(\lambda_A^1, \lambda_A^2, \cdots, \lambda_A^p)$ de $\mathbb{R}^p$ tel que

$$\nabla f(A) = \lambda_A^1\ \nabla h_1(A) + \lambda_A^2\ \nabla h_2(A) + \cdots + \lambda_A^P\ \nabla h_p(A).$$

Ceci donne encore $\forall i \in [\![1, p]\!],\ h_i(A) = 0$ et $\forall k \in [\![1, n]\!],\ \dfrac{\partial f}{\partial x_k}(A) - \displaystyle\sum_{i=1}^{p} \lambda_A^i\ \dfrac{\partial h_i}{\partial x_k}(X) = 0$.

D'ou l'idée de faire intervenir la fonction $\mathcal{L}$ de $n+p$ variables définie, avec quelques abus de notation, par :

$$\forall (X, \lambda_1, \lambda_2, \cdots, \lambda_p) \in \Omega \times \mathbb{R}^p,\ \mathcal{L}(X, \lambda_1, \lambda_2, \cdots, \lambda_p) = f(X) - \lambda_1\, h_1(X) - \lambda_2\, h_2(X) - \cdots - \lambda_p\, h_p(X).$$

Il est aisé de montrer que $\mathcal{L}$ est de classe $\mathcal{C}^1$ sur $\Omega \times \mathbb{R}^p$ et que :

$$\forall k \in [\![1, n]\!],\ \forall (X, \lambda_1, \lambda_2, \cdots, \lambda_p) \in \Omega \times \mathbb{R}^p,\ \frac{\partial \mathcal{L}}{\partial x_k}(X, \lambda_1, \lambda_2, \cdots, \lambda_p) = \frac{\partial f}{\partial x_k}(X) - \sum_{i=1}^{p} \lambda_i\ \frac{\partial h_i}{\partial x_k}(X)$$

$$\text{et} \qquad \forall i \in [\![1, p]\!],\ \forall (X, \lambda_1, \lambda_2, \cdots, \lambda_p) \in \Omega \times \mathbb{R}^p,\ \frac{\partial \mathcal{L}}{\partial \lambda_i}(X, \lambda_1, \lambda_2, \cdots, \lambda_p) = -h_i(X).$$

Rappelons que $\forall i \in [\![1, p]\!],\ h_i(A) = 0$ et $\forall k \in [\![1, n]\!],\ \dfrac{\partial f}{\partial x_k}(A) - \displaystyle\sum_{i=1}^{p} \lambda_A^i\ \dfrac{\partial h_i}{\partial x_k}(X) = 0$.

Alors on constate que $(A, \lambda_A^1, \lambda_A^2, \cdots, \lambda_A^p)$ est un point critique de $\mathcal{L}$...

▶ 6. Le Lagrangien (hors programme).

Th. 67 **le cas $\mathbf{p = 1}$!**

Ω est un ouvert de $\mathbb{R}^n$. f et g sont deux applications de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$. $\mathcal{C} = \{X \in \Omega \mid g(X) = 0\}$. A est un point de $\mathcal{C}$.

Si f possède un extremum local en A sous la contrainte $g(X) = 0$ ou sous la contrainte $\mathcal{C}$ (autrement dit si la restriction de f à $\mathcal{C}$ possède un extremum local en A) et si le gradient de g en A n'est pas nul alors il existe un réel λ_A tel que :

$$\nabla f(A) = \lambda_A\ \nabla g(A) \quad \text{ou} \quad \forall k \in [\![1, n]\!],\ \frac{\partial f}{\partial x_k}(A) = \lambda_A\ \frac{\partial g}{\partial x_k}(A).$$

Déf. 48 Ω est un ouvert de $\mathbb{R}^n$. f et g sont deux applications de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$. $\mathcal{C} = \{X \in \Omega \mid g(X) = 0\}$.

On pose : $\forall (x_1, x_2, \ldots, x_n, \lambda) \in \Omega \times \mathbb{R},\ \mathcal{L}(x_1, x_2, \ldots, x_n, \lambda) = f(x_1, x_2, \ldots, x_n) - \lambda\, g(x_1, x_2, \ldots, x_n)$.

$\mathcal{L}$ est le **Lagrangien** de f sous la contrainte $g(x_1, x_2, \ldots, x_n) = 0$ ou sous la contrainte $\mathcal{C}$.

Th. 68 Ω est un ouvert de $\mathbb{R}^n$. f et g sont deux applications de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$. $\mathcal{C} = \{X \in \Omega \mid g(X) = 0\}$. $A = (a_1, a_2, \ldots, a_n)$ est un point de $\mathcal{C}$.

$\mathcal{L}$ est le lagrangien de f sous la contrainte $g(X) = 0$.

On suppose que f admet un extremum local en A sous la contrainte $g(X) = 0$ et que le gradient de g en A n'est pas nul.

1. Il existe un réel λ_A tel que : $\nabla f(A) = \lambda_A \nabla g(A)$.
2. $\mathcal{L}$ est de classe $\mathcal{C}^1$ sur $\Omega \times \mathbb{R}$ et $(a_1, a_2, \ldots, a_n, \lambda_A)$ est un point critique de $\mathcal{L}$.
3. A est un encore un point critique de $\mathcal{L}_A : (x_1, x_2, \ldots, x_n) \to \mathcal{L}(x_1, x_2, \ldots, x_n, \lambda_A)$.
4. Si $\mathcal{L}_A$ possède un maximum local (resp. global) en A alors f possède un maximum local (resp. global) en A sous la contrainte $g(X) = 0$. Même chose pour un minimum.

Th. 69 Ω est un ouvert de $\mathbb{R}^n$. p est un élément de $\mathbb{N}^*$ .f, g_1, g_2, ..., g_p sont $p+1$ applications de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$. $\mathcal{C} = \{X \in \Omega \mid \forall i \in [\![1, p]\!],\ g_i(X) = 0\}$. A est un point de $\mathcal{C}$.

Si f possède un extremum local en A sous la contrainte $\mathcal{C}$ (autrement dit si la restriction de f à $\mathcal{C}$ possède un extremum local en A) et si la famille $\big(\nabla g_1(A), \nabla g_2(A), \cdots, \nabla g_p(A)\big)$ est libre il existe p réels λ_A^1, λ_A^2, ..., λ_A^p tels que :

$$\nabla f(A) = \lambda_A^1 \nabla g_1(A) + \lambda_A^2 \nabla g_2(A) + \cdots + \lambda_A^p \nabla g_p(A).$$

$(A, \lambda_A^1, \lambda_A^2, \ldots, \lambda_A^p)$ est alors un point critique de la fonction numérique $\mathcal{L}$ de $n+p$ variables définie par :

$\forall (X, \lambda_1, \lambda_2, \ldots, \lambda_n) \in \Omega \times \mathbb{R}^p,\ \mathcal{L}(X, \lambda_1, \lambda_2, \ldots, \lambda_n) = f(X) - \displaystyle\sum_{i=1}^{p} \lambda_i\, g_i(X)$.

VI SAVOIR FAIRE

- 1. Savoir montrer qu'une partie de $\mathbb{R}^n$ est un ouvert (resp. fermé) ;
- 2. Savoir montrer qu'une partie de $\mathbb{R}^n$ est convexe ;
- 3. Montrer qu'une partie est fermée et bornée.
- 4. Montrer la continuité d'une fonction "assez simple" de $\mathbb{R}^n$ dans $\mathbb{R}$.
- 5. Montrer qu'une fonction a ou n'a pas de limite en un point.
- 6. Montrer l'existence de dérivées partielles premières (resp. seconde) d'une fonction "assez simple" de $\mathbb{R}^n$ dans $\mathbb{R}$.
- 7. Calculer des dérivées partielles premières et secondes.
- 8. Calculer des dérivées directionnelles.
- 9. Ecrire le gradient en un point d'une fonction de classe $\mathcal{C}^1$.
- 10. Ecrire la hesienne en un point d'une fonction de classe $\mathcal{C}^2$.
- 11. Ecrire un dl1 en un point d'une fonction de classe $\mathcal{C}^1$.
- 12. Ecrire un dl2 en un point d'une fonction de classe $\mathcal{C}^2$.
- 13. Calculer la dérivée ou les dérivées partielles d'une "fonction composée".
  - $\mathbb{R}^n \to \mathbb{R} \to \mathbb{R}$. n dérivées partielles de $(x_1, x_2, \ldots, x_n) \to \varphi\Big(f(x_1, x_2, \ldots, x_n)\Big)$.
  - $\mathbb{R} \to \mathbb{R}^n \to \mathbb{R}$. Dérivée de $t \to f(u_1(t), u_2(t), \ldots, u_n(t))$.
- 14. Etudier les extremums d'une fonction assez simple.
- 15. Etudier la position du graphe d'une fonction par rapport à un hyperplan tangent.
- 16. Etudier des extremums sous contrainte d'égalités linéaires.

VII COMPLÉMENTS

▶ 1. Caractérisation des fermés.

Prop. 21 Soit F une partie non vide de $\mathbb{R}^n$. F est un fermé si et seulement si toute suite de F qui converge a sa limite dans F.

▶ 2. Fonctions convexes.

Déf. 49 D est un convexe de $\mathbb{R}^n$ et f est une application de D dans $\mathbb{R}$.

On dit que **f est convexe sur D** si : $\forall \lambda \in [0,1],\ \forall (A,B) \in D^2,\ f(\lambda A + (1-\lambda)B) \leqslant \lambda f(A) + (1-\lambda) f(B)$.

On dit que **f est concave sur D** si $-f$ est convexe sur D.

Prop. 22 Ω est un ouvert convexe de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$.

On pose : $\forall (A,B) \in \Omega^2,\ \forall t \in [0,1],\ \varphi_{AB}(t) = f(B + t(A-B))$.

f est convexe si et seulement si, pour tout couple (A,B) d'éléments de Ω, la fonction numérique de la variable réelle φ_{AB} est convexe sur $[0,1]$.

Th. 70 Ω est un ouvert convexe de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$. Les assertions suivantes sont équivalentes :

i) f est convexe sur Ω.

ii) $\forall (A,X) \in \Omega^2,\ f(X) \geqslant f(A) + < \nabla f(A), X - A >$.

iii) $\forall (A,X) \in \Omega^2,\ < \nabla f(X), X - A > \geqslant < \nabla f(A), X - A >$.

Th. 71 Ω est un ouvert convexe de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^1$.

1. Si f est convexe sur Ω et si A est un point critique de f alors f possède un minimum global en A.

1. Si f est concave sur Ω et si A est un point critique de f alors f possède un maximum global en A.

Th. 72 Ω est un ouvert convexe de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$ de classe $\mathcal{C}^2$. Les assertions suivantes sont équivalentes :

i) f est convexe sur Ω.

ii) $\forall B \in \Omega,\ \forall H \in \mathbb{R}^n,\ q_B(H) \geqslant 0$.

VIII EXEMPLES ET CONTRE-EXEMPLES USUELS

Les apllications partielles sont continues pas l'application.

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = 0 \end{cases}.$$

$A = (0,0)$. $f_{A,1}$ et $f_{A,2}$ sont continues en 0 mais f n'est pas continue en A.

---

f possède des dérivée partielles en A mais n'est pas continue en A

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = 0 \end{cases}$$

f admet des dérivées partielles premères en $(0,0)$ sans être continue en ce point.

---

f admet un dl 1 mais n'est pas de classe $\mathcal{C}^1$.

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} (x^2+y^2)\,\sin\left(\dfrac{1}{x^2+y^2}\right) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

f possède un développement limité d'ordre 1 en tout point de $\mathbb{R}^2$ mais n'est pas de classe $\mathcal{C}^1$ en tout point de $\mathbb{R}^2$.

---

f est continue en A, possède des dérivée partielles en A mais n'admet pas de dl 1 au voisinage de A.

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} \dfrac{xy}{\sqrt{x^2+y^2}} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

f est continue en $(0,0)$ admet des dérivées partielles en ce point mais ne possède pas de développement limité au voisinage de ce point.

$\dfrac{\partial^2 f}{\partial x_i\,\partial x_j}(A)$ et $\dfrac{\partial^2 f}{\partial x_i\,\partial x_j}(A)$ existent et sont distinctes.

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} \dfrac{xy^3}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

$\dfrac{\partial^2 f}{\partial x_i\,\partial x_j}(0,0)$ et $\dfrac{\partial^2 f}{\partial x_i\,\partial x_j}(0,0)$ esistent et sont distinctes.

---

f admet un dl 2 au voisinage de A sans que toutes les dérivées partielles d'ordre 2 existent en A.

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} (x^2+y^2)^2\,\sin\left(\dfrac{1}{x^2+y^2}\right) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

f possède un développement limité d'ordre 2 au voisinage de $(0,0)$ mais $\dfrac{\partial^2 f}{\partial x^2}(0,0)$ n'existe pas.
