Leçon 06 – Cours : Formes quadratiques

Objectif: Dans cette leçon nous présentons les formes quadratiques et les propriétés des matrices symétriques réelles qui y sont étroitement associées. Ces notions sont introduites dans le but de traiter en application les problèmes d'extrema libres et liés, très importants en économie, c'est l'objet du dernier paragraphe.

Les formes quadratiques jouent un rôle important dans la résolution de certains problèmes économiques comme ceux qui consistent à maximiser une fonction à plusieurs variables, mais aussi en statistique et en économétrie lorsque par exemple on fait de l'analyse factorielle ou quand on s'intéresse aux propriétés des vecteurs gaussiens.

1. Matrices symétriques

Définition : Une matrice A est symétrique si et seulement si : ${}^tA = A$. (où ${}^tA$ désigne la matrice transposée de A)

Si $A = (a_{ij})$ et si A est symétrique, on a donc : $\forall(i,j)\ a_{ij} = a_{ji}$. Deux termes symétriques par rapport à la diagonale sont égaux.

Théorème : Les valeurs propres d'une **matrice symétrique réelle** sont réelles.

Démonstration: Soit λ une valeur propre de A et P un vecteur propre associé à λ. Si λ n'est pas réelle, $\overline{\lambda}$ est aussi une valeur propre de A et $\overline{P}$ est un vecteur propre associé à $\overline{\lambda}$.

Donc $AP = \lambda P$ et en multipliant à gauche par ${}^t\overline{P}$:

$$ {}^t\overline{P}\,AP = \lambda({}^t\overline{P}\,P) \quad (1).$$

D'autre part $A\overline{P} = \overline{\lambda}\,\overline{P}$ et en prenant la transposée, on obtient : ${}^t\overline{P}\ A = \overline{\lambda}\,({}^t\overline{P})$

Si on multiplie à droite par P l'égalité devient : ${}^t\overline{P}\ AP = \overline{\lambda}\,({}^t\overline{P}\,P) \quad (2)$

$$(1) \text{ et } (2) \text{ donne : } \lambda({}^t\overline{P}\ P) = \overline{\lambda}\,({}^t\overline{P}\,P).$$

Si $P = \begin{pmatrix} x_1 \\ . \\ . \\ x_n \end{pmatrix}$ ${}^t\overline{P}\,P = \overline{x_1}\ x_1 + . + . + \overline{x_n}\ x_n \geq 0$. Or P est un vecteur propre de A donc P est différent du vecteur nul et ${}^t\overline{P}\,P \neq 0$ donc $\lambda = \overline{\lambda}$ et λ est donc réelle. C.Q.F.D.

Or si λ est réelle et A aussi on peut lui associer un vecteur propre réel. On supposera donc par la suite que tous les vecteurs propres sont réels.

Théorème : Si deux vecteurs propres d'une matrice symétrique sont associés à deux valeurs propres distinctes, ils sont **orthogonaux**.

Démonstration : Soit P_1 et P_2 les vecteurs propres associés respectivement aux deux valeurs propres distinctes λ_1 et λ_2 . $AP_1 = \lambda_1 P_1$ et en multipliant à gauche par ${}^tP_2$
on a donc : ${}^tP_2\,AP_1 = \lambda_1 {}^tP_2\,P_1$. Et transposant chaque membre on a

$$ {}^tP_1\,AP_2 = \lambda_1\,{}^tP_1\,P_2 \quad (1)$$

Et puisque P_2 est vecteur propre associé à λ_2, $AP_2 = \lambda_2 P_2$ et l'égalité (1) devient
$\lambda_2\,{}^tP_1\,P_2 = \lambda_1 {}^tP_1\,P_2$ ou $(\lambda_1 - \lambda_2){}^tP_1P_2 = 0$.
Or $\lambda_1 \neq \lambda_2$ donc nécessairement ${}^tP_2\,P_1 = 0$ et P_1 et P_2 sont orthogonaux. C.Q.F.D.

D'autre part si un sous espace propre est de dimension supérieure à un, on peut toujours trouver une base orthonormée de vecteurs propres pour ce sous espace. Donc si A est symétrique réelle et possède n vecteurs propres indépendants, on peut considérer qu'il existe une base orthonormée de $\mathbf{R}^n$ formée de vecteurs propres de A. Le théorème suivant affirme l'existence de ces n vecteurs propres dans le cas où A est symétrique réelle.

Théorème : Toute matrice réelle symétrique est diagonalisable sur $\mathbf{R}^n$ et admet une base orthonormée de vecteurs propres.

La démonstration de ce théorème se fait par récurrence et est assez fastidieuse. Aussi, bien que ce résultat soit fondamental, nous l'admettrons.

2. Formes quadratiques

2.1. Définition

Définition : On appelle **forme quadratique réelle** à n variables, toute application q de $\mathbf{R}^n$ dans $\mathbf{R}$ telle que : $q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \quad (a_{ij} \in \mathbf{R})$

Si $X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ x_n \end{pmatrix}$ et si la matrice $A = (a_{ij})$, $q(X) = {}^tXAX$.

On dit alors que A représente la forme quadratique q. Remarquons que si $a_{ij} \neq a_{ji}$ alors dans A on peut remplacer a_{ij} et a_{ji} par $\frac{a_{ij}+a_{ji}}{2}$ et qu'alors A est symétrique. Dorénavant, on considèrera qu'une forme quadratique est représentée par une matrice symétrique puisque c'est toujours possible. Pour chaque forme quadratique, une telle matrice est alors unique.

Théorème : Toute forme quadratique peut s'écrire comme une somme de carrés de combinaisons linéaires de ses variables, pondérée par les valeurs propres de la matrice symétrique qui lui est associée.

Démonstration : Soit A la matrice symétrique qui représente la forme quadratique q. Puisque A est diagonalisable et que la matrice P des vecteurs propres peut être choisie orthogonale : $P^{-1} = {}^tP$, $A = PD{}^tP$ et $q(X) = {}^tXPD{}^tPX$.

Donc $q(X) = {}^t({}^tPX)D({}^tPX)$ avec D matrice diagonale. Et si on pose $Z = {}^tPX = \begin{pmatrix} z_1 \\ z_2 \\ . \\ . \\ z_n \end{pmatrix}$,

les composantes z_i de Z sont des combinaisons linéaires des composantes de X et

$$q(X) = {}^tZDZ = \sum_{i=1}^{n} \lambda_i z_i^2 \qquad (1) \qquad\qquad \text{C.Q.F.D.}$$

Cette forme de q(X) montre le rôle important joué par les valeurs propres de A quand on s'intéresse au signe de q(X).

2.2. Nature d'une forme quadratique q

Définitions : Une forme quadratique est :

***définie positive** si : $\forall X \neq 0,\ q(X) > 0$,

***définie négative** si : $\forall X \neq 0,\ q(X) < 0$,

* **indéfinie** si elle est tantôt positive tantôt négative.

Une forme quadratique est :

***semi-définie positive** (ou *définie non-négative*) si : $\forall X\ \ q(X) \geq 0$, et q s'annule pour un vecteur non nul.

***semi-définie négative** (ou *définie non-positive*) si : $\forall X\ \ q(X) \leq 0$, et q s'annule pour un vecteur non nul.

Remarque : q est définie négative (resp. semi-définie négative) si et seulement si -q est définie positive (resp. semi-définie positive).

De la précédente décomposition (1) de q(X) vient immédiatement :

Théorème : Si q est une forme quadratique représentée par la matrice symétrique A :

*q est définie positive si et seulement si toutes les valeurs propres de A sont strictement positives.

*q est définie négative si et seulement si toutes les valeurs propres de A sont strictement négatives.

*q est indéfinie si et seulement si A a des valeurs propres non nulles et de signes différents.

*q est semi-définie positive ou négative si et seulement si toutes les valeurs propres sont de même signe et si au moins l'une d'entre elles est nulle.

Remarque : La méthode utilisée pour arriver aux résultats de l'"Exercez-vous 2" comporte un calcul long et fastidieux. Dans la plupart des cas le calcul des valeurs propres est compliqué. Aussi on va donner d'autres méthodes beaucoup plus performantes pour déterminer la nature de q et une forme de q analogue.

Lorsque q(X) s'écrit comme somme de termes dont chacun est le produit d'un réel par le carré d'une forme linéaire (combinaison linéaire des composantes de X), toutes les formes étant **indépendantes**, on dit que q(X) est mise sous **forme réduite**.

Cette forme existe d'après ce qui précède, mais elle n'est pas unique. Elle dépend de la méthode utilisée. L'une d'entre elles a pour coefficients les valeurs propres de la matrice A de q.

Néanmoins, toutes ces formes ont un point commun comme le précise le théorème suivant :

Théorème d'inertie de Sylvester : Pour une même forme quadratique q, toutes les formes réduites ont le même nombre s de coefficients positifs (c'est donc le nombre de valeurs propres positives) et le même nombre t coefficients négatifs (nombre de valeurs propres négatives).

Définition : La **signature d'une forme quadratique** est (*s*,*t*).

Remarque importante : Les formes intervenant sous les carrés étant indépendantes, une forme réduite d'une forme quadratique q dans $\mathbf{R}^n$ comporte au plus n carrés. Elle en a exactement n si la matrice symétrique A représentant q n'a pas de valeur propre nulle (ou si détA ≠0 ou si rang de A = n), on dit alors que q est **non dégénérée**. Si A a la valeur propre 0 avec un ordre de multiplicité r, les formes réduites de q comportent n-r carrés.
n-r est le *rang* de q (c'est aussi celui de A). On a s + t + r = n

2.3. Méthode de Gauss

La **méthode de Gauss** permet de trouver autrement une telle forme réduite (et donc la nature de q). Cette méthode a l'avantage d'être simple et de garantir des formes indépendantes sous les carrés. Il faudra la préférer à la diagonalisation de A.

Elle consiste tout d'abord à regrouper les termes contenant la première composante de X de la façon suivante :

$$q(x_1,x_2,\dots,x_n) = a_{11}x_1^2 + 2x_1 f(x_2,\dots,x_n) + R(x_2,\dots,x_n)$$

$$q(x_1,\dots,x_n) = a_{11}\left(x_1 + \frac{f(x_2,\dots,x_n)}{a_{11}}\right)^2 - \frac{f(x_2,\dots,x_n)^2}{a_{11}} + R(x_2,x_3,\dots,x_n)$$

Les deux derniers termes consistent en une forme quadratique des variables $x_2, x_3,\dots,x_n$. On itère le procédé avec x_2, puis x_3, … , puis x_n. A chaque étape, on supprime une composante.

Mais il peut arriver que q ne comporte pas de terme carré, on écrira alors $q(x_1,x_2,\dots,x_n)$ sous la forme : $a_{12}\, x_1x_2 + x_1 f(x_3,\dots,x_n) + x_2 g(x_3,\dots,x_n) + R(x_3,\dots,x_n)$ et

$$q(x_1,x_2,\dots,x_n) = \frac{a_{12}}{4}\left[\left(x_1+x_2+\frac{f(x_3,\dots,x_n)+g(x_3,\dots,x_n)}{a_{12}}\right)^2 - \left(x_1-x_2-\frac{f(x_3,\dots,x_n)-g(x_3,\dots,x_n)}{a_{12}}\right)^2\right] + R(x_3,\dots,x_n) - \frac{f(x_3,\dots,x_n)g(x_3,\dots,x_n)}{a_{12}}.$$

Reprenons l'"Exercez-vous 2" du paragraphe précédent :

$q(x,y,z) = 2xy + 2xz + 2yz = \frac{1}{2}[(x + y + 2z)^2 - (y - x)^2] - 2z^2$. Ce qui est cohérent avec le précédent résultat.

Les théorèmes suivants donnent des critères pour reconnaître la nature d'une forme quadratique (dégénérée ou non, définie positive, …) sans avoir recourt aux valeurs propres ou à une forme réduite.

Théorème : Soit q une forme quadratique représentée par la matrice symétrique $A = (a_{ij})$. q est **définie positive** si et seulement si :

$$\Delta_1 = a_{11} > 0,\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots,\ \Delta_i = \begin{vmatrix} a_{11} & . & . & a_{1i} \\ . & . & . & . \\ . & . & . & . \\ a_{i1} & . & . & a_{ii} \end{vmatrix} > 0, \ldots,\ \Delta_n = \text{dét}(A) > 0$$

N.B. : Les Δ_i sont appelés **mineurs principaux** de A, ce sont les déterminants obtenus en enlevant les n-i dernières colonnes et lignes de A.

Démonstration partielle (peut être sautée) : Supposons q définie positive. Alors d'après le théorème précédent, le déterminant de A qui est égal au produit des valeurs propres est strictement positif. D'autre part notons A_i la matrice obtenue à partir de A en supprimant les n-i dernières lignes et colonnes.

Soient les vecteurs X_i dont les n-i dernières composantes sont nulles :

$X_i = \begin{pmatrix} x_1 \\ . \\ x_i \\ 0 \\ . \\ 0 \end{pmatrix}$ et X_i' le vecteur de $\mathbf{R}^i$ ayant les i premières composantes de X_i : $X'_i = \begin{pmatrix} x_1 \\ . \\ x_i \end{pmatrix}$

Alors ${}^tX_i A X_i = {}^tX'_i A_i X'_i$. Or A_i est symétrique (puisque A l'est) et représente une forme quadratique q_i définie sur $\mathbf{R}^i$.

D'autre part $\forall X_i \neq 0\ q(X_i) > 0$ donc $\forall X'_i \in \mathbf{R}^i \setminus \{0\}\ q_i(X'_i) > 0$ et q_i est définie positive et d'après la première remarque dét$A_i > 0$.

Quant à la **réciproque**, elle se fait par récurrence mais elle est un peu fastidieuse, aussi nous l'admettrons (c.f. Algèbre pour économistes de B. GUERRIEN).

Remarquons que l'on obtient un théorème analogue en ce qui concerne les formes définies négatives. En effet q (de matrice A) est définie négative si et seulement si –q (de matrice –A) est définie positive. Et en appliquant le théorème précédent à -q on obtient une caractérisation des formes définies négatives :

$$\boxed{\text{q est définie négative si et seulement si } \Delta_1 < 0,\ \Delta_2 > 0,\ \ldots,\ (-1)^i\Delta_i > 0\ (i = 1,\ldots,n)}$$

Mais il est **inutile de retenir cet encadré**, il suffit d'étudier –q (de matrice -A). D'autre part on pourra aussi utiliser le fait que

détA = produit des valeurs propres de A et tr$A$ = somme des valeurs propres de A.

Donc si dét$A \neq 0$, q est non dégénérée et q est soit définie positive, soit définie négative, soit indéfinie. On peut donc utiliser le théorème précédent en l'appliquant à q et à -q et procéder par élimination.

Dans le cas des formes semi-définies, la caractérisation est beaucoup moins simple, on pourra se reporter au livre de B. GUERRIEN pour plus de détails.

3. Formes quadratiques et contraintes

On considère une forme quadratique q définie sur $\mathbf{R}^n$, représentée par une matrice symétrique A et des vecteurs $X = \begin{pmatrix} x_1 \\ . \\ x_n \end{pmatrix}$ de $\mathbf{R}^n$ soumis à p contraintes linéaires de la forme :

$\sum_{j=1}^{n} b_{ij}x_j = 0$ (i=1,…,p), ce qui peut s'écrire $BX = 0$, $B = (b_{ij})$ étant une matrice (p,n).

On peut considérer que ces contraintes sont **indépendantes** (donc $p \leq n$), quitte à éliminer celles qui sont combinaisons linéaires des autres. B est donc de rang p. D'autre part si p = n, seul $X = \mathbf{0}$ est soumis à toutes ces contraintes et ce cas est trivial.
On supposera donc $p < n$.

On notera B_p , la matrice carrée d'ordre p formée des p premières colonnes de B, on supposera cette matrice inversible.
Si ce n'est pas le cas, en changeant l'ordre des variables x_i, on peut toujours s'y ramener puisque rang B = p. Mais si on est obligé de changer l'ordre des variables, il faut aussi changer A en conséquence.

Résultat : Dans ces conditions, on montre qu'en fait, compte tenu des contraintes, q(X) est une forme quadratique q_1, ne dépendant que des (n-p) dernières variables :

$$q(x_1,....., x_n) = q_1(x_{p+1}, ..., x_n).$$

Démonstration (peut être sautée):
En décomposant B par blocs, on peut écrire B sous la forme $(B_p\ B_{n-p})$. On peut aussi décomposer X sous la forme $\begin{pmatrix} X_p \\ X_{n-p} \end{pmatrix}$ et : $(BX = 0) \Leftrightarrow (B_p X_p + B_{n-p} X_{n-p} = 0)$

B_p étant inversible la dernière égalité s'écrit $X_p = -(B_p)^{-1} B_{n-p} X_{n-p}$.

$$\text{Donc } (BX = 0) \Leftrightarrow \left(X = \begin{pmatrix} -(B_p)^{-1}B_{n-p}X_{n-p} \\ X_{n-p} \end{pmatrix} = B^*X_{n-p} \right) \text{ avec } B^* = \begin{pmatrix} -(B_p)^{-1}B_{n-p} \\ I_{n-p} \end{pmatrix}$$

Soumettre un vecteur de $\mathbf{R}^n$ à p contraintes indépendantes revient donc à supprimer p coordonnées et à considérer un vecteur dépendant uniquement de ses n-p dernières coordonnées.

Or $q(X) = {}^t XAX = {}^t(B^*X_{n-p})\,A(B^*X_{n-p}) = {}^t X_{n-p}\,EX_{n-p}$ avec $E = {}^t B^*AB^*$ symétrique. Ainsi l'étude de q(X), X étant soumis aux p contraintes précédentes revient à considérer une forme quadratique (représentée par une matrice symétrique E) dépendant de n-p variables, au lieu de n.

Exemple : Déterminer la nature de la forme quadratique q définie sur $\mathbf{R}^4$ et telle que

$q(X) = x^2 + 2yz + 2zt - 2t^2$ lorsque $X = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ est soumis aux deux contraintes : $\begin{cases} y+z+t = 0 \\ x+y+z+t = 0 \end{cases}$

*La façon la plus astucieuse d'aborder le problème est de constater que les contraintes sont équivalentes à $\begin{cases} x = 0 \\ y = -z-t \end{cases}$. Ainsi, on est ramené à l'étude du signe de : $q(X) = -2z^2 - 2t^2$ et $q(X) < 0$ si X est soumis aux contraintes et $X \neq 0$.

*Retrouvons ce résultat en utilisant la méthode du cours.

Ici $B = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = (B_2\ \ C_2)$ avec $B_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ de rang 2 et $C_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

(C_2 joue le rôle de B_{n-p})

Et $B^* = \begin{pmatrix} -B_2^{-1}C_2 \\ I_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ et la matrice $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}$ est la matrice symétrique qui représente q.

Et $E = {}^t(B^*)\,AB^*$ de q(X) sous les contraintes BX=0 s'écrit : $E = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$ et on a bien

$q(X) = q_1(z,t) = (z,t)\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} z \\ t \end{pmatrix} = -2z^2 - 2t^2$ si $BX = 0$.

L'étude de q sous les contraintes BX = 0 revient donc à celle de q_1. Une telle réduction n'est pas toujours aisée (la détermination de la matrice de q_1 (E dans l'exemple précédent) est souvent fastidieuse), aussi on va énoncer des théorèmes permettant de connaître la nature de q sous les p contraintes précédentes, sans être obligé de faire tous ces calculs.

Dans la pratique on appliquera ces théorèmes pour les exercices.

On utilisera la **matrice bordée** M suivante :

$M = \begin{pmatrix} A & {}^tB \\ B & 0 \end{pmatrix}$

et les **mineurs principaux bordés** : $M_i = \begin{vmatrix} a_{11} & . & a_{1i} & b_{11} & . & b_{p1} \\ . & . & . & . & . & . \\ a_{1i} & . & a_{ii} & b_{1i} & . & b_{pi} \\ b_{11} & . & b_{1i} & 0 & . & 0 \\ . & . & . & . & . & . \\ b_{p1} & . & b_{pi} & 0 & . & 0 \end{vmatrix}$ (**i=p+1,…,n**).

En fait $M_i = \begin{pmatrix} A_i & {}^tB_i \\ B_i & 0 \end{pmatrix}$ où A_i est obtenue à partir de la matrice A en supprimant les n-i dernières lignes et les n-i dernières colonnes, et B_i à partir de la matrice B en supprimant les n-i dernières colonnes.

On suppose là encore que les p premières colonnes de B sont linéairement indépendantes. Dans le cas contraire, on changera l'ordre des variables de façon à rendre B_p inversible, mais on n'oubliera pas de modifier de même A (en échangeant les mêmes variables que pour B).

Théorème : 1) q(X) est strictement positive quel que soit le vecteur $X\neq 0$ vérifiant BX = 0 si et seulement si :

$$(-1)^p \text{ dét}(M_i) > 0 \quad (i = p+1,\ldots,n)$$

2) q(X) est strictement négative quel que soit le vecteur $X\neq 0$ vérifiant BX = 0, si et seulement si :

$$(-1)^i \text{ dét}(M_i) > 0 \quad (i = p+1,\ldots,n)$$

Démonstration : Elle utilise un théorème du cours et la remarque qui suit ce théorème (en ce qui concerne les formes quadratiques strictement négatives) et le résultat suivant :
$\text{dét}(M_i) = (-1)^p \text{ dét}(B_p)^2 \text{ dét}(E_{i-p})$ $(i = p+1,\ldots,n)$ avec E_i la sous matrice de E formée par ses i premières lignes et colonnes.
En effet d'après un théorème du cours, E est définie positive si et seulement si dét $E_{i-p} > 0$ pour $i = p+1, .., n$. Soit détM_i, du sign de $(-1)^p$.
Et d'après la remarque suivant ce théorème, E est définie négative si et seulement si $(-1)^{i-p}$dét $E_{i-p} > 0$ pour $i = p+1, ..., n$. Soit dét M_i du signe de $(-1)^p(-1)^{i-p} = (-1)^i$.

Remarques *:* 1) On peut établir un résultat similaire pour les formes quadratiques semi-positives soumises à des contraintes en appliquant un théorème du cours, mais ce résultat est peu utilisable vu les calculs qu'il implique.
2) Parfois, lors de la résolution des exercices, la détermination du signe de q, avec ou sans contrainte, peut être faite directement, sans avoir à passer par le critère des mineurs principaux.

4. Application à l'étude des extrema d'une fonction de plusieurs variables

4.1. Extrema non liés

Rappelons tout d'abord quelques résultats d'analyse du cours de Mathématiques2 :

Soit f une fonction de $\mathbf{R}^n$ dans $\mathbf{R}$, ayant des dérivées partielles continues jusqu'à l'ordre 2 au voisinage V_A d'un point $A = (a_1, \ldots, a_n)$. On a donc d'après le théorème de symétrie de Schwarz : $\forall i,j=1,\ldots n \quad \dfrac{\partial^2 f}{\partial x_i \partial x_j}(A) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(A) \qquad (1).$

D'autre part f admet le développement de Taylor à l'ordre 2 suivant :

$$\forall X=(x_1,\ldots,x_n)\in V_A \quad f(X) = f(A) + (x_1 - a_1)\frac{\partial f}{\partial x_1}(A) + \ldots + (x_n - a_n)\frac{\partial f}{\partial x_n}(A) +$$

$$\frac{1}{2}\Big[\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-a_i)(x_j-a_j)\frac{\partial^2 f}{\partial x_i \partial x_j}(A)\Big] + ||X-A||^2\,\varepsilon(X-A) \quad \text{avec} \lim_{||X-A||\to 0}\varepsilon(X-A)=0.$$

On note souvent gradf(A) = $\left(\frac{\partial f}{\partial x_1}(A),\ldots,\frac{\partial f}{\partial x_n}(A)\right)$ (on dit : "**gradient de f** en A") . On notera aussi gradf(A) la matrice colonne correspondante $\begin{pmatrix} \frac{\partial f}{\partial x_1}(A) \\ . \\ . \\ \frac{\partial f}{\partial x_n}(A) \end{pmatrix}$.

Si on note : $\mathbf{F''}(A) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(A) & . & \frac{\partial^2 f}{\partial x_1 \partial x_n}(A) \\ . & . & . \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(A) & . & \frac{\partial^2 f}{\partial x_n^2}(A) \end{pmatrix}$ et $(X-A) = \begin{pmatrix} x_1-a_1 \\ . \\ x_n-a_n \end{pmatrix}$,

la formule de Taylor précédente peut s'écrire matriciellement sous la forme :

$$\boxed{f(X) = f(A) + {}^t\mathrm{grad}f(A)\,(X-A) + \frac{1}{2}\,{}^t(X-A)\mathbf{F''}(A)(X-A) + \|X-A\|^2\,\varepsilon(X-A)}$$

F"(A) est appelée **matrice hessienne** de f en A. D'après (1) cette matrice est symétrique réelle.

Or A correspond à un extremum local de f s'il existe un voisinage V_A de A dans $\mathbf{R}^n$ tel que : $\forall X \in V_A \setminus\{A\}$ f(X) - f(A) a un signe constant. Si ce signe est négatif, A correspond à un maximum, s'il est positif c'est un minimum.

On a déjà vu dans le cours de Mathématiques2 que si f est 2 fois continûment différentiable, pour que f admette un extremum en A, il faut nécessairement que A soit un **point stationnaire** de f, c'est à dire que gradf(A) = 0 (condition du premier ordre).

Dorénavant on supposera que A est un point stationnaire de f. f(X) - f(A) s'écrit alors :

$$f(X) - f(A) = \frac{1}{2}\,{}^t(X-A)\mathbf{F''}(A)(X-A) + \|X-A\|^2\,\varepsilon(X-A)$$

Pour X suffisamment proche de A, c'est à dire pour $\|X-A\|$ suffisamment proche de 0, $\|X-A\|^2\,\varepsilon(X-A)$ est négligeable devant les autres termes (sauf si ${}^t(X-A)\mathbf{F''}(A)(X-A)$ s'annule pour $X \neq A$). On en déduit alors que

$$\boxed{f(X) - f(A) \text{ est du signe de } {}^t(X-A)\mathbf{F''}(A)(X-A) \text{ sur } V_A.}$$

Aussi, A étant un **point stationnaire** de f, en considérant la forme quadratique définie par la matrice symétrique **F"**(A) (elle est symétrique étant donné le théorème de symétrie de Schwarz) et en utilisant les résultats du paragraphe précédent, on a :

Propriété :

*Si F"(A) est la matrice d'une forme quadratique définie positive, A correspond à un minimum de f sur V_A .

*Si F"(A) est la matrice d'une forme quadratique définie négative, A correspond à un maximum de f sur V_A .

*Si F"(A) est la matrice d'une forme quadratique indéfinie (f(X) - f(A)) change de signe sur V_A), A ne correspond pas à un extremum, c'est un point-col (ou un point-selle)

*Dans les autres cas on fera une étude directe.

Pour déterminer la nature de la forme quadratique de matrice **F"**(A), on pourra utiliser les critères du paragraphe précédent (mineurs principaux de la matrice hessienne des dérivées secondes de f, méthode de Gauss …).

4.2. Extrema sous contraintes

Soit f, une fonction de $\mathbf{R}^n$ dans **R** ayant des dérivées secondes partielles continues au voisinage V_A d'un point A.
On considère les points $X = (x_1,x_2,\dots,x_n)$ de V_A soumis aux p contraintes $g_j(X) = 0$ pour $j = 1, \dots,p$ ($p < n$), les fonctions g_j ayant des dérivées secondes partielles continues sur V_A.

On sait que f admet un extremum en A, sous les p contraintes $g_j(X) = 0$ si et seulement si, il existe p réels λ_{0j} tels que le **Lagrangien** de f, la fonction

$$L(X,\lambda) = f(X) + \sum_{j=1}^{p}\lambda_j g_j(X) \text{ admette un extremum en } (A,\lambda_0).$$

(on note $\lambda_0 = (\lambda_{01}, \dots, \lambda_{0p})$ et $\lambda = (\lambda_1, ..., \lambda_p)$)

Remarquons qu'alors $L(X,\lambda)$ vérifie $\mathrm{grad}L(A,\lambda_0) = 0$. Soit :

$$\begin{cases} \text{pour } i=1,\dots,n & \dfrac{\partial L}{\partial x_i}(A,\lambda_0) = 0 \\ \text{pour } j=1,\dots,p & \dfrac{\partial L}{\partial l_j}(A, \lambda_0) = 0 \end{cases} \quad \textit{ou} \quad \begin{cases} \text{pour } i=1,\dots,n & \dfrac{\partial f}{\partial x_i}(A) + \displaystyle\sum_{j=1}^{p} \lambda_{0j} \dfrac{\partial g_j}{\partial x_i}(A) = 0 \\ \text{pour } j=1,\dots,p & g_j(A) = 0 \end{cases}.$$

Supposons donc ces conditions vérifiées et notons $\mathbf{L}''(A,\lambda_0)$ la matrice hessienne de L en A pour λ_0 fixé (on considère ici les seules variables $(x_1 , x_2 , \dots x_n)$. On a alors

$$\mathbf{L}''(A,\lambda_0) = \begin{pmatrix} \dfrac{\partial^2 L}{\partial x_1^2}(A,\lambda_0) & . \; . & \dfrac{\partial^2 L}{\partial x_1 \partial x_n}(A,\lambda_0) \\ . & . \; . & . \\ . & . \; . & . \\ \dfrac{\partial^2 L}{\partial x_n \partial x_1}(A,\lambda_0) & . \; . & \dfrac{\partial^2 L}{\partial x_n^2}(A,\lambda_0) \end{pmatrix}$$

On note **Jg**(A), la **matrice Jacobienne** des contraintes, c'est à dire la matrice formée des dérivées partielles premières des g_j en A :

$$\mathbf{Jg}(A) = \left(\frac{\partial g_j}{\partial x_i}(A)\right) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(A) & . & \frac{\partial g_1}{\partial x_n}(A) \\ . & . & . \\ \frac{\partial g_p}{\partial x_1}(A) & . & \frac{\partial g_p}{\partial x_n}(A) \end{pmatrix}$$

On suppose que $\boxed{\mathbf{Jg}(A) \text{ est de rang } p}$ (c'est une matrice de format (p,n) avec p < n).

On a alors les résultats suivants.

En fait, en utilisant une version un peu compliquée du théorème des fonctions implicites, on montre qu'étant données les contraintes :

le problème revient à étudier la nature d'une forme quadratique de matrice symétrique **L"**(A,λ_0) sous les contraintes **Jg**(A)(X-A) = 0.

On utilise alors les résultats du théorème du paragraphe précédent et on détermine les signes des déterminants de la matrice bordée M(A) suivante et de ses mineurs $M_i(A)$ bordés (i = p+1, …,n) définies par

$$M(A) = \begin{pmatrix} L''(A,\lambda_0) & {}^t Jg(A) \\ Jg(A) & 0 \end{pmatrix}$$

$$M(A) = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2}(A,\lambda_0) & . & \frac{\partial^2 L}{\partial x_1 \partial x_n}(A,\lambda_0) & \frac{\partial g_1}{\partial x_1}(A) & . & \frac{\partial g_p}{\partial x_1}(A) \\ . & . & . & . & . & . \\ \frac{\partial^2 L}{\partial x_n \partial x_1}(A,\lambda_0) & . & \frac{\partial^2 L}{\partial x_n^2}(A,\lambda_0) & \frac{\partial g_1}{\partial x_n}(A) & . & \frac{\partial g_p}{\partial x_n}(A) \\ \frac{\partial g_1}{\partial x_1}(A) & . & \frac{\partial g_1}{\partial x_n}(A) & 0 & . & 0 \\ . & . & . & . & . & . \\ \frac{\partial g_p}{\partial x_1}(A) & . & \frac{\partial g_p}{\partial x_n} & 0 & . & 0 \end{pmatrix}$$

et pour **i = p+1, …,n**, comme au paragraphe précédent, on définit les mineurs bordés suivants :

$$M_i(A) = \begin{pmatrix} \frac{\partial^2 L}{\partial x_1^2}(A,\lambda_0) & . & \frac{\partial^2 L}{\partial x_1 \partial x_i}(A,\lambda_0) & \frac{\partial g_1}{\partial x_1}(A) & . & \frac{\partial g_p}{\partial x_1}(A) \\ . & . & . & . & . & . \\ \frac{\partial^2 L}{\partial x_i \partial x_1}(A,\lambda_0) & . & \frac{\partial^2 L}{\partial x_i^2}(A,\lambda_0) & \frac{\partial g_1}{\partial x_i}(A) & . & \frac{\partial g_p}{\partial x_i}(A) \\ \frac{\partial g_1}{\partial x_1}(A) & . & \frac{\partial g_1}{\partial x_i}(A) & 0 & . & 0 \\ . & . & . & . & . & . \\ \frac{\partial g_p}{\partial x_1}(A) & . & \frac{\partial g_p}{\partial x_i}(A) & 0 & . & 0 \end{pmatrix}.$$

En appliquant les deux théorèmes précédents on obtient :

Si $(-1)^p \text{dét}(M_i(A,\lambda_0)) > 0$ pour $i=p+1,\ldots,n$, f a un minimum strict en A sous les p contraintes g_j.

Si $(-1)^i \text{dét}(M_i(A,\lambda_0)) > 0$ pour $i=p+1,\ldots,n$, f a un maximum strict en A sous les p contraintes g_j.

Cas particulier : n=2 et p=1

On considère donc $f(x,y)$ sous la contrainte $g(x,y) = 0$.
$L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$.

Soient A et λ_0 tels que $(\text{gradL})(A,\lambda_0)=0$, $M(A) = \begin{pmatrix} L''_{x^2}(A,\lambda_0) & L''_{xy}(A,\lambda_0) & g'_x(A) \\ L''_{xy}(A,\lambda_0) & L''_{y^2}(A,\lambda_0) & g'_y(A) \\ g'_x(A) & g'_y(A) & 0 \end{pmatrix}$.

Ici i ne prend que la valeur 2 puisque $p+1 = n = 2$ et $M_2(A) = M(A)$.

Ainsi si $\text{dét}(M(A)) > 0$, $(-1)^i \text{dét}(M_i(A)) > 0$ donc A correspond à un maximum lié.
Et si $\text{dét}(M(A)) < 0$, $(-1)^p \text{dét}(M_i(A)) > 0$ et A correspond à un minimum lié.
Mais si $\text{dét}(M(A)) = 0$, on ne peut pas conclure et il faut étudier directement la différence $f(x_0+h,y_0+k) - f(x_0,y_0)$.

Exercices

Exercice 1 - Soit q la forme quadratique définie sur $\mathbf{R}^3$ par $q(x,y,z) = x^2 + 4xz + 2y^2 + z^2$.

1) Déterminer la matrice A symétrique associée à q dans la base canonique de $\mathbf{R}^3$.

2) Construire une base orthonormée P formée de vecteurs propres de A.

3) Déterminer la matrice B associée à q dans la base P. En déduire l'expression de la forme réduite de q(x,y,z).

4) Utiliser la méthode de Gauss pour trouver une forme réduite de q.

Exercice 2 – Déterminer la matrice symétrique représentant q, donner une forme réduite de forme réduite de q par la méthode de Gauss dans les cas :

1) $q(x,y,z) = x^2 + 5y^2 + 6z^2 + 2xy - 4xz$.

2) $q(x,y,z) = x^2 - y^2 + z^2 - xy - yz$

3) $q(x,y,z) = 2xy - x^2 - 2xz - 5y^2 + 6yz - 3z^2$

4) $q(x,y,z) = 2xy - 2yz$

Exercice 3 - Que doit vérifier le paramètre réel a pour que la forme quadratique suivante soit définie positive ?

$$q(x,y,z) = x^2 + ay^2 + z^2 - xy - yz$$

Exercice 4 – Pour chacune des formes quadratiques de l'exercice **2**, déterminer sa nature en utilisant les mineurs principaux et éventuellement la trace de la matrice symétrique qui la représente.

Exercice 5 - Soit $A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$. A est-elle la matrice d'une forme quadratique définie positive?

Exercice 6 - Déterminer et préciser la nature des extrema éventuels des fonctions f et g définies sur $\mathbf{R}^3$ par :

1) $f(x,y,z) = x^2 - 2xy + 2y^2 + 2y + z^2 - 4z + 5$

2) $g(x,y,z) = x^4 - 2x^2 y + 2y^2 - 2yz + 2z^2 - 4z + 5$

(indication : si on est bloqué dans le calcul des points stationnaires, on pourra admettre qu'il y a trois points stationnaires : $A(\sqrt{2}, 2, 2)$, $B(-\sqrt{2}, 2, 2)$ et $C(0, \frac{2}{3}, \frac{4}{3})$)

Exercice 7 - Déterminer et préciser la nature des extrema éventuels de la fonction f définie par $f(x,y,z) = x\ln y + z\ln x - y$

Exercice 8 - En utilisant les matrices hessiennes bordées, déterminer et préciser la nature des extrema éventuels des fonctions suivantes, les variables x et y étant liées par une contrainte :

1) $f(x,y) = xy$ avec $x^2 + y^2 = 2$ (remarquer qu'un point stationnaire du lagrangien correspond à $x = 0$ ou $\lambda = \pm\frac{1}{2}$.

2) $f(x,y) = \ln(x-y)$ avec $x^2 + y^2 = 2$

(indication : si on est bloqué dans le calcul des points stationnaires, on pourra admettre qu'il y a un seul point stationnaire : A(1,-1) et on donnera la valeur du multiplicateur de Lagrange correspondant)

Exercice 9 - 1) En utilisant les matrices hessiennes bordées, déterminer et préciser la nature des extrema éventuels de $f(x,y,z) = x + y + z$, les variables x, y et z étant liées par la contrainte (C) : $x^2 + y^2 + z^2 = 1$

2) Soit $f(x,y,z) = xyz$ sous la contrainte (C) : $\frac{x^2}{4} + \frac{y^2}{3} + z^2 = 1$.

Montrer que $\left(\frac{2\sqrt{3}}{3}, 1, \frac{\sqrt{3}}{3}\right)$ et $\left(\frac{2\sqrt{3}}{3}, -1, \frac{\sqrt{3}}{3}\right)$ sont des points stationnaires du Lagrangien, puis préciser si ce sont des maxima ou des minima de f sous (C).

Exercice 10 - Déterminer et préciser la nature des extrema éventuels de la fonctions f définie sur $\mathbf{R}^3$ par : $f(x,y,z) = x^2 + 2y^2 - 4xy + 12z + 6z^2$, les variables x, y et z étant soumises aux deux contraintes suivantes : $z - x = 3$ et $x - y = 0$.

(On vérifiera que $\left(-\frac{24}{5}, -\frac{24}{5}, -\frac{9}{5}\right)$ est le seul point stationnaire du Lagrangien).

Exercice 11 - Soit la forme quadratique : $q(x,y,z) = x^2 - 2xy + 3y^2 + 4yz - 2z^2$.

1) On considère (x,y,z) soumis à la contrainte $C_1 : x + 2y - z = 0$. Donner alors la nature de q.

2) On considère (x,y,z) soumis à la contrainte $C_2 : x + 2y - z = 0$ et $x - y + z = 0$
Donner la nature de q sous C_2.

Exercice 12 - Soit F la fonction de $\mathbf{R}^3$ dans $\mathbf{R}$ définie par :

$F(x,y,z) = 3x^2 + 2xy + y^2 + 2xz + \frac{2}{3} yz + \frac{4}{3} z^2 - 6x - 2y - 2z + 5$

Déterminer les points stationnaires de F. Pour chaque point stationnaire A, déterminer la matrice hessienne de A, F"(A) et en déduire la nature du point stationnaire A.

Exercice 13 – Soit $q(x,y,z) = 2x^2 – 4xy + 4xz + y^2 – 6yz + 3z^2$
1) Donner la matrice symétrique représentant q.
2) Donner la forme réduite de q et en déduire la nature de q.
3) On considère q munie de la contrainte : $2x + 2y + z = 0$. Donner, à l'aide de la matrice bordée, la nature de q sous cette contrainte.

Exercice 14 – Soit $f(x,y) = x\exp(-x+y^2) + xy - 3y^2 + 4y$ sous la contrainte, $x^2 + y^2 + xy = 3$.

1. Montrer que (1,1) est un point stationnaire du lagrangien pour lequel on précisera le multiplicateur de lagrange.
2. Donner la nature du point stationnaire (1,1) à l'aide de la matrice hessienne bordée.