## Expression covariante de l'optimisation sous contrainte, ou ...

## Kuhn et Tucker sans douleur.

Dans le présent article nous donnons une présentation et une démonstration des conditions de Kuhn et Tucker, basées sur un changement de repère, c'est à dire de variables, qui permet d'exploiter la notion tensorielle de covariance.

Nous n'avons jamais rencontré une telle idée dans les nombreux textes étudiés, sauf de manière indirecte et partielle dans certaines interprétations économiques. La raison en est probablement que les économistes n'ont pas tendance spontanément à introduire dans leurs calculs des "paniers" de biens ou de facteurs de production. Cela éviterait pourtant à nombre d'entre eux de faire apparaître des espaces euclidiens incongrus et des normes dans des raisonnements où ils n'ont pas leur place, puisque les espaces vectoriels considérés n'ont qu'une structure affine. On éviterait peut-être aussi de voir se multiplier les graphiques ou l'on confond dans un même $R^n$ "image" l'espace primal $E^n$ des variables et l'espace dual $E^{n*}$ des gradients de coûts ou des prix, avec des orthogonalités géométriques d'autant plus frappantes qu'elles sont trompeuses, car liées à une superposition arbitraire, sur les mêmes axes, d'unités de mesure disparates. Les professeurs courageux qui accepteraient de ne tracer leurs graphiques qu'en axes obliques rendraient un grand service à leurs élèves en les obligeant à réfléchir.

On nous objectera peut-être que tout cela n'est pas bien gênant. Toutefois ce manque de rigueur dans l'exposé des théories conduit souvent à des démonstrations inutilement compliquées et à masquer la véritable signification des propriétés démontrées.

Il en va ainsi pour l'étude des programmes mathématiques non linéaires, c'est à dire la maximisation d'une fonction de plusieurs variables assujetties à vérifier un système de relations d'inégalités ou d'égalités. Ceci donne lieu aux conditions K-T bien connues, dont nous allons approfondir la signification dans le cadre d'une démonstration étonnamment simple, ne faisant nullement appel à l'ingrat théorême de Farkas toujours cité, mais rarement détaillé (voir annexe 1 à titre d'exercice).

**1/ Les conditions nécessaires du premier ordre.**

Le problème posé est le suivant (pour simplifier l'exposé nous nous limitons à des inégalités) :

$$I \quad \left\| \begin{array}{ll} \quad \text{Max } f(x) & \\ \quad g^k(x) \geqslant 0 & \\ k \in (1,2,....,p) & \qquad x \in E^n \end{array} \right.$$

Soit $x_0$ un optimum local, ce qui implique $\quad f(x) \leq f(x_{0)} \quad$ en un voisinage compatible avec les contraintes.

Nous supposons que toutes les fonctions utilisées sont de classe $C^{(2)}$, c'est à dire possèdent des dérivées continues d'ordre 2.

Appelons $g^s(x)$ les contraintes saturées en $x_0$ , $g^r(x)$ celles non-saturées :

$$(1) \quad \left| \begin{array}{lll} g^s(x_0) = 0 & & \\ s \in \ (1,2,....,m) & \text{avec} & m \leqslant \min{(n,p)} \\ g^r(x_0) > 0 & & \\ r \in \ (m+1,.....,p) & & \end{array} \right.$$

Nous ne considérons que des points non singuliers du domaine des contraintes, c'est à dire pour lesquels la matrice $(m,n)$ $\|\partial g^s/\partial x^i\|$ [1] est de rang $m$ (indépendance des formes linéaires $dg^s$). Nous écartons ainsi certains points singuliers qui entrent dans le champ usuellement couvert par les conditions de qualification des contraintes ; ce fait a peu d'importance, car de tels points doivent de toute façon faire l'objet d'un traitement spécifique (voir les exemples donnés en annexe 2).

Définissons alors $(n-m)$ fonctions auxiliaires $h^{\bar{s}}(x)$ , ainsi que les $m$ fonctions $h^s(x)$ telles que :

$$(2)\quad \left|\begin{array}{l} h^s(x) \equiv g^s(x) \\ h^{\bar{s}}(x_0) = 0 \\ h^{\bar{s}}(x) \text{ de signe quelconque.} \end{array}\right.$$

Nous choisissons évidemment ces fonctions de manière telle que la matrice complète $(n,n)$ $\|\partial h^j/\partial x^i\|$ [1] soit effectivement de rang n, ce qui est toujours possible pour un point non singulier. (Nous désignons par $j$ un indice parcourant l'ensemble des indices $s$ et $\bar{s}$ ).

Considérons le changement de variables :

$$(3)\quad X^j = h^j(x)$$

$$(4)\quad dX^j = \frac{\partial h^j}{\partial x^i} dx^i$$

Posons :

$$(5)\quad \Phi(X) \equiv f[x(X)]$$

On en déduit :

$$d\Phi = df$$

$$\frac{\partial \Phi}{\partial X^j} dX^j = \frac{\partial f}{\partial x^i} dx^i$$

$$(6)\quad \frac{\partial \Phi}{\partial X^j}\frac{\partial h^j}{\partial x^i} dx^i = \frac{\partial f}{\partial x^i} dx^i$$

La relation (6) étant vérifiée quels que soient les $dx^i$ dans le domaine admissible, on a :

$$(7)\quad \frac{\partial \Phi}{\partial X^j}\frac{\partial h^j}{\partial x^i} = \frac{\partial f}{\partial x^i} \quad \text{ou encore} \quad \frac{\partial \Phi}{\partial X^j}\frac{\partial X^j}{\partial x^i} = \frac{\partial f}{\partial x^i}$$

Cette relation, ainsi établie par calcul différentiel, aurait pu être écrite directement puisqu'elle traduit simplement l'effet d'un changement de variables sur les composantes covariantes $\partial f/\partial x^i$ du vecteur gradient (élément de l'espace dual).

Soit en posant :

$$(8)\quad \lambda_j = -\frac{\partial \Phi}{\partial X^j}$$

$$(9)\quad \frac{\partial f}{\partial x^i} - \frac{\partial \Phi}{\partial X^j}\frac{\partial h^j}{\partial x^i} = 0$$

$$(10)\quad \frac{\partial f}{\partial x^i} + \lambda_j \frac{\partial h^j}{\partial x^i} = 0$$

Or dans les nouvelles coordonnées le problème d'optimisation s'écrit :

$$II \quad \left\|\begin{array}{l} Max\ \Phi(X) \\ X^s \geqslant 0 \end{array}\right.$$

Pour un déplacement admissible au voisinage de $x_0$ on a donc :

1. Toutes les dérivées ou différentielles sont, sauf exceptions signalées, supposées être prises au point $x_0$ examiné.

$$(11) \qquad \left| \begin{array}{l} d\Phi = \dfrac{\partial\Phi}{\partial X^s}\, dX^s + \dfrac{\partial\Phi}{\partial X^{\bar{s}}}\, dX^{\bar{s}} \\ dX^s \geqslant 0 \end{array} \right.$$

Pour avoir $d\Phi \leqslant 0$ il faut (condition nécessaire) avoir :

$$(12) \qquad \left| \begin{array}{l} \lambda_s = -\dfrac{\partial\Phi}{\partial X^s} \geqslant 0 \\ \lambda_{\bar{s}} = -\dfrac{\partial\Phi}{\partial X^{\bar{s}}} = 0 \end{array} \right.$$

En notant que les fonctions $h^{\bar{s}}$, a priori quelconques, peuvent dans l'expression finale tirée de (6) et (12) être remplacées formellement par les fonctions $g^r$ auxquelles on associe des multiplicateurs $\lambda_r$ nuls, on a donc retrouvé et démontré **les conditions nécessaires de Kuhn et Tucker**. Soit en revenant aux coordonnées initiales :

$$(13) \qquad \left| \begin{array}{l} \dfrac{\partial f}{\partial x^i} + \lambda_k \dfrac{\partial g^k}{\partial x^i} = 0 \qquad \text{pour} \qquad x = x_0 \\ \lambda_s \geqslant 0 \qquad \lambda_r = 0 \end{array} \right.$$

La première ligne traduit simplement l'effet du changement de coordonnées (curvilignes) en $x_0$ sur les fonctions $f$ et les fonctions $\Phi$, changement dans lequel les fonctions non saturées ne jouent aucun rôle. Ceci n'est pas perçu en général car les fonctions $\Phi$ sont dissimulées dans les coefficients $\lambda$ de la fonction de Lagrange ; elles n'apparaissent qu'indirectement par le biais d'interprétations économiques obtenues en desserrant des contraintes. En réalité elles sont au coeur même de la solution du problème posé.

Dans (13) ce sont les $\dfrac{\partial g^k}{\partial x^i}$ qui jouent le rôle de coefficients et non pas les $\lambda_k$ . Les conditions $\lambda_s \geqslant 0$ correspondent aux relations, évidentes dans les nouvelles coordonnées, $\dfrac{\partial\Phi}{\partial X^s} \leqslant 0$ .

Alors que dans certains exposés traditionnels l'introduction de la **fonction de Lagrange** peut paraître artificielle aux jeunes étudiants - un tour de passe-passe mathématique - , elle s'impose ici de manière naturelle du fait de la relation (13) obtenue on ne peut plus simplement.

Les calculs effectués ci-dessus ne sont d'ailleurs que l'expression algébrique d'une démonstration géométrique très simple et intuitive reposant sur les notions de cône tangent et de cône conjugué (dual). Mais ceci nécessiterait quelques définitions préalables. Notons que si nous avons évoqué le langage tensoriel nous n'avons pas eu besoin d'y faire vraiment appel ; toute notre démonstration repose sur des règles de calcul différentiel élémentaire.

La fonction de Lagrange du problème $I$ s'écrit donc :

$$(16) \qquad \mathcal{L}(x) = f(x) + \lambda_k\, g^k(x)$$

et nous avons pour un optimum local :

$$(17) \qquad \frac{\partial\mathcal{L}}{\partial x^i} = \frac{\partial f}{\partial x^i} + \lambda_k \frac{\partial g^k}{\partial x^i} = 0 \qquad \text{avec} \qquad \lambda_s \geqslant 0$$

**2/Les conditions du deuxième ordre.**

Nous pouvons maintenant essayer de déterminer des **conditions suffisantes** pour l'existence d'un optimum local en $x_0$ . Pour cela développons la fonction de Lagrange jusqu'au second ordre au voisinage de $x_0$ (voir justification de ce calcul en espace affine, en annexe 3) :

$$(18) \qquad \left| \begin{array}{l} l^i = x^i - x_0^i \\ lim_{x \longrightarrow x_0}\, \varepsilon_{ij} = 0 \\ \mathcal{L}(x) - \mathcal{L}(x_0) = \frac{1}{2} \dfrac{\partial^2\mathcal{L}}{\partial x^i \partial x^j}\, l^i\, l^j + \varepsilon_{ij}\, l^i\, l^j \end{array} \right.$$

Soit puisque $\lambda_k g^k(x_0)=0$ et $\lambda_k g^k(x)=\lambda_s g^s(x)$ :

$$(19) \qquad f(x)-f(x_0)=-\lambda_s g^s(x)+\frac{1}{2}\frac{\partial^2\mathcal{L}}{\partial x^i\partial x^j}l^i l^j+\varepsilon_{ij}l^i l^j$$

Il faut bien entendu tenir compte du fait que :

$$(20) \qquad \left| \begin{array}{l} g^s(x)\geqslant 0 \quad \text{et} \quad g^s(x_0)=0 \\ \lambda_s\geqslant 0 \end{array}\right.$$

Notons que $\lambda_s=0$ veut dire $\frac{\partial\Phi}{\partial X^s}=0$ , c'est à dire que la ligne de coordonnées $(X^s)$ est tangente en $x_0$ à l'hypersurface d'équation $\Phi(X)=\Phi(0)=f(x_0)$ .

Nous aurions pu mener directement ces calculs dans les nouvelles coordonnées, en développant la fonction $\Phi(X)$ au voisinage de $x_0$, c'est à dire pour $X=0$ . En désignant par $t$ et $q$ deux indices parcourant l'ensemble $(s,\bar{s})$, on obtient :

$$(21) \qquad \Phi(X)-\Phi(0)=\frac{\partial\Phi}{\partial X^s}X^s+\frac{1}{2}\frac{\partial^2\Phi}{\partial X^t\partial X^q}X^tX^q+\varepsilon'_{tq}X^tX^q$$

On vérifie sans difficulté, par un calcul différentiel simple, que $\frac{\partial^2\mathcal{L}}{\partial x^i\partial x^j}l^i l^j=\frac{\partial^2\Phi}{\partial X^t\partial X^q}X^tX^q$ .

On note immédiatement sur l'équation (21) que, $\frac{\partial\Phi}{\partial X^s}X^s$ étant négatif ou nul, la condition

$$(22) \qquad \frac{\partial^2\Phi}{\partial X^t\partial X^q}X^tX^q<0$$

est suffisante pour assurer que $x_0$ est un optimum local.

Elle signifie que la fonction $\Phi(X)$ est concave dans le nouveau système de coordonnées (caractérisé par des contraintes linéaires). Un examen plus détaillé est cependant nécessaire pour déterminer dans quelles situations les conditions du deuxième ordre sont non seulement suffisantes, mais aussi nécessaires.

Il est instructif d'examiner **un cas extrême** avant de traiter le cas général.

A priori un nombre quelconque de $\lambda_s$ , compris entre 0 et $m$ , peuvent être nuls , et $m$ peut éventuellement être égal à $n$ . Considérons le cas où effectivement $n$ contraintes sont saturées et où les $\lambda_s$ correspondants sont tous strictement positifs. Dans les nouvelles coordonnées on peut écrire :

$$(23) \qquad \Phi(X)-\Phi(0)=(-\lambda_s+\varepsilon_s)X^s$$

Quand $X$ tend vers 0 par n'importe quel chemin dans le domaine admissible, les $\varepsilon_s$ peuvent être rendus aussi petits que l'on veut tout en maintenant au moins un $X^s$ non nul. Donc l'on a $\Phi(X)<\Phi(0)$ sans condition supplémentaire du second ordre. L'optimum se situe en un sommet du domaine acceptable (l' hypersurface $f(x)=f(x_0)$ n'est tangente à aucune arête de ce sommet). Les conditions K-T du premier ordre sont alors suffisantes.

Dans le **cas général** nous supposons qu'il existe $m$ contraintes saturées $g^s(x)$ :

$$(24) \qquad \left| \begin{array}{l} g^s(x_0)=0 \\ s\in(1,2,....,m) \qquad \text{avec} \qquad m\leqslant\min(n,p) \end{array}\right.$$

Nous supposons d'autre part que $t'$ coefficients $\lambda_s$ associés à ces contraintes sont nuls. Nous distinguerons les contraintes $g^\sigma(x)$ et $g^{\bar{\sigma}}(x)$ telles que :

$$
\begin{array}{llll}
 & | \quad g^{\sigma}(x_0)=0 & \lambda_{\sigma}=0 \\
(25) & | \quad \sigma \in 1,2,......,t' & t' \leqslant m \\
 & | \quad g^{\bar\sigma}(x_0)=0 & \lambda_{\bar\sigma}>0 \\
 & | \quad \bar\sigma \in t'+1,....,m &
\end{array}
$$

La relation (19) devient donc :

$$(26) \qquad f(x)-f(x_0)=-\lambda_{\bar\sigma}\, g^{\bar\sigma}(x)+\frac{1}{2}\frac{\partial^2\mathcal{L}}{\partial x^i\partial x^j}\, l^i\, l^j+\varepsilon_{i\,j}\, l^i\, l^j$$

Bien entendu l'on a :

$$(27) \qquad -\lambda_{\bar\sigma}\, g^{\bar\sigma}(x)\leqslant 0$$

l'égalité stricte étant réalisée lorsque $x$ satisfait à $g^{\bar\sigma}(x)=0$ pour tous les $\bar\sigma$ . Il est clair donc que la condition :

$$(28) \qquad \frac{\partial^2\mathcal{L}}{\partial x^i\partial x^j}\, l^i\, l^j<0$$

est à la fois **suffisante et nécessaire**. Cette condition peut s'écrire sous la forme :

$$(29) \qquad \frac{\partial^2 f}{\partial x^i\partial x^j}+\lambda_k\frac{\partial^2 g^k}{\partial x^i\partial x^j}=\frac{\partial^2 f}{\partial x^i\partial x^j}+\lambda_{\bar\sigma}\frac{\partial^2 g^{\bar\sigma}}{\partial x^i\partial x^j}<0$$

On retrouve bien le cas classique où la concavité à la fois de la fonction économique et des fonctions représentant les contraintes saturées est une condition suffisante pour l'optimum (si par ailleurs les conditions nécessaires K-T sont satisfaites), mais aussi un cas plus complexe de concavité relative entre ces fonctions [2], où l'optimum peut exister bien que certaines de ces fonctions soient convexes. La conclusion générale a une expression plus simple dans les nouvelles coordonnées où la concavité de la fonction $\Phi(X)$ est donc à la fois suffisante et nécessaire.

Un **exemple tridimensionnel** permet de bien illustrer tout ceci :

On peut supposer, par isomorphisme entre espaces affines, que les nouvelles coordonnées sont les coordonnées affines d'un tel espace. La représentation géométrique des contraintes se fait donc de manière simple par des plans.

- Supposons d'abord, dans un espace $E^3$, qu'il y ait trois contraintes actives, et que l'optimum local se situe au sommet du trièdre ainsi défini, dont l'intérieur constitue le domaine admissible. Si les trois coefficients $\lambda$ sont positifs la surface $\Phi(X)=\Phi(0)$ ne recoupe pas l'intérieur du trièdre, quelle que soit sa courbure. Les conditions K-T nécessaires sont aussi suffisantes.

- Si l'un des coefficients $\lambda$ est nul cela signifie que la surface $\Phi$ est tangente en 0 à l'une des arêtes du trièdre (celle portant la ligne de coordonnées associée au $\lambda$ nul). Si deux coefficients sont nuls cela veut dire que la surface $\Phi$ est tangente en 0 à l'une des faces du trièdre, et donc à deux arêtes. Les trois coefficients ne peuvent être nuls ensemble.

- Supposons que deux contraintes seulement soient actives. Elles définissent l'arête d'un dièdre passant par 0 (point optimal examiné). Il y a alors obligatoirement un $\lambda$ nul, correspondant à la contrainte non active, et cela se traduit par $\Phi$ tangente à l'arête en 0. Si deux coefficients sont nuls cela veut dire que $\Phi$ est tangente à la fois à l'arête et à l'une des faces qui la définit.

- Supposons enfin qu'une seule contrainte soit active. Il y a alors obligatoirement deux $\lambda$ nuls (contraintes non actives) et cela veut dire que $\Phi$ est tangente en 0 à la seule face active.

Dans tous les cas autres que l'optimum au sommet (sans tangence) il faut faire appel à la condition du second ordre pour s'assurer que la surface $\Phi$ ne pénètre pas partiellement dans le domaine admissible au voisinage du point de tangence.

2. Une image plus géométrique est qu'il s'agit de la courbure relative des surfaces représentatives de ces fonctions.

## Annexe I

### Enoncé et démonstration du théorème de Farkas.

(1) $\quad x \in E^n \qquad v^{(\cdot)}, w \in E^{n*}$

(2) $\quad x \in \mathcal{D} \qquad \mathcal{D}\colon \quad v^{(s)}.\,x = 0 \qquad v^{(p)}.\,x \geqslant 0$

On suppose les vecteurs $v$ linéairement indépendants, ce qui implique en particulier que leur nombre soit inférieur ou égal à $n$. Par commodité de notation nous écrivons $v.x$ la forme linéaire $v(x) = v_i\, x^i$ .

(3) $\quad w.x \geqslant 0$

(4) $\quad w = \lambda_s\, v^{(s)} + \lambda_p\, v^{(p)} \qquad \lambda_p \geqslant 0$

Le théorème de Farkas dit que (2)+(3) implique (4) et que (2)+(4) implique (3).

La deuxième partie de la proposition est évidente. La première, réputée difficile, se démontre comme suit par un changement de variables :

Nous complétons les vecteurs $v^{(\cdot)}$ à $n$ en ajoutant le nombre nécessaire de vecteurs $v^{(t)}$ linéairement indépendants des précédents.

L'ensemble des vecteurs $v$ définit dans $E^{n*}$ une nouvelle base, et dans $E^n$ une nouvelle base duale que nous n'avons pas besoin d'expliciter. Soient $X^j$ les nouvelles coordonnées dans $E^n$, $W_j$ les nouvelles coordonnées dans $E^{n*}$. Les règles tensorielles de transformations des coordonnées permettent d'écrire :

(5) $\quad X^j = v^{(j)}.\,x = v_i^{(j)} x^i \qquad w_i = v_i^{(j)} W_j \qquad$ avec $\quad X^s = 0 \qquad X^p \geqslant 0$

On note que :

(6) $\quad w.x = w_i\, x^i = W_j\, X^j \geqslant 0$

ce qui implique donc :

(7) $\quad W_t = 0 \qquad W_p \geqslant 0 \qquad W_s$ de signe quelconque

Enfin la deuxième relation (5) permet d'écrire :

(8) $\quad w = v^{(j)}\, W_j = v^{(s)}\, W_s + v^{(p)}\, W_p$

ce qui est la relation cherchée.

## Annexe II

### Examen de points singuliers.

Les conditions usuelles de qualification des contraintes acceptent certains points que nous considérons comme singuliers en vertu de la condition d'indépendance des formes linéaires $g^s(x_0)$ .

Un exemple est fourni par le problème suivant :

$$
(1) \qquad
\begin{array}{l}
\| \ \text{Max } f(x) = -\,(x^1)^2 - 2(x^2) \\
\| \ g^1(x) = +\,(x^1)^2 - (x^2) \geqslant 0 \\
\| \ g^2(x) = +\,(x^1)^2 + (x^2) \geqslant 0
\end{array}
$$

Le cône tangent en O s'identifie avec l'axe des $(x^1)$ et les arcs tangents appartenant au domaine défini existent.

Il est facile de vérifier graphiquement que le point origine est l'optimum et que les conditions nécessaires K-T sont remplies, puisque :

(2) $\quad \operatorname{grad} f(0) + 0 \operatorname{grad} g^1(0) + 2 \operatorname{grad} g^2(0) = 0$

Notons cependant que les coefficients K-T ne sont pas uniques.

Considérons en revanche le domaine défini par les contraintes suivantes :

(3) $$\left\| \begin{array}{l} g^1(x) = -(x^1)^3 + (x^2) \geqslant 0 \\ g^2(x) = +(x^1)^4 - (x^2) \geqslant 0 \end{array} \right.$$

Les arcs tangents n'existent pas pour les points du cône tangent situés à droite de l'origine. On constate graphiquement que le point origine peut être un optimum local pour de nombreuses fonctions économiques de gradient quelconque.

Il nous semble que ces exemples élémentaires montrent que les points pour lesquels l'indépendance des formes linéaires $g^s(x_0)$ n'est pas respectée doivent faire l'objet d'un examen particulier, et qu'en conséquence les contraintes usuelles de qualification des contraintes n'ont qu'un intérêt théorique. Une méthode pratique, valable dans les cas complexes, pourrait être d'éliminer certaines contraintes puis d'examiner si les optima ainsi trouvés satisfont néanmoins à ces contraintes.

## Annexe III

### Développements limités en espace affine.

Le fait de travailler en espace simplement affine empêche évidemment d'utiliser toute démonstration faisant intervenir des normes. C'est ainsi que pour définir un voisinage du point $x_0$ il faut introduire $n$ relations du type :

(1) $$x^i - x_0^i \leqslant \mu^i$$

Il ne s'agit pas là d'une norme car cela n'aurait pas de sens d'ordonner les $\mu^i$ . Dire que $x$ tend vers $x_0$ veut dire qu'avec un système de vecteurs de base donné on fait tendre simultanément et de manière indépendante tous les $\mu^i$ vers 0.

Considérons la fonction :

(2) $$\varphi(t) = f(x_0 + th) \qquad h \in E^n$$

On peut écrire le développement limité suivant :

(3) $$\varphi(1) = \varphi(0) + \varphi'(0) + \tfrac{1}{2}\varphi''(\theta) \qquad 0 \leqslant \theta \leqslant 1$$

On obtient donc :

(4) $$f(x_0 + h) = f(x_0) + \frac{\partial f(x_0)}{\partial x^i} h^i + \frac{\partial^2 f(x_0 + \theta h)}{\partial x^i \partial x^j} h^i h^j$$

Posons :

(5) $$f(x_0 + h) = f(x_0) + \frac{\partial f(x_0)}{\partial x^i} h^i + \frac{\partial^2 f(x_0)}{\partial x^i \partial x^j} h^i h^j + \varepsilon_{ij} h^i h^j$$

La comparaison de ces deux expressions conduit à l'égalité tensorielle suivante :

(6) $$\varepsilon_{ij} h^i h^j = \left(\frac{\partial^2 f(x_0 + \theta h)}{\partial x^i \partial x^j} - \frac{\partial^2 f(x_0)}{\partial x^i \partial x^j}\right) h^i h^j$$

Cette égalité devant être vraie quel que soit $h$ , on obtient en définitive :

(7) $$\varepsilon_{ij} = \frac{\partial^2 f(x_0 + \theta h)}{\partial x^i \partial x^j} - \frac{\partial^2 f(x_0)}{\partial x^i \partial x^j}$$

(8) $\quad lim_{h \longrightarrow 0} \varepsilon_{ij} = 0$

Ces calculs justifient les équations (18) et similaires de la note principale. Il n'est donc point besoin de faire usage de normes même pour étudier les variations du second ordre.

Remarque : En calcul tensoriel on peut généraliser ce développement à des coordonnées curvilignes en remplaçant les dérivées secondes par les dérivées covariantes correspondantes.

G.Ringeisen

juillet 2004