Formes quadratiques

Jean-François Burnol, 15 octobre 2010

Soit $N \geq 1$ et $V = \mathbf{R}^N$ l'espace euclidien à N dimensions. À toute matrice symétrique $M = (m_{ij})_{1\leq i,j\leq N}$ est associée la forme bilinéaire symétrique $\phi(x,y) = \sum_{i,j} m_{ij}x_i y_j$ et la forme quadratique $q(x) = \sum_{i,j} m_{ij}x_i x_j$.

Tout d'abord nous nous intéressons aux formes quadratiques positives et aux moyens de les reconnaître.

Définition 1. *On dit que q est positive si $\forall x\ q(x) \geq 0$ et définie positive si elle est positive et si $q(x) = 0 \implies x = 0$.*

Théorème 1 (Critère de Sylvester)**.** *Pour que q soit définie positive il est nécessaire et suffisant que les N mineurs diagonaux partant du coin supérieur gauche de la matrice M soient strictement positifs.*

Il est plus commode pour la preuve de montrer l'énoncé plus complet suivant :

Théorème 2. *Les conditions suivantes sont équivalentes :*

1. *les N mineurs diagonaux partant du coin supérieur gauche de la matrice M sont tous strictement positifs,*
2. *il existe une matrice P avec ${}^tPMP = I$,*
3. *q est définie positive.*

(1) $\implies$ (2). On le fait par récurrence. Vrai pour $N = 1$. Supposons $N > 1$ et vrai pour $N - 1$. On prend P_0 qui convient pour le bloc M_0 de taille $N - 1$ en haut à gauche de M, notez que $\det(P_0)^2 \det(M_0) = 1$ donc $\det(P_0) \neq 0$. Un calcul de multiplication par blocs donne (le résultat doit être une matrice symétrique) :

$$\begin{pmatrix} {}^tP_0 & 0 \\ 0 & 1 \end{pmatrix} M \begin{pmatrix} P_0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} & & & a_1 \\ & I_{N-1} & & \vdots \\ & & & a_{N-1} \\ a_1 & \cdots & a_{N-1} & a_N \end{pmatrix}$$

Des manipulations simples de colonnes montrent que le déterminant de la matrice de droite vaut :

$$a_N - \sum_{1\leq j\leq N-1} a_j^2$$

Cette expression est $\det(P_0)^2 \det(M)$ donc strictement positive, notons-la δ^2. On a l'identité, dont la vérification par le calcul est laissée au lecteur :

$$\begin{pmatrix} & & & 0 \\ & I_{N-1} & & \vdots \\ & & & 0 \\ -\frac{1}{\delta}a_1 & \cdots & -\frac{1}{\delta}a_{N-1} & \frac{1}{\delta} \end{pmatrix} \begin{pmatrix} & & & a_1 \\ & I_{N-1} & & \vdots \\ & & & a_{N-1} \\ a_1 & \cdots & a_{N-1} & a_N \end{pmatrix} \begin{pmatrix} & & & -\frac{1}{\delta}a_1 \\ & I_{N-1} & & \vdots \\ & & & -\frac{1}{\delta}a_{N-1} \\ 0 & \cdots & 0 & \frac{1}{\delta} \end{pmatrix} = I_N$$

Cela nous donne au final une matrice P avec ${}^tPMP = I$ comme promis. $\square$

$(2) \implies (3)$. Comme ${}^tPMP = I$, on a $\det(P) \neq 0$. Soit $x \in \mathbf{R}^N$, vu comme un vecteur colonne. Soit q_0 la forme quadratique canonique (somme des carrés des coordonnées canoniques). Alors $q_0(x) = {}^txx = {}^t(Px)M(Px) = q(Px)$. Pour $y \neq 0$, il existe un unique x, non nul, avec $Px = y$, donc $q(y) = q(Px) = q_0(x) > 0$. □

$(3) \implies (1)$. Ici encore on fait une récurrence sur N. Vrai pour $N = 1$. Supposons $N > 1$ et vrai pour $N - 1$. On note que le bloc diagonal M_0 de taille $(N-1) \times (N-1)$ en haut à gauche de la matrice M est la matrice de la forme quadratique q restreinte au sous-espace $\mathbf{R}^{N-1} \subset \mathbf{R}^N$ engendré par les $N-1$ premiers vecteurs de la base canonique. Donc par l'hypothèse de récurrence tous les mineurs diagonaux $k \times k$, partant du haut gauche sont strictement positifs. On sait déjà $(1) \implies (2)$ donc on a à nouveau un P_0 (forcément inversible) tel que ${}^tP_0M_0P_0 = I_{N-1}$. Avec les mêmes notations que dans la preuve de $(1) \implies (2)$ on en déduit que $\det M$ est du signe de $a_N - \sum_{1\leq j<N} a_j^2$. Calculons alors :

$$\begin{aligned}
&\begin{pmatrix} -a_1 & \cdots & -a_{N-1} & 1 \end{pmatrix} \begin{pmatrix} {}^tP_0 & 0 \\ 0 & 1 \end{pmatrix} M \begin{pmatrix} P_0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -a_1 \\ \vdots \\ -a_{N-1} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} -a_1 & \cdots & -a_{N-1} & 1 \end{pmatrix} \begin{pmatrix} & & & a_1 \\ & I_{N-1} & & \vdots \\ & & & a_{N-1} \\ a_1 & \cdots & a_{N-1} & a_N \end{pmatrix} \begin{pmatrix} -a_1 \\ \vdots \\ -a_{N-1} \\ 1 \end{pmatrix} \\
&= \begin{pmatrix} -a_1 & \cdots & -a_{N-1} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_N - \sum_{1\leq j<N} a_j^2 \end{pmatrix} = a_N - \sum_{1\leq j<N} a_j^2
\end{aligned}$$

Ce calcul prouve que $a_N - \sum_{1\leq j<N} a_j^2$ est $q(x)$ pour un certain x, d'ailleurs non nul. Il est donc strictement positif, ainsi $\det M > 0$. □

Théorème 3. *Pour que q soit positive il est nécessaire et suffisant que tous les mineurs diagonaux (il y en a $2^N - 1$) soient positifs ou nuls.*

Preuve. Soit q_0 la forme quadratique canonique, de matrice I. Si q est positive, alors $q + \epsilon q_0$, de matrice $M + \epsilon I$, est strictement positive pour tout $\epsilon > 0$. Sa restriction à n'importe quel sous-espace de $\mathbf{R}^N$ engendré par k parmi les N vecteurs de la base canonique est donc strictement positive et la matrice associée a donc un déterminant strictement positif par le théorème précédent. En prenant la limite pour $\epsilon \to 0$ on obtient donc que tous les mineurs diagonaux de M sont positifs ou nuls.

Réciproquement supposons que cela soit le cas et considérons à nouveau la matrice $M + \epsilon I$. Notons $c_1, \ldots, c_N$ les colonnes de M, $e_1, \ldots, e_N$ les vecteurs canoniques en forme de colonnes, alors par multilinéarité du déterminant :

$$\det(M + \epsilon I) = \sum_{0\leq k\leq N} \sum_{1\leq i_1<i_2<\cdots<i_k\leq N} \epsilon^k d(i_1, i_2, \ldots, i_k)$$

où $d(i_1, i_2, \ldots, i_k)$ se calcule en remplaçant dans $\det M$ chacune des colonnes c_{i_j} par e_{i_j} ($d(\emptyset) = \det M$). En développant par rapport à la colonne i_1 puis la colonne i_2 etc..., on se convainc (enfin, moi je me convaincs) que $d(i_1, i_2, \ldots, i_k)$ n'est pas autre chose que le mineur diagonal de taille $N-k$ correspondant aux indices complémentaires de $\{i_1, i_2, \ldots, i_k\}$ dans $\{1, \ldots, N\}$. Il est donc par hypothèse positif ou nul. Donc $\det(M + \epsilon I) = \epsilon^N + \ldots$ a ses autres coefficients positifs ou nuls, et par conséquent $\det(M + \epsilon I) > 0$ pour tout $\epsilon > 0$. De plus le même résultat vaut pour les mineurs diagonaux de $M + \epsilon I$, en particulier ceux ancrés en haut à gauche. Par le théorème précédent $q + \epsilon q_0$ est strictement positive. En particulier elle est positive, $q(x) \geq -\epsilon q_0(x)$ et en prenant $\epsilon \to 0$ on obtient que q est positive. □

On notera que d'après le Théorème 2 pour établir que q est strictement positive il suffit de montrer que les N mineurs diagonaux principaux sont strictement positifs. Mais alors en fait **tous** les mineurs diagonaux seront strictement positifs (car la restriction de q à n'importe quel sous-espace de $\mathbf{R}^N$ est encore strictement positive). Mais l'exemple de $\mathrm{diag}(0, \ldots, 0, -1)$ montre que l'analogue est faux pour montrer qu'une q est positive.

Ceci donne donc au passage un exemple très intéressant du fait suivant : dans un espace topologique X, un ouvert U défini par N équations $f_1 > 0$, ..., $f_N > 0$ avec f_1, ..., $f_N : X \to \mathbf{R}$ des fonctions continues, peut avoir une adhérence $F = \overline{U}$ plus petite que le fermé défini par les équations $f_1 \geq 0$, ..., $f_N \geq 0$. Le fermé $\overline{U}$ est sûrement inclus dans celui défini par ces équations, mais cette inclusion peut être stricte.

Dans ce qui précède je n'ai pas fait appel aux résultats généraux sur les formes quadratiques. Je rattrape partiellement cette lacune. Le résultat fondamental est le suivant :

Théorème 4. *Toute matrice symétrique réelle M est diagonalisable dans une base orthonormée.*

On voit M aussi comme un endomorphisme de $\mathbf{R}^N$.

Lemme 1. *Les valeurs propres de M sont toutes réelles. De plus les espaces propres associés sont mutuellement orthogonaux.*

Preuve. Soit $\mathbf{C}^N$, muni du produit scalaire hermitien canonique $(x, y) = \sum_{1 \leq j \leq N} \overline{x_j} y_j = {}^t\overline{x} y$. On peut voir M comme un endomorphisme de $\mathbf{C}^N$. On a $(Mx, y) = {}^t\overline{Mx} y = {}^t\overline{x} M y = (x, My)$ car $M = {}^t\overline{M}$. Soit λ une valeur propre, x un vecteur propre associé (non nul). Alors $(Mx, x) = (x, Mx)$ donne $\overline{\lambda}(x, x) = \lambda(x, x)$ donc λ est en fait réel.

Si maintenant, en revenant à $\mathbf{R}^N$ muni de la forme quadratique canonique q_0, on a λ_1 et λ_2 deux valeurs propres distinctes et x_1 et x_2 des vecteurs propres respectifs, alors $\lambda_1(x_1, x_2) = (Mx_1, x_2) = (x_1, Mx_2) = \lambda_2(x_1, x_2)$ donc $(x_1, x_2) = 0$. □

Preuve du théorème. Soit λ_1, ..., λ_k les valeurs propres distinctes de M, d'espaces propres associés E_1, ..., E_k. Ils sont en somme directe. Soit $E = E_1 \oplus \cdots \oplus E_k$ et $F = E^\perp$. Comme E est stable par M, il en est de même de F : si $f \in F$ et $e \in E$ alors $(e, Mf) = (Me, f) = 0$ donc $Mf \perp e$ donc $Mf \in F$. Si F n'est pas nul, prenons-y une base orthonormée $(e_1, \ldots, e_n)$. Alors la matrice $N = (a_{ij})$ de l'endomorphisme M dans cette base a pour j^{e} colonne les coordonnées de Me_j dans cette base et elles se calculent par des produits scalaires : $a_{ij} = (e_i, Me_j)$. Donc $a_{ij} = a_{ji}$ et N est une matrice symétrique. Mais alors par le Lemme, l'endomorphisme M sur F a au moins un vecteur propre, contradiction. Donc $F = \{0\}$. Prenons maintenant dans chaque E_j une base orthonormée. Comme ils sont mutuellement orthogonaux, nous obtenons une base orthonormée de $\mathbf{R}^N = E_1 \perp E_2 \cdots \perp E_k$. Et cette base est formée de vecteurs propres pour l'endomorphisme M. □

Notons P la matrice exprimant la nouvelle base (formée de vecteurs propres) dans l'ancienne. Comme la nouvelle base est orthonormée, P est une matrice orthogonale ${}^tPP = I$. La matrice diagonale D de l'endomorphisme M dans la nouvelle base est $P^{-1}MP$, qui est aussi ${}^tPMP$. Ainsi :

Théorème 5. *Pour toute matrice symétrique réelle, il existe une matrice orthogonale P telle que ${}^tPMP$ est diagonale.*

Notons, avec répétitions éventuelles, $\lambda_1, \ldots, \lambda_r$ les valeurs propres positives et $-\mu_1, \ldots, -\mu_s$ les valeurs propres négatives, avec $N-r-s$ la multiplicité de la valeur propre nulle (si 0 est valeur propre). On peut choisir P de sorte que $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_r, -\mu_1, \ldots, -\mu_s, \ldots, 0)$. Posons alors

$$Q = P\,\mathrm{diag}(\lambda_1^{-\frac{1}{2}}, \ldots, \lambda_r^{-\frac{1}{2}}, \mu_1^{-\frac{1}{2}}, \ldots, \mu_s^{-\frac{1}{2}}, 1, \ldots, 1)$$

On obtient ${}^tQQ = \mathrm{diag}(\lambda_1^{-1}, \ldots, \lambda_r^{-1}, \mu_1^{-1}, \ldots, \mu_s^{-1}, 1, \ldots, 1)$, et

$${}^tQMQ = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$$

La matrice Q est inversible, et si on écrit $x = Qy$, ce qui signifie que $(y_1, \ldots, y_N)$ sont les coordonnées de x dans la base formée par les colonnes de Q, qui est une base orthogonale (pas orthonormée) de $\mathbf{R}^N$, on obtient

$$q(x) = {}^txMx = {}^ty\,{}^tQMQy = y_1^2 + \cdots + y_r^2 - y_{r+1}^2 - \cdots - y_{r+s}^2$$

Cette réduction est le fameux théorème de Jacobi, qui nous dit plus encore : r et s ne dépendent pas de la base, sans même supposer celle-ci orthogonale ; précisément : si l'on peut trouver une matrice inversible R de sorte qu'en posant $x = Rz$ ce qui relie les nouvelles coordonnées $z_1, \ldots, z_N$ aux anciennes $x_1, \ldots, x_N$, on obtienne

$$q(x) = z_1^2 + \cdots + z_u^2 - z_{u+1}^2 - \cdots - z_{u+v}^2$$

alors $u = r$ et $v = s$.

Preuve. Supposons $u < r$. Considérons le sous espace K de dimension r, défini par les équations $\{y_{r+1} = \cdots = y_N = 0\}$. La forme quadratique q est définie positive sur K. Mais on peut trouver dans K un vecteur x non nul avec $z_1 = \cdots = z_u = 0$, puisque $u < \dim K$. On a alors à la fois $q(x) > 0$ et $q(x) \leq 0$, contradiction. Donc $u \geq r$ et en intervertissant les rôles $r \geq u$, donc $u = r$. En appliquant ceci à $-q$ on obtient aussi $v = s$. □

La forme quadratique q est positive si et seulement si $s = 0$ (c'est-à-dire si et seulement si les valeurs propres de la matrice M sont toutes positives) et elle est définie positive si et seulement si $s = 0$ et $r = n$, c'est-à-dire si et seulement si les valeurs propres de M sont toutes strictement positives. Une forme quadratique positive est strictement positive si et seulement si $\det M > 0$.

Bien sûr si on avait disposé de ces résultats on aurait pu faire plus rapidement certaines des preuves des Théorèmes 2 et 3.