Gram et Hadamard

Jean-François Burnol, 24 mars 2011

Matrices de Gram

Définition. Soit E un espace vectoriel euclidien. La *matrice de Gram* associée à n vecteurs $x_1, \ldots, x_n$ est la matrice symétrique $G(\mathbf{x}) = (g_{ij})_{1\leq i,j\leq n}$ avec $g_{ij} = (x_i, x_j)$. Cette matrice symétrique définit une forme quadratique q sur l'espace vectoriel $\mathbf{R}^n$ et

$$q(\mathbf{t}) = \left\| t_1 x_1 + \cdots + t_n x_n \right\|_E^2 \geq 0$$

pour $\mathbf{t} = {}^t(t_1, \ldots, t_n) \in \mathbf{R}^n$. [1] Cette formule montre que la forme quadratique q est positive, et qu'elle est définie positive si et seulement si les vecteurs $x_1, \ldots, x_n$ sont linéairement indépendants dans E. Posons

$$g(\mathbf{x}) \coloneqq \det G(\mathbf{x}) \ .$$

Ce déterminant $g(\mathbf{x})$ est le *gramien* du système $\mathbf{x}$.

Si le système $\mathbf{x}$ est linéairement dépendant l'un des vecteurs, disons x_1, s'exprime comme combinaison $\lambda_2 x_2 + \cdots + \lambda_n x_n$ des autres. La première ligne de la matrice $G(\mathbf{x})$ est donc combinaison linéaire des autres lignes, et $g(\mathbf{x})$ est par conséquent nul.

Réciproquement si $g(\mathbf{x}) = 0$ soit $\mathbf{t} = {}^t(t_1, \ldots, t_n) \in \mathbf{R}^n \setminus \{0\}$ de sorte que $G(\mathbf{x}) \cdot \mathbf{t} = 0$. Alors ${}^t\mathbf{t} \cdot G(\mathbf{x}) \cdot \mathbf{t} = 0$ et donc $t_1 x_1 + \cdots + t_n x_n$ est le vecteur nul. Ce qui montre que $\mathbf{x}$ est linéairement dépendant.

On a compris ce qui se passe lorsque le gramien est nul, on va s'intéresser maintenant à son signe.

Transformation. Supposons que l'on passe à un nouveau système de vecteurs $\mathbf{y} = (y_1, \ldots, y_m)$ dépendant linéairement du premier. Soit T la matrice exprimant $\mathbf{y}$ dans $\mathbf{x}$. C'est une matrice avec n lignes et m colonnes. On calcule :

$$G(\mathbf{y})_{ij} = g(y_i, y_j) = g(\sum_k t_{ki} x_k, \sum_l t_{lj} x_l) = \sum_{k,l} t_{ki} G(\mathbf{x})_{kl} t_{lj} = \left({}^tT \cdot G(\mathbf{x}) \cdot T\right)_{ij}$$

1. J'utiliserai des caractères gras pour désigner des n-uplets que ceux-ci soient composés de scalaires ou de vecteurs. Et j'identifie $\mathbf{R}^n$ aux colonnes à n entrées.

soit encore :

$$G(\mathbf{y}) = {}^{t}T \cdot G(\mathbf{x}) \cdot T$$

et donc, lorsque $m = n$:

$$g(\mathbf{y}) = \big(\det T\big)^2 \; g(\mathbf{x}) \, .$$

Calcul via Gram-Schmidt. Lorsque $\mathbf{x}$ est un système linéairement indépendant, la méthode de Gram-Schmidt nous permet de construire récursivement un système orthonormé $\mathbf{e}$ tel que pour chaque k les vecteurs $x_1, \ldots, x_k$ et $e_1, \ldots, e_k$ engendrent le même espace. La matrice T exprimant $\mathbf{x}$ dans $\mathbf{e}$ est donc triangulaire supérieure et ses éléments diagonaux sont les (x_i, e_i). Nous obtenons, puisque le gramien d'un système orthonormé vaut 1 :

$$g(\mathbf{x}) = \big(\det T\big)^2 \; g(\mathbf{e}) = \Big(\prod_{1 \le i \le n} (x_i, e_i) \Big)^2 \, .$$

Or, $\|x_i\|_E^2 = \sum_{1 \le j \le i} |(x_i, e_j)|^2 \ge |(x_i, e_i)|^2$, avec égalité seulement si x_i est proportionnel à e_i. Ceci nous donne :

Théorème. *Pour tout système* $\mathbf{x}$ *de* n *vecteurs* $x_1, \ldots, x_n$ *dans un espace euclidien* E *le gramien* $g(\mathbf{x})$ *vérifie :*

$$0 \le g(\mathbf{x}) \le (\|x_1\| \cdots \|x_n\|)^2 \, ,$$

avec égalité pour la majoration si et seulement si les x_i *sont soit linéairement indépendants et mutuellement orthogonaux, soit non tous non nuls.*

Déterminant d'une matrice positive. Soit maintenant $M = (m_{ij})$ une matrice de $\mathcal{S}_n(\mathbf{R})$, positive. Cela signifie que la forme quadratique q_M correspondant à M sur l'espace vectoriel $\mathbf{R}^n$ est positive. Prenons $\varepsilon > 0$ et considérons $q_{M+\varepsilon I}$ qui est définie positive et fait de $\mathbf{R}^n$ un espace euclidien E. La matrice $M + \varepsilon I$ est la matrice de Gram des n vecteurs de la base canonique $\mathbf{e} = (e_1, \ldots, e_n)$.[2] Son déterminant est donc strictement positif. En passant à la limite pour $\varepsilon \to 0^+$ on obtient le résultat bien connu :

$$\det(M) \ge 0 \qquad \textit{pour une matrice symétrique positive.}$$

2. attention aux notations : dans le paragraphe précédent on a utilisé les e_j pour désigner une base orthonormée ; ici ce sont juste les vecteurs de la base canonique de $\mathbf{R}^n$.

On peut aussi passer à la limite dans la majoration

$$\det(M + \varepsilon I) \leq \|e_1\|^2 \cdots \|e_n\|^2 = (m_{11} + \varepsilon) \cdots (m_{nn} + \varepsilon) \,,$$

ce qui donne :

$$\det(M) \leq m_{11} \cdots m_{nn} \qquad \textit{pour une matrice symétrique positive.}$$

Cas d'égalité. Si $\det(M) = 0$ il y aura égalité pour la majoration si et seulement si l'un des éléments diagonaux est nul. Si $\det(M) > 0$ alors la forme quadratique q_M est définie positive et il n'y a pas besoin dans le raisonnement précédent du ε. La matrice M est la matrice de Gram des vecteurs canoniques $\mathbf{e}$ de $\mathbf{R}^n$ muni de q_M. Ainsi il y a égalité si et seulement si ces vecteurs sont mutuellement orthogonaux, c'est-à-dire si et seulement si M est diagonale. Récapitulons :

Théorème. *Une matrice* M *symétrique et positive a un déterminant qui est positif et qui est majoré par le produit des éléments diagonaux. Cette majoration est une égalité si et seulement si*

- *soit l'un des éléments diagonaux est nul,*
- *soit* M *est une matrice diagonale.*

Preuve. Il faut peut-être rajouter la remarque que si l'un des éléments diagonaux, disons m_{11}, est nul alors toute la première ligne et toute la première colonne sont nulles et automatiquement $\det(M) = 0$. En effet lorsque l'on a une forme bilinéaire $(\cdot, \cdot)$ associée à une forme quadratique positive q, alors, sans avoir à supposer que q est un produit scalaire, on a toujours l'inégalité de Cauchy-Schwarz $|(x, y)|^2 \leq q(x)q(y)$, qui montre l'implication $q(x) = 0 \Rightarrow \forall y\ (x, y) = 0$. Dire $m_{11} = 0$ c'est dire $q_M(e_1) = 0$ d'où la conclusion. Ainsi pour la preuve du théorème, on commence par dire que si l'un des éléments diagonaux est nul alors la matrice a une colonne nulle et donc un déterminant nul, et il y a égalité ; puis que si tous les éléments diagonaux sont non nuls il ne peut y avoir égalité que si q_M est définie positive et la conclusion suit alors par ce qui précède. □

Inégalité de Hadamard

Considérons maintenant une matrice réelle carrée N. On lui associe la matrice symétrique $M = {}^tN \cdot N$. Cette matrice est automatiquement positive car ${}^t\mathbf{t} \cdot M \cdot \mathbf{t} = {}^t(N\mathbf{t}) \cdot N\mathbf{t} \geq 0$. Son déterminant est positif (bon c'était clair car il vaut $\det(N)^2$) et il est majoré par le produit de ses éléments diagonaux. Mais cela n'est autre que :

Théorème (Hadamard)**.** *Pour toute matrice réelle carrée* $N = (n_{ij})_{1\leq i,j\leq n}$ *on a la majoration*

$$\left|\det N\right| \leq \prod_{1\leq j\leq n} \sqrt{\sum_{1\leq i\leq n} |\, n_{ij}\,|^2}$$

L'égalité n'a lieu que si : soit l'une des colonnes de N *est identiquement nulle, soit elles sont non nulles et mutuellement orthogonales pour le produit scalaire canonique.*

Le cas complexe ... est plus complexe ! Il faut alors en passer par des produits scalaires complexes (et les formes sesquilinéaires associées) et faire attention à la façon dont on définit les matrices de Gram, pour que la loi de transformation dans un changement de base prenne une forme sympathique (c'est-à-dire il est peu élégant si, comme dans énormément de livres, on voit apparaître séparément des ${}^tT$ et des $\overline{T}$ plutôt qu'uniquement la matrice adjointe T^*). Je laisse les détails au lecteur déterminé.

Décomposition d'Iwasawa

Finalement la méthode pour prouver l'inégalité de Hadamard a été :

1. je prends ma matrice N que je peux supposer inversible,
2. je forme la matrice symétrique $M = {}^tN \cdot N$,
3. je prouve qu'il existe une matrice triangulaire supérieure T telle que l'on ait aussi $M = {}^tT \cdot T$: cette étape peut se faire de différentes façons, celle suivie ici a été de l'identifier au procédé de Gram-Schmidt dans les espaces euclidiens,
4. je déduis de $M = {}^tT \cdot T$ que $\det(M)$ est positif et majoré par le produit des éléments diagonaux de M,
5. et c'est cela l'inégalité de Hadamard, en fait.

Maintenant de ${}^tN \cdot N = {}^tT \cdot T$ il vient

$$({}^tT)^{-1}\ {}^tN \cdot NT^{-1} = I = {}^t(NT^{-1}) \cdot NT^{-1}$$

donc $NT^{-1} = K$ est une matrice orthogonale. On peut écrire $T = AN$ avec A une matrice diagonale, et N une matrice triangulaire supérieure avec des 1 sur la diagonale (matrice dite *unipotente*), et au final toute matrice inversible peut s'écrire sous la forme :

$$N = K \cdot A \cdot N$$

C'est ce qu'on appelle une décomposition d'Iwasawa, non pas qu'Iwasawa ait obtenu ce résultat particulier très ancien, mais parce qu'il en a généralisé l'énoncé à d'autres groupes de matrices que celui de toutes les matrices inversibles (il s'agit alors d'une décomposition au sein d'un groupe de Lie réel semi-simple connexe ; le point est que les K, A, et N de la décomposition sont encore dans le groupe de Lie).

Pour nous, la formule avec les notations $N = O \cdot T$ conviendra très bien ! (O matrice orthogonale, T matrice triangulaire supérieure).

Discussion de l'unicité : il n'y a pas unicité, puisque si je prends D une des 2^n matrices diagonales avec uniquement des ± 1 je peux toujours remplacer O par OD et T par DT. Montrons que ce sont là en fait toutes les possibilités. Cela revient à établir :

Théorème. *Toute matrice réelle inversible* N *peut s'écrire de manière unique sous la forme* $N = O \cdot T$ *avec* O *orthogonale et* T *triangulaire supérieure avec une diagonale strictement positive.*

Preuve. On a l'existence d'une écriture $N = O \cdot T$ et pour imposer la positivité sur la diagonale on remplace T par DT avec D formée en prenant les signes des éléments diagonaux de T (comme $\det(T) \neq 0$ on sait que ces éléments sont non nuls). Supposons données deux telles décompositions $N = O \cdot T = O_1 \cdot T_1$ de sorte que $O_1^{-1}O$ est orthogonale et aussi égale à la matrice triangulaire supérieure T_1T^{-1}. La première colonne a donc ± 1 en haut et des zéros en-dessous ; ses produits scalaires avec les autres colonnes doivent être nuls, ce qui montre que la première ligne n'a que des zéros après le ± 1. On itère le raisonnement, et au final on constate que $D = T_1T^{-1}$ est une matrice diagonale composée de ± 1. Comme $DT = T_1$ et que à la fois T et T_1 n'ont que des éléments strictement positifs sur la diagonale c'est qu'en fait D est la matrice identité. Ainsi $T = T_1$ et $O = O_1$. □

Les matrices non-inversibles : si $\det(N) = 0$, alors pour $n \gg 1$ la matrice $N + \frac{1}{n}I$ est inversible (les zéros du polynôme caractéristique sont en nombre fini, donc isolés). Pour ces n on écrit $N + \frac{1}{n}I = O_n \cdot T_n$ avec O_n et T_n comme dans le théorème précédent. Le groupe orthogonal est un fermé borné dans $\mathrm{Mat}_{n\times n}(\mathbf{R}) \cong \mathbf{R}^{n^2}$, donc un compact, et il existe une suite extraite (n_k) avec l'existence de la limite $O = \lim O_{n_k}$. Il existe donc aussi $T = O^{-1}N = \lim O_{n_k}^{-1}(N + \frac{1}{n_k}I) = \lim T_{n_k}$ qui est nécessairement triangulaire supérieure avec des éléments diagonaux positifs ou nuls. Donc $N = O \cdot T$ avec O orthogonale et T triangulaire supérieure avec une diagonale positive.

L'exemple

$$\begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 0 & \cos\theta + \sin\theta \\ 0 & -\sin\theta + \cos\theta \end{pmatrix}$$

avec $-\frac{3\pi}{4} \leq \theta \leq \frac{\pi}{4}$ montre que l'on n'a plus unicité.