

YVES HAUGAZEAU

**Application du théorème de Sylvester à la
localisation des valeurs propres de $AX = \lambda BX$
dans le cas symétrique**

RAIRO. Analyse numérique, tome 14, n° 1 (1980), p. 25-41

http://www.numdam.org/item?id=M2AN_1980__14_1_25_0

© AFCET, 1980, tous droits réservés.

L'accès aux archives de la revue « RAIRO. Analyse numérique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

**APPLICATION DU THÉORÈME DE SYLVESTER
A LA LOCALISATION DES VALEURS PROPRES DE $AX = \lambda BX$
DANS LE CAS SYMÉTRIQUE (*)**

par Yves HAUGAZEAU ⁽¹⁾

Communiqué par F ROBERT

Résumé — Après avoir établi le lien entre, d'une part le théorème de Sylvester sur la décomposition en somme et différence de carrés des formes quadratiques et, d'autre part, la localisation des valeurs propres de $AX = \lambda BX$, par la méthode des suites de Sturm appliquée aux mineurs de $A - \lambda B$, on présente un algorithme de localisation qui généralise cette dernière méthode. Fondé sur une factorisation LDL cet algorithme est applicable, de manière stable, pour toute valeur réelle λ donnée.*

Abstract — Having established the connexion between Sylvester's theorem, about decomposition of quadratic forms in sum and difference of squares, and, on the other hand, the localization of eigen values of a problems $AX = \lambda BX$, using the method of Sturm's sequences, one present an algorithm wich generalizes this last method. Founded on LDL factorization, that algorithm is able, in a stable way, for any given real λ .*

1 VALEURS PROPRES, ET DÉCOMPOSITION EN CARRÉS DES FORMES QUADRATIQUES

1. 1. Rappel élémentaire et notations

Il est bien connu que toute matrice réelle symétrique A est diagonalisable, c'est-à-dire qu'il existe une matrice unitaire Q (composée de vecteurs propres de A) telle que l'on ait

$$A = QDQ^*,$$

(*) Reçu novembre 1978

(1) Université de Bordeaux I, U E R de Mathématiques et Informatique, Talence (France)

où D est une matrice diagonale dont les coefficients diagonaux sont les valeurs propres λ_i de A , et où Q^* , la transposée de Q , est aussi l'inverse de Q .

En posant alors $Y = Q^* X$, la forme quadratique Φ associée à A , admet les deux expressions ci-dessous :

$$\Phi(X) = \langle AX, X \rangle = \langle DY, Y \rangle. \quad (1)$$

Dans ces expressions, la notation $\langle \dots \rangle$ désigne la somme des produits terme à terme, de deux vecteurs de \mathbb{R}^n .

Dans ce paragraphe on appelle \mathcal{L}_i la forme linéaire qui associe à X la i -ième coordonnée de Y .

Les n formes linéaires \mathcal{L}_i sont linéairement indépendantes.

De la deuxième expression de (1) on tire

$$\Phi(X) = \sum_{i=1}^n \lambda_i \mathcal{L}_i(X)^2.$$

On suppose que p_1 et p_2 sont des entiers positifs ou nuls tels que

$$\begin{aligned} \lambda_i < 0 & \quad \text{pour } p_1 < i \leq p_1 + p_2, \\ \lambda_i > 0 & \quad \text{pour } 1 \leq i \leq p_1, \\ \lambda_i = 0 & \quad \text{pour } p_1 + p_2 < i \leq n, \end{aligned}$$

avec la convention habituelle, que l'un de ces ensembles de valeurs de i est vide si les inégalités qui le définissent sont incompatibles.

Le fait d'avoir ainsi groupé en tête les valeurs propres positives, puis les valeurs propres négatives et enfin les valeurs propres nulles est toujours possible, en modifiant au besoin la matrice unitaire Q (en la multipliant à droite par une matrice de permutation convenable).

Il vient :

$$\Phi(X) = \sum_{i=1}^{p_1} |\lambda_i| \mathcal{L}_i(X)^2 - \sum_{i=p_1+1}^{p_1+p_2} |\lambda_i| \mathcal{L}_i(X)^2. \quad (2)$$

La décomposition en carrés donnée par (2) a donc la propriété suivante :

Le nombre p_1 des carrés précédés du signe plus est égal au nombre des valeurs propres strictement positives de A .

Le nombre p_2 des carrés précédés du signe moins est égal au nombre des valeurs propres strictement négatives de A .

Dans cet énoncé les valeurs propres multiples éventuelles ont été comptées avec leur multiplicité.

1.2. Le problème étudié

Soient A et B deux matrices réelles symétriques, la seconde étant définie positive. On s'intéresse au problème classique de valeurs propres :

(P) Trouver λ tel qu'il existe au moins un vecteur X non nul, pour lequel

$$AX = \lambda BX$$

LEMME 1 : Si λ est un nombre réel donné, quelconque, alors la forme quadratique Φ_λ définie par

$$\Phi_\lambda(X) = \langle (A - \lambda B)X, X \rangle, \quad (3)$$

admet une décomposition en carrés pour laquelle :

Le nombre p_1 des carrés précédés du signe plus, est égal au nombre des valeurs propres (distinctes ou non) du problème (P) qui sont strictement supérieures à λ .

Le nombre p_2 des carrés précédés du signe moins, est égal au nombre des valeurs propres du problème (P) qui sont strictement inférieures à λ .

Démonstration : Soit $\{V_i\}$ ($i \leq i \leq n$) une base de vecteurs propres du problème (P). On sait, par la théorie de Courant et Fischer [1] qu'une telle base existe, et de plus que $\langle BV_i, V_j \rangle = 0$ si les valeurs propres λ_i et λ_j ont des valeurs distinctes. La condition $\langle BV_i, V_j \rangle = 0$ pour $i \neq j$ est toujours réalisable, en modifiant éventuellement le choix de la base dans chacun des sous-espaces propres, dans le cas de valeurs propres multiples. Enfin $\langle BV_i, V_i \rangle$ est strictement positif pour tout i puisque B est définie positive.

Avec ces rappels, l'identité (3) s'écrit :

$$\Phi_\lambda(X) = \sum_{i=1}^n \alpha_i^2 (\lambda_i - \lambda) \langle BV_i, V_i \rangle,$$

où α_i désigne la i -ième coordonnée de X sur la base $\{V_i\}$.

Associer à X cette i -ième coordonnée définit une forme linéaire en X que l'on note ici \mathcal{L}_i . Les n formes \mathcal{L}_i sont linéairement indépendantes. Et on a, par conséquent :

$$\Phi_\lambda(X) = \sum_{i=1}^n (\lambda_i - \lambda) \langle BV_i, V_i \rangle \mathcal{L}_i(X)^2.$$

Sur cette expression se lit l'énoncé du lemme 1, puisque chaque nombre $\langle BV_i, V_i \rangle$ est positif.

PROPOSITION 1 : *Si la forme quadratique $\Phi_\lambda(X)$, définie par (3), admet une décomposition en carrés*

$$\Phi_\lambda(X) = \sum_{i=1}^{p_1} \mathcal{L}_i(X)^2 - \sum_{i=p_1+1}^{p_1+p_2} \mathcal{L}_i(X)^2, \quad (4)$$

où les formes linéaires \mathcal{L}_i sont linéairement indépendantes, alors, le problème (P) admet exactement p_1 valeurs propres (distinctes ou non) strictement supérieures à λ et p_2 valeurs propres strictement inférieures à λ .

N. B. : Cet énoncé implique que la valeur λ elle-même est valeur propre de multiplicité $p_0 = n - p_1 - p_2$ dès que p_0 est différent de zéro.

Démonstration : Il suffit de rapprocher le théorème d'inertie de Sylvester [4] et le lemme 1.

Le théorème de Sylvester assure que p_1 et p_2 , définis comme nombres de carrés respectivement précédés du signe plus et du signe moins, sont intrinsèques pour la forme quadratique Φ_λ , c'est-à-dire indépendants de la décomposition. Le lemme 1 établit que l'une de ces décompositions a la propriété annoncée, ce qui prouve la proposition 1.

REMARQUE : La proposition 1 est l'un des énoncés possibles du théorème du Sylvester, et n'est donc pas un résultat nouveau. Mais regarder le résultat fondamental sous cet angle conduit à la conséquence suivante.

Conséquence : Toute méthode de décomposition en carrés de la forme quadratique Φ_λ , induit (par le dénombrement des signes plus et des signes moins) une méthode pour connaître la position de λ par rapport aux valeurs propres du problème (P).

2. LOCALISATION DES VALEURS PROPRES

2.1. La méthode des suites de Sturm, dans le présent contexte

Cette méthode classique consiste à compter le nombre de changements de signe dans la suite des valeurs numériques des mineurs principaux de $A - \lambda B$. Le nombre de changements de signe est égal au nombre de valeurs propres de (P) strictement inférieures à λ .

Toutefois, le nombre de changements de signe n'est pas défini si certains mineurs principaux sont nuls. Il en résulte que la méthode n'est pas applicable,

pour *tout* λ , dans le cas général (bien que dans certains cas, celui des matrices tridiagonales en particulier, on sache néanmoins conclure).

La méthode des suites de Sturm repose sur la propriété d'entrelacement des valeurs propres d'une matrice symétrique réelle A (de dimension $k \times k$) et des valeurs propres d'une matrice symétrique réelle \tilde{A} (de dimension $(k + 1) \times (k + 1)$) liée à A par

$$\tilde{A} = \begin{bmatrix} A & a \\ a^* & \alpha \end{bmatrix}. \tag{5}$$

Dans (5), la notation a désigne un vecteur réel de \mathbb{R}^k , a^* désigne son transposé et α est un nombre réel.

On démontre ci-dessous la propriété d'entrelacement au sens large, pour le problème à deux matrices symétriques dans le cas général c'est-à-dire pour le problème (P). Là encore le résultat n'est pas nouveau [3] mais la démonstration donnée met en lumière le rôle de la décomposition en carrés des formes quadratiques, dans cette propriété classiquement rattachée aux propriétés extrémales.

Soit (\tilde{P}) le problème du type (P) relatif aux matrices

$$\tilde{A} = \begin{bmatrix} A & a \\ a^* & \alpha \end{bmatrix} \quad \text{et} \quad \tilde{B} = \begin{bmatrix} B & b \\ b^* & \beta \end{bmatrix}.$$

Les matrices A et B sont de dimension $k \times k$.

Les valeurs propres respectives des problèmes (P) et (\tilde{P}) sont rangées, respectivement, dans l'ordre croissant,

$$\begin{aligned} \lambda_{i-1} \leq \lambda_i \leq \lambda_{i+1} & \quad (2 \leq i \leq k-1), \\ \tilde{\lambda}_{i-1} \leq \tilde{\lambda}_i \leq \tilde{\lambda}_{i+1} & \quad (2 \leq i \leq k). \end{aligned}$$

Le vecteur cursif de \mathbb{R}^{k+1} est $\tilde{X} = \begin{pmatrix} X \\ x \end{pmatrix}$.

PROPOSITION 2 : Les valeurs propres $\{\lambda_i\}$ de (P) et $\{\tilde{\lambda}_i\}$ de (\tilde{P}) sont entrelacées, au sens large, c'est-à-dire que

$$\tilde{\lambda}_1 \leq \lambda_1 \leq \dots \leq \tilde{\lambda}_i \leq \lambda_i \leq \tilde{\lambda}_{i+1} \leq \dots \leq \lambda_k \leq \tilde{\lambda}_{k+1}. \tag{6}$$

Démonstration : On pose

$$A_\lambda = A - \lambda B; \quad a_\lambda = a - \lambda b; \quad \alpha_\lambda = \alpha - \lambda \beta; \quad A_\lambda = \tilde{A} - \lambda \tilde{B}$$

on appelle en outre \mathcal{L}_{a_λ} la forme linéaire sur \mathbb{R}^k définie par

$$\mathcal{L}_{a_\lambda} : X \rightarrow \langle a_\lambda^*, X \rangle. \quad (7)$$

Avec ces définitions, les formes quadratiques Φ_λ et $\tilde{\Phi}_\lambda$ (notations évidentes) sont liées par

$$\tilde{\Phi}_\lambda(\tilde{X}) = \Phi_\lambda(X) + 2x \mathcal{L}_{a_\lambda}(X) + \alpha_\lambda x^2. \quad (8)$$

On considère une décomposition en carrés de Φ_λ :

$$\Phi_\lambda(X) = \sum_{i=1}^{p_1} \mathcal{L}_i(X)^2 - \sum_{i=p_1+1}^{p_1+p_2} \mathcal{L}_i(X)^2. \quad (9)$$

Les formes linéaires \mathcal{L}_i ($1 \leq i \leq p_1 + p_2$) sont indépendantes.

On veut construire une décomposition en carrés de $\tilde{\Phi}_\lambda$ en partant de (8) et (9).

Deux cas peuvent se présenter :

HYPOTHÈSE H : La forme linéaire \mathcal{L}_{a_λ} est combinaison linéaire des formes \mathcal{L}_i , c'est-à-dire que

$$\mathcal{L}_{a_\lambda} = \sum_{i=1}^{p_1+p_2} \gamma_i \mathcal{L}_i.$$

Premier cas : L'hypothèse H est satisfaite.

On écrit alors :

$$\begin{aligned} \tilde{\Phi}_\lambda(\tilde{X}) = \sum_{i=1}^{p_1} [\mathcal{L}_i(X) + \gamma_i x]^2 - \sum_{i=p_1+1}^{p_1+p_2} [\mathcal{L}_i(X) - \gamma_i x]^2 \\ + \left(\alpha_\lambda - \sum_{i=1}^{p_1} \gamma_i^2 + \sum_{i=p_1+1}^{p_1+p_2} \gamma_i^2 \right) x^2. \end{aligned}$$

Il est immédiat de vérifier que les formes linéaires sur \mathbb{R}^{k+1} définies par

$$\begin{aligned} \tilde{\mathcal{L}}_i(\tilde{X}) &= \mathcal{L}_i(X) + \gamma_i x \quad \text{pour } 1 \leq i \leq p_1, \\ \tilde{\mathcal{L}}_i(\tilde{X}) &= \mathcal{L}_i(X) - \gamma_i x \quad \text{pour } p_1 < i \leq p_1 + p_2, \\ \tilde{\mathcal{L}}_{p_1+p_2+1}(\tilde{X}) &= x. \end{aligned}$$

sont linéairement indépendantes : en effet, si une combinaison linéaire de ces formes est nulle sur \mathbb{R}^{k+1} , elle est nulle sur le sous-espace \mathbb{R}^k des vecteurs \tilde{X} tels que $x=0$. L'indépendance linéaire des \mathcal{L}_i prouve que les coefficients des $\tilde{\mathcal{L}}_i$ pour $i \leq p_1 + p_2$ sont tous nuls.

Alors la combinaison linéaire est proportionnelle à la seule forme $\mathcal{L}_{p_1+p_2+1}$ et ce dernier coefficient est nécessairement nul.

Par conséquent si le nombre

$$\left(\alpha_\lambda - \sum_{i=1}^{p_1} \gamma_i^2 + \sum_{i=p_1+1}^{p_1+p_2} \gamma_i^2 \right),$$

est nul on a

$$\tilde{p}_1 = p_1 \quad \text{et} \quad \tilde{p}_2 = p_2,$$

s'il est positif on a

$$\tilde{p}_1 = p_1 + 1 \quad \text{et} \quad \tilde{p}_2 = p_2,$$

s'il est négatif on a

$$\tilde{p}_1 = p_1 \quad \text{et} \quad \tilde{p}_2 = p_2 + 1.$$

REMARQUE : Lorsque $p_1 + p_2 = k$ l'hypothèse H est nécessairement satisfaite puisque les formes \mathcal{L}_i constituent alors une base du dual de \mathbb{R}^k .

Deuxième cas : L'hypothèse H n'est pas satisfaite.

Dès lors les formes linéaires sur \mathbb{R}^k : \mathcal{L}_i et \mathcal{L}_{a_i} sont indépendantes. En particulier \mathcal{L}_{a_i} n'est pas nulle.

Si, de plus, α_λ est nul on écrira :

$$\tilde{\Phi}_\lambda(\tilde{X}) - \Phi_\lambda(X) = 2x \mathcal{L}_{a_i}(X) = \frac{1}{2} [(\mathcal{L}_{a_i}(X) + x)^2 - (\mathcal{L}_{a_i}(X) - x)^2],$$

sinon, on écrira :

$$\tilde{\Phi}_\lambda(\tilde{X}) - \Phi_\lambda(X) = 2x \mathcal{L}_{a_i}(X) + \alpha_\lambda x^2 = \frac{1}{\alpha_\lambda} [(\mathcal{L}_{a_i}(X) + \alpha_\lambda x)^2 - \mathcal{L}_{a_i}(X)^2].$$

Ainsi, lorsque l'hypothèse H n'est pas satisfaite, on a toujours $\tilde{p}_1 = p_1 + 1$ et $\tilde{p}_2 = p_2 + 1$, car on ajoute toujours deux carrés précédés de signes opposés, à l'expression de $\Phi_\lambda(X)$ pour obtenir $\tilde{\Phi}_\lambda(\tilde{X})$.

Bien entendu, il convient de s'assurer que les formes linéaires sur \mathbb{R}^{k+1} définies par

$$\begin{aligned} \tilde{\mathcal{L}}_i(\tilde{X}) &= \mathcal{L}_i(X); & \tilde{\mathcal{L}}_{p_1+p_2+1}(\tilde{X}) &= \mathcal{L}_{a_i}(X) + x; \\ \tilde{\mathcal{L}}_{p_1+p_2+2}(\tilde{X}) &= \mathcal{L}_{a_i}(X) - x, \end{aligned}$$

son linéairement indépendantes, ainsi que sont indépendantes les formes

linéaires sur \mathbf{R}^{k+1} définies par

$$\begin{aligned}\tilde{\mathcal{L}}_i(\tilde{X}) &= \mathcal{L}_i(X); & \tilde{\mathcal{L}}_{p_1+p_2+1}(\tilde{X}) &= \mathcal{L}_{a_k}(X) + \alpha_\lambda X; \\ \tilde{\mathcal{L}}_{p_1+p_2+2}(\tilde{X}) &= \mathcal{L}_{a_k}(X).\end{aligned}$$

La vérification est analogue à celle du premier cas.

REMARQUE : Dans le second cas $p_1 + p_2$ est inférieur à k . La valeur λ est valeur propre de (P) avec la multiplicité $p_0 = k - (p_1 + p_2)$, elle est (si $p_0 \geq 2$) valeur propre de (\tilde{P}) avec la multiplicité $[(k+1) - (\tilde{p}_1 + \tilde{p}_2)] = p_0 - 1$.

Pourtant il serait faux de croire que la diminution d'une unité entre la multiplicité des valeurs propres de (P) et celle de ces mêmes valeurs propres (lorsqu'elles sont aussi valeurs propres) dans (\tilde{P}) a nécessairement lieu. En effet, dans le premier cas (celui où l'hypothèse H est satisfaite) rien ne s'oppose à ce que $p_0 = k - (p_1 + p_2)$ soit positif. Lorsqu'il en est ainsi on a $\tilde{p}_0 = p_0$.

Une conséquence, commune, à tous les cas possibles est celle-ci :

$$(\tilde{p}_1 - p_1) \in \{0, 1\} \quad \text{et} \quad (\tilde{p}_2 - p_2) \in \{0, 1\}.$$

Or, cette propriété est exactement celle de l'entrelacement dès qu'on sait (cf. proposition 1) que p_1 (resp. p_2) est aussi le nombre de valeurs propres de (P) qui sont strictement supérieures (resp. inférieures) à λ , et de même pour \tilde{p}_1 et \tilde{p}_2 relativement au problème (\tilde{P}) .

COROLLAIRE : Si λ n'est valeur propre d'aucun des problèmes (P) relatifs aux tronquées des matrices A et B au rang k , le nombre de changements de signe dans la suite (de Sturm) des mineurs principaux de $(A - \lambda B)$ est égal au nombre p_2 des valeurs propres du problème (P) du rang n , qui sont strictement inférieures à λ .

Démonstration : Chaque déterminant du mineur d'ordre k est égal au produit

$$\text{dét}(B^{(k)}) \prod_{i=1}^k (\lambda_i^{(k)} - \lambda),$$

où $B^{(k)}$ désigne la tronquée de B au rang k et $\{\lambda_i^{(k)}\}$ l'ensemble des valeurs propres du problème (P) au rang k .

En effet

$$\text{dét}(A^{(k)} - \lambda B^{(k)}) = \text{dét} B^{(k)} \cdot \text{dét}(B^{(k)^{-1}} A^{(k)} - \lambda I).$$

Sous l'hypothèse du corollaire, tous les mineurs principaux sont donc différents de zéro et leur signe est celui de

$$(-1)^{p_2^{(k)}}$$

où $p_2^{(k)}$ est le nombre de valeurs propres du k -ième problème (P), qui sont strictement inférieures à λ .

La connaissance du signe du mineur de rang $k + 1$ permet ainsi de connaître la parité de $p_2^{(k+1)} = \tilde{p}_2^{(k)}$. D'où, connaissant $p_2^{(k)}$, la valeur exacte de $p_2^{(k+1)}$ est déterminée, puisque $(\tilde{p}_2^{(k)} - p_2^{(k)}) \in \{0, 1\}$ d'après l'entrelacement.

Le dénombrement des changements de signe, dans la suite des mineurs principaux n'est donc que le calcul de la suite des valeurs de $p_2^{(k)}$ pour les problèmes (P) des rangs successifs ($1 \leq k \leq n$).

2.2. Proposition d'une méthode fondée sur la factorisation LDL.

Dans [7], G. Peters et J. H. Wilkinson font appel à la méthode des suites de Sturm pour la phase de localisation. L'article traite d'une méthode d'interpolation entre des valeurs λ pour lesquelles la suite de Sturm a été calculée, afin de limiter autant que possible le nombre de calculs de telles suites. Car il demeure que le calcul d'une suite de Sturm, principalement au voisinage des valeurs propres des problèmes (P^(k)) est exposé aux instabilités dues aux pivots « petits ». La méthode de factorisation LDL* (Martin, Peters et Wilkinson [6]) est même impossible en présence d'un pivot nul, et elle est écartée par les auteurs de l'article [7].

Il est bien clair que, en ayant pour but de localiser les valeurs propres, les valeurs λ pour lesquelles il convient de calculer les suites de Sturm se situent entre les valeurs propres, voire même en les valeurs propres du problème (P) et que le résultat d'entrelacement des racines laisse prévoir une fréquence importante de pivots « petits ».

Notre proposition ici, est de factoriser, dans ces cas, une autre matrice que la matrice $(A - \lambda B)$ mais pour laquelle p_1 et p_2 [c'est-à-dire $p_1^{(n)}$ et $p_2^{(n)}$] sont les mêmes que pour $(A - \lambda B)$. La seule perte est que l'on n'obtient pas toujours chacun des $p_1^{(k)}$ et $p_2^{(k)}$ relatifs à la matrice $(A - \lambda B)$, et seulement le couple $p_1^{(n)}$, $p_2^{(n)}$ correct. Mais ce résultat suffit pour permettre la localisation.

THÉORÈME : *Quelle que soit la matrice symétrique réelle A donnée, il existe une matrice \hat{A} , symétrique réelle, semblable à A par une isométrie, et qui admet une factorisation LDL*.*

Démonstration : La méthode de démonstration est constructive en ce sens qu'elle suit la démarche de l'algorithme LDL*. Elle construit une matrice unitaire Q, une matrice diagonale D et une matrice L, triangulaire inférieure, ayant des coefficients diagonaux égaux à 1, telles que la matrice symétrique

$$\hat{A} = QAQ^*,$$

admettre la factorisation

$$\hat{A} = LDL^*.$$

On rappelle — pour introduire les notations appropriées, en quoi consiste l'étape k de la factorisation LDL^* .

NOTATIONS POUR LA FACTORISATION ORDINAIRE : En commençant l'étape k , on dispose d'une factorisation

$$A = \begin{bmatrix} L^{(k)} & 0 \\ G^{(k)} & I \end{bmatrix} \times \begin{bmatrix} D^{(k)} & 0 \\ 0 & B^{(k)} \end{bmatrix} \times \begin{bmatrix} L^{(k)*} & G^{(k)*} \\ 0 & I \end{bmatrix} \quad (10)$$

où

(i) $L^{(k)}$ est un bloc de dimension $k \times k$, triangulaire inférieur, et dont la diagonale n'a que des coefficients égaux à 1;

(ii) $D^{(k)}$ est un bloc de dimension $k \times k$, qui est diagonal;

(iii) $G^{(k)}$ est un bloc de dimension $(n-k) \times k$;

(iv) $B^{(k)}$ est un bloc de dimension $(n-k) \times (n-k)$.

La notation $*$ désigne toujours la transposition.

D'après l'identité (10), l'identification par blocs conduit à ceci : en posant

$$A = \begin{bmatrix} A^{(k)} & M^{(k)*} \\ M^{(k)} & R^{(k)} \end{bmatrix},$$

on obtient :

$$\left. \begin{aligned} A^{(k)} &= L^{(k)} D^{(k)} L^{(k)*}, \\ M^{(k)} &= G^{(k)} D^{(k)} L^{(k)*}, \\ R^{(k)} &= B^{(k)} + G^{(k)} D^{(k)} G^{(k)*} \end{aligned} \right\} \quad (11)$$

Dans (11), les deux premières identités ont servi au calcul de $L^{(k)}$, $D^{(k)}$ et $G^{(k)}$, et la dernière *définit* $B^{(k)}$ par

$$B^{(k)} = R^{(k)} - G^{(k)} D^{(k)} G^{(k)*}. \quad (12)$$

Toutefois, en pratique, la matrice $B^{(k)}$ n'est pas calculée dans sa totalité, car cela est nullement nécessaire.

Lorsque $k=0$ on a $B^{(0)} = A$. Les matrices $L^{(0)}$, $D^{(0)}$ et $G^{(0)}$ qui ont au moins une dimension nulle, sont inexistantes (ensemble vide).

Le premier coefficient de la matrice $B^{(k)}$, dans le cas ordinaire, est non nul. On le pose égal à d_{k+1} et l'identité (10) fournit, sans calcul, la première des identités (11) relatives au rang $k+1$.

Vient ensuite la construction de $G^{(k+1)}$. La première ligne de $G^{(k)}$ est entrée dans la composition de $L^{(k+1)}$. Il reste alors un bloc de dimension $(n-k-1) \times k$

qui est commun à $G^{(k)}$ et $G^{(k+1)}$. On obtient G^{k+1} en bordant, à droite, ce bloc par un vecteur $l^{(k+1)}$ construit comme suit :

La première colonne de $B^{(k)}$ est un vecteur de dimension $(n-k)$ dont le premier coefficient est égal à d_{k+1} . Le quotient de ce vecteur par d_{k+1} est un vecteur, dont la première coordonnée est égale à 1, et dont les $(n-k-1)$ coordonnées suivantes composent le vecteur $l^{(k+1)}$ cherché.

Cela termine l'étape k , dans le cas ordinaire.

MODIFICATION DE L'ALGORITHME LDL^* : Supposons qu'on ait pu factoriser sous la forme (10) une matrice \hat{A}_k de (dimension $n \times n$) associée à A par

$$\hat{A}_k = Q_k A Q_k^*,$$

où Q_k est une matrice unitaire. On a

$$A_k = \begin{bmatrix} L^{(k)} & 0 \\ G^{(k)} & I \end{bmatrix} \times \begin{bmatrix} D^{(k)} & 0 \\ 0 & B^{(k)} \end{bmatrix} \times \begin{bmatrix} L^{(k)*} & G^{(k)} \\ 0 & I \end{bmatrix}. \quad (13)$$

Et $B^{(k)}$ est lié au bloc $\hat{R}_k^{(k)}$ [« en bas à droite » dans \hat{A}_k et de dimension $(n-k) \times (n-k)$] par l'identité

$$B^{(k)} = \hat{R}_k^{(k)} - G^{(k)} D^{(k)} G^{(k)*}. \quad (14)$$

Si le premier coefficient de $B^{(k)}$ est non nul, on effectue l'étape k , comme il est décrit ci-dessus dans le cas ordinaire et on obtient une factorisation au rang $k+1$ de \hat{A}_k qui prendra donc le nom de \hat{A}_{k+1} . On aura

$$\hat{A}_{k+1} = \hat{A}_k \quad \text{et} \quad Q_{k+1} = Q_k.$$

Étudions la situation dans laquelle le premier coefficient de $B^{(k)}$, est nul. Le bloc $B^{(k)}$ a été défini par (14) conformément à l'identité (13) d'où part l'étape k .

Un premier cas — très particulier — est celui où le bloc $B^{(k)}$ est entièrement nul. Dans ce cas, l'identité (13) est la factorisation cherchée. En effet il suffit de poser $\hat{A} = \hat{A}_k$ et

$$L = \begin{bmatrix} L^{(k)} & 0 \\ G^{(k)} & I \end{bmatrix}; \quad D = \begin{bmatrix} D^{(k)} & 0 \\ 0 & 0 \end{bmatrix},$$

pour avoir le résultat cherché.

La matrice D ainsi construite possède la propriété suivante : Sa diagonale a $(p_1 + p_2)$ termes non nuls et $p_0 = (n - p_1 + p_2)$ termes nuls. Les p_0 termes nuls sont groupés à la fin tandis que les p_1 termes positifs et p_2 termes négatifs apparaissent (dans un ordre non privilégié) aux $(p_1 + p_2)$ premières positions.

Hors le cas où $B^{(k)}$ est identiquement nul, il existe une matrice $Q^{(k)}$, unitaire, telle que

$$\tilde{B}^{(k)} = Q^{(k)} B^{(k)} Q^{(k)*},$$

ait son premier terme différent de zéro.

L'existence d'une telle matrice $Q^{(k)}$ sera démontrée dans le lemme 2. En l'admettant, actuellement, on pose

$$Q_{k+1} = \begin{bmatrix} I & 0 \\ 0 & Q^{(k)} \end{bmatrix} Q_k,$$

et il vient

$$\hat{A}_{k+1} = \begin{bmatrix} I & 0 \\ 0 & Q^{(k)} \end{bmatrix}, \quad \hat{A}_k \begin{bmatrix} I & 0 \\ 0 & Q^{(k)} \end{bmatrix}.$$

On tire alors de (13) :

$$\hat{A}_{k+1} = \begin{bmatrix} L^{(k)} & 0 \\ \tilde{G}^{(k)} & I \end{bmatrix} \times \begin{bmatrix} D^{(k)} & 0 \\ 0 & \tilde{B}^{(k)} \end{bmatrix} \times \begin{bmatrix} L^{(k)*} & \tilde{G}^{(k)*} \\ 0 & I \end{bmatrix} \quad (15)$$

et de (14) :

$$\tilde{B}^{(k)} = \hat{R}_{k+1}^{(k)} - \tilde{G}^{(k)} D^{(k)} \tilde{G}^{(k)*}, \quad (16)$$

en ayant posé $\tilde{G}^{(k)} = Q^{(k)} G^{(k)}$.

La notation $\hat{R}_{k+1}^{(k)}$ désigne le bloc « en bas à droite » dans \hat{A}_{k+1} et dont la dimension est $(n-k) \times (n-k)$.

Pour la matrice \hat{A}_{k+1} obtenue, l'étape $k+1$ de la factorisation LDL^* est possible.

Procédant ainsi, itérativement, l'algorithme se termine toujours par une factorisation (13) à un rang k_0 tel que :

ou bien $B^{(k_0)} = 0$ (cas particulier rencontré plus haut), ou bien $k_0 = n$.

La démonstration du théorème est donc terminée (sous réserve de la démonstration du lemme 2) car, dans chacun des cas, on a une factorisation

$$\hat{A} = LDL^*,$$

d'une matrice \hat{A} de la forme

$$\hat{A} = QAQ^*.$$

De plus, dans D , les p_0 termes nuls de la diagonale ($p_0 \geq 0$) sont groupés à la fin.

LEMME 2 : Si B est une matrice symétrique réelle, non nulle, dont le premier coefficient est égal à zéro, il existe une matrice unitaire Q telle que

$$\tilde{B} = QBQ^*$$

ait son premier coefficient différent de zéro.

La démonstration consiste en la construction d'une matrice Q ayant la propriété demandée.

(a) S'il y a, dans B , un terme diagonal non nul, soit i son rang (s'il y en a plusieurs l'un quelconque d'entre eux convient). La matrice de permutation Q qui échange les coordonnées de rang i et de rang 1, en laissant invariantes toutes les autres coordonnées, répond à la question.

(b) Si la diagonale de B est composée de termes tous nuls, l'hypothèse B n'est pas nulle assure l'existence d'un terme b_{ij} (au moins) différent de zéro. On prendra alors pour Q la matrice $Q = Q_2 \cdot Q_1$ ainsi construite : Q_1 est la rotation d'angle $(\pi/4)$ dans le plan des coordonnées i et j , les autres coordonnées restant inchangées, et Q_2 est la permutation entre les coordonnées de rang i et de rang 1.

Cela termine la démonstration du lemme 2.

THÉORÈME 2 : Deux matrices réelles symétriques A et B étant données, la seconde étant définie positive, on pose :

$$A_\lambda = A - \lambda B$$

la matrice réelle symétrique construite à l'aide d'une valeur réelle λ donnée.

L'algorithme de factorisation LDL^* modifié (cf. th. 1) fournit, quel que soit λ , la factorisation

$$\hat{A}_\lambda = LDL^*$$

d'une matrice semblable à A_λ par une isométrie

$$\hat{A}_\lambda = QA_\lambda Q^*.$$

Le nombre des termes positifs (resp. négatifs) de la diagonale D est égal au nombre des valeurs propres supérieures (resp. inférieures) à λ dans le problème (P) (chacune étant comptée avec sa multiplicité). Le nombre de termes égaux à zéro, dans la diagonale de D est la multiplicité de la valeur propre λ . Enfin le produit des termes diagonaux de D est égal au déterminant de $(A - \lambda B)$.

Démonstration : On s'appuie sur la proposition 1 et on considère la forme quadratique Φ_λ définie par

$$\Phi_\lambda(X) = \langle A_\lambda X, X \rangle.$$

L'algorithme LDL^* modifié donne la factorisation

$$\hat{A}_\lambda = LDL^*$$

de la matrice

$$\hat{A}_\lambda = QA_\lambda Q^*,$$

Évidemment Q , L et D dépendent aussi de λ .

Alors :

$$\Phi_\lambda(X) = \langle Q^* \hat{A}_\lambda QX, X \rangle = \langle DL^* QX, L^* QX \rangle.$$

Ainsi, par le changement de base, de matrice L^*Q , la forme quadratique Φ_λ admet une décomposition en carrés, dont les coefficients sont les termes diagonaux de D . Cela démontre le résultat relatif aux nombres de coefficients respectivement positifs, négatifs et nuls.

Enfin, la matrice Q étant unitaire, et le déterminant de L étant égal à 1 on a

$$\det A_\lambda = \det(\hat{A}_\lambda) = \det(D)$$

ce qui termine la démonstration du théorème 2.

2.3. Application au calcul des vecteurs propres

Soit λ une valeur propre du problème (P). Soit p_0 sa multiplicité.

La factorisation LDL^* modifiée fournit un triplet Q, L, D où Q est unitaire, où L est triangulaire inférieure avec les coefficients diagonaux égaux à 1 et où D est diagonale avec ses $n - p_0$ premiers coefficients non nuls et les p_0 derniers nuls.

La matrice $A_\lambda = A - \lambda B$ est lié à Q, L, D par l'existence d'une matrice \hat{A}_λ telle que

$$\hat{A}_\lambda = QA_\lambda Q^*$$

et

$$\hat{A}_\lambda = LDL^*.$$

PROPOSITION 3 : *Le sous-espace propre relatif à la valeur propre λ est l'image du sous-espace engendré par les p_0 derniers vecteurs de la base canonique de \mathbf{R}^n , par l'inverse de L^*Q .*

Démonstration : Il s'agit de caractériser les vecteurs X tels que :

$$A_\lambda X = 0.$$

Or L, L^*, Q, Q^* sont inversibles.

La caractérisation s'écrit encore :

$$Q^* LDL^* QX = 0$$

ce qui équivaut à

$$DL^* QX = 0.$$

La proposition 3 est démontrée.

3. QUELQUES REMARQUES EN VUE DE L'IMPLEMENTATION

3.1. Encombrement de l'enregistrement de L et D

Pour les matrices bandes, et plus généralement pour les matrices profils (ou « Skyline ») la factorisation LDL^* — lorsqu'elle est possible — respecte la structure bande ou profil. Hélas il n'en est plus nécessairement de même avec la modification introduite dans l'algorithme LDL^* afin de rendre sa portée générale.

Ici intervient le choix particulier des matrices Q utilisées (cf. lemme 2) qui limite à la fois l'inconvénient du point de vue de l'encombrement (deux lignes seulement peuvent être altérées à l'étape k) et du temps nécessaire au calcul.

Une remarque du même type s'impose pour les matrices provenant de problèmes d'éléments finis plans lorsqu'on a recours à une numérotation optimale au sens de A. George [2] ou de R. J. Lipton et R. E. Tarjan [5]. Là encore le caractère optimal de la numérotation initiale peut être perdu par le changement de base Q_k . Mais l'optimal initial se heurte, dans le cas des pivots nuls, à une impossibilité de factorisation qui fait que l'optimum attendu repose sur une hypothèse non réalisée.

Les calculs relatifs à la modification de l'algorithme sont uniquement ceux de $\tilde{G}^{(k)} = Q^{(k)} G^{(k)}$ et de $\hat{R}_{k+1}^{(k)} = Q^{(k)} \hat{R}_k Q^{(k)*}$ à chaque étape k , relative à un pivot rencontré égal à zéro.

3.2. Encombrement de l'enregistrement de Q

Lorsque le but du calcul de factorisation LDL^* est seulement l'application du théorème 2, il est inutile de conserver la mémoire des opérateurs $Q^{(k)}$ qui définissent, de proche en proche,

$$Q = Q_n \equiv \prod_{k=0}^{n-1} \begin{pmatrix} I & 0 \\ 0 & Q^{(k)} \end{pmatrix},$$

où le produit est effectué à gauche, pour k variant de 0 à $n-1$.

La seule utilité de conserver la connaissance de Q est le calcul d'un sous-espace propre (cf. 2.3). Dans ce cas les systèmes à résoudre sont de la forme

$$L^* Y_i = e_i,$$

$$Q X_i = Y_i,$$

où e_i désigne le i -ième vecteur de la base canonique avec i appartient à l'ensemble des p_0 derniers indices.

Le système $L^* Y_i = e_i$ est résolu simplement avec la matrice triangulaire L^* , de diagonale égale à I .

Pour calculer ensuite X_i , il suffit de faire agir sur Y_i l'opérateur $\begin{bmatrix} I & 0 \\ 0 & Q^{(n)*} \end{bmatrix}$ puis sur le résultat $\begin{bmatrix} I & 0 \\ 0 & Q^{(n-1)*} \end{bmatrix}$, etc. pour réaliser l'opérateur inverse de Q , c'est-à-dire $Q^{-1} = Q^*$.

Et pour cela, il suffit d'avoir gardé mémoire, pour chaque k , de l'opérateur $Q^{(k)}$ utilisé (permutation de deux axes, éventuellement précédée d'une rotation d'angle $\pi/4$ dans le plan de deux axes) ce qui peut être fait sous une forme très peu encombrante.

3.3. L'instabilité de la méthode LDL^* .

Dans [7] les auteurs écartent la méthode LDL^* pour des raisons d'instabilité. Ce qui précède ne répond qu'aux cas d'impossibilité de factorisation (pivots nuls). En fait, l'instabilité éventuelle ne peut provenir que de la présence de pivots « petits ». Aussi, en effectuant le même type de modification de l'algorithme LDL^* , en présence de pivots « petits » qu'en présence de pivots nuls, la modification proposée doit permettre de répondre à l'obstacle d'instabilité.

Il y a lieu d'élargir le test de nullité du pivot pour le remplacer par un test de petitesse. Des expérimentations sont en cours, dans le cadre d'une implémentation de la méthode. On ne peut donner prématurément de conclusions pratiques, car l'exploration de coefficients dans $B^{(k)}$ (donc aussi leur calcul) est en opposition avec l'avantage essentiel de la méthode LDL^* ordinaire où (14) ne fait que l'objet que du calcul d'une colonne de $B^{(k)}$. Il s'agit donc, comme c'est souvent le cas en matière d'implémentation, de choix judicieux du test de branchement et de la procédure précise d'exploration de termes, autres que la première colonne, dans $B^{(k)}$. Le compromis à trouver concerne aussi l'effet de la remarque faite en 3.1.

3 4. Calcul des sous-espaces propres en présence de valeurs propres voisines

Lorsque plusieurs valeurs propres sont voisines, c'est-à-dire que les arrondis rendent impossible de déterminer si une valeur λ est valeur propre de multiplicité p_0 ou si p_0 valeurs propres (distinctes ou non) sont proches de la valeur λ , la méthode LDL^* modifiée [avec un test convenable de « nullité à ε près » sur le premier coefficient de $B^{(k)}$] fournira, de manière stable, le sous-espace de dimension p_0 engendré par l'ensemble des vecteurs propres relatifs aux valeurs propres « voisines » de λ .

Cette remarque renforce l'intérêt de l'étude des tests à retenir pour réaliser une implémentation efficace

BIBLIOGRAPHIE

- 1 R COURANT et D HILBERT, *Methods of Mathematical Physics*, vol I, chap VI, Interscience, 1953
- 2 A GEORGE, *Nested Dissection of a Regular Finite Element Mesh*, S I A M J Numer Anal vol 10 n°2, 1973, p 345-363
- 3 I GLAZMAN et Y LIUBITCH, *Analyse lineaire dans les espaces de dimensions finies*, Mir, Moscou 1972, p 161-162
- 4 A LENTIN et J RIVAUD, *Leçons d'algebre moderne* Vuibert, Paris, 1961, p 382-383
- 5 R J LIPTON et R E TARIAN, *A Separator Theorem for Planar Graphs* Conference on Theoretical Computer Science, University of Waterloo, 15-17 août 1977
- 6 R S MARTIN, G PETERS et J H WILKINSON *Symmetric Decomposition of a Positive Definite Matrix*, Num Math, vol 7, 1965 p 362-383
- 7 G PETERS et J H WILKINSON, *Eigenvalues of $Ax = \lambda Bx$ with Band Symmetric A and B*, Comp J, vol 12, 1969 p 398-404