Revue française d'automatique informatique recherche opérationnelle. Mathématique

F. Chatelin

Brève communication. Sur la convergence de l'algorithme *QR* pour une matrice hermitienne

Revue française d'automatique informatique recherche opérationnelle. Mathématique, tome 7, nº R1 (1973), p. 57-61

<http://www.numdam.org/item?id=M2AN_1973__7_1_57_0>

R.A.I.R.O.
(7e année, avril 1973, R-1, p. 57-61)

SUR LA CONVERGENCE DE L'ALGORITHME *QR* POUR UNE MATRICE HERMITIENNE

par F. CHATELIN (1)

Résumé. — *On applique à une matrice tridiagonale hermitienne l'algorithme QR avec les translations d'origine définies ci-dessous. La convergence de l'algorithme avec la première translation est cubique, on montre qu'avec les deux autres translations la convergence est* au moins *cubique.*

1. INTRODUCTION

Soit A une matrice hermitienne tridiagonale. On suppose qu'aucun des éléments sous-diagonaux n'est nul, sinon A pourrait se décomposer en somme directe de deux (ou plusieurs) matrices tridiagonales. Alors toutes les valeurs propres de A sont simples, on les note $|\lambda_1| \geqslant |\lambda_2| \geqslant \dots \geqslant |\lambda_n|$. S'il existe des valeurs propres de même module, ce sont des paires de valeurs propres opposées.

L'algorithme QR avec translation d'origine k_s [2] peut s'écrire :

$$A_1 = A$$
$$\begin{cases} Q_s(A_s - k_s I) = R_s \\ R_s Q_s^H + k_s I = A_{s+1} \end{cases} \qquad s = 1, 2, \dots$$

où les Q_s sont unitaires, et les R_s triangulaires supérieures.

Les éléments suivants de A_s sont notés :

$$a_{nn}^{(s)} = a_n^{(s)} \quad , \quad a_{n-1\,n-1}^{(s)} = a_{n-1}^{(s)}$$
$$a_{nn-1}^{(s)} = \varepsilon_s \qquad a_{n-1\,n-2}^{(s)} = \eta_s.$$

(1) Institut de Recherches en Mathématiques Avancées, Grenoble.

Les deux translations d'origine de [2] sont :

(*a*) $k_s = a_n^{(s)}$

(*b*) $k_s = \alpha_s$, valeur propre de $C_s = \begin{pmatrix} a_{n-1}^{(s)} & \bar{\varepsilon}_s \\ \varepsilon_s & a_n^{(s)} \end{pmatrix}$

la plus proche de $a_n^{(s)}$.

J. H. Wilkinson a montré [3] qu'avec l'une ou l'autre translation, QR est globalement convergent (i.e. l'algorithme converge avec la translation effectuée dès la première itération). Il a aussi montré [2] qu'avec la translation (*a*) la convergence asymptotique est cubique.

Nous étudions la convergence asymptotique de l'algorithme avec la translation (*b*).

2. CONVERGENCE ASYMPTOTIQUE DE QR AVEC LA TRANSLATION (*b*)

Le terme $a_{nn-1}^{(s)}$ tend vers 0 comme $\left(\frac{\lambda_n - k_s}{\lambda_{n-1} - k_s}\right)^s$. Lorsque $s \to \infty$, $k_s \to \lambda_n$, mais $\lambda_{n-1} - k_s$ reste $\neq 0$ puisque les valeurs propres sont simples. Donc, ce terme valant ε_s à la $s^{\text{ème}}$ itération, est de l'ordre $\varepsilon_s \frac{\lambda_n - k_s}{\lambda_{n-1} - k_s}$ à la $(s+1)^e$.

Il suffit donc d'étudier $\lambda_n - \alpha_s$ pour obtenir l'ordre de convergence avec la translation (*b*).

On omet l'indice s dans ce qui suit.

On pose $d = a_n - a_{n-1}$, $r = 2\frac{|\varepsilon|}{d}$, si $d \neq 0$.

Les deux valeurs propres α, β de C s'écrivent :

$$\alpha = a_n + \frac{d}{2}(\sqrt{1+r^2} - 1)$$

$$\beta = a_{n-1} - \frac{d}{2}(\sqrt{1+r^2} - 1)$$

Si $d = 0$, $\alpha = a_n + |\varepsilon|$. On suppose d $\neq 0$ (remarque 1).

C est diagonalisable dans la base orthonormée représentée par la matrice :

$$P = \frac{1}{\nu}\begin{bmatrix} 1 & \frac{\bar{\varepsilon}}{|\varepsilon|}\frac{1}{r}(\sqrt{1+r^2} - 1) \\ -\frac{\bar{\varepsilon}}{|\varepsilon|}\frac{1}{r}(\sqrt{1+r^2} - 1) & 1 \end{bmatrix}$$

où $$\nu^2 = 1 + \frac{1}{r^2}(\sqrt{1+r^2} - 1)^2.$$

La matrice :

$$A = \left(\begin{array}{cc|cc} B & & O & \\ & & \bar{\eta} & 0 \\ \hline & \eta & & \\ O & 0 & & C \end{array}\right) \begin{array}{l} \updownarrow\ n-2 \\ \\ \updownarrow\ 2 \end{array}$$

est semblable, par transmutation par $X = \left(\begin{array}{c|c} I & O \\ \hline O & P \end{array}\right)$, à :

$$A' = \frac{1}{\nu^2} \left(\begin{array}{cc|cc} B & & O & \\ & & \bar{\eta} & \bar{\delta} \\ \hline O & \eta & \beta & 0 \\ & \delta & 0 & \alpha \end{array}\right)$$

où
$$\delta = \frac{\varepsilon}{|\varepsilon|} \frac{1}{r} (\sqrt{1 + r^2} - 1)\eta.$$

Puisque l'algorithme converge, la norme de la partie hors-diagonale de A, donc de A', tend vers 0. Dès que cette norme est assez petite, on peut majorer $|\lambda_n - \alpha|$ à l'aide de la théorie des perturbations [1] :

$$|\lambda_n - \alpha| \leqslant \frac{1}{\nu^4} \frac{1}{r^2} (\sqrt{1 + r^2} - 1)^2 \frac{2}{|\beta - \alpha|} |\eta|^2.$$

Soit :

$$|\lambda_n - \alpha| \leqslant C\, |\eta|^2 \frac{|\varepsilon|^2}{|d|^3}\,; C \text{ constante.}$$

L'ordre de convergence de l'algorithme QR avec la translation (b) est donc :

$$\frac{|\varepsilon|^3}{d^4} |\eta|^2$$

La convergence est au moins cubique.

Avec la translation (*a*), l'ordre de convergence est $\dfrac{|\varepsilon|^3}{d^2}$.

Le gain obtenu avec la translation (*b*) est donc de l'ordre de $\dfrac{1}{d^2} |\eta|^2$. $|\eta|$ est en général bien supérieur à $|\varepsilon|$, mais tend vers 0.

3. UNE TRANSLATION EQUIVALENTE

On peut penser à approcher λ_n, non par α_s mais par le développement limité au deuxième ordre (par rapport à la partie hors-diagonale de A) de λ_n au voisinage de $a_n^{(s)}$.

On obtient la translation :

$$(c) \qquad \begin{cases} k_s = a_n^{(s)} + \dfrac{|\varepsilon_s|^2}{d_s}, & \text{si} \quad d_s = a_n^{(s)} - a_{n-1}^{(s)} \neq 0 \\ \text{sinon } a_n^{(s)} + |\varepsilon_s| \end{cases}$$

La formule (c) ne contient pas de calcul de $\sqrt{\ }$.

Pour comparer à (b), il est facile de voir, à l'aide de [1], que :

$$|\alpha_s - a_n^{(s)}| \leqslant \frac{|\varepsilon_s|^2}{|d_s|}, \quad |\lambda_n - a_n^{(s)}| \leqslant 2 \frac{|\varepsilon_s|^2}{|d_s|}.$$

Les ordres de convergence obtenus sont les mêmes car la différence entre les approximations (b) et (c) de λ_n, est de l'ordre de $\dfrac{|\varepsilon|^4}{d^3}$.

REMARQUES

1. Le cas $d_s = 0$ est, à partir d'un certain rang, un « accident » puisque les valeurs propres sont simples.

2. Quoique les algorithmes soient globalement convergents, on peut dans bien des cas augmenter sensiblement le nombre d'itérations en démarrant la translation dès la première itération.

EXEMPLE :

Pour une matrice hermitienne d'ordre 10, de valeurs propres

$$\lambda_i = 10 - i + 1 \quad , \quad i = 1, \ldots, 10,$$

les valeurs propres ont été calculées à une précision inférieure ou égale à 10^{-5}, avec un nombre d'itérations égal à :

— 23, avec la translation (a), appliquée dès la première itération,

— 18, avec la translation (a), effectuée à partir du moment où le terme $a_n^{(s)}$ commence à se stabiliser : $\left| 1 - \dfrac{a_n^{(s)}}{a_n^{(s-1)}} \right| < 0{,}1$,

— 15, avec les translations (b) ou (c), effectuées à partir du moment où $a_n^{(s)}$ commence à se stabiliser.

REFERENCES

[1] F. Chatelin-Laborde, *Perturbation d'une matrice hermitienne ou normale*, Num Mat 17 (1971), 318-337.

[2] J. H. Wilkinson, *The algebraic eigenvalue problem*, Oxford : Clarendon Press (1965).

[3] J. H. Wilkinson, *Global convergence of QR algorithm*, Edinburgh, IFIP 68.