

C. LEBAUD

Remarques sur la convergence de la méthode Q.R.

Revue française d'informatique et de recherche opérationnelle, série rouge, tome 2, n^o 3 (1968), p. 103-112.

http://www.numdam.org/item?id=M2AN_1968__2_3_103_0

© AFCET, 1968, tous droits réservés.

L'accès aux archives de la revue « Revue française d'informatique et de recherche opérationnelle, série rouge » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). 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/>

REMARQUES SUR LA CONVERGENCE DE LA METHODE Q.R.

par C. LEBAUD (1)

Résumé. — *L'auteur démontre la convergence de la méthode Q.R. pour une certaine classe de matrices. Il en déduit une application aux matrices de type Hessenberg supérieur et le comportement asymptotique de l'erreur.*

On démontre dans cette note des théorèmes de convergence de la méthode QR qui contiennent des résultats de B. Parlett [3]. Les démonstrations en sont plus simples et adaptées directement de J. H. Wilkinson [5].

1. — DEFINITION DE L'ALGORITHME QR [1]

On rappelle que toute matrice carrée A d'ordre n à éléments complexes peut être exprimée comme le produit, QR , d'une matrice unitaire Q et d'une matrice triangulaire supérieure R . Si A est inversible, cette factorisation est unique, si l'on convient de choisir les éléments diagonaux de R réels et positifs.

L'algorithme consiste, en posant $A = A_1$, à engendrer une suite A_s de matrices unitairement semblables à la matrice A_1 . A la s -ième étape, on factorise la matrice A_s :

$$A_s = Q_s R_s \quad (1,1)$$

Si Q^* désigne la matrice adjointe de Q , on a alors :

$$Q_s^* A_s Q_s = R_s Q_s \quad (1,2)$$

On définit

$$A_{s+1} = R_s Q_s \quad (1,3)$$

(1) Faculté des Sciences de Rennes.

Les itérés successifs satisfont aux équations :

$$A_{s+1} = Q_s^* A_s Q_s = Q_s^* Q_{s-1}^* Q_{s-2}^* \dots Q_1^* A_1 Q_1 Q_2 \dots Q_s \quad (1,4)$$

En posant

$$P_s = Q_1 Q_2 \dots Q_s \quad (1,5)$$

$$U_s = R_s R_{s-1} \dots R_1 \quad (1,6)$$

On a :

$$A_1^s = P_s U_s \quad (1,7)$$

$$A_{s+1} = P_s^* A_1 P_s \quad (1,8)$$

D'où le

Théorème 1.1. — *Si A est une matrice carrée d'ordre n à éléments complexes, alors la matrice A_{s+1} construite par l'algorithme QR satisfait à la relation :*

$$A_{s+1} = P_s^* A P_s \quad (1,9)$$

où P_s est une matrice d'ordre n , unitaire telle que la matrice

$$P_s^* A^s = U_s \quad (1,7')$$

soit triangulaire supérieure.

De plus, si la matrice A est inversible et si toutes les matrices R_s sont choisies telles que leurs éléments diagonaux soient des nombres réels positifs, alors la matrice U_s a tous ses éléments diagonaux réels positifs.

2. — DEFINITION DE LA CONVERGENCE

Nous avons pour but la recherche des valeurs propres d'une matrice A . Il serait donc souhaitable que la matrice A_s , semblable à la matrice A , prenne une forme triangulaire supérieure; ceci n'est malheureusement pas toujours réalisables. Nous sommes amenés à donner la définition suivante de la convergence.

On ne considérera dans la suite que des partitionnements de matrices carrées d'ordre n tels que les blocs diagonaux soient des matrices carrées.

Définition 2.1. — *On dit que l'algorithme QR appliqué à une matrice A carrée d'ordre n , à éléments complexes, converge pour un partitionnement donné s'il existe des matrices B , B_s , V_s telles que :*

$$A_s = V_s^* B_s V_s$$

et satisfaisant aux propriétés :

i) par rapport à ce partitionnement, la matrice B est bloc triangulaire supérieure, les matrices V_s sont bloc diagonales et unitaires,

ii) les matrices B_s convergent au sens habituel vers la matrice B .

3. — ETUDE DE LA CONVERGENCE

Soit A une matrice carrée inversible d'ordre n à éléments complexes, soit J une matrice de Jordan inférieure associée à la matrice A :

$$J = \begin{pmatrix} J_1 & & & & \\ & J_2 & & \bigcirc & \\ & & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & & \\ & & & & \\ & & & & \\ & & & & J_m \end{pmatrix} \quad (3,1)$$

Les matrices J_i sont carrées d'ordre n_i et de l'une des deux formes suivantes :

$$J_i = [\lambda_i] \quad \text{si} \quad n_i = 1$$

$$J_i = \begin{pmatrix} \lambda_i & & & & \\ 1 & & & \bigcirc & \\ & \lambda_i & & & \\ & 1 & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & \cdot & \\ & & & & \\ & & & & 1 \lambda_i \end{pmatrix} \quad \text{si } n_i \neq 1 \quad (3,2)$$

La notation $\rho(B)$ désignant le rayon spectral d'une matrice B , supposons que les matrices J_i soient ordonnées de façon que les deux conditions suivantes soient réalisées :

- 1) $i \leq j \Rightarrow \rho(J_i) \geq \rho(J_j)$
- 2) si $\rho(J_{i-1}) > \rho(J_i) = \dots = \rho(J_{j-1}) > \rho(J_j)$

alors on a $n_i \geq n_{i+1} \geq \dots \geq n_{j-1}$.

En regroupant dans la matrice J , les matrices J_i de même rayon spectral, on peut écrire :

$$J = \begin{pmatrix} \Lambda_1 & & & & \\ & \Lambda_2 & & \bigcirc & \\ & & \cdot & & \\ & & & \cdot & \\ \bigcirc & & & & \\ & & & & \\ & & & & \\ & & & & \Lambda_k \end{pmatrix} \quad (3,3)$$

où les matrices Λ_i sont carrées et telles que :

$$\Lambda_i = \begin{pmatrix} J_{i_1} & & & & \\ & J_{i_2} & & \bigcirc & \\ & & \ddots & & \\ \bigcirc & & & \ddots & \\ & & & & J_{i_p} \end{pmatrix} \quad (3,4)$$

$$\rho(J_{i_{1-1}}) > \rho(J_{i_1}) = \rho(J_{i_2}) = \dots = \rho(J_{i_p}) > \rho(J_{i_{p+1}})$$

Posons $p_i = n_{i_1}$ (3.5). C'est l'ordre le plus élevé des sous-matrices de Jordan intervenant dans la matrice Λ_i .

Alors les matrices Λ_i satisfont à la condition :

$$\rho(\Lambda_1) > \rho(\Lambda_2) > \dots > \rho(\Lambda_k) > 0 \quad (3,6)$$

On sait qu'il existe une matrice carrée Y inversible telle que :

$$A = Y^{-1} J Y \quad (3,7)$$

Notons

$$X = Y^{-1} \quad (3,8)$$

Procédons comme dans [5].

3.1. Définitions et lemmes préliminaires

On se fixe un partitionnement des matrices.

Soit Y une matrice carrée inversible d'ordre n à éléments complexes, définie par blocs $Y = (Y_{ij})_{i,j=1,2,\dots,k}$, pour ce partitionnement.

Définition 3.1. — On appelle *décomposition de Gauss par blocs de la matrice Y ainsi partitionnée*, la décomposition de Y en le produit de deux matrices L et U telles que

$$Y = LU$$

et par rapport à ce partitionnement :

i) la matrice L soit bloc triangulaire inférieure, ses blocs diagonaux étant des matrices unité ;

ii) la matrice U soit bloc triangulaire supérieure.

Lemme 3.1. — Une décomposition de Gauss par blocs d'une matrice Y inversible ($Y = (Y_{ij})_{i,j=1,2,\dots,k}$) existe et est unique si et seulement si pour tout entier p compris entre 1 et k , les déterminants des matrices

$$Y = (Y_{ij})_{i,j=1,2,\dots,p}$$

sont non nuls.

La démonstration est analogue à celle de la décomposition de Gauss habituelle.

Lemme 3.2 (cf. [5]). — Soit F_s une matrice d'ordre n à éléments complexes, inversible, factorisée en le produit d'une matrice \tilde{Q}_s unitaire et d'une matrice \tilde{R}_s triangulaire supérieure dont les éléments diagonaux sont des nombres réels positifs. Si la matrice F_s tend vers la matrice unité I quand s tend vers l'infini, alors les matrices \tilde{Q}_s et \tilde{R}_s tendent aussi vers la matrice unité I .

Lemme 3.3. — Soit A une matrice carrée d'ordre n , inversible, factorisée en le produit d'une matrice unitaire Q et d'une matrice bloc triangulaire supérieure R . Une telle décomposition existe, et toutes les décompositions sont de la forme

$$A = QR = (QV)(V^*R)$$

où V est une matrice bloc diagonale unitaire quelconque.

En effet une telle décomposition existe, il suffit de faire la décomposition élément par élément de A . Soit donc deux telles décompositions :

$$A = Q_1R_1 = Q_2R_2$$

Les différentes matrices étant inversibles, on peut écrire :

$$Q_2^{-1}Q_1 = R_2R_1^{-1} = V$$

Alors d'après cette égalité, V est une matrice unitaire et bloc triangulaire supérieure. De l'équation $V^* = V^{-1}$, on déduit que la matrice V^{-1} est bloc triangulaire supérieure et bloc triangulaire inférieure, donc bloc diagonale, d'où le lemme.

3.2. Etude de A^s

Nous considérons le partitionnement des matrices carrées d'ordre n tel que les blocs-diagonaux de la matrice J soient les matrices Λ_i .

Supposons que la matrice Y définie en (3.7) admette une décomposition de Gauss par blocs : (définition 3.1)

$$Y = LU \tag{3,9}$$

On a alors

$$\begin{aligned} A^s &= XJ^sLU \\ &= X(J^sLJ^{-s})J^sU \end{aligned} \tag{3,10}$$

Posons

$$M^{(s)} = J^sLJ^{-s} \tag{3,11}$$

En faisant le produit par blocs on a :

$$M_{ij}^{(s)} = \Lambda_i^s L_{ij} \Lambda_j^{-s} \tag{3,12}$$

La matrice $M^{(s)}$ est bloc triangulaire inférieure, les blocs diagonaux étant des matrices unité (cf. lemme 3.1) et pour :

$$i > j \quad \|M_{ij}^{(s)}\| \leq \|\Lambda_i^s\| \|L_{ij}\| \|\Lambda_j^{-s}\| \quad (3,13)$$

$$\|M_{ij}^{(s)}\| \leq \|L_{ij}\| (\|\Lambda_i^s\|^{1/s} \|\Lambda_j^{-1}\|^{1/s})^s \quad (3,14)$$

Or pour toute matrice A et pour toute norme matricielle, le rayon spectral $\rho(A)$ vérifie :

$$\rho(A) = \lim_{s \rightarrow +\infty} \|A^s\|^{\frac{1}{s}}$$

Par suite :

$$\lim_{s \rightarrow +\infty} \|\Lambda_i^s\|^{\frac{1}{s}} \|(\Lambda_j^{-1})^s\|^{\frac{1}{s}} = \frac{\rho(\Lambda_i)}{\rho(\Lambda_j)} \quad (3,15)$$

Or par construction pour $i > j$ on a $\rho(\Lambda_i) < \rho(\Lambda_j)$ donc pour $i > j$

$$\|M_{ij}^{(s)}\| \rightarrow 0 \text{ quand } s \rightarrow +\infty.$$

On peut donc écrire :

$$M^{(s)} = I + E_s \quad (3,16)$$

où E_s est une matrice bloc triangulaire strictement inférieure tendant vers zéro.

Soit $X = QR$ la décomposition de la matrice X en une matrice Q unitaire et une matrice R triangulaire supérieure dont les éléments diagonaux sont des nombres réels positifs.

On a alors :

$$\begin{aligned} A^s &= QR(I + E_s)(J^s U) \\ &= Q[I + RE_s R^{-1}][RJ^s U] \end{aligned} \quad (3,17)$$

soit

$$I + RE_s R^{-1} = \tilde{Q}_s \tilde{R}_s, \quad (3,18)$$

la décomposition de la matrice $I + RE_s R^{-1}$ en une matrice \tilde{Q}_s unitaire et une matrice \tilde{R}_s triangulaire supérieure dont les éléments diagonaux sont des nombres réels positifs.

Or la matrice E_s tendant vers zéro, la matrice $I + RE_s R^{-1}$ tend vers la matrice unité I et d'après le lemme (3.2)

$$\lim_{s \rightarrow +\infty} \tilde{Q}_s = \lim_{s \rightarrow +\infty} \tilde{R}_s = I \quad (3,19)$$

Les équations (3.17) et (3.18) entraînent :

$$A^s = (Q\tilde{Q}_s)(\tilde{R}_s R J^s U) \quad (3,20)$$

où — $Q\tilde{Q}_s$ est une matrice unitaire,

— $\tilde{R}_s R J^s U$ est une matrice bloc triangulaire supérieure.

3.3. Etude de la matrice P_s ; convergence de la matrice A_{s+1}

D'après l'équation (1.7) et le lemme (1.3), l'équation (3.20) permet de dire qu'il existe une matrice bloc diagonale unitaire V_s telle que :

$$P_s = Q\tilde{Q}_sV_s \tag{3,21}$$

Alors d'après les équations (1.8) et (3.7) on a :

$$A_{s+1} = (V_s^*\tilde{Q}_s^*Q^*)(XJY)(Q\tilde{Q}_sV_s)$$

or

$$X = Y^{-1} = QR$$

donc

$$A_{s+1} = V_s^*\tilde{Q}_s^*(RJR^{-1})\tilde{Q}_sV_s \tag{3,22}$$

or :

- i) la matrice RJR^{-1} est bloc triangulaire supérieure, les blocs diagonaux étant semblables aux matrices Λ_i ,
- ii) la matrice \tilde{Q}_s tend vers la matrice unité I ,
- iii) les matrices V_s sont blocs diagonales unitaires, d'où le

Théorème 3.1. — *Avec les notations précédentes, et le partitionnement considéré, si la matrice Y admet pour ce partitionnement une décomposition de Gauss par blocs, alors l'algorithme QR appliqué à la matrice A converge pour ce partitionnement.*

Corollaire 3.1. — *Avec les mêmes hypothèses et les mêmes notations, la matrice A_s tend à prendre une forme bloc triangulaire supérieure :*

$$(\tilde{A}_{ij})_{i,j} = 1, 2 \dots k.$$

Le i -ième bloc diagonal étant la matrice \tilde{A}_{ii} ayant pour valeurs propres, les valeurs propres de i -ième plus grand module de la matrice A .

3.4. Application

Définition 3.2. — *On appelle matrice Hessenberg supérieure non réduite, une matrice Hessenberg supérieure dont les éléments de la sous-diagonale sont des nombres complexes non nuls.*

Théorème 3.2 (cf. [2] et [5]). — *Si H désigne une matrice Hessenberg supérieure non réduite inversible, pour toute réduite de Jordan J inférieure de la matrice H , il existe une matrice Y inversible telle que :*

$$H = Y^{-1}JY$$

et la matrice Y admet une décomposition de Gauss.

Donc, en particulier, si nous prenons une réduite de Jordan J qui ordonne les modules des valeurs propres de H en décroissant, si nous considérons le

partitionnement de la matrice J qui sépare les modules des valeurs propres, la matrice Y admet une décomposition de Gauss par blocs par rapport à ce partitionnement, d'où le :

Théorème 3.3. — *L'algorithme QR appliqué à une matrice Hessenberg supérieure non réduite inversible H est convergente pour le partitionnement associé à celui qui ordonne les valeurs propres de H par ordre de module strictement décroissant.*

Corollaire 3.2. — *Le transformé H_s d'une matrice Hessenberg supérieure non réduite H tend à prendre une forme bloc triangulaire supérieure $(H_{ij})_{i,j=1,2,\dots,k}$. Les valeurs propres du i -ième bloc diagonal étant les valeurs propres de i -ième plus grand module de la matrice H .*

4. — COMPORTEMENT ASYMPTOTIQUE DE L'ERREUR

4.1. Lemme préliminaire

Nous désignerons par $\|\cdot\|_2$ la norme euclidienne, soit :

$$\|A\|_2^2 = \rho(A^*A)$$

L'expression $\binom{s}{p}$ désigne le coefficient binomial $\frac{s!}{p!(s-p)!}$.

Soit A une matrice carrée d'ordre n à coefficients complexes, soit J une réduite de Jordan inférieure de A dont les matrices diagonales seront appelées J_i .

Soit $\lambda_1, \lambda_2, \dots, \lambda_k$ les valeurs propres de la matrice A telles que :

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_k|$$

Parmi les blocs J_i dont la valeur propre a pour module $|\lambda_1|$ (resp. $|\lambda_k|$) il en existe un au moins dont l'ordre est maximum, soit p (resp. q) cet ordre.

Lemme 4.1. — *Avec les notations précédentes, si A est inversible, il existe deux constantes $\nu(A)$ et $\nu(A^{-1})$ telles que*

$$\frac{\|A^s\|_2}{\nu(A)\binom{s}{p-1}[\rho(A)]^{s-(p-1)}} \rightarrow 1 \text{ quand } s \rightarrow +\infty \quad (4,1)$$

$$\frac{\|(A^{-1})^s\|_2}{\nu(A^{-1})\binom{s}{p-1}[\rho(A^{-1})]^{s-(q-1)}} \rightarrow 1 \text{ quand } s \rightarrow +\infty \quad (4,2)$$

Démonstration

La relation (4.1) est une application directe de [4]. Vérifions la seconde.

Il existe une matrice Y carrée d'ordre n inversible telle que :

$$A = Y^{-1}JY$$

d'où

$$A^{-1} = Y^{-1}J^{-1}Y$$

Il faut donc écrire une forme de Jordan inférieure de J^{-1} , donc de chacun de ses blocs diagonaux et appliquer [4].

Or si B désigne un bloc de Jordan de valeur propre $\lambda \neq 0$:

$$B = \begin{pmatrix} \lambda & & & & \\ 1 & \lambda & \circ & & \\ & \cdot & \cdot & & \\ & & \cdot & & \\ \circ & & & \cdot & \\ & & & & 1 & \lambda \end{pmatrix}$$

d'ordre b , il est immédiat de vérifier que la matrice B^{-1} admet $\frac{1}{\lambda}$ comme valeur propre d'ordre b et que ses vecteurs propres sont les mêmes que ceux de la matrice B , donc il existe une matrice S d'ordre b inversible telle que :

$$B^{-1} = S^{-1} \begin{pmatrix} \frac{1}{\lambda} & & & & \\ & \circ & & & \\ & & \cdot & & \\ 1 & & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ \circ & & & & 1 & \frac{1}{\lambda} \end{pmatrix} S.$$

Alors la relation (4.2) s'en déduit immédiatement.

4.2. Etude de $M_{ij}^{(s)}$

On a vu (3.13) que : si $i > j$ alors

$$\|M_{ij}^{(s)}\|_2 \leq \|L_{ij}\|_2 \|\Lambda_i^s\|_2 \|\Lambda_j^{-s}\|_2$$

Appliquons le lemme (4.1) aux matrices Λ_i et Λ_j^{-1} . Avec les notations définies en (3.4) et (3.5), on a :

$$\frac{\|\Lambda_i^s\|_2}{v(\Lambda_i)_{(p_i-1)}^s [\rho(\Lambda_i)]^{s-(p_i-1)}} \rightarrow 1 \text{ quand } s \rightarrow +\infty \quad (4.3)$$

$$\frac{\|\Lambda_j^{-s}\|_2}{v(\Lambda_j^{-1})_{(p_j-1)}^s [\rho(\Lambda_j^{-1})]^{s-(p_j-1)}} \rightarrow 1 \text{ quand } s \rightarrow +\infty \quad (4.4)$$

Théorème 4.1. — Pour tout entier i et j l'expression $\|\Lambda_i^s\|_2 \|\Lambda_j^{-s}\|_2$ a le même comportement asymptotique que :

$$s^{p_i + p_j - 2} \left(\frac{\rho(\Lambda_i)}{\rho(\Lambda_j)} \right)^s.$$

Démonstration

Les relations (4.3) et (4.4) entraînent :

$$\frac{\|\Lambda_i^s\|_2 \|\Lambda_j^{-s}\|_2}{\nu(\Lambda_i)\nu(\Lambda_j^{-1}) \frac{(\rho(\Lambda_j))^{p_j-1}}{(\rho(\Lambda_i))^{p_i-1}} \binom{s}{p_i-1} \binom{s}{p_j-1} \left(\frac{\rho(\Lambda_i)}{\rho(\Lambda_j)} \right)^s} \rightarrow 1 \text{ quand } s \rightarrow +\infty$$

d'où le théorème (4.1).

Théorème 4.2. — Pour $i > j$, il existe une constante K telle que l'on a :

$$\|M_{ij}^{(s)}\|_2 \leq K s^{p_i + p_j - 2} \left(\frac{\rho(\Lambda_i)}{\rho(\Lambda_j)} \right)^s \quad (4,5)$$

pour s assez grand.

Démonstration

Il suffit d'écrire la relation (3,13) et d'appliquer le théorème 4.1.

4.3. Comportement asymptotique de l'erreur

C'est par définition celui de la matrice E_s , donc de chacun des blocs $M_{ij}^{(s)}$ pour $i > j$; il suffit alors d'appliquer le théorème (4.2).

L'inégalité (4,5) montre que :

i) si tous les diviseurs élémentaires de la matrice A sont linéaires, soit $p_i = p_j = 1$, on a :

$$\text{pour } i > j \quad , \quad \|M_{ij}^{(s)}\|_2 \leq K \left(\frac{\rho(\Lambda_i)}{\rho(\Lambda_j)} \right)^s$$

pour s assez grand;

ii) si la matrice A a des diviseurs élémentaires non linéaires, la convergence est « lente », ce qui est confirmé par les calculs numériques.

BIBLIOGRAPHIE

- [1] J. G. F. FRANCIS, The QR transformation, parts I and II, *Comp. Jour.*, 4, pp. 265-322, 1960-1961.
- [2] B. PARLETT, Canonical decomposition of Hessenberg matrices, *Mathematics of Computation*, vol. XXI, n° 98, 1967.
- [3] B. PARLETT, Necessary and sufficient conditions for convergence of the QR algorithm on the Hessenberg matrices, *Proceedings A.C.M. National Meeting*, 1966.
- [4] R. S. VARGA, *Matrix iterative Analysis*, Prentice Hall, 1962 (Th. 3.1, p. 65).
- [5] J. H. WILKINSON, *The algebraic eigenvalue problem*, Oxford University Press, 1965.