REVUE FRANÇAISE D'AUTOMATIQUE, D'INFORMATIQUE ET DE RECHERCHE OPÉRATIONNELLE. RECHERCHE OPÉRATIONNELLE

B. LEMAIRE

Régularité des matrices à diagonale dominante. Applications à l'absorption dans les chaînes et processus de Markov

Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle, tome 19, nº 3 (1985), p. 233-241.

<http://www.numdam.org/item?id=RO_1985__19_3_233_0>

R.A.I.R.O. Recherche opérationnelle/Operations Research
(vol. 19, n° 3, août 1985, p. 233 à 241)

RÉGULARITÉ DES MATRICES A DIAGONALE DOMINANTE. APPLICATIONS A L'ABSORPTION DANS LES CHAÎNES ET PROCESSUS DE MARKOV (*)

par B. LEMAIRE [1]

Résumé. — *Les théorèmes précisant la localisation dans le plan complexe des valeurs propres d'une matrice, ont des corollaires permettant d'affirmer la régularité de matrices à diagonales dominantes, sous certaines hypothèses. Celles-ci, cependant, sont rarement vérifiées dans nombre d'applications en R.O. : ainsi, pour des calculs de probabilités ou de temps d'absorption dans les chaînes et processus de Markov qui modélisent des problèmes de fiabilité, d'évolution de populations scolaires, etc. Aussi, propose-t-on, après un bref rappel des résulats classiques, un corollaire nouveau applicable à ces cas.*

Mots clés : Matrices à diagonale dominante; chaînes de Markov, processus de Markov; fiabilité.

Abstract. — *Corollaries of Gerschgorin theorem allow to determine classes of diagonally dominant matrices that are non-singular. However, most matrices obtained in O.R. problems do not belong to these classes: for example, in the evaluation of absorption probabilities or times in Markov chains or processes, resulting from the modelisation of reliability problems (M.T.T.F. calculation) or of the evolution of students populations. Also we propose — after a short recall of classical results — a new corollary and its application.*

Keywords: Diagonally dominant matrices; Markov chains; Markov processes; reliability.

I.1. Localisation des valeurs propres d'une matrice dans le plan complexe

Soit $A=[a_{ij}]$ une matrice $n \times n$ à éléments complexes; on pose :

$$R_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \qquad i=1, 2, \ldots, n.$$

On considère aussi, dans le plan complexe, les disques $D_i = \{ z \mid |z-a_{ii}| \leqq R_i \}$, centrés en a_{ii} et de rayon R_i;

(*) Reçu octobre 1984.
[1] Conservatoire National des Arts et Métiers, Paris.

R.A.I.R.O. Recherche opérationnelle/Operations Research, 0399-0559/85/03 233 09/$ 2.90

Soit $\mathscr{D}$ la réunion des disques D_i : $\mathscr{D}=\bigcup_{i=1}^{n} D_i$.

THÉORÈME 1 : *Toute valeur propre de A appartient à $\mathscr{D}$.*

En effet, soit λ une valeur propre de A et v un vecteur propre associé à λ. Soit $|v_r|=\max_i |v_i|$ $(|v_r|>0)$. La ligne r du système $A.v=\lambda v$ s'écrit :

$$\sum_{\substack{j=1\\ j\neq r}}^{n} a_{rj}v_j=v_r(\lambda-a_{rr});$$

d'où :

$$|\lambda-a_{rr}|\leqq\sum_{\substack{j=1\\ j\neq r}}^{n}|a_{rj}|\frac{|v_j|}{|v_r|}\leqq R_r. \quad \blacksquare$$

Ce théorème est dû, notamment, à Gerschgorin (1931) [1].

On rappelle qu'une matrice A d'ordre $n\geqq 2$ est « réductible » s'il existe une matrice de permutation P telle que :

$$P^t A.P.=\left[\begin{array}{c|c} A_{11} & A_{12}\\ \hline 0 & A_{22}\end{array}\right] \qquad (A_{11} \text{ et } A_{22} \text{ sont carrées}).$$

Sinon A est « irréductible ».

Le graphe $G(A)=(X, U)$ associé à A est défini par :

$$X=\{i\,|\,i=1, 2, \ldots, n\} \qquad \text{et} \qquad U=\{(i, j)\,|\,a_{ij}\neq 0;\ i, j\in X\};$$

on montre aisément que A est irréductible si et seulement si $G(A)$ est fortement connexe.

THÉORÈME 2 : *Si A est irréductible et s'il existe une valeur propre λ sur la frontière de $\mathscr{D}$, alors les n cercles d'équation $|z-a_{ii}|=R_i$ passent par λ.*

Rappelons la démonstration tout en l'exprimant en termes de graphes :

Démonstration : Soient λ une valeur propre de A, située sur la frontière de $\mathscr{D}$, v un vecteur propre associé à λ et $|v_r|=\max_i |v_i|$.

λ étant sur la frontière $\mathscr{D}$ ne peut être à l'intérieur de D_r. D'où :

$$R_r=|\lambda-a_{rr}|\leqq\sum_{j\neq r}|a_{rj}|\frac{|v_j|}{|v_r|}\leqq\sum_{j\neq r}|a_{rj}|=R_r;$$

donc :

$$\sum_{j \neq r} |a_{rj}| \left(1 - \frac{|v_j|}{|v_r|}\right) = 0.$$

Comme chaque terme est positif ou nul, on en déduit que $|v_j| = |v_r|$ pour tout j tel que $a_{rj} \neq 0$.

Puisque A est irréductible, il existe au moins un indice $p \neq r$ tel que $a_{rp} \neq 0$; λ appartient à la frontière de D_p. Soit k quelconque ($1 \leqq k \leqq n$): $G(A)$ étant fortement connexe, il existe un chemin de r à k, soit (r, r_1, r_2, . . . , r_m, k) associé à une séquence d'éléments non nuls : a_{rr_1}, $a_{r_1 r_2}$, . . . , $a_{r_m k}$. Donc λ se trouve à la frontière de D_{r_1}, et de D_{r_2}, . . . , et de D_{r_m}, donc de D_k : tout cercle frontière de $D_k (1 \leqq k \leqq n)$ passe par λ. ■

En fait, on suppose seulement, ici, que r est une racine du graphe (dans un graphe fortement connexe, tout sommet est une racine) : mais r n'est pas connu à l'avance. Ce théorème est dû à Taussky (1948) [1].

I.3. Matrices à diagonale dominante

Une matrice carrée A d'ordre n est dite :

– à diagonale dominante, si :

$$\sum_{\substack{j=1 \\ j \neq i}}^{n} |a_{ij}| \leqq |a_{ii}| \qquad (1 \leqq i \leqq n);$$

– à diagonale strictement dominante, si l'inégalité est stricte pour tout i;

– à diagonale quasi strictement dominante si l'inégalité est stricte pour au moins un i.

Les matrices à diagonale dominante sont d'usage fréquent : en électrostatique (matrice de capacités); en analyse numérique (dans la résolution d'équations aux dérivées partielles discrétisées); en économie (avec les matrices intersectorielles « input-output » de Leontiev); en recherche opérationnelle avec la notion d'absorption dans les chaînes et processus de Markov (Voir plus bas).

COROLLAIRE 1 : *Toute matrice à diagonale strictement dominante est régulière* [1].

COROLLAIRE 2 : *Toute matrice irréductible à diagonale quasi strictement dominante est régulière* [1].

Ces propriétés ont été aussi établies directement par nombre d'auteurs, dont Hadamard [1].

I.3.

Dans de nombreux exemples on a affaire à des matrices certes à diagonale quasi strictement dominante, mais réductibles : le corollaire ci-dessous, plus général que le corollaire 2, détermine une classe de matrices régulières parmi les matrices réductibles à diagonale quasi strictement dominante.

Soit $\mathscr{C} = \{\mathscr{C}_1, \mathscr{C}_2, \ldots, \mathscr{C}_p\}$ l'ensemble des composantes fortement connexes de $G(A)$. On peut ordonner $\mathscr{C}$ par la relation :

$$\mathscr{C}_k \leqq \mathscr{C}_l \quad \Leftrightarrow \quad [\mathscr{C}_k = \mathscr{C}_l \qquad \text{ou} \quad \exists\, i_k \in \mathscr{C}_k, \exists\, j_l \in \mathscr{C}_l \qquad \text{tels que} \quad j_l \in \hat{\Gamma}\, i_k)].$$

Tout élément maximal de l'ensemble ordonné $\mathscr{C}$, sera nommé « classe terminale ». $\mathscr{C}$ étant fini, il existe au moins une telle classe.

On numérote les lignes et colonnes de A en attribuant des indices consécutifs aux sommets d'une même composante fortement connexe. On commence par attribuer les indices les plus élevés aux sommets d'une classe terminale; puis on supprime celle-ci dans $G(A)$: dans le sous-graphe obtenu, on attribue les indices immédiatement inférieurs aux sommets d'une classe terminale du sous-graphe; et ainsi de suite tant que tous les sommets n'ont pas été supprimés. Après reclassement selon ces nouveaux indices, des lignes et des colonnes de A, on obtient la matrice A' qui a la forme (« triangulaire supérieure d'Hessenberg ») ci-dessous :

$$A' = \left[\begin{array}{c|c|c|c|c}
A_1 & A_{12} & A_{13} & \cdots & A_{1p} \\ \hline
0 & A_2 & A_{23} & \cdots & A_{2p} \\ \hline
0 & 0 & A_3 & \cdot & A_{3p} \\ \hline
\vdots & \vdots & \vdots & \ddots & \vdots \\ \hline
0 & 0 & 0 & \cdots & A_p
\end{array}\right]$$

A' est composée de blocs carrés non nuls le long de la diagonale principale (un par composante fortement connexe) et de blocs nuls au dessous de celle-ci.

En effet, soit $\mathscr{C}_{m_k}$ la composante associée à A_k et $\mathscr{C}_{m_l}$, à A_l; si $k<l$, $\mathscr{C}_{m_k}<\mathscr{C}_{m_l}$ par suite le bloc A_{lk} est nul (sinon il existerait $i_l \in \mathscr{C}_{m_l}$ et $j_k \in \mathscr{C}_{m_k}$ tels que $a_{i_l j_k} \neq 0$, d'où $\mathscr{C}_{m_l}<\mathscr{C}_{m_k}$).

Cette permutation ne change pas le spectre de A et donc sa régularité, en effet, soit ρ la permutation des indices et P la matrice de permutation associée ($p_{ij}=1$ si $j=\rho(i)$, $p_{ij}=0$ sinon) : $A'=P^t.A.P$ et par suite $A'=\det A$.

COROLLAIRE 3 : *Toute matrice à diagonale dominante telle que les sous matrices des classes terminales soient à diagonale quasi strictement dominante est régulière.*

Démonstration : dét $A=\prod_{k=1}^{p}$ dét A_k;

pour k fixé :

– si A_k est associée à une classe terminale, A_k est à diagonale quasi strictement dominante par hypothèse, et irréductible (puisque c'est une composante fortement connexe). D'après le corollaire 2, A_k est régulière : dét $A_k \neq 0$;

– si A_k est associée à une classe non terminale $\mathscr{C}_{m_k}$, il existe au moins dans $G(A)$ un arc allant d'un sommet $i_k \in \mathscr{C}_{m_k}$ à un sommet j_h d'une classe $\mathscr{C}_{m_h}$ telle que $\mathscr{C}_{m_k}<\mathscr{C}_{m_h}$: $a_{i_k, j_h} \neq 0$. Par suite A_k est à diagonale quasi strictement dominante et dét $A_k \neq 0$.

Ainsi dét $A=\prod_{k=1}^{p}$ dét $A_k \neq 0$: A est régulière. ■

Remarque : Dans les applications de ce corollaire, on n'aura pas à effectuer la décomposition en classe; la nature des phénomènes étudiés fera que l'hypothèse est satisfaite.

II. APPLICATIONS A DES CALCULS D'ABSORPTION DANS LES CHAÎNES ET PROCESSUS DE MARKOV

II.1. Calcul des probabilités d'absorption dans une chaîne de Markov

Soit une chaîne de Markov constituée d'un ensemble $\mathscr{T}$ de t états transitoires (formant une ou plusieurs classes) et d'un ensemble $\bar{\mathscr{T}}$ de r états récurrents, tous absorbants. La matrice $M=[P_{ij}]$ des probabilités de transition de cette chaîne a l'allure suivante :

$$\begin{array}{r} 1 \\ \vdots \\ t \\ M=t+1 \\ \vdots \\ t+r \end{array} \left[\begin{array}{c|c} Z & \mathscr{D} \\ \hline 0 & I_r \end{array}\right]$$
$$1 \ldots t \quad t+1 \ldots t+r$$

Z est de dimension $t \times t$; $\mathscr{D}$ est de dimension $t \times r$; 0 est la matrice nulle $r \times t$; I_r est la matrice unité $r \times r$.

Calculons le nombre moyen $\bar{n}_{ij}$ de passages par l'état transitoire j, connaissant l'état initial i (transitoire), avant l'absorption; rappelons que :

$$\bar{n}_{ij}=\sum_{m=0}^{\infty} 1 . p_{ij}^{(m)}=\delta_{ij}+\sum_{m=1}^{\infty} p_{ij}^{(m)} \qquad (i, j \in \mathscr{T})$$

($\delta_{ij}=1$ si $i=j$, $\delta_{ij}=0$ sinon); $p_{ij}^{(m)}$ est la probabilité de passage de E_i à E_j en m transitions; on a : $p_{ij}^{(0)}=\delta_{ij}$ et $p_{ij}^{(1)}=p_{ij}$.

L'équation de Chapman-Kolmogorov peut s'écrire : $p_{ij}^{(m)}=\sum_{k \in \mathscr{T}} p_{ik}\, p_{kj}^{(m-1)}$.

D'où :

$$\bar{n}_{ij}=\delta_{ij}+\sum_{k \in \mathscr{T}} p_{ik} \sum_{m=1}^{\infty} p_{kj}^{(m-1)}=\delta_{ij}+\sum_{k \in \mathscr{T}} p_{ik} . \bar{n}_{kj}$$

Soit matriciellement : $N=I_t+Z . N$, où : $N=[\bar{n}_{ij}]$, c'est-à-dire :

$$(I_t-Z) . N=I_t.$$

Prouvons que la matrice $A=I_t-Z$ est régulière : puisque $\sum_{j \in \mathscr{T} \cup \bar{\mathscr{T}}} p_{ij}=1$ pour tout $i \in \mathscr{T} \cup \bar{\mathscr{T}}$, $\sum_{j \in \mathscr{T}} p_{ij} \leqq 1$: A est à diagonale dominante. Elle comporte au moins une classe terminale; toute classe terminale étant transitoire, il existe au moins un arc d'un état de cette classe vers un état récurrent : par suite, la sous matrice carrée associée à toute classe terminale est à diagonale quasi strictement dominante. D'après le corollaire 3, A est régulière.

Notons que le corollaire 2 ne permettrait de conclure que si la chaîne comportait une seule classe d'états transitoires (A irréductible).

On a donc alors simplement :

$$N=(I_t-Z)^{-1}.$$

Calculons maintenant la probabilité a_{ij} d'absorption par l'état absorbant $j\,(j\in\bar{\mathcal{T}})$ lorsque l'état initial est $i\,(i\in\mathcal{T})$. Rappelons que :

$$a_{ij}=p_{ij}+\sum_{m=2}^{\infty} f_{ij}^{(m)},$$

où $f_{ij}^{(m)}$ est la probabilité de premier passage en j, partant de i, au bout de m transitions.

Ici : $f_{ij}^{(m)}=\sum_{k\in\mathcal{T}} p_{ik}\,f_{kj}^{(m-1)}$ $(m\geqq 2)$, d'où :

$$a_{ij}=p_{ij}+\sum_{k\in\mathcal{T}} p_{ik}\left(\sum_{m=2}^{\infty} f_{kj}^{(m-1)}\right)=p_{ij}+\sum_{k\in\mathcal{T}} p_{ik}\,a_{kj}.$$

Soit matriciellement, avec $\mathcal{A}=[a_{ij}]$:

$$\mathcal{A}=\mathcal{D}+Z\,.\,\mathcal{A},$$

d'où :

$$\mathcal{A}=(I_t-Z)^{-1}\,.\,\mathcal{D}$$

puisque cet inverse existe.

Pour les applications à l'étude de populations scolaires se reporter à [8], p. 23-27 et à [3].

II.2. Absorption dans un processus de Markov ([2])

Considérons un processus de Markov homogène, à espace d'états fini, comportant des états transitoires $\mathcal{T}=\{1, 2, \ldots, k\}$ et des états absorbants $\bar{\mathcal{T}}=\{k+1, \ldots, n\}$. Le générateur A du processus de Markov est donc de la forme :

$$\begin{array}{c} \\ 1 \\ \vdots \\ k \\ k+1 \\ \vdots \\ n \end{array}
\begin{array}{c} \begin{array}{cc} 1\ldots k & k+1\ldots n \end{array} \\ \left[\begin{array}{c|c} A_1 & A_2 \\ \hline 0 & 0 \end{array}\right] \end{array}=A,\qquad
\begin{cases} a_{ij}=\lambda_{ij}\geqq 0 \quad \text{pour } i\neq j, \\ a_{ii}=-\displaystyle\sum_{\substack{j=1\\ j\neq i}}^{n}\lambda_{ij}. \end{cases}$$

([2]) Cette partie a été rédigée avec S. Natkin.

On suppose connue une répartition initiale à $t=0$ des probabilités des états : $P(0)=[p_1(0), \ldots, p_n(0)]$; posons $P_{tr}(t)=[p_1(t), \ldots, p_k(t)]$ et $P_a^{(t)}=[p_{k+1}^{(t)}, \ldots, p_n^{(t)}]$. Le vecteur-ligne des probabilités des états à t, connaissant $P(0)$, est : $P(t)=[P_{tr}(t), P_a(t)]$.

Puisque $P'(t)=P(t).A$, la multiplication par blocs fournit :

$$P'_{tr}(t)=P_{tr}(t).A_1 \tag{1}$$

et

$$P'_a(t)=P_{tr}(t).A_2. \tag{2}$$

Soit T_i la variable aléatoire « durée de séjour dans l'état transitoire i, avant absorption », connaissant $P(0)$:

$$\bar{T}_i=E(T_i)=\int_0^\infty p_i(t)\,dt.$$

En intégrant chaque membre de (1) entre $t=0$ et $t=\infty$:

$$\int_0^\infty dP_{tr}(t)=\left(\int_0^\infty P_{tr}(t).dt\right).A_1,$$

d'où en posant $\bar{T}=[\bar{T}_1, \ldots, \bar{T}_k]$:

$$-P_{tr}(0)=\bar{T}.A_1,$$

A est une matrice à diagonale dominante; A_2 n'est pas identiquement nulle (sinon on ne pourrait atteindre les états absorbants) : A_1 est donc à diagonale strictement dominante, mais peut être réductible.

Toutefois, comme au paragraphe précédent, les sous matrices des classes terminales de A_1 sont à diagonale quasi strictement dominante : d'après le corollaire 3, on en déduit que A_1 est régulière; d'où :

$$\bar{T}=-P_{tr}(0).A_1^{-1}.$$

Une application à l'étude d'un système informatique redondant figure dans [8].

BIBLIOGRAPHIE

1. R. S. VARGA, *Matrix Iterative Analysis*, 322 p., Prentice hall, 1962.
2. GANTMACHER, *Théorie des matrices*, t. 2, 268 p., Dunod, Paris, 1966.
3. S. LEWKOWIEZ, *Étude de l'évolution de la population scolaire du second degré en France à l'aide d'un modèle statistique*, Exploitation Informatique, Mémoire d'ingénieur, CNAM, Paris, 1981.

4. B. Kraft et F. Thomas, Rapport de stage, CERCI, 1981.

5. S. Natkin, *Quelques aspects de la sûreté de fonctionnement des systèmes informatiques*, Mémoire d'ingénieur, CNAM, Paris, 1978.

6. S. Natkin, *Les réseaux de Petri stochastiques et leur application à l'évaluation des systèmes informatiques*, 1980, Thèse Docteur-Ingénieur, CNAM, Paris, 1980.

7. Roseaux, *Exercices et problèmes résolus de recherche opérationnelle*, t. 2 : problèmes aléatoires en R.O., Masson, Paris, 1983.

8. B. Lemaire, *Régularité des matrices à diagonale dominante. Applications et ex applications et exemples à l'absorption dans les chaînes et processus de Markov*, Note de recherche CNAM, R.O., n° 1, 1984.

9. J. G. Kemeny et J. L. Snell, *Finite Markov Chains*, Springer-Verlag, 1976.