

INTRODUCTION À L'ANALYSE NUMÉRIQUE MATRICIELLE ET À L'OPTIMISATION

INTRODUCTION À L'ANALYSE NUMÉRIQUE MATRICIELLE ET À L'OPTIMISATION

Cours et exercices corrigés

P.G. Ciarlet

Membre de l'Institut,
Professeur à l'Université Pierre et Marie Curie

Ouvrage publié sous la direction de

P.G. Ciarlet et J.-L. Lions

Membres de l'Institut

DUNOD

Traduction

- en anglais, Cambridge University Press, 1989
- en italien, Masson, S.p.a., Milan, 1989
- en chinois, Presses de l'Université de Nankin, 1990
- en japonais, Mihon Hyoron, Tokyo (à paraître)

NOUS NOUS ENGAGEONS EN FAVEUR DE L'ENVIRONNEMENT :



Nos livres sont imprimés sur des papiers certifiés pour réduire notre impact sur l'environnement.



Le format de nos ouvrages est pensé afin d'optimiser l'utilisation du papier.



Depuis plus de 30 ans, nous imprimons 70% de nos livres en France et 25% en Europe et nous mettons tout en œuvre pour augmenter cet engagement auprès des imprimeurs français.



Nous limitons l'utilisation du plastique sur nos ouvrages (film sur les couvertures et les livres).

© Dunod, 1998, 2024 pour la nouvelle présentation
© Masson, Paris 1990, pour la 1^{re} édition
11 rue Paul Bert, 92240 Malakoff
www.dunod.com
ISBN 978-2-10-087442-2

PRÉSENTATION DE LA COLLECTION “MATHÉMATIQUES APPLIQUÉES POUR LA MAÎTRISE”

La Collection *Mathématiques appliquées pour la maîtrise* a pour but de présenter les principales théories mathématiques générales directement orientées vers les applications, de les développer de manière rigoureuse, et d’indiquer explicitement et avec précision la très grande variété de leurs applications.

Des *théories mathématiques générales orientées vers les applications* sont, notamment, les fondements de l’analyse des *équations différentielles et aux dérivées partielles*, linéaires ou non, qui “gouvernent” tellement de situations en Physique, en Mécanique, en Chimie, etc., et jusqu’en Économétrie ! Ce sont aussi les outils principaux de l’*Analyse Numérique*, préalables obligés au traitement sur ordinateur : analyse numérique matricielle, méthodes de l’optimisation, méthodes de différences finies ou d’éléments finis pour l’approximation des solutions d’équations aux dérivées partielles ; c’est aussi la *Statistique*, dont les applications sont universelles, et où l’ordinateur a apporté, là encore, une impulsion nouvelle considérable ; c’est aussi la *Mécanique des Solides* et la *Mécanique des Fluides* dont une connaissance déjà sérieuse est indispensable à tout mathématicien appliqué.

Ces théories générales sont, dans la Collection, développées de manière rigoureuse, par le biais des solutions les plus synthétiques, les plus élégantes et les plus “confirmées” ; elle fournissent ainsi tous les outils nécessaires pour aborder la grande majorité des problèmes posés quotidiennement par les applications. Les théories générales présentées dans cette Collection ont d’ailleurs été élaborées pour faire face précisément aux *applications*, c’est-à-dire à des problèmes posés dans des disciplines parfois très éloignées des mathématiques mais néanmoins susceptibles d’être formalisés de façon mathématique.

Ces mêmes théories devraient également servir de point de départ pour l’étude des *nouveaux* problèmes posés par les applications ; il est en effet essentiel de savoir que ces nouveaux problèmes, d’importance fondamentale, se présentent sous la forme de questions complètement “ouvertes”. Après le préalable d’une modélisation mathématique souvent déjà imparfaite, la *seule* façon de les aborder réside alors dans un traitement “massif” sur ordinateur, à l’aide *précisément des méthodes et des outils fondamentaux présentés dans cette Collection*.

C'est pourquoi cette Collection, qui s'adresse à tous les étudiants du Deuxième Cycle de Mathématiques dites "appliquées", mais aussi (au moins pour certains de ses volumes) aux étudiants du Deuxième Cycle de Mathématiques dites "pures", de Mécanique, de Physique, aux élèves des Grandes Ecoles d'Ingénieurs, . . . , devrait non seulement initier ses lecteurs à des théories rigoureuses et élégantes, tout en leur fournissant un outil déjà utilisable dans de très nombreuses applications, mais aussi, nous l'espérons. leur donner le désir d'aller bien au-delà.

Pour l'accueil compréhensif qu'elle a bien voulu réserver à cette Collection, il nous est particulièrement agréable de remercier la maison Masson, en la personne notamment de M. J. F. Le Grand. Nous tenons également à remercier bien vivement M. A. Warusfel, dont l'activité et le dévouement ont beaucoup contribué à la conception et à l'élaboration de cette Collection.

P. G. CIARLET J. L. LIONS

PRÉFACE

L'objet essentiel de cet ouvrage est de donner, tout en restant dans des limites raisonnables, une description et une analyse relativement complètes des méthodes les plus couramment utilisées en Analyse Numérique Matricielle et en Optimisation.

A l'heure où, dans tous les domaines scientifiques, apparaît la nécessité d'utiliser les méthodes numériques les mieux adaptées aux performances spectaculaires des ordinateurs, je souhaite que cet ouvrage contribue non seulement à montrer l'efficacité des méthodes d'emploi universel qui y sont décrites, mais aussi à mettre en évidence l'intérêt que peut présenter leur analyse mathématique. Si le premier aspect intéresse surtout les utilisateurs éventuels, et le second les lecteurs épris de rigueur mathématique, il n'est pas interdit d'espérer que ce livre pourrait susciter, chez les uns comme chez les autres, un intérêt *simultané* pour ces deux caractéristiques complémentaires de l'Analyse Numérique.

*
* *

Cet ouvrage s'adresse tout particulièrement aux étudiants de première année de la Maîtrise de Mathématiques Pures, de la Maîtrise de Mathématiques et Applications Fondamentales, de la Maîtrise de Mécanique, ainsi qu'aux élèves de première année des Grandes Écoles. Mais il devrait également être utile aux Ingénieurs, Physiciens, Mécaniciens, Biologistes, Économistes, . . . , qui souhaitent avoir une idée des méthodes numériques constamment utilisées aujourd'hui, et être en mesure de mettre en œuvre certaines d'entre elles.

En raison du niveau mathématique relativement modéré que requiert cet ouvrage (surtout dans sa première partie), la plupart des matières qui y sont traitées pourraient “glisser” progressivement dans l’enseignement mathématique des Classes Préparatoires aux Grandes Écoles et des Diplômes d’Études Universitaires Générales. C’est là un de mes vœux les plus chers, qui, s’il correspond avant tout à une inclination personnelle, me semble aussi répondre à un simple souci *d’équité* : est-il en effet raisonnable de se limiter aux formules de Cramer et d’ignorer les méthodes de Gauss et de Cholesky ? Est-il en effet raisonnable de se limiter au polynôme caractéristique et à la réduite de Jordan et d’ignorer l’existence de la méthode QR et de la méthode de Givens-Householder ? Est-il en effet raisonnable de rejeter les coniques et autres quadriques dans les ténèbres extérieures et d’ignorer l’intérêt de la notion de direction conjuguée pour la minimisation d’une fonctionnelle quadratique ?

*
* *

Dans tous les cas, l’enseignant devrait pouvoir adapter facilement l’ouvrage au gré de ses besoins et au niveau de son auditoire. A titre indicatif, un enseignement semestriel de trois heures hebdomadaires correspond au contenu des chapitres 1 à 6, ou à celui (de niveau un peu plus élevé) des chapitres 7 à 10, ou encore à celui des chapitres 4 à 8.

On a supposé les lecteurs déjà familiarisés avec les principales propriétés des matrices (notamment le calcul matriciel), les espaces vectoriels normés de dimension finie (continuité et dérivabilité des fonctions de plusieurs variables, compacité, applications linéaires, etc.). S’il est vrai que dans la deuxième partie, il est fait usage des espaces de Banach, des espaces de Hilbert, et du calcul différentiel dans les espaces vectoriels normés quelconques, les définitions et résultats correspondants sont toujours rappelés avec précision, et surtout, les lecteurs peu enclins à ce genre de généralisations pourront *sans aucun inconvénient* continuer à “rester en dimension finie”. C’est dans le même esprit qu’on s’est volontairement abstenu d’utiliser la notion de convergence faible, sauf dans une démonstration ; mais il s’agit d’un résultat d’existence en dimension infinie, dont l’analogie en dimension finie est démontré par ailleurs de façon élémentaire.

Les caractéristiques suivantes de l’ouvrage méritent, croyons-nous, d’être signalées :

- *réunion dans un même volume de l’Analyse Numérique Matricielle et de l’Optimisation*, avec passage progressif, et nombreuses références, de l’une à l’autre ;
- ordre choisi pour les matières traitées et les préliminaires correspondants, qui correspond à un *niveau mathématique généralement croissant* ;
- importance de la place accordée aux *rappels et compléments* mathématiques nécessaires ;
- description de problèmes variés, issus de la Physique, de la Mécanique, de l’Économie, etc., conduisant à la résolution de problèmes d’Analyse Numérique Matricielle ou d’Optimisation ;
- démonstrations *complètes*, et aussi *simples* que possible ;
- souci *pédagogique* : place accordée aux aspects descriptifs, présence de Commentaires Bibliographiques, d’un Index, existence d’un *Recueil d’Exercices* très complet.

*
* *

D’une façon plus précise, dans la *première partie* (chapitres 1 à 6), plus spécialement consacrée à l’*Analyse Numérique Matricielle*, on trouvera :

- les rappels et compléments nécessaires sur les *matrices* et sur les *normes* (vectorielles ou matricielles), ces dernières étant d’un usage constant dans toute la suite (chapitre 1) ;
- quelques notions sur le *conditionnement* d’un problème d’Analyse Numérique Matricielle (chapitre 2) ;

— quelques aperçus sur la variété des méthodes de l'Analyse Numérique (méthodes de différences finies ou d'approximation variationnelle pour les problèmes aux limites, d'interpolation, d'approximation au sens des moindres carrés, etc.) qui conduisent à un *système linéaire* ou à un problème de *valeurs propres* (chapitre 3) ;

— la description et l'analyse des principales méthodes *directes* (de Gauss, de Cholesky, de Householder ; cf. chapitre 4) et *itératives* (de Jacobi, de Gauss-Seidel, de relaxation ; cf. chapitre 5) de *résolution des systèmes linéaires* ;

— la description et l'analyse des principales méthodes de *calcul des valeurs propres* (de Jacobi, de Givens-Householder, QR) et des *vecteurs propres* des matrices.

Dans la *deuxième partie* (chapitre 7 à 10), plus spécialement consacrée à l'*Optimisation*, on trouvera :

— les rappels et compléments nécessaires sur le *calcul différentiel* dans les espaces vectoriels normés (chapitre 7) et sur les *espaces de Hilbert* (chapitre 8) ;

— une introduction progressive à l'Optimisation, par le biais de l'étude des *multiplieurs de Lagrange*, des *extremums* des fonctions réelles, de la *convexité*, et de la *méthode de Newton* (chapitre 7) ;

— la description de problèmes variés (approximation de problèmes aux limites linéaires et non linéaires, problèmes d'origine "économique") qui conduisent à la *minimisation de fonctionnelles*, avec ou sans contraintes (chapitres 8 et 10) ;

— la description et l'analyse des principales méthodes de l'Optimisation : méthode de *relaxation*, méthodes de *gradient* à pas optimal, à pas fixe ou variable, méthode du *gradient conjugué*, méthode de *pénalisation* (chapitre 8), méthode d'*Uzawa* (chapitre 9), méthode du *simplexe* (chapitre 10) ;

— une introduction à la *dualité* : lemme de Farkas-Minkowski, *relations de Kuhn et Tucker*, Lagrangiens et points-selles, application à la *programmation linéaire* (chapitres 9 et 10).

Pour une description plus complète des matières traitées, les lecteurs se reporteront aux *introductions* des chapitres.

*
* *

Les résultats importants sont énoncés sous forme de *théorèmes*, les plus importants d'entre eux étant repérés par une flèche dans la marge. Un "Z" dans la marge signale généralement une difficulté, ou un "piège" inattendu, qui peuvent être de nature très variée. De nombreuses *remarques* accompagnent le texte. Bien qu'elles puissent être en principe sautées en première lecture, elles ne doivent pas être négligées ; leur rôle est en effet d'aider les lecteurs, en situant mieux certains résultats, en signalant des cas particuliers intéressants, en mentionnant la possibilité ou l'impossibilité de certaines extensions, etc.

De nombreux exercices, de difficulté très variable, complètent ce présent ouvrage dans un *Recueil d'Exercices* de la même collection. Certains, qui se présentent sous la forme de *problèmes* apportent des compléments parfois très importants et il y est fait référence dans le texte.

On trouvera deux types *d'informations bibliographiques* : il s'agit soit de références "ponctuelles" (à propos d'un résultat précis, d'un exercice, etc.), qui apparaissent sous forme de notes en bas de page, soit de références de caractère plus général. Ces dernières sont classées *par sujet* sous la rubrique "Commentaires Bibliographiques", placée en fin d'ouvrage. Les lecteurs intéressés par la *programmation effective* des méthodes étudiées au fil des chapitres y trouveront à cet égard des indications bibliographiques précieuses.

*
* *

Lors de la rédaction, de nombreux collègues et étudiants ont bien voulu me faire part de leurs avis, observations, suggestions, etc. A ce propos, je tiens tout particulièrement à exprimer ma reconnaissance à Jean-Marie THOMAS qui a lu de façon détaillée la totalité du manuscrit, et dont les remarques sont à l'origine de nombreuses améliorations.

Ma gratitude s'adresse aussi à Claude BASDEVANT, Michel CROUZEIX, David FEINGOLD, Srinivasan KESAVAN, Colette LEBAUD, Jean MEINGUET, Annie PUECH-RAOULT, Pierre-Arnaud RAVIART, François ROBERT, Ulrich TULOWITZKI, Lars WAHLBIN, qui à des titres et degrés divers, m'ont fait profiter de leurs conseils éclairés.

D'un point de vue plus personnel, il m'est également agréable de remercier Richard S. VARGA, dont l'enseignement et l'enthousiasme très communicatif ont su me donner le goût de l'Analyse Numérique.

Enfin, c'est un grand plaisir pour moi que de remercier très sincèrement, une nouvelle fois, Hélène BUGLER pour la diligence et la qualité de son travail.

*
* *

Je dédie ce livre à Hélène et Gaston CIARLET.

Philippe G. CIARLET
Décembre 1980

TABLE DES MATIÈRES

Présentation de la collection	v
Préface	vii

PREMIÈRE PARTIE

Analyse numérique matricielle

1. Rappels et compléments sur les matrices	3
Introduction	3
1.1. Principales notations et définitions	3
1.2. Réduction des matrices	8
1.3. Propriétés particulières aux matrices symétriques et hermitiennes	11
1.4. Normes vectorielles et normes matricielles	14
1.5. Suites de vecteurs et de matrices	21
2. Généralités sur l'analyse numérique matricielle	23
Introduction	23
2.1. Les deux problèmes fondamentaux ; généralités sur les méthodes employées	23
2.2. Conditionnement d'un système linéaire	27
2.3. Conditionnement d'un problème de valeurs propres	34
3. Origine des problèmes de l'analyse numérique matricielle	37
Introduction	37
3.1. La méthode des différences finies pour un problème aux limites en dimension un	38
3.2. La méthode des différences finies pour un problème aux limites en dimension deux	45
3.3. La méthode des différences finies pour les problèmes aux limites d'évolution	48
3.4. Approximation variationnelle d'un problème aux limites en dimension un	53
3.5. Approximation variationnelle d'un problème aux limites en dimension deux	60
3.6. Problèmes de valeurs propres	62
3.7. Problèmes d'interpolation et d'approximation	66
4. Méthodes directes de résolutions de systèmes linéaires	71
Introduction	71
4.1. Deux remarques concernant la résolution des systèmes linéaires	72
4.2. La méthode de Gauss	73
4.3. La factorisation LU d'une matrice	82
4.4. La factorisation et la méthode de Cholesky	87
4.5. La factorisation QR d'une matrice et la méthode de Householder	90

5. Méthodes itératives de résolution des systèmes linéaires	95
Introduction	95
5.1. Généralités sur les méthodes itératives	95
5.2. Description des méthodes de Jacobi, de Gauss-Seidel, de relaxation	97
5.3. Convergence des méthodes de Jacobi, de Gauss-Seidel, de relaxation	102
6. Méthodes de calcul des valeurs propres et des vecteurs propres	110
Introduction	110
6.1. La méthode de Jacobi	111
6.2. La méthode de Givens-Householder	118
6.3. La méthode QR	123
6.4. Calcul des vecteurs propres	129

DEUXIÈME PARTIE

Optimisation

7. Rappels et compléments de calcul différentiel. Premières applications	135
Introduction	135
7.1. Dérivées première et seconde d'une application	137
7.2. Extremums des fonctions réelles: multiplicateurs de Lagrange	146
7.3. Extremums des fonctions réelles: prise en compte des dérivées secondes .	151
7.4. Extremums des fonctions réelles: prise en compte de la convexité	153
7.5. La méthode de Newton	158
8. Généralités sur l'optimisation. Premiers algorithmes	167
Introduction	167
8.1. Le théorème de projection: premières conséquences	168
8.2. Généralités sur les problèmes d'optimisation	173
8.3. Exemples de problèmes d'optimisation	179
8.4. Méthodes de relaxation et de gradient pour des problèmes sans contraintes	182
8.5. Méthodes de gradient conjugué pour des problèmes sans contraintes ...	194
8.6. Méthodes de relaxation, de gradient, et de pénalisation, pour des problèmes avec contraintes	201
9. Introduction à la programmation non linéaire	207
Introduction	207
9.1. Lemme de Farkas-Minkowski	208
9.2. Les relations de Kuhn et Tucker	211
9.3. Lagrangiens et points-selles. Introduction à la dualité	219
9.4. La méthode d'Uzawa	226
10. Programmation linéaire	231
Introduction	231
10.1. Généralités sur la programmation linéaire	232
10.2. Exemples de problèmes de programmation linéaire	235
10.3. La méthode du simplexe	237
10.4. Dualité et programmation linéaire	252
Commentaires bibliographiques	259
Références	262
Principales notations utilisées	266
Index	271

PREMIÈRE PARTIE

ANALYSE NUMÉRIQUE
MATRICIELLE

RAPPELS ET COMPLÉMENTS SUR LES MATRICES

Introduction

Ce chapitre a pour but de rappeler, et de démontrer, un certain nombre de résultats relatifs aux matrices et aux espaces vectoriels de dimension finie, et dont un usage constant sera fait dans toute la suite de l'ouvrage.

On suppose les lecteurs déjà familiarisés avec les propriétés élémentaires des espaces vectoriels de dimension finie (le calcul matriciel notamment), pour lesquelles on renvoie au paragraphe 1.1 les principales notations et définitions, ainsi que la notion de *décomposition par blocs d'une matrice*, qui est à signaler pour son importance en Analyse Numérique Matricielle.

Afin de rendre l'ouvrage aussi "autonome" que possible, tous les résultats importants pour la suite sont démontrés, en particulier *la réduction d'une matrice quelconque à la forme triangulaire*, *la diagonalisation des matrices normales* (théorème 1.2-1), et *l'équivalence d'une matrice à la matrice diagonale de ses valeurs singulières* (théorème 1.2-2). A cet égard, il convient de signaler que nous n'aurons pas à utiliser le théorème de Jordan. Nous examinons ensuite (théorème 1.3-1) les *caractérisations des valeurs propres des matrices symétriques ou hermitiennes* par l'intermédiaire du *quotient de Rayleigh*, notamment les caractérisations par "min-max" et par "max-min".

On passe ensuite en revue les *normes vectorielles* les plus couramment utilisées en Analyse Numérique Matricielle, qui sont des cas particuliers des "*normes l_p* " (théorème 1.4-1), puis on calcule les *normes matricielles subordonnées* correspondantes (théorème 1.4-2), un exemple de norme matricielle non subordonnée à une norme vectorielle étant donné au théorème 1.4-4. On rappelle également (théorème 1.4-5) les conditions d'invertibilité de matrices de la forme $(I+B)$, et on montre (théorème 1.4-3) que *le rayon spectral d'une matrice est la borne inférieure des valeurs de ses normes*, ce dernier résultat servant ensuite à démontrer *deux résultats relatifs à la suite des puissances successives d'une même matrice* (théorème 1.5-1 et 1.5-2), qui jouent un rôle fondamental dans l'étude des méthodes itératives de résolution de systèmes linéaires étudiées au chapitre 5.

1.1. Principales notations et définitions

Soit V un espace vectoriel de dimension finie n , sur le corps \mathbf{R} des nombres réels, ou sur le corps \mathbf{C} des nombres complexes ; s'il n'y a pas lieu de distinguer, on dit qu'il s'agit du corps \mathbf{K} des *scalaires*.

Une *base* de V est un ensemble $\{e_1, e_2, \dots, e_n\}$ de n vecteurs linéairement indépendants de V , qu'on notera $(e_i)_{i=1}^n$, ou simplement (e_i) si aucune confusion n'est à craindre. Tout vecteur $v \in V$ admet alors une décomposition unique

$$v = \sum_{i=1}^n v_i e_i,$$

les scalaires v_i , que nous noterons parfois $(v)_i$, étant les *composantes* du vecteur v sur la base (e_i) . *Lorsqu'une base est fixée sans ambiguïté, on peut ainsi identifier V à \mathbf{K}^n* ; c'est pourquoi il nous arrivera également de noter $v = (v_i)_{i=1}^n$, ou simplement (v_i) , un vecteur de composante v_i .

En notation matricielle, le vecteur $v = \sum_{i=1}^n v_i e_i$ sera *toujours* représenté par le *vecteur colonne*

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

et on désignera par v^T et v^* les *vecteurs lignes* suivants :

$$v^T = (v_1 v_2 \dots v_n), \quad v^* = (\bar{v}_1 \bar{v}_2 \dots \bar{v}_n),$$

où, en général, $\bar{\alpha}$ désigne le nombre complexe conjugué du nombre α . Le vecteur ligne v^T est le *vecteur transposé* du vecteur colonne v , et le vecteur ligne v^* est le *vecteur adjoint* du vecteur colonne v .

L'application $(\cdot, \cdot) : V \times V \rightarrow \mathbf{K}$ définie par

$$(u, v) = v^T u = u^T v = \sum_{i=1}^n u_i v_i \quad \text{si } \mathbf{K} = \mathbf{R},$$

$$(u, v) = v^* u = \overline{u^* v} = \sum_{i=1}^n u_i \bar{v}_i \quad \text{si } \mathbf{K} = \mathbf{C},$$

est appelée *produit scalaire euclidien* si $\mathbf{K} = \mathbf{R}$, *hermitien* si $\mathbf{K} = \mathbf{C}$, ou *canonique* si l'on ne précise pas le corps des scalaires. Si l'on souhaite rappeler la dimension de l'espace, on écrira

$$(u, v) = (u, v)_n.$$

Soit V un espace muni de son produit scalaire canonique. Deux vecteurs u et v de V sont *orthogonaux* si $(u, v) = 0$. Par extension, on dit qu'un vecteur v est *orthogonal à une partie* U de V , et on note $v \perp U$, lorsque le vecteur v est orthogonal à tous les vecteurs de U . Enfin, un ensemble $\{v_1, \dots, v_k\}$ de vecteurs de l'espace V est dit *orthonormal* si

$$(v_i, v_j) = \delta_{ij}, \quad 1 \leq i, j \leq k,$$

où δ_{ij} est le *symbole de Kronecker* : $\delta_{ij} = 1$ si $i = j$, $\delta_{ij} = 0$ si $i \neq j$.

Soit V et W deux espaces vectoriels sur le même corps, munis de bases $(e_j)_{j=1}^n$ et $(f_i)_{i=1}^m$ respectivement. *Relativement à ces bases*, une application linéaire

$$\mathcal{A} : V \rightarrow W$$

est représentée par la *matrice à m lignes et n colonnes* :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

les éléments a_{ij} de la matrice A étant définis de façon unique par les relations

$$Ae_j = \sum_{i=1}^m a_{ij} f_i, \quad 1 \leq j \leq n.$$

Autrement dit, le j -ème vecteur colonne

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

de la matrice A représente le vecteur Ae_j dans la base $(f_i)_{i=1}^m$. On appelle

$$(a_{i1} a_{i2} \dots a_{in})$$

le i -ème vecteur ligne de la matrice A .

Une matrice à m lignes et n colonnes est appelée *matrice de type (m, n)* , et on note $\mathcal{A}_{m, n}(\mathbf{K})$, ou simplement $\mathcal{A}_{m, n}$, l'espace vectoriel sur le corps \mathbf{K} formé par les matrices de type (m, n) à éléments dans \mathbf{K} . Un vecteur colonne est donc une matrice de type $(m, 1)$ et un vecteur ligne une matrice de type $(1, n)$. Une matrice est dite *réelle* ou *complexe* selon que ses éléments sont dans le corps \mathbf{R} ou dans le corps \mathbf{C} .

Une matrice A d'éléments a_{ij} est notée

$$A = (a_{ij}),$$

le premier indice i étant *toujours* celui de la ligne et le second, j , celui de la colonne. Étant donné une matrice A , on désigne par $(A)_{ij}$ l'élément de la i -ème ligne et de la j -ème colonne.

La *matrice nulle* et le *vecteur nul* sont désignés par la même lettre 0 .

Étant donné une matrice $A \in \mathcal{A}_{m, n}(\mathbf{C})$, on note $A^* \in \mathcal{A}_{n, m}(\mathbf{C})$ la *matrice adjointe* de la matrice A , définie de façon unique par les relations

$$(Au, v)_m = (u, A^*v)_n \quad \text{pour tout } u \in \mathbf{C}^n, \quad v \in \mathbf{C}^m,$$

qui entraînent $(A^*)_{ij} = \bar{a}_{ji}$. De la même façon, étant donné une matrice $A = \mathcal{A}_{m, n}(\mathbf{R})$, on note $A^T \in \mathcal{A}_{n, m}(\mathbf{R})$ la *matrice transposée* de la matrice A , définie de façon unique par les relations

$$(Au, v)_m = (u, A^T v)_n \quad \text{pour tout } u \in \mathbf{R}^n, \quad v \in \mathbf{R}^m,$$

qui entraînent $(A^T)_{ij} = a_{ji}$.

REMARQUES. (1) On peut encore définir la matrice transposée d'une matrice complexe, mais c'est une notion d'intérêt moindre, l'application $u, v \rightarrow \sum_{i=1}^n u_i v_i$ n'étant pas un produit scalaire dans \mathbf{C}^n .

(2) On a préféré la notation A^T à la notation habituelle tA , cette dernière étant davantage adaptée à la notion de base duale ; la notation A^T rappelle la dépendance de la notion de matrice transposée sur un produit scalaire particulier, le produit scalaire canonique en l'occurrence. ■

A la composition des applications linéaires correspond la multiplication des matrices : Si $A = (a_{ik})$ est une matrice de type (m, l) et $B = (b_{kj})$ de type (l, n) , leur produit AB est la matrice de type (m, n) définie par

$$(AB)_{ij} = \sum_{k=1}^l a_{ik} b_{kj}.$$

On rappelle que $(AB)^T = B^T A^T$, $(AB)^* = B^* A^*$.

Soit $A = (a_{ij})$ une matrice de type (m, n) . On appelle *sous-matrice* de la matrice A toute matrice de la forme

$$\begin{pmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_q} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_q} \\ \vdots & \vdots & & \vdots \\ a_{i_p j_1} & a_{i_p j_2} & \dots & a_{i_p j_q} \end{pmatrix},$$

les entiers i_k et j_l vérifiant

$$1 \leq i_1 < i_2 < \dots < i_p \leq m; \quad 1 \leq j_1 < j_2 < \dots < j_q \leq n.$$

Soit $A = (a_{ij})$ la matrice représentant une application linéaire de V dans W , et soit

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_N, \quad W = W_1 \oplus W_2 \oplus \dots \oplus W_M$$

des décompositions des espaces V et W en sommes directes de sous-espaces V_J et W_I , de dimensions n_J et m_I respectivement, engendrés par des vecteurs de base. A ces décompositions des espaces V et W , on associe la *décomposition par blocs de la matrice* A :

$$A = \begin{pmatrix} \begin{array}{|c|c|c|c|} \hline A_{11} & A_{12} & \dots & A_{1N} \\ \hline A_{21} & A_{22} & \dots & A_{2N} \\ \hline \vdots & \vdots & & \vdots \\ \hline A_{M1} & A_{M2} & \dots & A_{MN} \\ \hline \end{array} \end{pmatrix} = (A_{IJ})$$

Z chaque sous-matrice A_{IJ} , de type (m_I, n_J) , représentant une application linéaire de l'espace V_J dans l'espace W_I . L'intérêt de telles décompositions par blocs est que certaines des opérations définies sur les matrices restent *formellement* les mêmes, "les coefficients a_{ij} étant remplacés par les sous-matrices A_{IJ} ". Mais attention à l'ordre des facteurs !

Ainsi, soit $A = (A_{IK})$ et $B = (B_{KJ})$ deux matrices, de type (m, l) et (l, n) respectivement, décomposées par blocs, la *décomposition correspondant à l'indice K étant la même* pour chaque matrice. Alors la matrice AB admet comme *décomposition par blocs*

$$AB = (C_{IJ}), \quad \text{avec} \quad C_{IJ} = \sum_K A_{IK} B_{KJ},$$

et on dit qu'on a effectué le *produit par blocs* des deux matrices.

De la même façon, soit v un vecteur de l'espace V , et soit $v = \sum_{J=1}^N v_J$, $v_J \in V_J$, sa décomposition (unique) associée à la décomposition de l'espace V en somme directe. Le vecteur $Av \in W$ admet alors

$$Av = \sum_{I=1}^M w_I, \quad \text{avec} \quad w_I = \sum_{J=1}^N A_{IJ} v_J,$$

comme décomposition unique associée à la décomposition de l'espace W en somme directe. Il est équivalent de considérer que *les vecteurs v et Av sont décomposés en blocs* :

$$v = \begin{pmatrix} \overline{v_1} \\ \overline{v_2} \\ \vdots \\ \overline{v_N} \end{pmatrix}, \quad Av = \begin{pmatrix} \overline{w_1} \\ \overline{w_2} \\ \vdots \\ \overline{w_M} \end{pmatrix}, \quad w_I = \sum_{J=1}^N A_{IJ} v_J,$$

et que l'on a effectué le *produit par blocs de la matrice A et du vecteur v* .

Une matrice de type (n, n) est dite matrice *carrée*, ou *matrice d'ordre n* si l'on veut préciser l'entier n ; il est alors commode de dire qu'une matrice est *rectangulaire* lorsqu'elle n'est pas nécessairement carrée. On note

$$\mathcal{A}_n = \mathcal{A}_{n,n} \quad \text{ou} \quad \mathcal{A}_n(\mathbf{K}) = \mathcal{A}_{n,n}(\mathbf{K}),$$

l'*anneau des matrices carrées d'ordre n , à éléments dans le corps \mathbf{K}* .

Sauf mention du contraire, *les matrices considérées jusqu'à la fin de ce paragraphe sont carrées*.

Si $\mathbf{A} = (a_{ij})$ est une matrice carrée, les éléments a_{ii} sont appelés *éléments diagonaux*, et les éléments a_{ij} , $i \neq j$, sont appelés *éléments hors-diagonaux*. La *matrice unité* est la matrice

$$\mathbf{I} = (\delta_{ij}).$$

Une matrice \mathbf{A} est *inversible* s'il existe une matrice (unique si elle existe), notée \mathbf{A}^{-1} et appelée *matrice inverse* de la matrice \mathbf{A} , telle que $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. Dans le cas contraire, on dit que la matrice est *singulière*. On rappelle que, si \mathbf{A} et \mathbf{B} sont des matrices inversibles

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}, \quad (\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T, \quad (\mathbf{A}^*)^{-1} = (\mathbf{A}^{-1})^*.$$

Une matrice \mathbf{A} est :

symétrique si \mathbf{A} est réelle et $\mathbf{A} = \mathbf{A}^T$;

hermitienne si $\mathbf{A} = \mathbf{A}^*$;

orthogonale si \mathbf{A} est réelle et $\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A} = \mathbf{I}$;

unitaire si $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A} = \mathbf{I}$;

normale si $\mathbf{A}\mathbf{A}^* = \mathbf{A}^*\mathbf{A}$.

Une matrice $\mathbf{A} = (a_{ij})$ est *diagonale* si $a_{ij} = 0$ pour $i \neq j$; on la note

$$\mathbf{A} = \text{diag}(a_{ii}) = \text{diag}(a_{11}, a_{22}, \dots, a_{nn}).$$

La *trace* d'une matrice $\mathbf{A} = (a_{ij})$ est définie par

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}.$$

Soit \mathfrak{S}_n le groupe des permutations de l'ensemble $\{1, 2, \dots, n\}$. A tout élément $\sigma \in \mathfrak{S}_n$, on associe la *matrice de permutation*

$$\mathbf{P}_\sigma = (\delta_{i\sigma(j)}).$$

On notera qu'une *matrice de permutation est orthogonale*.

Le *déterminant* d'une matrice \mathbf{A} est défini par

$$\det(\mathbf{A}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon_\sigma a_{\sigma(1)1} a_{\sigma(2)2} \dots a_{\sigma(n)n},$$

où ε_σ désigne la signature de la permutation σ .

Les *valeurs propres* $\lambda_i = \lambda_i(\mathbf{A})$, $1 \leq i \leq n$, d'une matrice \mathbf{A} d'ordre n sont les n racines, réelles ou complexes, distinctes ou confondues, du *polynôme caractéristique*

$$p_{\mathbf{A}} : \lambda \in \mathbf{C} \rightarrow p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$$

de la matrice \mathbf{A} . La *spectre* de la matrice \mathbf{A} est le sous-ensemble

$$\text{sp}(\mathbf{A}) = \bigcup_{i=1}^n \{\lambda_i(\mathbf{A})\}$$

du plan complexe. On rappelle les relations

$$\begin{aligned} \text{tr}(\mathbf{A}) &= \sum_{i=1}^n \lambda_i(\mathbf{A}), & \det(\mathbf{A}) &= \prod_{i=1}^n \lambda_i(\mathbf{A}), \\ \text{tr}(\mathbf{A}\mathbf{B}) &= \text{tr}(\mathbf{B}\mathbf{A}), & \text{tr}(\mathbf{A} + \mathbf{B}) &= \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}), \\ \det(\mathbf{A}\mathbf{B}) &= \det(\mathbf{B}\mathbf{A}) = \det(\mathbf{A}) \det(\mathbf{B}). \end{aligned}$$

Le *rayon spectral* d'une matrice A est le nombre ≥ 0 défini par

$$\rho(A) = \max\{|\lambda_i(A)|; 1 \leq i \leq n\}.$$

A toute valeur propre λ d'une matrice A est associé (au moins) un vecteur p tel que

$$p \neq 0 \quad \text{et} \quad Ap = \lambda p,$$

appelé *vecteur propre* de la matrice A , *correspondant à la valeur propre* λ . Si $\lambda \in \text{sp}(A)$, le sous-espace vectoriel

$$\{v \in V; Av = \lambda v\}$$

(de dimension au moins égale à 1) est appelé *sous-espace propre, correspondant à la valeur propre* λ .

On conviendra que dans la décomposition par blocs $A = (A_{IJ})$ d'une matrice *carrée*, les *sous-matrices diagonales* A_{II} sont toujours *carrées*.

Étant donné deux espaces vectoriels de dimensions finies (mais non nécessairement égales) V et W , le *rang d'une application linéaire* $\mathcal{A} : V \rightarrow W$ est égal à la dimension du sous-espace vectoriel

$$\text{Im}(\mathcal{A}) = \{\mathcal{A}v \in W; v \in V\}.$$

Si les espaces V et W sont munis de bases, vis-à-vis desquelles l'application \mathcal{A} est représentée par une matrice A , le rang de \mathcal{A} est aussi égal au plus grand ordre des sous-matrices (carrées) inversibles de la matrice A . C'est pourquoi le rang de \mathcal{A} est aussi appelé *rang de la matrice* A . On le note $r(A)$.

Faisons enfin une remarque générale, valable pour toute la suite : *Toutes les fois que ce sera "raisonnablement" clair*, on ne mentionnera pas les ensembles d'indices. C'est ainsi que si $A = (a_{ij})$ est une matrice de type (m, n) on se contentera d'écrire

$$\max_i \{\min_j a_{ij}\}, \quad \text{au lieu de} \quad \max_{1 \leq i \leq m} \{\min_{1 \leq j \leq n} a_{ij}\};$$

c'est ainsi qu'on écrira seulement

$$p_i^* p_j = \delta_{ij}, \quad \text{au lieu de} \quad p_i^* p_j = \delta_{ij}, \quad 1 \leq i, j \leq n,$$

s'il est clair que les indices i et j décrivent le même ensemble $\{1, 2, \dots, n\}$, etc.

1.2. Réduction des matrices

Soit V un espace vectoriel de dimension finie n , et soit $\mathcal{A} : V \rightarrow V$ une application linéaire, représentée par une matrice (carrée) $A = (a_{ij})$ relativement à une base (e_i) . Relativement à une autre base (f_i) , la même application est représentée par la matrice

$$B = P^{-1}AP,$$

où P est la matrice inversible dont le j -ème vecteur colonne est formé des composantes du vecteur f_j dans la base (e_i) . La matrice P est appelée *matrice de passage, de la base* (e_i) *dans la base* (f_i) .

Une même application linéaire \mathcal{A} étant ainsi représentée par différentes matrices selon la base choisie, le problème se pose de trouver une base vis-à-vis de laquelle la matrice représentant l'application soit "aussi simple que possible". De façon équivalente, étant donné une matrice A , il s'agit de trouver parmi toutes les *matrices semblables* à la matrice A , c'est-à-dire de la forme $P^{-1}AP$, P : matrice inversible, celles qui ont une forme "aussi simple que possible" : c'est le problème de la *réduction d'une matrice*.

Le cas le plus "favorable" est celui où il existe une matrice inversible P telle que la matrice $P^{-1}AP$ soit diagonale, auquel cas on dit que la matrice A est *diagonalisable*.

On notera que, dans ce cas, les éléments diagonaux de la matrice $P^{-1}AP$ sont les valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_n$ de la matrice A , et que le j -ème vecteur colonne de la matrice P est formé des composantes (relativement à la même base que pour la matrice A) d'un vecteur propre correspondant à λ_j ; on a en effet l'équivalence

$$P^{-1}AP = \text{diag}(\lambda_i) \Leftrightarrow Ap_j = \lambda_j p_j, \quad 1 \leq j \leq n.$$

Autrement dit, une matrice est diagonalisable si, et seulement si, il existe une base de vecteurs propres.

Il existe des matrices qui ne sont pas diagonalisables (exercice 1.2-1). Pour de telles matrices, le théorème de Jordan donne la forme la plus simple des matrices semblables ; nous renvoyons le lecteur intéressé par ce résultat aux Commentaires Bibliographiques donnés en fin d'ouvrage. Pour ce qui nous concerne, le résultat qui suit, de démonstration beaucoup plus simple, suffit pour tous les besoins ultérieurs.

Nous rappelons tout d'abord les définitions suivantes : une matrice $A = (a_{ij})$ d'ordre n est triangulaire supérieure si $a_{ij} = 0$ pour $i > j$, et triangulaire inférieure si $a_{ij} = 0$ pour $i < j$. S'il n'y a pas lieu de distinguer, on dit que la matrice est triangulaire.

→ **Théorème 1.2-1. (1) Étant donné une matrice carrée A , il existe une matrice unitaire U telle que la matrice $U^{-1}AU$ soit triangulaire.**

(2) Étant donné une matrice normale A , il existe une matrice unitaire U telle que la matrice $U^{-1}AU$ soit diagonale.

(3) Étant donné une matrice symétrique A , il existe une matrice orthogonale O telle que la matrice $O^{-1}AO$ soit diagonale.

DÉMONSTRATION. (i) Démontrons d'abord la propriété (1) pour des matrices de passage non nécessairement unitaires. Cette propriété est vraie pour $n = 1$; supposons la démontrée pour les matrices d'ordre $(n - 1)$. Notons $\mathcal{A} : V \rightarrow V$ l'application linéaire associée à la matrice A . Cette application possède au moins un vecteur propre f_1 , correspondant à une valeur propre λ . Soit alors $\varepsilon_2, \dots, \varepsilon_n$ des vecteurs tels que $(f_1, \varepsilon_2, \dots, \varepsilon_n)$ soit une base de V . De la sorte,

$$\mathcal{A}f_1 = \lambda f_1, \quad \mathcal{A}\varepsilon_j = \alpha_j f_1 + \mathcal{B}\varepsilon_j, \quad 2 \leq j \leq n,$$

où \mathcal{B} est une application linéaire du sous-espace W engendré par les vecteurs $\varepsilon_2, \dots, \varepsilon_n$.

D'après l'hypothèse de récurrence, il existe une base $(f_i)_{i=2}^n$ de W , avec $f_i = \sum_{j=2}^n \gamma_{ij} \varepsilon_j$, dans laquelle l'application \mathcal{B} est représentée par une matrice triangulaire supérieure. Des égalités

$$\mathcal{A}f_1 = \lambda f_1, \quad \mathcal{A}f_i = \left(\sum_{j=2}^n \alpha_j \gamma_{ij} \right) f_1 + \mathcal{B}f_i, \quad 2 \leq i \leq n,$$

on déduit que l'application \mathcal{A} est représentée par une matrice triangulaire supérieure dans la base (f_i) .

(ii) Utilisant le procédé de Gram-Schmidt, on construit une base $(u_i)_{i=1}^n$, orthonormale, et telle que

$$u_j = \sum_{k=1}^j \gamma_{kj} f_k, \quad 1 \leq j \leq n; \quad f_i = \sum_{l=1}^i \beta_{il} u_l, \quad 1 \leq i \leq n.$$

Comme par ailleurs

$$\mathcal{A}f_j = \sum_{i=1}^j b_{ij} f_i, \quad 1 \leq j \leq n,$$

d'après (i), il s'ensuit que $\mathcal{A}u_j$ est une combinaison linéaire des vecteurs u_1, \dots, u_j , ce qui montre que l'application \mathcal{A} est encore représentée par une matrice triangulaire

supérieure par rapport à la base (u_i) . Comme la base (u_i) est orthonormale, la matrice de passage correspondante est unitaire.

(iii) Posons

$$T = (t_{ij}) = U^{-1}AU = U^*AU.$$

Si la matrice A est normale ($A^*A = AA^*$), la matrice T l'est aussi, puisque

$$T^*T = U^*A^*UU^*AU = U^*A^*AU.$$

La matrice T étant triangulaire supérieure,

$$\sum_{k=1}^n |t_{1k}|^2 = (TT^*)_{11} = (T^*T)_{11} = |t_{11}|^2, \quad \text{d'où } t_{1k} = 0, \quad 2 \leq k \leq n,$$

$$\sum_{k=2}^n |t_{2k}|^2 = (TT^*)_{22} = (T^*T)_{22} = |t_{22}|^2, \quad \text{d'où } t_{2k} = 0, \quad 3 \leq k \leq n,$$

etc., ce qui montre que la matrice T est diagonale.

(iv) Si la matrice A est symétrique, le vecteur propre f_1 et la valeur propre λ de (i) sont réels, et les raisonnements précédents sont encore valables, en remplaçant partout "unitaire" par "orthogonal" et "matrice adjointe" par "matrice transposée". ■

REMARQUES. (1) Les matrices de passage vérifiant les conditions de l'énoncé ne sont pas uniques (considérer par exemple $A = I$).

(2) Les éléments diagonaux de la matrice triangulaire $U^{-1}AU$ de (1), ou de la matrice diagonale $U^{-1}AU$ de (2), ou de la matrice diagonale de (3), sont les valeurs propres de la matrice A . En conséquence, ce sont des nombres réels si A est une matrice hermitienne ou symétrique, et des nombres complexes de module 1 si la matrice A est unitaire ou orthogonale.

(3) Il résulte de (2) que toute matrice hermitienne ou unitaire est diagonalisable par une matrice de passage unitaire.

(4) Si O est une matrice orthogonale, le raisonnement précédent montre l'existence d'une matrice unitaire U telle que la matrice $D = U^*OU$ soit diagonale (les éléments diagonaux de D étant de module 1), mais la matrice U n'est pas en général réelle, c'est-à-dire orthogonale. On trouvera des indications à ce sujet à l'exercice 1.2-2. ■

On appelle *valeurs singulières* d'une matrice A carrée les racines carrées positives des valeurs propres de la matrice hermitienne A^*A (ou $A^T A$ si la matrice A est réelle). Ces dernières sont toujours ≥ 0 , puisque de la relation $A^*Ap = \lambda p$, $p \neq 0$, on déduit $(Ap)^*Ap = \lambda p^*p$. On notera également que *les valeurs singulières sont toutes > 0 si et seulement si la matrice A est inversible*. En effet

$$Ap = 0 \Rightarrow A^*Ap = 0 \Rightarrow p^*A^*Ap = (Ap)^*Ap = 0 \Rightarrow Ap = 0.$$

Deux matrices A et B de type (m, n) sont dites *équivalentes* s'il existe une matrice inversible Q d'ordre m et une matrice inversible P d'ordre n telles que

$$B = QAP.$$

Z Naturellement, il s'agit d'une notion *plus générale* que celle de la similitude des matrices. On peut d'ailleurs démontrer que toute matrice carrée est équivalente à une matrice diagonale :

Théorème 1.2-2. *Si A est une matrice réelle carrée, il existe deux matrices orthogonales U et V telles que*

$$U^TAV = \text{diag}(\mu_i),$$

et si A est une matrice complexe carrée, il existe deux matrices unitaires U et V telles que

$$U^*AV = \text{diag}(\mu_i).$$

Dans les deux cas, les nombres $\mu_i \geq 0$ sont les valeurs singulières de la matrice A .

DÉMONSTRATION. Pour fixer les idées, supposons la matrice A complexe. D'après le théorème 1.2-1, il existe une matrice unitaire V telle que

$$V^*A^*AV = \text{diag}(\mu_i^2),$$

les nombres $\mu_i \geq 0$ étant les valeurs singulières de la matrice A . Notant f_j le j -ème vecteur colonne de la matrice AV , cette égalité matricielle s'écrit encore

$$f_i^*f_j = \mu_i^2\delta_{ij}, \quad 1 \leq i, j \leq n.$$

Soit $\{\mu_1, \mu_2, \dots, \mu_r\}$ l'ensemble (peut-être vide !) des valeurs singulières nulles ; par conséquent,

$$f_j = 0, \quad 1 \leq i \leq r.$$

Posons

$$u_j = \mu_j^{-1}f_j, \quad r+1 \leq j \leq n,$$

de sorte que l'on a déjà

$$u_i^*u_j = \delta_{ij}, \quad r+1 \leq i, j \leq n.$$

Ensuite, on "complète" par des vecteurs u_i , $1 \leq i \leq r$, tels que

$$u_i^*u_j = \delta_{ij}, \quad 1 \leq i, j \leq n.$$

Alors la matrice U de j -ème vecteur colonne u_j répond à la question : d'une part, la relation ci-dessus montre que c'est une matrice unitaire et, d'autre part,

$$(U^*AV)_{ij} = u_i^*f_j = \begin{cases} 0 = \mu_i\delta_{ij}, & 1 \leq j \leq r, \\ \mu_j u_i^*u_j = \mu_i\delta_{ij}, & r+1 \leq j \leq n. \end{cases}$$

La démonstration est analogue si la matrice A est réelle. ■

1.3. Propriétés particulières aux matrices symétriques et hermitiennes

Pour fixer les idées, nous allons considérer dans ce qui suit le cas des matrices hermitiennes, mais il est entendu que tout le contenu de ce paragraphe s'applique aussi bien au cas des matrices symétriques, en remplaçant partout "hermitien", "unitaire", "complexe", "matrice adjointe" par "symétrique", "orthogonal", "réel", "matrice transposée", respectivement.

Rappelons que toutes les valeurs propres d'une matrice hermitienne sont réelles, et que toute matrice hermitienne est diagonalisable, la matrice de passage étant unitaire (théorème 1.2-1). Il existe, de surcroît, diverses caractérisations remarquables des valeurs propres d'une matrice hermitienne, qui font l'objet du théorème 1.3-1 ci-dessous. Pour les énoncer, il nous faut tout d'abord une définition.

Soit A une matrice carrée représentant une application linéaire d'un espace V sur le corps C , muni de son produit canonique. Le quotient de Rayleigh de la matrice A est l'application

$$R_A : V - \{0\} \rightarrow C$$

définie par

$$R_A(v) = \frac{(Av, v)}{(v, v)} = \frac{v^*Av}{v^*v}, \quad v \neq 0.$$