UNIVERSITE PAUL SABATIER

Notes du cours de C.A. ROCHE

(D'après le cours de D. AZE) Août 2009

1. Préliminaires algébriques

1.1. **Introduction.** Ce cours est un cours d'algèbre linéaire, avec la mise en place d'algorithmes efficaces pour la résolutions de systèmes linéaires. Il est orienté vers le calcul numérique, le calcul scientifique. Il fait suite à un cours formel d'algèbre linéaire. Ainsi l'on s'attachera essentiellement à développer la version matricielle des concepts de l'algèbre linéaire, traités en L1 et L2. Il suppose donc, une connaissance de l'algèbre linéaire jusqu'à la décomposition de Dunford, avec une sommaire description des formes normales de Jordan.

Commentaire bibliographique.

Les références pour ce cours abondent. Nous retiendrons pour les étudiants désirant se préparer à l'AGREG les libres de D. Serre [1] et celui de Luca Amodei et Jean-Pierre Dedieu [3]. Pour d'autres points de vue : Ph. Ciarlet [2][4] où [5]. Le classique et indémodable Gantmacher [6] ou celui de N. Gastinel [7].

Il existe plusieurs textes en ligne qui sont très adaptés à notre programme. Par exemple celui du parisien Yves Achdou [9] et pour certaines questions dont on n'aura que peu de temps à consacrer, lire le Montpellierin Michel Cuer [8].

N'oublions pas qu'il est exigible d'acquérir un peu d'aisance dans l'utilisation de MATLAB. On pourra consulter à ce sujet [10]. Il est utile d'installer dans son ordinateur personnel une version de SCILAB (à partir de par exemple [11]), libre de droits, et avoir une notice d'utilisation, par exemple [12] (malheureusement trop axé sur le traitement du signal, à mon avis).

Problématique : Tout d'abord : Résoudre le système : $AX = B$ où X, Y sont des vecteurs de $\mathbb{R}^n$ et A une matrice carrée réelle.

Méthode naïve : calcul de A^{-1}. Elle marche si A est inversible. Or en calcul numérique il y a deux problèmes additionnels d'importance capitale : La matrice A doit être considérée comme étant une représentation approchée des données du problème qui nous occupe réellement. D'une part dû aux erreurs de mesure du scientifique qui nous fournis les données, et d'autre part dû à la représentation en virgule flottante (simple ou double précision) en machine.

Ensuite, toute méthode effective de calcul, doit être mesurée au regard de sa complexité, du nombre d'opérations à effectuer. En effet, il n'est pas possible d'attendre indéfiniment les résultats d'un calcul, par exemple lors de l'utilisation dans un robot industriel, où les résultats doivent être disponibles "en temps réel", disons en quelques dixièmes de seconde. De plus, dû aux arrondis, le nombre d'opérations à effectuer rends le résultat de plus en plus éloigné du vrai résultat à fur et à mesure que les étapes de calcul s'empilent. Ainsi le calcul de A^{-1} selon la structure de la matrice A peut être la mauvaise méthode... sauf... sauf si par exemple, l'on doit résoudre le problème $AX_n = B_n$ toujours avec la même matrice A mais pour une quantité très très importante de deuxièmes membres B_n.

Associé à une matrice, il y a plusieurs problèmes où l'analyse numérique peut apporter des réponses d'intérêt. Par exemple, le problème de la recherche de valeurs propres.

L'exemple le plus évident d'intérêt de la détermination des valeurs propres d'une matrice est celui de l'étude asymptotique des suites dites définies par une récurrence linéaire. Ces modèles sont courants dans les approximations linéaires des études de plusieurs populations mises en concurrence dans le même biotope.

Si le nombre d'individus de trois populations sont x_1, y_1 et z_1 et l'on sait que (en première approximation) $x_n = a_1x_{n-1} + a_2y_{n-1} + a_3z_{n-1}, y_n = b_1x_{n-1} + b_2y_{n-1} + b_3z_{n-1}; z_n = c_1x_{n-1} +$

$c_2y_{n-1} + c_3z_{n-1}$. Nous nous intéressons évidement à la version matricielle :

$$AX_{n+1} = X_n, \quad A = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \quad X_n = \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$$

avec x_0, y_0, z_0 les tailles des populations au départ. Evidement $X_n = A^n X_0$ mais connaître les tailles des populations après 10000 générations, dû aux erreur d'arrondi, ne doit pas être effectué par le simple calcul de A^{10000} en principe trivial ! Il est souvent meilleur d'étudier la structure spectrale de A. Car si par exemple A est diagonalisable, $A = P^{-1}\Delta A$ avec Δ une diagonale. Puisque $A^n = P^{-1}\Delta^n A$ nous en tirons des conséquences claires, dans certains cas. Par exemple si toutes les valeurs propres de A sont en module inférieures strictement à 1, asymptotiquement, toutes les populations étudiées disparaissent ! Si exactement une des valeurs propres est de module strictement supérieur à 1, la population totale atteint, *in fine*, une taille supérieure au nombre des particules de l'univers !

Nous faire penser que des modèles linéaires de la sorte, puissent être applicables aux populations humaines, et disons, quelques autres "populations" telles que la disponibilité des resources naturelles (pétrole, minéraux, nourriture, eau, air respirable) et déduire que l'on cours à la catastrophe s'appelle "faire une analyse Malthusienne" de l'écologie de la planète. L'exemple de base est le *Rapport Meadows*, commandé par le Club De Rome, dans les années 60, années fastes de l'analyse matricielle. Le rapport entre le comportement asymptotique d'une approximation linéaire, et l'asymptotique réelle d'une variable est toujours un problème scientifique majeur. Il n'est envisageable que lorsque l'on dispose des "estimations à priori" sur les non linéairités. Nous n'aborderons pas ces aspects de la modélisation, nous nous en tiendront à l'analyse matricielle, même si comprendre ses limites est le premier pas du réalisme dans le génie mathématique. Quelques aspects de ces problématiques seront abordés en deuxième semestre, dans le cours de Calcul Differentiel, et celui d'Equations Différentielles.

1.2. **Rappel sur les matrices.** Dans la suite $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.

Si $\mathbf{A}$ est un anneau, on notera $M_{m,n}(\mathbf{A})$ l'ensemble des matrices à m lignes et n colonnes à valeurs dans A; si $n = m$, $M_n(\mathbf{A}) = M_{m,n}(\mathbf{A})$. L'anneau $\mathbf{A}$ sera dit ensemble de scalaires, souvent ce sera le corps $\mathbb{K}$ mais il peut aussi être l'ensembles des matrices carrées $M_n(\mathbb{K})$ pour raisonner "par blocs". Il sera important de toujours écrire les éléments des matrices avec la convention LICO, indice de ligne – indice de colonne.

$$A = (a_{ij})_{i\ j} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in M_{m,n}(\mathbf{A})$$

avec i l'indice de ligne, j l'indice de colonne, et les $a_{ij} \in \mathbf{A}$ dits "ses éléments".

Nous ne rappellerons pas les opérations matricielles : addition de matrices, produit matriciel. Nous nous contenterons d'attirer l'attention sur les conditions de cohérence. En effet, pour que la somme ou le produit puisse être effectué, il faut que les tailles soient "compatibles, sinon, l'opération n'est pas définie. Ainsi, dans ce texte, l'affirmation $AB = CD + E$ voudra dire, "les tailles des matrices concernées sont telles que les opérations sont définies, et de plus le résultats des calculs donnent l'égalité signalée". Ces tailles sont :

$$M_{m,n}(\mathbf{A}) \times M_{s,t}(\mathbf{A}) \to M_{l,r}(\mathbf{A}) : (A, B) \mapsto A + B \quad \textbf{si} \quad (m = s = l, n = t = r);$$
$$M_{m,n}(\mathbf{A}) \times M_{s,t}(\mathbf{A}) \to M_{l,r}(\mathbf{A}) : (A, B) \mapsto AB \quad \textbf{si} \quad (n = s, l = m, r = t).$$

Puisque l'on définit (toujours) le produit d'un élément de $\mathbf{A}$ par une matrice, nous avons une structure de $\mathbb{K}$-espace vectoriel sur chaque $M_{m,n}(\mathbb{K})$. Cette structure sur $M_{m,n}(\mathbf{A})$ s'appelle "structure de $\mathbf{A}$-module" lorsque $\mathbf{A}$ n'est pas un corps commutatif. De plus, remarquons que l'on peut représenter la multiplication par un scalaire par le produit par la matrice diagonale, où tous les éléments de la diagonale sont égaux au scalaire, et nous dirons que nous avons une structure de $\mathbf{A}$-algèbre. Nous rappelons que le produit matriciel n'est pas commutatif, à partir de la taille 2.

Les matrices élémentaires suivantes, nous donnent une base d'espace vectoriel :

$$\mathbb{E}_{ij} = (\delta_{ls}^{ij}) \quad \delta_{ls}^{ij} = 1 \Leftrightarrow (l = i \textbf{ et } s = j) \text{ 0 sinon.}$$

La quantité de zéros à ajouter, est à déterminer d'après le contexte, de sorte que les calculs soient cohérents. En clair, nous utiliserons cette notation, même si la taille de la matrice n'est pas indiquée, tout à plus nous savons que le nombre de colonnes dépasse j et le nombre de lignes i.

Vecteurs colonne. Nous considérons les éléments de $\mathbf{A}^m$ comme "des colonnes", ainsi, il y a "égalité" entre $\mathbf{A}^m$ et $M_{m,1}(\mathbf{A})$ malgrès la représentation habituelle, qui nous fait penser aux éléments d'un produit cartesien comme des lignes. Nous aurons souvent usage de la transposition $X \in \mathbb{K}^m \mapsto x = X^T \in M_{1,m}(\mathbb{K})$ ainsi un élément de $\mathbb{K}^m$ est en correspondance biunivoque (et linéaire) avec la matrice ligne de ses coordonnées :

$$x = (x_1, x_2, \ldots, x_m) \leftrightarrow X = x^T = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \sum_{i=1}^{m} x_i \mathbb{E}_{i1}.$$

Ceci, en plus de correspondre à une pratique dans la litérature, nous permets d'épargner de l'espace dans un texte, par exemple.

Pour chaque matrice $A \in M_{m,n}(\mathbb{K})$ nous associons, canoniquement une application linéaire (toujours notée avec la minuscule correspondante) :

$$M_{m,n}(\mathbb{K}) \ni A \mapsto a \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) : X = (x_1, x_2, \ldots, x_n)^T \mapsto a(X) = AX.$$

Ces définitons marchent parfaitement même avec des matrices à coefficients dans un anneau (unitaire) mais nous n'aurons guerre l'usage de la théorie des modules (abordée en mastère) sauf en ce que l'on reconnaîtra les caractères propres à la linéarité. Mais sans doute le premier problème sérieux provient de la réciproque d'une application linéaire.

Une matrice A est dite inversible, si c'est une matrice carrée et s'il existe une deuxième matrice B telle que $AB = BA = I$, la matrice identité. Dans ce cas, B est notée A^{-1} et elle est unique. Le déterminant, fonction polynomiale de ces éléments, permet de déterminer si une matrice carrée est inversible ou non *dans le cas d'un corps*; A est inversible si et seulement si $\det(A) \neq 0$. Pour un anneau commutatif quelconque il faudra demander que le déterminant soit inversible dans l'anneau. Une matrice qui n'est pas inversible, est dite singulière.

Nous n'aurons pas l'opportunité de faire usage de cette remarque, sauf dans le cas des matrices à coefficients entiers relatifs, avec un déterminant 1 ou -1. En effet, les matrices "plus générales" qui nous intéressent, sont à coefficients matrices carrées, et dans ce cas, la commutativité fait défaut, ce qui est plus grave pour la formule de Cramer.

Remarquons $(AB)^{-1} = B^{-1}A^{-1}$ et que $\det(AB) = \det(A)\det(B)$.

Lorsque une matrice est inversible, l'application linéaire canoniquement associée est un isomorphisme linéaire, dont la réciproque est canoniquement associé à la matrice inverse.

Il est utile de relier,

Transposition. Notation : Si $A = (a_{ji}) \in M_{m,n}(A)$, ${}^t\!A = A^T$ est la matrice dont l'élément en ligne i et colonne j, noté a_{ij}^T est le scalaire a_{ji}.

Les deux formules suivantes sont élémentaires, mais très importantes ! $(A + B)^T = A^T + B^T, \quad (AB)^T = B^T A^T$.

Trace d'une matrice. Pour chaque matrice $A \in M_{m,n}(\mathbf{A})$ nous définisons sa trace, comme la somme des coefficients dans la diagonale.

$$\operatorname{tr}(A) = \sum_{i=1}^{\min(n,m)} a_{ii}$$

où $\min(n, m)$ est le minimum de n et m et c'est le nombre d'éléments de la diagonale.

Les matrices triangulaires supérieures se définissent sans peine, en disant que c'est une matrice carrée $A = (a_{ij})$ qui vérifie : si $j < i \Rightarrow a_{ij} = 0$. Tout ce qui est strictement en-dessous de la diagonale est nul. Il s'agit d'un sous-espace vectoriel de l'espace de toutes les matrices carrées, en fait, puisque le produit de deux matrices triangulaires supérieurs est encore triangulaire supérieur, il s'agit d'une sous-algèbre.

Les matrices triangulaires inférieures sont les transposées des matrices triangulaires supérieures. Il s'agit encore, d'une sous algèbre. Nous aurons peut-être l'opportunité d'utiliser des matrices strictement triangulaires, ce sont celles où en plus la diagonale est nulle.

Posons une fois pour toutes la notation suivante, si X est, soit un vecteur (colonne ou ligne) ou une liste ordonnée de quantités, $X = (d_1, d_2, \dots, d_n)$,

$$diag(X) = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{pmatrix} \in M_{n,n}(\mathbf{A})$$

matrice diagonale, avec sur sa diagonale les quantités données.

1.3. **Opérations par blocs.** Il y a une structure, très difficile à décrire formellement, mais d'usage constant et indispessable pour ce cours, c'est les décompositions par blocs des matrices, et les opérations que l'on éffectue selon ces décomposition. L'idée est simple, nous rassemblons les éléments d'une matrice, en petites sous-matrices, ainsi une matrice est considérée comme "matrice de matrices". Nous donnerons juste un exemple sans justification. Si

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} B & Y \\ W^T & a \end{pmatrix} \quad B = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},\ Y = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}, W = \begin{pmatrix} a_{31} \\ a_{32} \end{pmatrix},\ a = a_{33}$$

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} = \begin{pmatrix} D & Z \\ U^T & c \end{pmatrix} \quad D = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix},\ Z = \begin{pmatrix} c_{13} \\ c_{23} \end{pmatrix}, U = \begin{pmatrix} c_{31} \\ c_{32} \end{pmatrix},\ c = c_{33}$$

Alors le produit matriciel se calcule comme suit :

$$AC = \begin{pmatrix} B & Y \\ W^T & a \end{pmatrix} \begin{pmatrix} D & Z \\ U^T & c \end{pmatrix} = \begin{pmatrix} BD + YU^T & BZ + cY \\ W^T D + aU^T & W^T Z + ac \end{pmatrix}.$$

Il est très important de ne pas faire commuter les facteurs, qui maintenant sont matriciels. Pour effectuer un produit, la partition des colonnes du premier facteur doit être la même que la partition des lignes du deuxième facteur. Ici, nous avons vu les colonnes partitionnées en $(2, 1)$ comme pour les lignes du deuxième facteur. Bien d'autres décompositions en blocs sont d'usage courrant, on attire l'attention en particulier sur les matrices diagonales par blocs.

En exercice nous traiterons la formule du complément de Schur, qui dit que si $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ est une matrice carrée, avec A carrée inversible, alors $\det(M) = \det(A)\det(D - CA^{-1}B)$.

1.4. **Structures additionnelles.** Notion de produit scalaire sur un $\mathbb{K}$-vectoriel. Soit E un $\mathbb{K}$-espace vectoriel où $\mathbb{K}$ est $\mathbb{R}$ ou $\mathbb{C}$ (dans la cas de $\mathbb{C}$ on dira produit hermitien). Une application

$$< \cdot, \cdot >: E \times E \to \mathbb{K}$$

est un produit scalaire (hermitien) si elle vérifie

1- $\forall Y \in E \ \ < \cdot, Y >$ est $\mathbb{K}$-linéaire ;

2- $\forall X, Y \in E, \ < X, Y >= \overline{< Y, X >}$.

3- $\forall X \in E, \ < X, X > \geq 0$ et $< X, X >= 0 \Rightarrow X = 0$.

Il s'en suivent les notions usuelles : $X \perp Y \Leftrightarrow < X, Y >= 0$; et $X^{\perp}$ le sous-espace des vecteurs perpendiculaires (au sens hermitien) au vecteur X.

Exemples standards. La dualité $* : \mathbb{K}^n \to \mathbb{K}^n$ nous donne le produit standard :

$$< X, Y >:= Y^* X.$$

Espace de Hilbert. Nous avons que $< X, X >$ est un nombre réel positif, dès lors, $||X|| = \sqrt{< X, X >}$ nous donne une application $|| \cdot || : E \to \mathbb{R}$. C'est une norme. La seule chose à vérifier est l'inégalité de Cauchy-Schwarz. L'espace vectoriel normé qui en résulte est, en dimension finie, complet, dans le sens que toute suite de vecteurs de Cauchy est convergente. On dira que c'est un espace de Hilbert.

L'on définit les basses orthonormales. Le théorème de la base incomplète est ici amélioré par :

Théorème 1 (Procédé de Gram-Schmidt)**.** *Soit E un espace de Hilbert de dimension finie, $Y_1, Y_2, \ldots, Y_n$ une base indicée, il existe une base orthonormale de E, $(X_1, X_2, \ldots, X_n)$ telle que pour tout $0 \leq i \leq n$, $vec(Y_1, Y_2, \ldots, Y_i) = vec(X_1, X_2, \ldots, X_i)$. Les vecteurs s'obtiennent selon l'algorithme :*

$$U_1 = Y_1; \quad \alpha_1 = ||U_1||; \quad X_1 = \frac{1}{\alpha_1} U_1.$$

$$U_2 = Y_2 - < Y_2, X_1 > X_1 \qquad (\Rightarrow U_2 \perp X_1)$$

$$\alpha_2 = ||U_2||; \quad X_2 = \frac{1}{\alpha_2} U_2 \qquad (\Rightarrow ||X_2|| = 1)$$

$$\cdots$$

$$U_i = Y_i - \sum_{l=1}^{i-1} < Y_2, X_l > X_l \qquad (\Rightarrow U_2 \perp X_l, \ 1 \leq l < i)$$

$$\alpha_i = ||U_i||; \quad X_i = \frac{1}{\alpha_i} U_i \qquad (\Rightarrow ||X_i|| = 1).$$

Il est important de remarquer ici, que l'on a utilisé la norme associée à la structure hilbertienne. Dans les cas standard, $\mathbb{C}^n$ c'est la racine carrée de la somme des carrés des modules des composantes. Pour un espace vectoriel normé arbitraire, ce résultat n'existe pas : l'orthogonalité des vecteurs n'est même pas définie !

Matrice adjointe : la matrice adjointe de A est $A^* = \bar{A}^T$.

Une matrice A est symétrique si $A = A^T$.

Elle est hermitienne si $A = A^*$.

Les matrices hermitiennes à coefficients réels sont les matrices réelles symétriques. Souvent on dira "symétrique" pour symétrique réelle, en délaissant la notion de symétrie dans le cas complexe.

Elle est orthogonale, si c'est une matrice réelle et $A^T = A^{-1}$. On en déduit que c'est une matrice carrée.

La matrice $A \in M_n(\mathbb{C})$ est unitaire si $A^* = A^{-1}$.

normale si elle commute avec son adjointe (c'est-à-dire $AA^* = A^*A$).

orthogonale $\Rightarrow$ unitaire $\Rightarrow$ normale.

symétrique réelle $\Rightarrow$ hermitienne $\Rightarrow$ normale.

Spectre. Définition . Un scalaire λ est dit valeur spectrale de la matrice carrée A, de taille n, et l'on écrit $\lambda \in sp(A)$, si la matrice $A-\lambda I$ est singulière. Leur ensemble est noté $sp(A)$, $\quad sp(A) \subset \mathbb{C}$. Il faudra prendre garde, de toujours penser pour les problèmes spectraux, les matrices réelles, comme des matrices complexes.

Bien entendu, en dimension finie, $A-\lambda I$ singulière, équivaut à $\det(A-\lambda I)=0$. Or $\det(A-\lambda I)$ est un polynôme de degré n en λ, le polynôme caractéristique de A, Ses racines sont exactement les valeurs spectrales. Puisque l'endomorphisme associé n'est alors pas injectif, on a que chaque valeur spectrale est une valeur propre.

Plus précisement, qu'il existe un vecteur **non nul**, $X \in \mathbb{C}^n \setminus \{0\}$ dans ce noyau, c'est-à-dire solution de $AX=\lambda X$, il est dit **vecteur propre** associé à λ. L'ensemble de tels vecteurs est dit, l'**espace propre associé**. Sa dimension est la **multiplicité géométrique** de la valeur propre de A. L'expression $P(\lambda)=\det(A-\lambda I)$ **polynôme caractéristique** de A, de degré n, a valeur en 0 égale au déterminant de A. Le corps $\mathbb{C}$ étant algébriquement clos, le polynôme caractéristique d'une matrice numérique, admet n racines **comptées selon leur multiplicité**. Ce sont les valeurs propres de la matrice. Leurs multiplicité, en tant que racine de cette équation polynomiale, est dite la multiplicité (algébrique) de la valeur propre de A.

Le spectre est (dans le cas de la dimension finie) exactement l'ensemble des valeurs propres. Notons, une fois pour toutes $\rho(A) := \max\{|\lambda| / \ \lambda \in sp(A)\}$ dit **rayon spectral** de la matrice. Remarquons, (Exercice), que si $Q(t)$ est un polynôme complexe quelconque, pour chaque $\lambda \in sp(A)$, $Q(\lambda) \in sp(Q(A))$.

LEMME

Les valeurs propres d'une **matrice triangulaire** sont les éléments de sa diagonale. En particulier c'est la cas pour les matrices diagonales.

Toutes les valeurs propres d'une **matrice hermitienne** sont **réelles**.

Toutes les valeurs propres d'une matrice **symétrique réelle** sont réelles.

Toutes les valeurs propres d'une matrice unitaire sont de module 1.

Dém. Pour la première affirmation, il suffit de remarquer que la matrice $A-\lambda I$ est aussi triangulaire puis calculer le déterminant $\det(A-\lambda I)$ en développant par colonnes, ce qui donne, le produit des $a_{ii}-\lambda$. C'est bien un polynôme en λ dont les racines sont les a_{ii}.

Pour le deuxième point, si M est hermitienne ($M=M^*$) et $\lambda \in \mathbb{C}$ est une de ses valeurs propres, il existe $X \in \mathbb{C}^n \setminus \{0\}$ tel que $MX=\lambda X$. Alors $\lambda < X,X >=< MX,X >=< X,M^*X >=< X,MX >=< X,\lambda X >= \bar{\lambda} < X,X >$ et $\lambda=\bar{\lambda}$ qui est donc réel. Enfin un calcul similaire donne que les valeurs propres d'une matrice anti-hermitienne sont imaginaires purs.

Les matrices symétriques réelles, étant juste les matrices hermitiennes d'éléments réels, elles ont aussi ses valeurs propres toutes réelles. Il est important de comprendre que ceci veut dire que : puisque, calculés dans le complexe, nous avons un système complet de valeurs propres, et comme par ailleurs ces racines sont nécessairement réelles, l'ensemble des valeurs propres réelles d'une matrice symétrique réelle à pour somme de ses multiplicités algébriques le maximum possible, c'est-à-dire la taille de la matrice.

Nous verrons que les matrices hermitiennes, en particullier symétriques réelles sont diagonalisables.

Montrons que les valeurs propres d'une matrice unitaire U sont de module 1. En effet, s'il existe $X \in \mathbb{C}^n \setminus \{0\}$ tel que pour $\lambda \in \mathbb{C}$, $UX=\lambda X$, alors $X=\lambda U^{-1}X=\lambda U^*X \ : U^*X=\lambda^{-1}X$. Mais aussi $\lambda < X,X >=< UX,X >=< X,U^*X >=< X,\lambda^{-1}X >= \bar{\lambda}^{-1} < X,X >$. Ainsi $\lambda\bar{\lambda} < X,X >=< X,X >$, et comme $||X||^2 \neq 0$, $\lambda\bar{\lambda}=|\lambda|=1$. Il en va évidement de même pour les matrices orthogonales.

Soit μ un nombre complexe de module 1. Posons $\mu = a + ib$ avec $a, b \in \mathbb{R}$. La matrice

$$M_\mu = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

est orthogonale de valeurs propres $\mu, \bar{\mu}$. Il en va de même de la matrice diagonale par blocs $dia(I, M)$; qui a taille n si I est l'identité en taille $n-2$. Pour chaque taille ≥ 2, **tout nombre complexe de module 1 est la valeur propre d'une matrice orthogonale**, de cette taille.

Exercice : Si $A \in M_n(\mathbb{C})$, AA^* est hermitienne, toutes ses valeurs propres sont réelles positives ou nulles. Solution :

Puisque $(A^*A)^* = A^*A^{**} = A^*A$ c'est une matrice hermitienne de $M_m(\mathbb{C})$. Montrons que les valeurs propres de A^*A qui sont donc réelles, sont de plus positives ou nulles. Soit $\lambda \in sp(A^*A)$ et $V \in \mathbb{C}^m$ un vecteur propre associé à λ. On a $\lambda ||V||^2 = \lambda V^*V = V^*(\lambda V) = V^*(A^*AV) = (AV)^*AV = ||AV||^2$, donc $\lambda = \frac{||AV||^2}{||V||^2} \geq 0$. Ici, l'on utilise de manière essentielle que $V \neq 0$.

Notons $U_n(\mathbb{C})$ l'ensemble des matrices carrées complexes de taille n qui sont unitaires.

Théorème de Schur Soit $M \in M_n(\mathbb{C})$, il existe $U \in U_n(\mathbb{C})$ telle que U^*MU soit triangulaire supérieure.

Dém. Soit $\lambda \in \mathbb{C}$ une valeur propre de M, et $X \in \mathbb{C}^n \setminus \{0\}$ un vecteur propre associé. Le vecteur $X_1 = \frac{1}{||X||}X$ est de norme 1 et peut être complété par des vecteurs $X_2, \ldots, X_n$ en une base orthonormale de $\mathbb{C}^n$. (Gram-Schmidt). Formons la matrice W dont les colonnes sont les vecteurs $X_1, X_2, \ldots, X_n$: $W = (X_1\ X_2\ \cdots\ X_n)$. Elle est unitaire, son inverse est W^*. Calculons la matrice de l'endomorphisme canoniquement associé à M dans la base orthonormale ci-dessus, c'est-à-dire W^*MW. Notons ses colonnes $V_1, \ldots, V_n$.

Sa première colonne est $V_1 = W^*MWX_1 = \frac{1}{||X||}W^*MX = \frac{1}{||X||}W^*\lambda X = \frac{1}{||X||}\lambda W^*X = \lambda W^*X_1 = \lambda$. Pour les autres, nous n'avons, pour l'instant, aucun renseignement. Ainsi :

$$M' = W^*MW = \begin{pmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{pmatrix}$$

Où la matrice N est une matrice carrée complexe de taille $n-1$. Une récurrence se dessine, puisque en taille $n = 1$ le résultat est trivial.

Il existe, par récurrence, une matrice unitaire $\tilde{W}$ de taille $n-1$ telle que $N' = (\tilde{W})^*N\tilde{W}$ est triangulaire supérieure. Notons V la matrice diagonale par blocs $V = dia(1, \tilde{W})$:

$$V = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & \tilde{W} & \\ 0 & & & \end{pmatrix}$$

visiblement $V^* = dia(1, \tilde{W}^*)$ et un calcul par blocs donne :

$$V^*M'V = \begin{pmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & N' & \\ 0 & & & \end{pmatrix}.$$

Or, puisque la matrice N' est triangulaire supérieur, on a fini, $V^*MV = (WV)^*MWV$ est triangulaire supérieur. ($U = WV$ est unitaire, comme produit de deux matrices unitaires).

Le théorème suivant n'est qu'une version amélliorée du précèdent, une sorte de théorème de Schur tordu, qui ne marche que pour les matrices normales. Ce qui est déjà beaucoup ! Ce sera notre Théorème Fondamental.

Théorème 2 (Décomposition Spectrale des Matrices NORMALES)**.** *Soit $M \in M_n(\mathbb{C})$, vérifiant $M^*M = MM^*$ (M normale) alors il existe $U \in M_n(\mathbb{C})$ vérifiant $U^* = U^{-1}$ (U est unitaire) et des scalaires $d_1, d_2, \ldots, d_n \in \mathbb{C}$ tels que*

$$M = U^{-1}\text{diag}(d_1, d_2, \ldots, d_n)U.$$

Nous pouvons écrire aussi : $M = U^*diag(d_1, d_2, \ldots, d_n)U$. En clair : "Toute matrice normale est unitairement semblable à une matrice diagonale." Comme toute matrice unitairement semblable à une matrice diagonale est normale, ce théorème caractérise les matrices normales. Les d_i sont évidement les valeurs propres. Les matrices unitaires, hermitiennes ou anti-hermitiennes sont normales, elles se diagonalisant donc par un changement de base unitaire. Ces familles de matrices se distinguent d'après la nature de leur valeurs propres : les matrices unitaires ont leur valeurs propres de module 1; les matrices hermitiennes ont leurs valeurs propres réelles et enfin, les anti-hermitiennes ont leurs valeurs propres imaginaires pures.

Enfin, remarquons que la même matrice de passage diagonalise l'adjointe, trouvant, naturellement les valeurs propres conjuguées (en transposant-conjugant) :

$$M^* = U^{-1}\text{diag}(\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_n)U.$$

La preuve ci-dessous, si bien est algorithmique, ne donne pas la meilleure manière de procéder du point de vue numérique. Ceci pourra être fait (et nous le feront) dans certains cas précis.

Attention : ce résultat est faux en calculant en réel.

Démonstration. Nous procédons exactement comme pour démontrer le Théorème de Schur. Et rapidement l'on voit qu'est-ce qu'il faut changer pour aboutir à une diagonalisation plutôt qu'une simple triangulation. De plus on verra que l'hypothèse (M normale) est bien adaptée. Ainsi, l'on prend un vecteur propre X de M, on le norme en X_1 puis on le complète en une base orthonormée, donnant les colonnes de notre premier changement de base : $W = (X_1\ X_2\ \cdots\ X_n)$; $\quad M' = W^*MW$. Nous avons que M' a la même forme que dans la preuve du Théorème de Schur, triangulaire par blocs, un (petit) bloc de taille 1 donné par la valeur propre associée au vecteur X_1 puis le deuxième bloc, fait d'une ligne inconnue, et la matrice carrée $(n-1) \times (n-1)$ qui se prêterai bien à une récurrence. Tout d'abord, si l'on veut procéder comme ça pour une diagonalisation, la ligne inconnue devra s'annuller. Or, calculons W^*M^*W qui est $(W^*MW)^*$:

$$(M')^* = W^*M^*W = \begin{pmatrix} \lambda & * & \cdots & * \\ 0 & & & \\ & & & \\ \vdots & & N & \\ & & & \\ 0 & & & \end{pmatrix}^* = \begin{pmatrix} \bar{\lambda} & 0 & \cdots & 0 \\ \bar{*} & & & \\ & & & \\ \vdots & & N^* & \\ & & & \\ \bar{*} & & & \end{pmatrix}.$$

Si l'on veut procéder ainsi par récurrence, c'est-à-dire obtenir que les $*$ soient des zéros, la première colonne de W doit être aussi vecteur propre de M^*. C'est exactement ce qu'il faut obtenir, le reste est identique au cas "super général' de Schur.

Un petit Lemme, bien connu nous fournit le moyen d'achèver.

Lemme 1 (Spectre de matrices commutantes)**.** *Soient $A, B \in M_n(\mathbb{C})$, vérifiant $AB = BA$ alors A et B ont un vecteur propre commun.*

Deux endomorphismes qui commutent dans un espace vectoriel complexe de dimension finie, ont un vecteur propre commun.

Il est important ici, que l'on travaille sur $\mathbb{C}$, ce qui nous assure que l'ensemble des valeurs propres dans le corps des coefficients utilisé, est non vide.

Démonstration. Il est plus aisé de fournir une preuve de la deuxième assertion. On procède par récurrence sur la dimension, en dimension 1, puisque tout vecteur non nul est un vecteur propre pour tout endomorphisme, il n'y a rien à démontrer. Même si le corps n'est pas algébriquement clos, tous les endomorphismes d'un espace vectoriel de dimension 1 sont des homothéties. Soient $u, v : E \mapsto E$ deux endomorphismes du $\mathbb{C}$-espace vectoriel E de dimension $n \geq 2$ vérifiant $u \circ v = v \circ u$.

L'endomorphisme u possède une valeur propre, $\lambda_1 \in \mathbb{C}$, notons l'espace propre associé E_{u,λ_1}. Si cet espace est de dimension n, tout vecteur de E est propre pour u, et comme à son tour v possède au moins un vecteur propre, u et v possèdent un vecteur propre commun (dans le cas $E_{u,\lambda_1} = E,\ u = \lambda_1 \mathrm{I}$ où I est l'identité, u est scalaire).

Ainsi l'on peut supposer que la dimension de E_{u,λ_1} est au plus $n-1$. Il suffit de démontrer ici, que v se restreint à E_{u,λ_1} en un endomorphisme, et que tout vecteur propre de cette restriction est un vecteur propre commun de u et v.

Or si $X \in E_{u,\lambda_1},\ v(X)$ vérifie $u(v(X)) = v(u(X)) = v(\lambda_1 X) = \lambda_1 v(X)$. Qui veut dire exactement que $v(X)$ est soit nul, soit propre pour u, associé à la valeur propre λ_1. En tout état de cause, $v(X) \in E_{u,\lambda_1}$. Cet espace est stable pour l'endomorphisme v, nous pouvons restreindre v (à la source et au but) et obtenir un endomorphisme de E_{u,λ_1}. Cet endomorphisme, ainsi restreint, possède un vecteur propre (car nous travaillons sur $\mathbb{C}$). C'est donc un vecteur propre de v qui est propre aussi pour u. On utilise l'évidence qui consiste à constater qu'un vecteur propre d'une restriction d'un endomorphisme, est à même temps un vecteur propre de l'endomorphisme sans le restreindre. La récurrence est achevée. Par ailleurs nous constatons, que rien n'est dit sur la valeur, de la valeur propre. Si u et v partagent un vecteur propre au moins, les valeurs propres respectivement associées n'ont aucune relation. Tout peut arriver, comme on le voit, en considérant les endomorphismes scalaires $\lambda\mathrm{I}$ et $\mu\mathrm{I}$.

Rédigeons une démonstration de ce Lemme très utile, qui soit purement matricielle. Cela nous permettra de prendre du savoir faire sur les triangulations par blocs des matrices.

Soient donc A, B deux matrices carrées de taille n qui commutent. Soit λ_1 une valeur propre de la matrice A (sur $\mathbb{C}$, ça existe toujours). Soit $X_1, X_2, \ldots, X_l$ une base de l'espace propre associé. Complétons ces vecteurs en une base de $\mathbb{C}^n$ et formons la matrice de passage, P, dont les colonnes sont les vecteurs X_i.

Il est facile de constater que la matrice $P^{-1}AP$ est triangulaire par blocs et que son premier bloc est scalaire : $\lambda_1 \mathrm{I}_l$:

$$P^{-1}AP = \left(\begin{array}{ccccccccc}
\lambda & 0 & & & \cdots & 0 & * & \cdots & * \\
0 & \lambda & 0 & & \cdots & 0 & * & \cdots & * \\
0 & \cdots & \lambda & 0 & \cdots & 0 & * & \cdots & * \\
\vdots & & & & \ddots & \vdots & \vdots & & \vdots \\
 & & & & & 0 & & & \\
0 & & & \cdots & & \lambda & * & \cdots & * \\
0 & & & & & 0 & * & \cdots & * \\
\vdots & & & & & \vdots & & N & \\
0 & & & & \cdots & 0 & & &
\end{array}\right) = \begin{pmatrix} \lambda_1 \mathrm{I}_l & C \\ 0 & N \end{pmatrix}$$

En décomposant $P^{-1}BP$ en blocs (selon la partition $(l, n-l)$) nous pouvons calculer $L := P^{-1}ABP = P^{-1}APP^{-1}BP$ et $L' := P^{-1}BAP = P^{-1}BPP^{-1}AP$ avec :

$$P^{-1}BP = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

ce qui donne

$$L = \begin{pmatrix} \lambda_1 B_{11} + CB_{21} & \lambda_1 B_{12} + CB_{22} \\ NB_{21} & NB_{22} \end{pmatrix} \quad L' = \begin{pmatrix} \lambda_1 B_{11} & B_{11}C + B_{12}N \\ \lambda_1 B_{21} & B_{21}C + B_{22}N \end{pmatrix}.$$

Exprimer la commutation $AB = BA$, c'est-à-dire, $L = L'$, entraîne $CB_{21} = 0,\ NB_{22} = B_{21} + B_{22}N$ et $NB_{21} = \lambda_1 B_{21}$. On constate alors, que

$$\begin{pmatrix} \lambda_1 \mathrm{I}_l & C \\ 0 & N \end{pmatrix} \begin{pmatrix} 0 \\ B_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_1 B_{21} \end{pmatrix}$$

c'est-à-dire que les colonnes de cette dernière matrice sont des vecteurs propres de $P^{-1}AP$. Or, puisque les l premières colonnes de P sont une base de l'espace propre de A associé à λ_1, les vecteurs propres de $P^{-1}AP$ sont exactement les vecteurs de $\mathbb{C}^n$ qui ont leurs $l-n$ dernières composantes nulles. Ainsi $B_{21} = 0$. Il s'en suit que la matrice $P^{-1}BP$ est aussi triangulaire par blocs, et la commutation $B_{22}N = NB_{22}$. Par un changement de base parmis les l premiers vecteurs de la base canonique, l'on peut donner à $P^{-1}BP$ une première colonne colinéaire à e_1 ce qui signale le vecteur propre commun. Cet argument, difficile à suivre et à rédiger, montre l'intérêt que l'on a parfois à raisonner avec des endomorphismes et espaces vectoriels, et non pas se cantonner au calcul matriciel brut.

Ceci achève la (les) démonstrations du Lemme, somme toute élémentaire.

Reprenant la preuve du théorème fondamental, nous prenons (et normons) un vecteur propre simultané pour M et M^* : X_1 complétant en une base orthonormale, pour former la matrice W. L'on a que W^*MW est la diagonale (par blocs) $dia(\lambda_1, N)$. L'espace vectoriel engendré par les vecteurs $X_2, \ldots, X_n$ est exactement $X_1^{\perp}$.

Il est facile de voir maintenant que N et N^* commutent, simplement en calculant $W^*MM^*W = W^*MWW^*M^*W = dia(\lambda_1, N)dia(\bar{\lambda_1}, N^*) = dia(|\lambda_1|^2, NN^*)$ qui doit être égal à $W^*M^*MW = W^*M^*WW^*MW = dia(\bar{\lambda_1}, N)dia(\lambda_1, N^*) = dia(|\lambda_1|^2, N^*N)$ ce qui évidement donne $NN^* = N^*N$. La matrice de taille $n-1$, N est normale, par hypothèse de récurrence, elle se diagonalise en conjuguant avec une matrice unitaire $\tilde{W}$, en posant $V = dia(1, \tilde{W})$ on termine en constatant que WV est unitaire et que V^*W^*MWV est diagonale.

La preuve matricielle du Théorème fondamental est achevée.

Voici la rédaction, sans matrices, du théorème fondamental. Raisonnons, avec une base fixée de départ, qui peut être la base canonique de $\mathbb{C}^n$, à l'aide de l'endomorphisme m de $\mathbb{C}^n$ de matrice M. Notons m^* l'endomorphisme qui a M^* comme matrice dans cette base, c'est donc l'adjoint. Raisonnons par récurrence sur la dimension n. Pour $n = 1$ l'énoncé est évident, car toute matrice de taille 1×1 est diagonale ($U = I_1$). Montrons d'abord que m et m^* ont un vecteur propre commun. Nous avons que $m \circ m^* = m^* \circ m$ et c'est un fait important, que lorsque deux endomorphismes "commutent" ils ont un vecteur propre en commun. Voici la démonstration pour $u, v : \mathbb{C}^n \to \mathbb{C}^n$, $\mathbb{C}-$linéaires vérifiant $u \circ v = v \circ u$. Soit λ une valeur propre de u et $F = \ker(u - \lambda \mathrm{id}\ )$ son espace propre, qui est de dimension au moins 1. Evidement $u(F) \subset F$ et si $x \in F$ $u(v(x)) = v(u(x)) = v(\lambda x) = \lambda v(x)$ ainsi $v(x)$ est propre, pour u, associé à λ, ce qui fait que $v(x) \in F$. Ainsi, la restriction de v à l'espace F est un endomorphisme de F. Cette restriction, comme tout endomorphisme, à au moins une valeur propre et un vecteur propre associé, un tel vecteur de F est donc propre pour v et pour u. Evidement, les valeurs propres associées ne sont pas les mêmes, il ne fallait pas s'attendre à tant !

Enfin, ainsi m et m^* ont un vecteur propre $x \in \mathbb{C}^n$ commun. Considérons l'orthogonal hermitien standard : $H = x^{\perp} = \{y \in \mathbb{C}^n\ / y^*x =< x, y >= \sum_{i=0}^{n} x_i \bar{y}_i = 0\}$. Cet $\mathbb{C}$-espace vectoriel est de dimension exactement $n-1$ et est laissé stable par m^*. En effet, si $y \in H$, $< x, m^*(y) >=<$

$m(x), y >= \lambda_m < x, y >= 0$ et $m^*(H) \subset H$. Nous avons noté λ_m la valeur propre de m associée à x.

Un calcul similaire nous dit que H est aussi laissé stable par m.

Notons s la restriction de m à cet espace H. Il s'agit d'un endomorphisme. Démontrons que l'adjoint s^* est la restriction de m^*. Par définition d'adjoint, si $y \in H,\ s^*(y)$ est l'unique vecteur de H vérifiant que pour tout $z \in H,\ < z, s^*(y) >=< s(z), y >$. Or le vecteur $m^*(y)$ fait l'affaire, car si $z \in H\ s(z) = m(z)$ et $< s(z), y >=< m(z), y >=< z, m^*(y) >$ l'adjoint de s est bien la restriction de m^*.

Ainsi, s et s^* commutent aussi, s est un endomorphisme normal de l'espace H. Par hypothèse de récurrence lorsque $n > 1$, il existe une base orthonormale de H formée de vecteurs propres de s. Ce sont donc des vecteurs propres de m. En complétant cette base, à l'aide du normalisé $\frac{x}{||x||}$ on obtient une base orthonormale de $\mathbb{C}^n$ formée de vecteurs propres de m. Ecrire la matrice de m dans cette base, reviens bien à obtenir l'expression de diagonalisation souhaitée. $\diamondsuit$

Le théorème spectral pour les matrices normales possède une version réelle.

Théorème 3 (Décomposition Spectrale des Matrices normales RÉELLES, par blocs)**.** *Soit $M \in M_n(\mathbb{R})$, vérifiant $M^TM = MM^T$ (M normale) alors il existe $O \in M_n(\mathbb{R})$ vérifiant $O^T = O^{-1}$ (O est une matrice orthogonale) et des blocs de taille ≤ 2 $B_1, B_2, \ldots, B_m$ tels que*

$$M = O^{-1}\mathrm{diag}(B_1, B_2, \ldots, B_m)O.$$

Ces blocs sont tantôt des scalaires 1×1 tantôt des matrices de similitudes directes $M_\mu,\ \mu \notin \mathbb{R}$.

La preuve est plus délicate ([1]).

Définition 1. *Soit $M \in M_n(\mathbb{C})$, une matrice hermitienne, nous dirons que M est hermitienne définie positive si pour tout $X \in \mathbb{C}^n \setminus \{0\}$ nous avons*

$$X^*MX > 0.$$

Nous dirons que M est hermitienne semi-définie positive si seulement $X^*MX \geq 0$ est toujours assuré. Evidement, lorsque M est à coefficients réels, elle est symétrique et l'on dira "symétrique définie (ou semi-définie) positive".

Evidement, il faut remarquer que X^*MX est un nombre réel dès que M est hermitienne. De par ses dimensions, c'est un scalaire, de plus, son transposé-conjugué (c'est-à-dire son conjugué) :

$$(X^*MX)^* = X^*M^*X = X^*MX, \text{ i.e. } X^*MX \in \mathbb{R}.$$

En complément nous signalons quelques "curiosités" :

Définition 2. *Vecteurs positifs Un vecteur X de $\mathbb{R}^n$ sera dit positif, si toutes ses coordonnées sont des nombres positifs ou nuls, on le notera $X \geq 0$ (Ou $0 \leq X$). Il est dit strictement positif, si toutes ses coordonnées le sont (noté $0 < X$). Si $X \in \mathbb{C}^n$ on posera $|X|$ le vecteurs de $\mathbb{R}^n$ dont les coordonnées sont les modules des coordonnées de X. C'est un vecteur positif. On posera, pour $X, Y \in \mathbb{R}^n,\ X \leq Y$ si $Y - X \geq 0$. De même, nous dirons qu'une matrice $A \in M_{mn}(\mathbb{R})$ (peut-être ni carrée ni symétrique) est positive, si chaque élément l'est; notant $A \geq 0$ (avec $A > 0$ avec le sens correspondant, ainsi que $|A|$ pour $A \in M_{mn}(\mathbb{C})$. Nous avons $A \leq B$ si $0 \leq B - A$, pour la matrice nulle 0 de la bonne taille.*

Lemme 2. *1- Pour une matrice A et un vecteur X (tels que AX est défini), $|AX| \leq |A| \cdot |X|$.*

2- Une matrice A est positive si et seulement si, $X \geq 0$ entraîne $AX \geq 0$.

3- Une matrice A est strictement positive si et seulement si, $X \geq 0$ et $X \neq 0$ entraînent $AX > 0$.

Démonstration. Soit $A \in M_{m,n}(\mathbb{C}), X \in \mathbb{C}^n$. L'inégalité triangulaire pour le module donne :

$$|(AX)_i| = |\sum_{j=1}^{n} a_{ij}X_j| \leq \sum_{j=1}^{n} |a_{ij}||X_j| = (|A| \cdot |X|)_i$$

d'où l'on déduit l'inégalité entre vecteurs $|AX| \leq |A| \cdot |X|$ énoncée en 1.

Si la matrice A est réelle, et tous ses éléments sont positifs ou nuls, pour chaque vecteur à coefficients réels positifs ou nuls, nous avons que les coefficients de AX sont somme de nombres positifs ou nuls, dès-lors $AX \geq 0$. Si de plus A est strictement positive, et X est positif non nul, X a un (au moins) de ses coefficients strictement positif, comme le sont tous les éléments de A. Dans le calcul de chaque coefficient de AX tous les termes sont positif ou nuls, et l'un (au moins) d'entre eux est strictement positif : la somme est strictement positive. Il en résulte que chaque coefficient de AX est strictement positif, c'est-à-dire $AX > 0$. Ceci établis la nécessité des condition dans les points 2 et 3.

Pour les réciproques. Remarquons que les vecteurs de la base canonique de $\mathbb{R}^n$ vérifient que $e_i \geq 0, e_i \neq 0$. Alors, si pour tout vecteur $X \geq 0$ nous avons $AX \geq 0$ à fortiori, la i-ème colonne de A qui est Ae_i est formée de nombres positifs ou nuls, ainsi $A \geq 0$.

Si l'on sait que $X \geq 0$ *et* $X \neq 0$ entrain $AX > 0$, avec les vecteurs de la base canonique, on prouve que chaque élément de la matrice A est strictement positif. $\diamondsuit$

Théorème 4 (Théorème de Perron-Frobenius)**.** *Soit $A \in M_n(\mathbb{R})$ une matrice positive et $\rho(A)$ son rayon spectral. Le nombre $\rho(A)$ est valeur propre de A, il est associé à un vecteur propre positif.*

Démonstration. Laissée pour plus tard, puisque l'on a besoin de discuter des normes.

L'on définit le concept de "matrice irréductible" en disant qu'une matrice $A \in M_n(\mathbb{C})$ est *réductible* s'il est possible de trouver une partition $I \cup J = [1 : n]$ telle que si $(i, j) \in I \times J$ alors $a_{ij} = 0$. Les matrices irréductibles sont celles qui ne sont pas réductibles. Tout ceci mérite des éclaircissements. Profitant des connaissances de topologie, construissons un graphe où l'on met des n sommets (nommés avec les nombres de $[1 : n]$) et on relie le sommet i au sommet j par un segment (de courbe) si et seulement si, l'élément a_{ij} est non nul. Les différents segments ainsi tracés ne se rencontrent jamais en dehors des sommets (pensez à faire des petits ponts). C'est tout de même un sous-ensemble de $\mathbb{R}^3$, il acquiert la topologie induite. La matrice est irréductible si et seulement si cet espace est connexe. En clair, on peut aller de i à j, "en passant" par des k intermédiaires lorsque les éléments de la matrice qui font les ponts, ne sont pas nuls.

La forme forte du Théorème de Perron-Frobenius qui suit, a trouvé des nombreuses applications en sciences. Lorsque les coefficients d'un vecteur sont des masses, ou des dénombrements d'objets, la stricte positivité d'un vecteur à un sens existentiel particulièrement aigu.

Théorème 5 (Théorème de Perron-Frobenius, forme forte [1])**.** *Soit $A \in M_n(\mathbb{R})$ une matrice positive et irréductible, $\rho(A)$ son rayon spectral. Le nombre $\rho(A)$ est valeur propre simple de A, il est associé à un vecteur propre strictement positif. De plus $\rho(A) > 0$.*

Références

[1] D. Serre. Les Matrices, théorie et pratique. Dunod, Paris, 2001.

[2] P. Ciarlet. Introduction à l'analyse numérique matricielle et à l'optimisation. Masson, Paris 1982.

[3] L. Amodei, J.-P. Dedieu. Analyse Numérique Matricielle. Dunod, Paris 2008.

[4] P. Ciarlet et J.-M. Thomas. Exercices d'analyse numérique matricielle et d'optimisation. Masson, Paris 1982.

[5] P. Lascaux et R. Théodor. Analyse numérique matricielle appliquée à l'art de l'ingénieur. Masson, Paris, 1987.

[6] F. R. Gantmacher. Théorie des matrices I et II. Dunod, Paris, 1966.

[7] N. Gastinel. Analyse numérique linéaire. Hermann, Paris, 1970.

[8] M. Cuer.`http ://montellier??`. Montpellier, 2006.

[9] Y. Achdou. `http ://www.ann.jussieu.fr/~achdou/node7.html`. Paris, 2005.

[10] J.Prado. Introduction à MATLAB `http ://www.tsi.enst.fr/~prado/enseignement/polys/matlab.html` Paris, 1998.

[11] Consortium SCILAB.`http ://www.scilab.org/`, INRIA.

[12] B. Pinçon. Une introduction à Scilab. `http ://www.iecn.u-nancy.fr/~pincon/scilab/docA4.pdf` . Nancy, 2005.