

CAHIERS DU BURO

G. KREWERAS

Premiers éléments du calcul matriciel

Cahiers du Bureau universitaire de recherche opérationnelle.
Série Recherche, tome 8 (1966), p. 3-41

http://www.numdam.org/item?id=BURO_1966__8__3_0

© Institut Henri Poincaré — Institut de statistique de l'université de Paris, 1966,
tous droits réservés.

L'accès aux archives de la revue « Cahiers du Bureau universitaire de recherche opérationnelle. Série Recherche » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

PREMIERS ÉLÉMENTS DU CALCUL MATRICIEL

par

G. KREWERAS

Le texte ci-après reproduit, dans leurs grandes lignes, des exposés sur le calcul matriciel qui ont été faits au cours de plusieurs des stages organisés par le B.U.R.O. à l'usage des ingénieurs qui s'intéressent aux méthodes de la recherche opérationnelle.

Ce texte ne comporte que les indications nécessaires pour établir, d'une manière enchaînée, les principales propriétés des matrices. Le point de vue adopté a été celui du calcul matriciel formel, plutôt que celui des opérateurs sur les espaces vectoriels de dimension finie ; toutefois les propriétés élémentaires de ces espaces ont été utilisées dans l'établissement de la partie relative à l'analyse spectrale, laquelle suppose également connues les toutes premières notions relatives aux déterminants.

Il nous a paru utile d'indiquer assez complètement l'une des marches qui permettent d'introduire et de justifier la réduction à la forme canonique de Jordan. Par contre, certaines définitions, propriétés et méthodes n'ont pas été abordées, de sorte que le titre : "Premiers Eléments" doit être pris au pied de la lettre. Les lecteurs soucieux de se perfectionner dans le calcul matriciel trouveront des compléments dans les ouvrages spécialisés, et notamment dans les suivants :

- E. BODEWIG : *Matrix Calculus*, North-Holland Publ. Co., Amsterdam, 1956

- F.R. GANTMACHER : *The theory of matrices*, vol I and II Chelsea Publ. Co, N.Y., 1960

- A. LICHNEROWICZ : *Algèbre et analyse linéaires*, Masson, 1960

- D. PHAM : *Techniques du calcul matriciel*, Dunod 1962

- J.M. SOURIAU : *Calcul linéaire*, P.U.F., 1959.

I - OPERATIONS SUR LES MATRICES

1.1 - Définition formelle

La définition d'une matrice exige la donnée d'un premier ensemble fini $\{u_1, u_2, \dots, u_m\}$, appelé *entrée* d'un deuxième ensemble fini $\{v^1, v^2, \dots, v^n\}$, appelé *sortie* et d'un corps K (qui sera le plus souvent le corps des nombres rationnels, quelquefois celui des nombres réels ou celui des nombres complexes). Les éléments de K sont appelés des *scalaires*.

Une *matrice* est par définition une application du produit cartésien $\{u_1, \dots, u_m\} \times \{v^1, \dots, v^n\}$ dans K .

1.2 - Notations

Une matrice, une fois définie, peut bien entendu toujours se désigner par une lettre unique. Le signe = entre deux matrices indique qu'il s'agit de la même application du même produit cartésien dans le même corps.

Il est souvent commode de se représenter l'entrée $\{u_1, \dots, u_m\}$ comme l'ensemble des lignes, la sortie $\{v^1, \dots, v^n\}$ comme l'ensemble des colonnes, d'un tableau à double entrée ; le couple (u_i, v^j) est alors une case de ce tableau, dans laquelle on inscrit son image $a_i^j \in K$:

			v^j
u_i		a_i^j	

L'ordre d'écriture des lignes, ainsi que celui des colonnes, étant libre, il y a en principe $m! \times n!$ manières d'écrire ainsi une même matrice.

Le plus souvent cependant on prend comme entrée $\{u_1, \dots, u_m\}$ l'ensemble des "numéros" de 1 à m , et comme sortie $\{v^1, \dots, v^n\}$ l'ensemble des "numéros" de 1 à n . Il y a dans ce cas un ordre naturel d'écriture des lignes et des colonnes, et l'on peut se contenter, sans rappeler en tête de chaque rangée (ligne ou colonne) son nom, d'inscrire le scalaire (ou "terme") a_i^j dans la case définie

par la i -ième ligne et la j -ième colonne. On dit alors que l'on a une matrice de *format* $m \times n$, et l'on appelle a_i^j son "terme général". Exemple de matrice de format 3×4 :

$$A = \begin{array}{|cccc|} \hline -1 & 0 & 4 & 2 \\ \hline 0 & 1 & -2 & -2 \\ \hline 3 & 3 & 1 & 0 \\ \hline \end{array}$$

Dans les questions théoriques, il est quelquefois commode de désigner une matrice de terme général a_i^j par $\|a_i^j\|$, ou (a_i^j) , ou $\boxed{a_i^j}$ (la notation $|a_i^j|$ est à proscrire).

1.3 - Formats particuliers

Si $m = 1$ et $n > 1$, une matrice de format $m \times n$ est appelée *vecteur-ligne*. Si $m > 1$ et $n = 1$, elle est appelée *vecteur-colonne*. Plus généralement on dit qu'une matrice est *large* si $m < n$ et *haute* si $m > n$.

Si $m = n$, la matrice est dite *carrée* d'ordre m .

Enfin dans le cas trivial où $m = n = 1$, on assimile la matrice au scalaire inscrit dans son unique case.

1.4 - Combinaisons linéaires

Etant donné deux matrices A et B de même format $m \times n$ et de termes appartenant au même corps K , et deux scalaires λ et μ , on désigne par

$$C = \lambda A + \mu B$$

une matrice, de même format que A et B , dont le terme général c_i^j est défini par

$$c_i^j = \lambda a_i^j + \mu b_i^j$$

a_i^j et b_i^j étant les termes généraux respectifs de A et B .

Les matrices de format $m \times n$ donné forment ainsi un *espace vectoriel* de dimension $m \cdot n$. Dans celui-ci il existe notamment un "élément nul", qui est la matrice dont tous les termes sont égaux au "0" du corps K ; on la désigne également par 0 (ou par $0_{m \times n}$ si l'on craint une équivoque).

Remarque : lorsqu'on considère le produit d'une matrice A par un scalaire λ ("homothétie"), il est indifférent d'écrire λA ou $A\lambda$, du moins si le corps K est commutatif (ce qui est pratiquement toujours le cas).

1.5 - Multiplication

1/ Définition

Pour définir le produit (*non commutatif*) d'une matrice A par une matrice B, il faut supposer

a) que la sortie de A est formée des mêmes objets que l'entrée de B

b) que les termes de A et B appartiennent au même corps K.

Autrement dit, il faut avoir :

$$\begin{array}{ccc} \{u_1, \dots, u_m\} \times \{v^1, \dots, v^n\} & \xrightarrow{A} & K \\ \{v_1, \dots, v_n\} \times \{w^1, \dots, w^p\} & \xrightarrow{B} & K \end{array}$$

Pour des raisons de clarté on a désigné un même objet par v^j ou v_j suivant qu'on l'envisage dans la sortie de A ou dans l'entrée de B.

On désignera les termes généraux de A et B respectivement par a_i^j et b_j^k , avec :

$$i \in \{1, 2, \dots, m\}$$

$$j \in \{1, 2, \dots, n\}$$

$$k \in \{1, 2, \dots, p\}$$

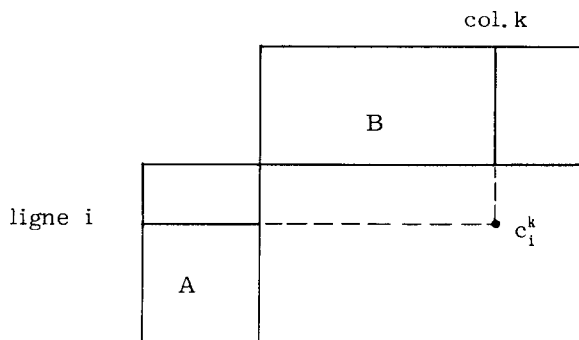
Par définition le produit $C = AB$ sera alors une matrice telle que

$$\{u_1, \dots, u_m\} \times \{w^1, \dots, w^p\} \xrightarrow{C} K$$

et dont le terme général sera

$$c_j^k = \sum_j a_i^j b_j^k \text{ (somme de n produits).}$$

Dans le langage des formats, la multiplication se présente donc comme suit :



Elle s'effectue, comme on dit en abrégé, "ligne par colonne". En particulier si A est un vecteur-ligne ($1 \times n$) et B un vecteur-colonne ($n \times 1$), le produit AB désigne un scalaire (on l'appelle pour cette raison le "produit scalaire" de A par B).

Revenant au cas général où les matrices A et B sont de formats respectifs ($m \times n$) et ($n \times p$), il est commode de se représenter A comme un "empilement" de vecteurs-ligne A_i et B comme une juxtaposition de vecteurs-colonne B^k ; le terme c_i^k de la matrice produit AB n'est autre, alors, que le produit scalaire $A_i B^k$:

$$A = \begin{array}{|c} A_1 \\ A_2 \\ \cdot \\ \cdot \\ A_m \end{array} \quad B = \begin{array}{|c} B^1 \ B^2 \ \dots \ B^p \end{array}$$

$$c_i^k = A_i \ B^k$$

2/ Non-commutativité

La définition même de la multiplication exclut en général la commutativité, du fait des formats.

Cependant, même si A et B sont toutes deux des matrices carrées de même format, c'est-à-dire si AB et BA ont tous deux un sens et désignent des matrices de même format, on a *en général*

$$AB \neq BA$$

Il existe cependant des cas très particuliers (on en rencontrera plus loin) où l'on a $AB = BA$; on dit dans ces cas que les matrices A et B "commutent".

3/ Distributivité

La multiplication ainsi définie est manifestement distributive (à gauche et à droite) par rapport aux combinaisons linéaires, ce qui permet d'écrire :

$$\begin{aligned}(\lambda_1 A_1 + \lambda_2 A_2) B &= \lambda_1 (A_1 B) + \lambda_2 (A_2 B) \\ A (\mu_1 B_1 + \mu_2 B_2) &= \mu_1 (A B_1) + \mu_2 (A B_2)\end{aligned}$$

4/ Associativité

L'associativité est la propriété fondamentale de la multiplication matricielle. Pour l'établir, on considère trois matrices :

	A	B	C
de formats	$m \times n$	$n \times p$	$p \times q$
et de termes généraux	a_i^j	b_j^k	c_k^l

à partir desquelles on peut définir les produits :

	$D = AB$	$E = BC$
de formats	$m \times p$	$n \times q$
et de termes généraux	$d_i^k = \sum_j a_i^j b_j^k$	$e_j^l = \sum_k b_j^k c_k^l$

puis les produits $DC = (AB)C$ et $AE = A(BC)$, qui ont tous deux le même format $m \times q$. L'associativité sera établie s'il est vérifié que ces deux derniers produits ont même terme général. Or le premier donne :

$$\sum_k d_i^k c_k^l = \sum_k \left(\sum_j a_i^j b_j^k \right) c_k^l$$

et le second :

$$\sum_j a_i^j e_j^l = \sum_j \left(a_i^j \sum_k b_j^k c_k^l \right) ;$$

et les deux expressions obtenues sont bien égales puisqu'elles représentent deux manières de calculer une même somme de np produits, tels que $a_i^j b_j^k c_k^l$, obtenus en se fixant i et l et en prenant tous les couples d'indices (j, k) possibles.

On a donc bien associativité, et l'on peut écrire :

$$ABC = (AB)C = A(BC)$$

5/ Matrices unités

On appelle matrice unité d'ordre m , et l'on désigne souvent par I_m (ou simplement I si aucune équivoque n'est à craindre), une matrice carrée d'ordre m dans laquelle :

$$\begin{aligned} a_i^j &= 0 & \text{si } i \neq j \\ a_i^j &= 1 & \text{si } i = j \end{aligned}$$

Par exemple :

$$I_3 = \begin{array}{|ccc|} \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline \end{array}$$

L'application de la règle de multiplication montre immédiatement que, A étant une matrice quelconque de format $m \times n$, on a : $I_m A = A I_n = A$, ou plus simplement $I A = A I = A$. Ceci justifie le nom des matrices unités, puisqu'elles sont "neutres" pour la multiplication à gauche et à droite.

1.6 - Ecriture par blocs

Soit U un vecteur-ligne et V un vecteur-colonne, ayant tous deux le même nombre n de termes (ou "composantes") :

$$U = \begin{array}{|cccc|} \hline a^1 & a^2 & \dots & a^n \\ \hline \end{array} \qquad V = \begin{array}{|c|} \hline b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_n \\ \hline \end{array}$$

On a vu plus haut que le produit UV est le scalaire :

$$UV = a^1 b_1 + a^2 b_2 + \dots + a^n b_n$$

Si sur l'ensemble d'indices (commun par hypothèse à U et à V) on définit une relation classificatoire répartissant ces indices en (par exemple) 3 classes α β γ , la donnée de U revient à la donnée de 3 vecteurs-lignes U^α , U^β , U^γ , et de même la donnée de V revient à celle de 3 vecteurs-colonnes V_α , V_β , V_γ . En groupant convenablement les n termes tels que $a^j b_j$, on voit que l'on a :

$$UV = U^\alpha V_\alpha + U^\beta V_\beta + U^\gamma V_\gamma$$

On peut donc utiliser l'écriture :

$$U = \begin{bmatrix} U^\alpha & U^\beta & U^\gamma \end{bmatrix} \quad V = \begin{bmatrix} V_\alpha \\ V_\beta \\ V_\gamma \end{bmatrix}$$

en conservant la règle de multiplication "ligne par colonne" comme si U et V n'avaient plus que 3 "composantes", qui seraient des vecteurs et non plus des scalaires.

Cette convention, fort utile, de l'écriture par blocs s'étend aisément à un produit de matrices quelconque, AB ; la seule condition est que sur l'ensemble des numéros de colonne de A et sur l'ensemble des numéros de ligne de B (il s'agit nécessairement du même ensemble) on ait défini la même relation classificatoire. Les classes à définir pour les lignes de A et pour les colonnes de B sont arbitraires, et se retrouvent respectivement sur les lignes et sur les colonnes de la matrice produit AB ; on peut ainsi multiplier par la règle habituelle des matrices dont les "termes apparents" sont des matrices et non plus des scalaires.

Il ne faut pas confondre leur "format apparent" avec leur format réel, qui peut être beaucoup plus grand.

Exemple :

$$\begin{bmatrix} E & F \\ G & H \end{bmatrix} \begin{bmatrix} J & K \\ L & M \end{bmatrix} = \begin{bmatrix} EJ + FL & EK + FM \\ GJ + HL & GK + HM \end{bmatrix}$$

1.7 - Inversion

Si deux matrices A et B sont telles que $AB = I$ (matrice unité), A est dite "inverse à gauche" de B et B "inverse à droite" de A.

Une matrice A étant donnée, elle peut n'avoir aucune inverse ni à droite ni à gauche, elle peut n'avoir des inverses que d'un côté et non de l'autre, et elle peut en avoir des deux côtés.

Dans ce dernier cas, si X est une inverse à droite de A et Y une inverse à gauche de A , on a à la fois :

$$AX = I \quad \text{et} \quad YA = I$$

Mais alors nécessairement $X = Y$. En effet on a d'une part :

$$(YA) X = IX = X$$

et d'autre part :

$$Y (AX) = YI = Y$$

X et Y sont toutes deux égales au produit YAX .

Il y a donc dans ce cas une inverse unique, la même à droite et à gauche, de la matrice A ; on la désigne alors par A^{-1} , et l'on a :

$$AA^{-1} = A^{-1}A = I \quad (A^{-1})^{-1} = A$$

Les matrices possédant cette propriété d'avoir une inverse (unique) à droite et à gauche sont dites *inversibles*. Il sera établi plus loin que seules des matrices carrées peuvent être inversibles, mais cependant qu'elles ne le sont pas toutes.

Le produit AB de deux matrices inversibles A et B est une matrice inversible et l'on a :

$$(AB)^{-1} = B^{-1} A^{-1}$$

comme on le voit en effectuant les deux calculs ci-après :

$$(AB) (B^{-1} A^{-1}) = A (BB^{-1}) A^{-1} = AIA^{-1} = AA^{-1} = I$$

$$(B^{-1} A^{-1}) (AB) = B^{-1} (A^{-1}A) B = B^{-1} IB = B^{-1} B = I$$

1.8 - Quelques matrices particulièrement simples

1/ Matrices "presque nulles"

Pour un format $m \times n$ donné, appelons E_s^t la matrice dont tous les termes sont nuls sauf le terme de la case (s, t) , égal à 1. Si A est une matrice de terme général a_{ij} , on a :

$$A = \sum_{(i,j)} a_{ij}^j E_i^j$$

La règle de multiplication matricielle montre immédiatement que :

$$E_u^v \times E_s^t = \begin{cases} 0 & \text{si } v \neq s \\ E_u^t & \text{si } v = s \end{cases}$$

Il s'ensuit notamment que le produit AE_s^t a des termes tous nuls dans ses colonnes autres que la t-ième, et que sa t-ième colonne reproduit la s-ième colonne de A.

Il est clair qu'une matrice E_s^t , même si elle est carrée, ne peut pas être inversible. (*Exercice* : y a-t-il une exception à cet énoncé ?)

2/ Matrices diagonales

Pour un format carré $n \times n$ donné, on appelle *matrices diagonales* toutes les combinaisons linéaires du type

$$\lambda_1 E_1^1 + \lambda_2 E_2^2 + \dots + \lambda_n E_n^n \quad ;$$

une telle matrice diagonale sera parfois notée en abrégé $\langle \lambda_1 \lambda_2 \dots \lambda_n \rangle$ ou encore plus brièvement $\langle \lambda_i \rangle$. On a non seulement

$$\langle \lambda_i \rangle + \langle \mu_i \rangle = \langle \lambda_i + \mu_i \rangle$$

mais encore

$$\langle \lambda_i \rangle \times \langle \mu_i \rangle = \langle \lambda_i \mu_i \rangle \quad ;$$

les matrices diagonales se multiplient terme à terme. De plus si elles sont définies sur un corps commutatif, ce qui est le cas pratiquement le plus fréquent, leur multiplication est elle-même *commutative*.

(*Exercice* : formuler une condition nécessaire et suffisante pour qu'une matrice diagonale soit inversible).

1.9 - Exemples de matrices (carrées) inversibles

1/ Soit la matrice (dite parfois "de transposition")

$$E_{\{1,2\}} = I - E_1^1 - E_2^2 + E_1^2 + E_2^1 = \begin{array}{|ccc|} \hline 0 & 1 & \underline{0} \\ 1 & 0 & \underline{0} \\ \underline{0} & \underline{0} & I \\ \hline \end{array} \quad \begin{array}{l} \underline{0} : \text{vecteurs-lignes} \\ \underline{0} : \text{vecteurs-colonnes} \end{array}$$

On vérifie immédiatement que :

$$E_{\{1,2\}} \times E_{\{1,2\}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I \end{bmatrix} = I$$

la matrice $E_{\{1,2\}}$ est donc sa propre inverse.

Si l'on calcule le produit $E_{\{1,2\}} \cdot A$, A étant une matrice que l'on peut écrire par blocs sous la forme

$$A = \begin{bmatrix} U_1 \\ U_2 \\ A' \end{bmatrix},$$

on voit que

$$E_{\{1,2\}} \cdot A = \begin{bmatrix} U_2 \\ U_1 \\ A' \end{bmatrix}$$

De même si

$$A = \begin{bmatrix} V_1 & V_2 & A'' \end{bmatrix}, \text{ on a } A \cdot E_{\{1,2\}} = \begin{bmatrix} V_2 & V_1 & A'' \end{bmatrix}$$

On voit ainsi qu'à partir d'une matrice quelconque A , la multiplication à gauche (resp. à droite) par une matrice $E_{\{1,2\}}$ du format convenable revient à la permutation en bloc des deux premières lignes (resp. colonnes).

Résultats analogues pour les matrices :

$$E_{\{s,t\}} = I - E_s^s - E_t^t + E_s^t + E_t^s$$

(s -ième ligne : 1 dans la t -ième colonne, 0 partout ailleurs
 t -ième ligne : 1 dans la s -ième colonne, 0 partout ailleurs
autres lignes : mêmes termes que si la matrice était unité).

$E_{\{s,t\}}$ est sa propre inverse, et la multiplication d'une matrice à gauche (resp. à droite) par $E_{\{s,t\}}$ revient à permuter en bloc sa r -ième et sa s -ième lignes (resp. colonnes).

Il en résulte notamment que si A est une matrice carrée inversible d'ordre n , il en est de même de toute matrice obtenue en modifiant de n'importe quelle façon l'ordre d'écriture des lignes et l'ordre d'écriture des colonnes. En effet on peut passer de A à n'importe laquelle de ces matrices en procédant à des permutations successives de deux lignes ou de deux colonnes ; et puisque chacune de ces permutations revient à une multiplication par une matrice du type $E_{\{s,t\}}$, qui est inversible, le produit demeure toujours inversible.

2/ Soit la matrice diagonale particulière :

$$D_1(\lambda) = \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}, \text{ avec } \lambda \neq 0.$$

On vérifie que :

$$D_1(\lambda) \times D_1\left(\frac{1}{\lambda}\right) = D_1\left(\frac{1}{\lambda}\right) \times D_1(\lambda) = I$$

Si $A = \begin{bmatrix} U \\ A_1 \end{bmatrix}$, $D_1(\lambda) A = \begin{bmatrix} \lambda U \\ A_1 \end{bmatrix}$; multiplier une matrice A par $D_1(\lambda)$ à gauche (resp. à droite) revient donc à multiplier par λ toute sa première ligne (resp. colonne).

Résultats analogues pour les matrices $D_s(\lambda)$ que l'on peut obtenir à partir d'une matrice unité en remplaçant par λ le 1 qui figure dans la case de s -ième ligne et s -ième colonne. Multiplier A à gauche (resp. à droite) par $D_s(\lambda)$ revient à multiplier par λ toute la s -ième ligne (resp. colonne).

3/ Soit

$$F_s^t(\lambda) = I + \lambda E_s^t$$

On vérifie immédiatement que

$$F_s^t(\lambda) \cdot F_s^t(-\lambda) = F_s^t(-\lambda) \cdot F_s^t(\lambda) = I.$$

Si :

$$A = \begin{bmatrix} U_1 \\ U_2 \\ A \end{bmatrix}, \quad F_1^2(\lambda) A = \begin{bmatrix} U_1 + \lambda U_2 \\ U_2 \\ A \end{bmatrix}$$

Multiplier une matrice à gauche par $F_s^t(\lambda)$, c'est donc remplacer sa s -ième ligne par une ligne obtenue en y ajoutant le produit de la t -ième par λ ; de même multiplier une matrice à droite par $G_s^t(\lambda)$ revient à modifier sa t -ième colonne par addition de la s -ième multipliée par λ .

2 - RANG

2.1 - Annulateurs

Contrairement à ce qui se passe dans le calcul ordinaire, où un produit ne peut être nul sans que l'un au moins des facteurs le soit, on peut avoir avec des matrices :

$$AB = 0 \quad \text{avec} \quad A \neq 0 \quad \text{et} \quad B \neq 0$$

A est alors un annulateur à gauche de B , et B un annulateur à droite de A .

Les matrices *nulles* sont évidemment des annulateurs à droite et à gauche pour toute autre matrice (telle que la multiplication ait un sens), ou *annulateurs universels*.

Si une matrice donnée A , de format $m \times n$, ne possède aucun annulateur à droite autre que 0, on dit qu'elle est "*à colonnes indépendantes*". Cette propriété s'exprime par l'implication :

$$AX = 0 \implies X = 0$$

valable notamment si X est un vecteur-colonne à n composantes.

Si :

$$A = \begin{bmatrix} V_1 & V_2 & \dots & V_n \end{bmatrix} \quad \text{et} \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

la propriété des colonnes de A d'être indépendantes signifie que la combinaison linéaire $V_1 x_1 + V_2 x_2 + \dots + V_n x_n$ ne peut être nulle sans que tous ses coefficients scalaires x le soient. Cette propriété est alors manifestement commune à toutes les matrices obtenues à partir de A en changeant de n'importe quelle manière l'ordre d'écriture des lignes ou l'ordre d'écriture des colonnes ou ces deux ordres à la fois.

2.2 - Propriétés des matrices à colonnes indépendantes

1/ A partir de toute matrice à colonnes indépendantes, on en forme une autre, également à colonnes indépendantes, en supprimant des colonnes ou en adjoignant de nouvelles lignes.

En effet supposons que $\begin{bmatrix} A & B \end{bmatrix}$ soit à colonnes indépendantes et montrons que $\begin{bmatrix} A \\ C \end{bmatrix}$ l'est aussi :

$$\begin{array}{ccccccc} \begin{bmatrix} A \\ C \end{bmatrix} X = 0 & \Longrightarrow & \begin{bmatrix} AX \\ CX \end{bmatrix} = 0 & \Longrightarrow & AX = 0 & & \\ & \text{par multi-} & & \text{par défi-} & & & \\ & \text{tiplication} & & \text{nition de 0} & & & \\ & & \Longrightarrow & \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} X \\ O \end{bmatrix} = 0 & \Longrightarrow & \begin{bmatrix} X \\ O \end{bmatrix} = 0 & \Longrightarrow & X = 0 \\ & \text{par multi-} & & \text{par hypo-} & & \text{par défini-} & & \\ & \text{tiplication} & & \text{thèse} & & \text{tion de 0} & & \end{array}$$

Finalement $\begin{bmatrix} A \\ C \end{bmatrix} X = 0 \Longrightarrow X = 0$, ce qu'il s'agissait d'établir.

2/ Toute matrice carrée inversible est à colonnes indépendantes.

En effet, si A est carrée inversible, on a :

$$AX = 0 \Longrightarrow A^{-1}(AX) = 0 \Longrightarrow IX = 0 \Longrightarrow X = 0,$$

donc A est bien à colonnes indépendantes.

3/ Toute matrice carrée à colonnes indépendantes est inversible

Cette propriété peut se montrer par récurrence sur l'ordre de la matrice carrée.

Elle est triviale pour les "matrices d'ordre 1", qui sont, on l'a vu, de simples scalaires : si $ax = 0 \Longrightarrow x = 0$, c'est que $a \neq 0$, et son inverse est $\frac{1}{a}$.

Supposons-la démontrée jusqu'à l'ordre $n - 1$, et considérons une matrice carrée A d'ordre n , à colonnes indépendantes.

La matrice A n'a certainement aucune colonne nulle, et

comporte donc dans sa dernière colonne au moins un terme non nul $\alpha \neq 0$; sans restreindre la généralité, on peut supposer que ce terme apparaît à la dernière ligne. On a alors :

$$A = \begin{bmatrix} A' & V \\ U & \alpha \end{bmatrix},$$

où A' est une matrice carrée d'ordre $n - 1$.

Puisque A est à colonnes indépendantes, on a notamment :

$$\begin{bmatrix} A' & V \\ U & \alpha \end{bmatrix} \begin{bmatrix} X \\ -\frac{1}{\alpha} U X \end{bmatrix} = 0 \implies X = 0$$

ou encore :

$$\left. \begin{array}{l} A'X - \frac{1}{\alpha} VUX = 0 \\ UX - UX = 0 \end{array} \right\} \implies X = 0$$

c'est-à-dire finalement :

$$\left(A' - \frac{1}{\alpha} VU \right) X = 0 \implies X = 0$$

La matrice carrée d'ordre $n - 1$

$$B = A' - \frac{1}{\alpha} VU$$

est à colonnes indépendantes, donc inversible en vertu de l'hypothèse de récurrence, et d'inverse B^{-1} .

Mais A peut alors s'écrire :

$$A = \begin{bmatrix} B + \frac{1}{\alpha} VU & V \\ U & \alpha \end{bmatrix},$$

et une vérification directe établit immédiatement que A est inversible et a pour inverse :

$$A = \begin{bmatrix} B^{-1} & -\frac{1}{\alpha} B^{-1} V \\ -\frac{1}{\alpha} UB^{-1} & \frac{1}{\alpha} + \frac{1}{\alpha^2} UB^{-1} V \end{bmatrix}$$

L'ordre particulier des lignes ne changeant rien à l'inversibilité, la propriété est donc établie.

4/ Une matrice large ne peut être à colonnes indépendantes

En effet si une matrice large A était à colonnes indépendantes, il en serait de même de la matrice carrée $\begin{bmatrix} A \\ 0 \end{bmatrix}$ d'après la propriété

(1°), et celle-ci serait inversible d'après la propriété (3°). Mais une matrice ayant une ligne entière nulle ne peut manifestement pas avoir d'inverse à droite.

Une matrice à colonnes indépendantes est donc nécessairement ou haute ou carrée.

5/ Une matrice large ne peut être inversible

En effet si une matrice large A était inversible, elle aurait une inverse à gauche B, telle que BA = I. Mais alors on aurait AX = 0 \implies BAX = 0 \implies IX = 0 \implies X = 0, et la matrice large A serait à colonnes indépendantes, ce qui a été reconnu impossible.

2,3 - Matrices à colonnes liées

Etre à "colonnes liées" est pour une matrice la propriété *contraire* de celle d'être à colonnes indépendantes. Dire que A est à colonnes liées, c'est donc dire que A possède un annulateur à droite autre que 0, ou encore que AX = 0 est possible pour un vecteur-colonne X \neq 0.

Il existe alors dans A une colonne égale à une combinaison linéaire des autres. Supposons en effet que ce soit la k-ième composante de X qui soit non nulle ($x_k \neq 0$).

Si A = $\begin{bmatrix} V_1 & V_2 & \dots & V_n \end{bmatrix}$, on a bien

$$x_1 V_1 + x_2 V_2 + \dots + x_n V_n = 0,$$

d'où :

$$V_k = \sum_{j \neq k} \left(-\frac{x_j}{x_k} \right) V_j$$

Si la matrice de $n - 1$ colonnes qui reste en supprimant la colonne V_k est à son tour à colonnes liées, il y existe à nouveau une colonne V_1 égale à une combinaison linéaire des $n - 2$ restantes ; V_k et V_1 sont toutes deux des combinaisons linéaires des $n - 2$ autres colonnes.

De proche en proche on peut extraire de toute matrice A à colonnes liées une sous-matrice formée de r colonnes indépendantes, dont toutes les autres colonnes de A sont des combinaisons linéaires particulières.

2.4 - Rang d'une matrice

Le rang d'une matrice peut être défini comme le nombre maximum de colonnes indépendantes que l'on peut en extraire.

D'une manière "duale" de celle suivie précédemment, on peut définir pour une matrice donnée, suivant qu'elle possède ou non des annulateurs à gauche non nuls, les notions de lignes indépendantes ou liées. Les propriétés "duales" des propriétés (1°) à (5°) peuvent alors s'énoncer et s'établir immédiatement.

Soit alors A une matrice quelconque de format $m \times n$, et soit r son rang ($r \leq n$). Soit A' une sous-matrice de A formée de r colonnes indépendantes (éventuellement A elle-même si $r = n$). A' , on l'a vu, ne peut être que haute ou carrée. Si A' est carrée, elle est inversible en vertu de la propriété (3°), et elle est alors à lignes indépendantes en vertu de la propriété duale de (2°) ; les r lignes de A obtenues en complétant s'il y a lieu ces r lignes de A' sont donc encore indépendantes en vertu de la propriété duale de (1°).

A possède donc r lignes indépendantes, et n'en saurait posséder davantage ; car si A avait $r + 1$ lignes indépendantes, on pourrait, par une méthode duale, trouver dans A $r + 1$ colonnes indépendantes, ce qui est contraire à la définition du rang.

Le rang d'une matrice est donc, aussi bien, le nombre maximum de lignes indépendantes que l'on peut en extraire.

Enfin si une matrice A est de rang r , le produit PA , où P désigne une matrice (carrée) inversible, est aussi de rang r . Cela tient au fait que, si l'on désigne par A' une sous-matrice quelconque formée de colonnes de A , A' et PA' sont toutes deux à colonnes indé-

pendantes, ou toutes deux à colonnes liées ; en effet si $A'X' = 0$ entraîne $X' = 0$, on a :

$$PA'X' = 0 \implies P^{-1} PA'X = 0 \implies A'X' = 0 \implies X' = 0$$

et si $A'X' = 0$ est possible avec $X' = 0$, on a $PA'X' = 0$, avec le même X' .

De même (par dualité) la multiplication à droite de A par Q inversible ne change pas non plus le rang.

2.5 - Equivalence

Deux matrices A et B de même format sont dites équivalentes s'il existe deux matrices inversibles P et Q telles que $B = PAQ$.

Cette dénomination est légitime puisqu'il y a :

- a) *réflexivité* : $A = IAI$
- b) *symétrie* : $B = PAQ \implies A = P^{-1} BQ^{-1}$
- c) *transitivité* : $\left. \begin{array}{l} B = PAQ \\ C = RBS \end{array} \right\} \implies C = RPAQS$

(RP et QS sont inversibles si P , Q , R et S le sont)

De plus, on a vu à la fin du paragraphe précédent que l'équivalence conserve le rang.

Mais en outre on va montrer que deux matrices de même format et de même rang sont nécessairement équivalentes ; on établira pour cela que toute matrice non nulle A de format $m \times n$ et de rang r est équivalente à la matrice (dite *canonique*)

$$C_r = \begin{array}{cc|l} \boxed{\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array}} & \begin{array}{l} r \text{ lignes} \\ m - r \text{ lignes} \end{array} \\ \begin{array}{cc} r & n - r \end{array} & \begin{array}{l} \text{colonnes} \\ \text{colonnes} \end{array} \end{array}$$

2.6 - Réduction à la forme canonique

Partant d'une matrice A , on peut, par une succession d'équivalences, l'amener à la forme canonique ci-dessus : on utilisera pour cela des multiplications à droite ou à gauche par des matrices inversibles particulières, qui appartiendront aux types E , D , et F cités plus haut comme exemples de matrices inversibles.

A ayant un terme non nul, par exemple $a_s^t \neq 0$, on peut amener celui-ci dans la case (1, 1) en multipliant A à gauche par $E_{\{1,s\}}$ et à droite par $E_{\{1,t\}}$; puis on peut rendre ce terme égal à 1 en multipliant (à droite ou à gauche indifféremment) par $D_1\left(\frac{1}{a_s^t}\right)$. L'ensemble de la 1ère ligne et de la 1ère colonne a alors l'aspect suivant :

$$\begin{array}{ccccccc} 1 & b^2 & \dots & \dots & \dots & \dots & b^n \\ & a_2 & & & & & \\ & \cdot & & & & & \\ & \cdot & & & & & \\ & \cdot & & & & & \\ & \cdot & & & & & \\ & a_m & & & & & \end{array}$$

On peut ensuite remplacer a_i par 0 en multipliant à gauche par $F_i^1(-a_i)$, et cela successivement pour $i = 2, \dots, m$; puis remplacer b_j par 0 en multipliant à droite par $G_1^j(-b^j)$, et cela successivement pour $j = 2, \dots, n$. Au bout de ces opérations, qui sont toutes des équivalences, la première ligne et la première colonne, en dehors de leur case commune qui contient 1, ne contiennent que des zéros. C'est la première étape de la réduction.

Si les $(m - 1)(n - 1)$ autres termes sont tous nuls, la réduction est terminée. Sinon l'étape suivante consistera à nouveau, par des multiplications convenables par des matrices des types E, D et F, et sans rien changer à la 1ère ligne ni à la 1ère colonne, à amener 1 dans la case de 2ème ligne et 2ème colonne et 0 dans tout le reste de la 2ème ligne et de la 2ème colonne.

Le rang étant r au départ, et restant inchangé dans les équivalences successives pratiquées au cours de chaque étape, c'est au bout de r étapes de ce genre que l'on aura donné aux r premières lignes et colonnes l'aspect

$$\left[\begin{array}{cc} I_r & 0 \\ 0 & \end{array} \right],$$

et que les $(m - r)(n - r)$ termes restants seront tous nuls.

L'équivalence avec la forme canonique annoncée est ainsi démontrée.

Bien entendu dans les cas particuliers où l'on a $m = r$ ou $n = r$, on obtient les formes canoniques particulières

$$\boxed{\begin{array}{cc} I_r & 0 \end{array}} \quad \text{ou} \quad \boxed{\begin{array}{c} I_r \\ 0 \end{array}}$$

Si $m = n = r$, la matrice A est inversible, et c'est $\boxed{I_r}$ qui constitue la matrice canonique. Ceci prouve notamment que toute matrice inversible est décomposable en produit de matrices des types E, D et F ; le procédé qui a été indiqué pour la réduction à la forme canonique peut ainsi servir de base à un procédé pratique d'inversion.

Enfin l'équivalence d'une matrice de rang r avec une matrice canonique permet d'établir très simplement les propriétés suivantes :

1/ Le rang d'un produit de matrices est au plus égal au plus petit des rangs des facteurs.

2/ Le rang d'une matrice A de format $m \times n$ est le plus petit nombre ρ tel que A soit le produit d'une matrice de format $m \times \rho$ par une matrice de format $\rho \times n$ (cette propriété est parfois adoptée comme définition du rang).

2.7 - Systèmes d'équations linéaires

Le système le plus général de m équations linéaires à n inconnues peut s'écrire :

$$AX = B,$$

où A est la matrice des coefficients des inconnues, de format $m \times n$, X est le vecteur-colonne des n inconnues, B le vecteur-colonne des m "seconds membres".

Si A est de rang r , on a $A = PC_rQ$, avec P et Q inversibles. On a alors :

$$PC_rQX = B$$

$$\text{d'où} \quad C_r(QX) = P^{-1} B$$

$$\text{d'où} \quad C_r X' = B'$$

(à la condition de poser $QX = X'$ ou $X = Q^{-1} X'$)

$$\text{et} \quad P^{-1} B = B' \quad \text{ou} \quad B = PB'$$

L'équation $C_r X' = B'$, à laquelle on a ramené l'équation $AX = B$, peut s'écrire sous la forme développée :

$$\begin{array}{rcl}
 x_1' & = & b_1' \\
 x_2' & = & b_2' \\
 \vdots & & \vdots \\
 x_r' & = & b_r' \\
 \hline
 0 & = & b_{r+1}' \\
 \vdots & & \vdots \\
 0 & = & b_m'
 \end{array}$$

On voit que les $m - r$ dernières lignes expriment des conditions (linéaires) auxquelles doivent satisfaire les seconds membres pour que le système soit "compatible", c'est-à-dire possède au moins une solution.

Si ces $m - r$ conditions sont remplies, il subsiste $n - r$ indéterminations pour les inconnues.

Il ne suffit donc nullement, comme certains utilisateurs sont parfois tentés de le croire, de comparer le nombre d'équations et celui des inconnues pour voir si un système est "impossible", "déterminé" ou indéterminé". Exemples :

$$\begin{array}{ll}
 2x - 6y + 4z = 3 & \text{système impossible, bien qu'il y ait plus} \\
 - 5x + 15y - 10z = 7 & \text{d'inconnues que d'équations.} \\
 \\
 10x + 14y = 6 & \text{système à 1 indétermination, bien qu'il} \\
 35x + 49y = 21 & \text{y ait plus d'équations que d'inconnues.} \\
 15x + 21y = 9 &
 \end{array}$$

3 - ANALYSE SPECTRALE

3.1 - Préliminaires

1/ Déterminants

La notion de déterminant est ici supposée connue. Les propriétés utilisées seront les suivantes :

a) le déterminant d'une matrice carrée A de format $n \times n$ est *un nombre* $\det(A)$ égal à la somme (alternée) de $n!$ produits de n facteurs, chacun des produits faisant intervenir un seul terme par ligne et par colonne, et le signe étant $+$ pour le produit des termes diagonaux.

b) la condition nécessaire et suffisante pour qu'une matrice A soit inversible est que son déterminant $\det(A)$ ne soit pas nul.

c) le déterminant du produit de deux matrices est le produit de leurs déterminants : $\det(AB) = \det(A) \times \det(B)$.

2/ Similitude

Deux matrices carrées A et B de même format sont dites *semblables* s'il existe une matrice inversible P telle que :

$$A = PBP^{-1}$$

La relation ainsi définie est réflexive, symétrique et transitive. Elle est plus "fine" que la relation dite "équivalence" définie plus haut au § 2.5.

Si $A = PBP^{-1}$, on voit sans peine que pour tout exposant entier non-négatif k on a $A^k = PB^kP^{-1}$ (on pose par convention $A^0 = I$).

Deux matrices semblables ont des déterminants égaux, puisque $\det(P) \det(P^{-1}) = \det(I) = 1$.

3/ Polynômes matriciels

Etant donné un polynôme quelconque de degré k d'une variable x

$$\varphi(x) = x^k + c_1 x^{k-1} + \dots + c_{k-1} x + c_k,$$

dont les coefficients $c_1 \dots c_k$ appartiennent au corps K, et étant donné une matrice A de format $n \times n$ définie sur le corps K, on désigne par $\varphi(A)$ la matrice (de même format $n \times n$)

$$\varphi(A) = A^k + c_1 A^{k-1} + \dots + c_{k-1} A + c_k I,$$

appelée polynôme matriciel en A.

Il résulte de cette définition

a) qu'à la combinaison linéaire $\varphi(x) = p_1 \varphi_1(x) + p_2 \varphi_2(x)$ et au produit $\Psi(x) = \varphi_1(x) \times \varphi_2(x)$ de deux polynômes en x correspondent respectivement les polynômes matriciels

$$\varphi(A) = p_1 \varphi_1(A) + p_2 \varphi_2(A)$$

et $\Psi(A) = \varphi_1(A) \times (\varphi_2(A))$; c'est évident pour la combinaison linéaire, et pour le produit cela vient de ce que $A^{u+v} = A^u A^v = A^v A^u$ (dans le cas particulier de deux puissances d'une même matrice carrée A, le produit matriciel est commutatif).

b) que si $A = PBP^{-1}$, alors on a (quel que soit le polynôme $\varphi(x)$) $\varphi(A) = P\varphi(B)P^{-1}$.

4/ Objet de l'analyse spectrale

L'un des objets essentiels de l'analyse spectrale est l'étude des propriétés des puissances successives d'une matrice carrée A ("progressions géométriques matricielles").

Deux points de vue se prêtent dans cette étude un appui mutuel :

a) recherche de matrices *semblables* à A dont la forme soit assez simple pour se prêter à un calcul direct.

b) recherche, dans l'espace des vecteurs-colonne X, de *sous-espaces invariants* par la transformation $X' = AX$ (en abrégé : "sous-espaces invariants de A").

3.2 - Matrices nilpotentes

1/ Définition

Soit par exemple la matrice $B = \begin{bmatrix} -10 & 4 \\ -25 & 10 \end{bmatrix}$. On voit que $B^2 = 0$, et par conséquent $B^3 = B^4 = \dots = 0$.

Il existe donc des matrices qui, sans être nulles elles-mêmes, ont des puissances successives qui sont nulles à partir d'un certain rang. De telles matrices sont dites *nilpotentes*.

Si B est de format $n \times n$ et si l'on a $B^k = 0$ avec $B^{k-1} \neq 0$, k s'appelle l'*ordre de nilpotence*. Une matrice nulle a pour ordre de nilpotence 1.

2/ Une classe d'exemples

Appelons U_n la matrice, de format $n \times n$,

$$U_n = E_1^2 + E_2^3 + \dots + E_{n-1}^n$$

(termes "surdiagonaux" égaux à 1, autres termes nuls).

On a :

$$(U_n)^2 = E_1^3 + E_2^4 + \dots + E_{n-2}^n$$

$$(U_n)^3 = E_1^4 + E_2^5 + \dots + E_{n-3}^n$$

.....

$$(U_n)^{n-1} = E_1^n$$

$$(U_n)^n = 0$$

U_n est donc une matrice nilpotente, d'ordre de nilpotence n . Il résulte des propriétés des matrices E_s^t que, $X_1 X_2 \dots X_n$ étant des vecteurs-colonne quelconques, on a :

$$\boxed{X_1 X_2 \dots X_{n-1} X_n} \cdot U_n = \boxed{0 X_1 \dots X_{n-2} X_{n-1}}$$

Par convention complémentaire, on pose $U_1 = 0$.

3/ Sous-espaces invariants d'une matrice nilpotente

Soit B une matrice $n \times n$ nilpotente, d'ordre de nilpotence k .

Si Z est un vecteur quelconque de l'espace vectoriel S à n dimensions des vecteurs-colonne à n composantes, appelons BZ le *conséquent* de Z . Pour tout Z de S , on a $B^k Z = 0$ (puisque $B^k = 0$). Mais si i est un entier donné ($0 \leq i \leq k$), on a, pour certains Z de S , $B^i Z = 0$; les vecteurs correspondants forment manifestement un sous-espace vectoriel S_i de S . De plus, puisque $B^i Z = 0 \implies B^{i+1} Z = 0$, on a $S_i \subset S_{i+1}$. Enfin, puisque $B^{k-1} Z \neq 0$ par hypothèse, il existe un Z tel que $B^{k-1} Z \neq 0$, c'est-à-dire un Z qui est dans $S_k (= S)$ sans être dans S_{k-1} ; son conséquent BZ est donc dans S_{k-1} sans être dans S_{k-2} , et ainsi de suite jusqu'à $B^{k-1} Z$ qui est dans S_1 sans être dans $S_0 (= \{0\})$.

Il existe donc une suite de sous-espaces invariants

$$\{0\} = S_0 \subset S_1 \subset S_2 \subset \dots \subset S_{k-1} \subset S_k = S,$$

chacun d'eux ayant au moins une dimension de plus que le précédent, Comme S est à n dimensions, on en conclut que $k \leq n$: l'ordre de nilpotence d'une matrice $n \times n$ ne dépasse jamais n .

4/ Vecteurs de S_{i+1} indépendants par rapport à S_i

Soit dans $S_{i+1} - S_i$ (qui n'est pas un sous-espace vectoriel!) β vecteurs $Z^1 Z^2 \dots Z^\beta$ indépendants par rapport à S_i , c'est-à-dire tels que :

$$\lambda_1 Z^1 + \lambda_2 Z^2 + \dots + \lambda_\beta Z^\beta \in S_i \implies \lambda_1 = \lambda_2 = \dots = \lambda_\beta = 0$$

Leurs conséquents BZ^1, BZ^2, BZ^β appartiennent alors à $S_i - S_{i-1}$. Supposons que

$$(1) \quad \lambda_1 (BZ^1) + \lambda_2 (BZ^2) + \dots + \lambda_\beta (BZ^\beta) \in S_{i-1};$$

le produit par B^{i-1} du premier membre est nul par définition de S_{i-1} , d'où $B^i (\lambda_1 Z^1 + \dots + \lambda_\beta Z^\beta) = 0$; ceci équivaut à $\lambda_1 Z^1 + \dots + \lambda_\beta Z^\beta \in S_i$, ce qui entraîne :

$$(2) \quad \lambda_1 = \lambda_2 = \dots = \lambda_\beta = 0$$

De (1) on a déduit (2), ce qui montre finalement que les conséquents de β vecteurs de $S_{i+1} - S_i$ indépendants par rapport à S_i sont β vecteurs de $S_i - S_{i-1}$ indépendants par rapport à S_{i-1} .

Or si l'on appelle δ_i le nombre de dimensions de S_i , on peut trouver dans $S_{i+1} - S_i$ un nombre de vecteurs, indépendants par rapport à S_i , égal à $\delta_{i+1} - \delta_i = \beta_{i+1}$ (système *maximal*). Si l'on applique le résultat ci-dessus à $\beta = \beta_{i+1}$ vecteurs, on en conclut que :

$$\beta_{i+1} \leq \beta_i$$

La suite des k entiers positifs $\beta_1 \beta_2 \dots \beta_k$ est donc non-croissante ; et l'on a $\beta_1 + \beta_2 + \dots + \beta_k = \delta_k - \delta_0 = n$.

5/ Formation d'un tableau de vecteurs

Partons d'un système de β_k vecteurs de $S_k - S_{k-1}$, indépendants par rapport à S_{k-1} , soit $Z_k^1 Z_k^2 \dots Z_k^{\beta_k}$, et formons ligne par ligne le tableau suivant :

$$(T) \quad \begin{array}{cccc} Z_k^1 & Z_k^2 & \dots & Z_k^{\beta_k} & (\in S_k) \\ Z_{k-1}^1 & Z_{k-1}^2 & \dots & Z_{k-1}^{\beta_{k-1}} & (\in S_{k-1}) \\ \dots & \dots & \dots & \dots & \dots \\ Z_1^1 & Z_1^2 & \dots & \dots & Z_1^{\beta_1} (\in S_1) \end{array}$$

Chaque ligne commence par les conséquents respectifs des vecteurs écrits dans la ligne au-dessus, puis est complétée s'il y a lieu de manière à former un système *maximal* de vecteurs indépendants par rapport au sous-espace vectoriel suivant. Les n vecteurs de (T) sont indépendants : on l'établit aisément, par exemple en remontant le tableau à partir de la dernière ligne.

On désignera dans ce qui suit par $\alpha_1 \alpha_2 \dots \alpha_h$ les nombres de vecteurs écrits dans les *colonnes* successives de (T) : $\alpha_1 = k$, $\beta_1 = h$. La suite $\alpha_1 \alpha_2 \dots \alpha_h$, dite conjuguée de la suite $\beta_1 \beta_2 \dots \beta_k$, est elle aussi une suite d'entiers positifs non-croissante de somme n .

6/ Forme réduite de Jordan

La règle même de formation du tableau (T) entraîne que, pour les α_j vecteurs de la j -ième colonne, on a :

$$B. \quad \boxed{Z_1^j \ Z_2^j \ \dots \ Z_{\alpha_j}^j} = \boxed{0 \ Z_1^j \ \dots \ Z_{\alpha_j-1}^j} = \boxed{Z_1^j \ Z_2^j \ \dots \ Z_{\alpha_j}^j} \cdot U_{\alpha_j}$$

U_{α_j} étant une matrice du type défini en (II, 2°).

Plus brièvement

$$BQ_j = Q_j U_{\alpha_j} \quad (Q_j \text{ de format } n \times \alpha_j)$$

Les h égalités ci-dessus, pour $j = 1, 2, \dots, h$, se résument par l'écriture en blocs

$$B \cdot \begin{bmatrix} Q_1 & Q_2 & \dots & Q_h \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 & \dots & Q_h \end{bmatrix} \times \langle U_{\alpha_1} \ U_{\alpha_2} \ \dots \ U_{\alpha_h} \rangle$$

La matrice $Q = \begin{bmatrix} Q_1 & Q_2 & \dots & Q_h \end{bmatrix}$ est de format $n \times n$ et ses colonnes, qui sont les vecteurs du tableau (T), sont indépendantes; elle est donc inversible.

Quant à la matrice $J = \langle U_{\alpha_1} \ U_{\alpha_2} \ \dots \ U_{\alpha_h} \rangle$, elle est "diagonale par blocs": chacun des h blocs diagonaux est une matrice ayant ses termes surdiagonaux égaux à 1 et tous ses autres termes nuls, et chacun des $h(h-1)$ autres blocs est entièrement nul.

On a donc montré que

$$B = QJQ^{-1}$$

toute matrice nilpotente A est semblable à une matrice du type J , appelée *forme réduite de Jordan* de la matrice B ; l'ordre de nilpotence k est la dimension (nombre de lignes) α_1 du plus grand bloc diagonal $U_{\alpha_1} = U_k$.

4 - REDUCTION CANONIQUE DES MATRICES CARREES QUELCONQUES

Soit M une matrice carrée quelconque, de format $m \times m$, définie sur le corps K , et soit E l'espace des vecteurs-colonne X à m composantes dans K .

4.1 - Droites invariantes de E ; valeurs propres; vecteurs propres

Une droite de E est par définition un sous-espace vectoriel à 1 dimension, c'est-à-dire l'ensemble des vecteurs X "homothétiques" d'un vecteur donné X non nul.

Dire que la droite définie par un vecteur non nul X est invariante revient à dire qu'il existe dans K un λ tel que

$$MX = \lambda X,$$

ce qui peut s'écrire

$$(3) \quad (\lambda I - M) X = 0 \quad X \neq 0$$

La recherche d'une droite invariante peut donc se conduire comme suit :

a) chercher un λ tel que la matrice $\lambda I - M$ soit à colonnes liées, c'est-à-dire non-inversible ; un tel λ sera appelé une *valeur propre* (ou valeur spectrale) de M .

b) une fois trouvée une valeur propre λ_0 , trouver un vecteur X_0 satisfaisant à (3) ; un tel vecteur sera appelé un *vecteur propre* de la valeur propre λ_0 .

4.2 - Polynôme caractéristique de M

La recherche d'une valeur propre de M conduit à exprimer que la matrice $\lambda I - M$ est non-inversible, c'est-à-dire que son déterminant est nul :

$$\det(\lambda I - M) = 0$$

Or ce déterminant, pour λ indéterminé, est un polynôme $f(\lambda)$ de degré m en λ , dont les coefficients peuvent s'exprimer à l'aide des termes de M :

$$f(\lambda) = \lambda^m + c_1 \lambda^{m-1} + \dots + c_m$$

Ce polynôme est appelé *polynôme caractéristique* de la matrice M . Les valeurs propres, s'il en existe, sont donc les zéros du polynôme caractéristique, ou les racines de l'*équation caractéristique* (ou équation spectrale) $f(\lambda) = 0$.

On supposera dorénavant que le corps K est celui des nombres complexes (C), de manière à bénéficier du théorème de d'Alembert, qui permet, dans C , de décomposer $f(\lambda)$ en facteur du premier degré du type $\lambda - a$. L'ensemble des valeurs propres (ou *spectre*) est alors non-vidé, et l'on a :

$$f(\lambda) = (\lambda - a_1)^{n_1} (\lambda - a_2)^{n_2} \dots (\lambda - a_p)^{n_p}$$

* En allemand "Eigenwert" et "Eigenvektor", d'où les expressions anglaises à connaître "eigen-value" et "eigen-vector".

a_1, a_2, \dots, a_p sont des valeurs propres distinctes et la somme $n_1 + n_2 + \dots + n_p$ de leurs "ordres de multiplicité" est égale à m .

4.3 - Cas de m valeurs propres distinctes

Dans ce cas, dit parfois "régulier", on a $p = m$, $n_1 = n_2 = \dots = n_p = 1$. Il y a m valeurs propres a_1, a_2, \dots, a_m distinctes, auxquelles on peut faire correspondre des vecteurs propres X_1, X_2, \dots, X_m (dont chacun est défini à un facteur non nul près).

On a alors :

$$MX_1 = a_1 X_1 \quad MX_2 = a_2 X_2 \quad \dots \quad MX_m = a_m X_m,$$

ce qui se résume par :

$$(4) \quad M \begin{bmatrix} X_1 & X_2 & \dots & X_m \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & \dots & X_m \end{bmatrix} \langle a_1 \ a_2 \ \dots \ a_m \rangle$$

Or les m vecteurs propres sont indépendants. Supposons en effet établi que les $i - 1$ premiers d'entre eux le sont, et supposons que

$$(5) \quad \lambda_1 X_1 + \dots + \lambda_i X_i = 0$$

On a alors, en prémultipliant par M ,

$$(6) \quad \lambda_1 a_1 X_1 + \dots + \lambda_i a_i X_i = 0$$

et, en combinant linéairement (5) et (6) avec les coefficients a_i et -1 ,

$$\lambda_1(a_i - a_1) X_1 + \dots + \lambda_{i-1}(a_i - a_{i-1}) X_{i-1} = 0$$

Les différences entre parenthèses sont toutes non nulles puisque les a_i sont distincts, et $X_1 \dots X_{i-1}$ sont indépendants, donc $\lambda_1 = \dots = \lambda_{i-1} = 0$, et alors $\lambda_i = 0$ par suite de (5). $X_1 \dots X_i$ sont donc indépendants. On établit ainsi de proche en proche que $X_1 X_2 \dots X_m$ sont indépendants.

La matrice $P = \begin{bmatrix} X_1 & X_2 & \dots & X_m \end{bmatrix}$ est donc inversible, et si l'on appelle Δ la matrice diagonale $\langle a_1 \ a_2 \ \dots \ a_m \rangle$, (4) s'écrit :

$$(7) \quad M = P \Delta P^{-1}$$

M est alors semblable à une matrice diagonale, dont les termes diagonaux sont les valeurs propres.

4.4 - Théorème de Cayley-Hamilton

1/ Cas régulier

Toujours dans le cas régulier, on peut déduire de (7), grâce à la remarque faite en 1,5°, (b), que :

$$f(M) = Pf(\Delta)P^{-1}$$

Or $f(\Delta)$ est une matrice diagonale dont les termes diagonaux sont $f(a_1) f(a_2) \dots f(a_m)$, expressions qui sont nulles puisque $a_1 a_2 \dots a_m$ sont précisément les zéros du polynôme caractéristique $f(\lambda)$. Donc $f(\Delta) = 0$, ce qui entraîne

$$f(M) = 0$$

La matrice M satisfait, comme on dit, à sa propre équation caractéristique. Ce résultat est connu sous le nom de théorème de Cayley-Hamilton.

2/ Exemple : m = 2

Par exemple si $m = 2$, on a :

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \lambda I - M = \begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix}$$

$$f(\lambda) = \lambda^2 - (a + d)\lambda + ad - bc$$

Pour que les valeurs propres soient distinctes il faut et il suffit que

$$(8) \quad (a + d)^2 - 4(ad - bc) \neq 0$$

Le théorème de Cayley-Hamilton nous apprend que sous cette condition (8), on a :

$$M^2 - (a + d)M + (ad - bc)I = 0$$

ce qui résume les quatre égalités numériques

$$(9) \quad \begin{array}{lll} a^2 + bc - (a + d)a + ad - bc = 0 & ab + bd - (a + d)b & = 0 \\ ca + dc - (a + d)c & = 0 & cb + d^2 - (a + d)d + ad - bc = 0 \end{array}$$

Ces égalités expriment que quatre polynômes particuliers des quatre variables a, b, c, d sont nuls pourvu qu'un cinquième polynôme, qui est celui du 1er membre de (8), ne le soit pas.

3/ Cas général

Il est manifeste, sur l'exemple ci-dessus, que la condition (8) ne joue en fait aucun rôle, c'est-à-dire que les égalités (9) sont des identités, valables même s'il y avait une valeur double, et même encore si M était défini sur un autre corps que C , qui ne permettrait pas nécessairement la décomposition de $f(\lambda)$ en facteurs du premier degré.

Il résulte des propriétés générales des polynômes de plusieurs variables qu'il en sera de même si M est une matrice $m \times m$ quelconque : le théorème de Cayley-Hamilton, établi ci-dessus dans le cas régulier, est vrai dans *tous les cas*, même s'il y a des valeurs spectrales multiples, et même si (comme cela peut arriver sur d'autres corps que C) il n'y a pas de spectre du tout : *toute matrice carrée satisfait à sa propre équation caractéristique.*

4.5 - Première réduction

On a vu, dans le cas général (sur le corps C), que l'équation caractéristique est :

$$f(\lambda) = (\lambda - a_1)^{n_1} (\lambda - a_2)^{n_2} \dots (\lambda - a_p)^{n_p} = 0. \quad i \neq j \implies a_i \neq a_j$$

$$n_1 + \dots + n_p = m$$

On a donc, en vertu du théorème de Cayley-Hamilton,

$$f(M) = (M - a_1 I)^{n_1} (M - a_2 I)^{n_2} \dots (M - a_p I)^{n_p} = 0,$$

d'où il résulte que pour tout vecteur X de l'espace E , on a

$$f(M). X = 0$$

Nous nous intéressons au sous-espace vectoriel E_i des vecteurs X tels que

$$(M - a_i I)^{n_i} X = 0$$

Nous aurons notamment besoin pour cela du lemme ci-dessous.

1/ *Lemme.* Si quatre matrices carrées ($m \times m$) H, H', K, K' sont deux à deux *permutables*, et sont telles que $H'H + K'K = I$, alors les es-

paces vectoriels \mathcal{H} et \mathcal{K} définis respectivement par $HX = 0$ et par $KX = 0$ ont pour *somme directe* l'espace vectoriel défini par $(HK)X = 0$.

Il s'agit d'établir d'une part que \mathcal{H} et \mathcal{K} sont indépendants, c'est-à-dire n'ont aucun vecteur non nul commun, d'autre part que tout vecteur X tel que $HKX = 0$ est somme d'un vecteur de \mathcal{H} et d'un vecteur de \mathcal{K} .

La première partie est pratiquement évidente puisque $HX = 0$ et $KX = 0$ entraînent $H'HX + K'KX = 0$, et que le 1er membre de cette égalité n'est autre que X par suite de l'hypothèse $H'H + K'K = I$.

Pour établir la 2ème partie, appelons X un vecteur tel que $HKX = 0$. Pour ce vecteur (comme d'ailleurs pour n'importe quel vecteur de E) on a :

$$H'HX + K'KX = X$$

Or le vecteur $H'HX$ appartient à \mathcal{H} puisque $KH'HX = 0$ (en effet $KH'H$ est égal, par suite des hypothèses de permutabilité, à $H'HK$, et $H'HKX = 0$ puisque $HKX = 0$) ; pour une raison analogue le vecteur $K'KX$ appartient à \mathcal{K} . Le lemme est donc établi.

2/ Invariance et indépendance des E_i

a) Si $X \in E_i$, $MX \in E_i$ (E_i est invariant par M). En effet les matrices M et $(M - a_i I)^{n_i}$ sont permutables puisque la seconde est une combinaison linéaire de puissances de la première. Donc $X \in E_i$ entraîne que $(M - a_i I)^{n_i} MX = M(M - a_i I)^{n_i} X = M \cdot 0 = 0$, et par conséquent $MX \in E_i$.

b) E est la somme directe de E_1, E_2, \dots, E_p .

Ceci résulte de $p - 1$ applications successives du lemme, pour $i = 1, 2, \dots, p - 1$, en prenant pour matrice H le produit des i premières matrices de $\frac{(M - a_1 I)^{n_1}}{(M - a_i I)^{n_i}}$ à $(M - a_1 I)^{n_1}$ et pour K la matrice $(M - a_{i+1} I)^{n_{i+1}}$. L'existence des matrices H' et K' est garantie par suite du fait que les deux polynômes $h(\lambda) = (\lambda - a_1)^{n_1} \dots (\lambda - a_i)^{n_i}$ et $k(\lambda) = (\lambda - a_{i+1})^{n_{i+1}}$ sont premiers entre eux ; on a en effet alors l'identité de Bezout, qui énonce l'existence de deux polynômes $h'(\lambda)$ et $k'(\lambda)$ tels que $h'(\lambda) h(\lambda) + k'(\lambda) k(\lambda) \equiv 1$; et les expressions matricielles $h'(M)$ et $k'(M)$ donnent bien des matrices H' et K' permutables avec H et K , et telles que $H'H + K'K = I$.

A la dernière étape ($i = p - 1$) l'ensemble des X tels que $HKH = 0$ est bien l'espace E tout entier en raison de la remarque initiale tirée du théorème de Cayley-Hamilton.

3/ Diagonalisation par blocs

En appelant provisoirement v_i le nombre de dimensions de E_i ($v_1 + v_2 + \dots + v_p = m$), soit $P_i = \begin{bmatrix} X_1 & X_2 & \dots & X_{v_i} \end{bmatrix}$ une matrice dont les v_i colonnes forment une *base* de E_i . Comme E_i est invariant par M , le transformé par M de chacun de ces v_i vecteurs sera une combinaison linéaire de ces v_i vecteurs, ce qui peut se résumer par :

$$(10) \quad MP_i = P_i A_i,$$

A_i étant une matrice carrée ($v_i \times v_i$) dont la j -ième colonne est constituée par les composantes de MX_j dans la base P_i ($j = 1, 2, \dots, v_i$).

Si l'on donne à i toutes les valeurs de 1 à p , l'ensemble des égalités matricielles $MP_i = P_i A_i$ peut se résumer par :

$$M \begin{bmatrix} P_1 & P_2 & \dots & P_p \end{bmatrix} = \begin{bmatrix} P_1 & P_2 & \dots & P_p \end{bmatrix} \langle A_1 \ A_2 \ \dots \ A_p \rangle$$

Comme les colonnes des blocs P_i sont des bases d'espaces E_i indépendants et que $v_1 + v_2 + \dots + v_p = m$, la matrice $P = \begin{bmatrix} P_1 & P_2 & \dots & P_p \end{bmatrix}$ est inversible ; en appelant \overline{M} la matrice $\langle A_1 \ A_2 \ \dots \ A_p \rangle$, on a donc

$$(11) \quad M = \overline{M}P^{-1}$$

M est donc semblable à une matrice diagonale par blocs, formée d'autant de blocs diagonaux qu'il y a de valeurs propres distinctes.

4.6 - Réduction canonique ultime

1/ Nilpotence des $A_i - a_i I$

L'égalité (10) entraîne que pour tout polynôme $\varphi(\lambda)$ on a $\varphi(M)P_i = P_i \varphi(A_i)$, puisque $M^h P_i = P_i A_i^h$. Notamment :

$$(M - a_i I)^{n_i} P_i = P_i (A_i - a_i I)^{n_i}.$$

Le premier membre est nul par définition de E_i , dont P_i est une base ; donc aussi le second. Mais, puisque la matrice P_i est à colonnes indépendantes,

$$P_i (A_i - a_i I)^{n_i} = 0 \implies (A_i - a_i I)^{n_i} = 0$$

$A_i - a_i I$ est donc une matrice nilpotente, de format $\nu_i \times \nu_i$ et d'ordre de nilpotence au plus égal à n_i .

2/ Polynôme caractéristique de A_i

$A_i - a_i I$, étant nilpotente, est semblable à sa forme réduite de Jordan J_i , définie en II, 6° ; d'où

$$(12) \quad \begin{aligned} A_i - a_i I &= Q_i J_i Q_i^{-1} \\ A_i &= Q_i (a_i I + J_i) Q_i^{-1} \end{aligned}$$

A_i et $a_i I + J_i$ étant semblables, il en est de même de $\lambda I - A_i$ et $\lambda I - (a_i I + J_i) = (\lambda - a_i)I - J_i$, qui ont par conséquent même déterminant (cf. I, 4°). Or le déterminant de $(\lambda - a_i)I - J_i$ est d'un calcul immédiat du fait de la nullité de tous les termes autres que ceux de la diagonale et de la surdiagonale : il est égal à $(\lambda - a_i)^{\nu_i}$.

C'est donc $(\lambda - a_i)^{\nu_i}$ qui est le polynôme caractéristique de A_i : $\det(\lambda I - A_i) = (\lambda - a_i)^{\nu_i}$.

3/ Egalité de ν_i et de n_i

$$\text{De } \overline{M} = \langle A_1 \ A_2 \ \dots \ A_p \rangle$$

on déduit :

$$\lambda I - \overline{M} = \langle \lambda I - A_1 \ \lambda I - A_2 \ \dots \ \lambda I - A_p \rangle$$

La forme, diagonale par blocs, du second membre permet de voir que son déterminant est le produit des déterminants des p blocs, d'où

$$(13) \quad \det(\lambda I - \overline{M}) = (\lambda - a_1)^{\nu_1} (\lambda - a_2)^{\nu_2} \dots (\lambda - a_p)^{\nu_p}$$

Mais d'autre part on a supposé que $f(\lambda) =$

$$(14) \quad \det(\lambda I - M) = (\lambda - a_1)^{n_1} (\lambda - a_2)^{n_2} \dots (\lambda - a_p)^{n_p}$$

Comme $\lambda I - \overline{M}$ et $\lambda I - M$, qui sont semblables en conséquence de (11), doivent avoir même déterminant, on en déduit, par comparaison de (13) et (14), que les ν_i sont respectivement égaux aux n_i .

4/ Forme canonique ultime

Les égalités (12), pour $i = 1, 2, \dots, p$, peuvent se résumer par l'écriture, diagonale par blocs et de format total $m \times m$:

$$\begin{aligned} (\overline{M} =) \langle A_i \rangle &= \langle Q_i \rangle \times \langle a_i I + J_i \rangle \times \langle Q_i \rangle^{-1}, \\ &= \langle Q_i \rangle \times \overline{M} \times \langle Q_i \rangle^{-1} \end{aligned}$$

qui montre que les deux matrices \overline{M} et $\overline{\overline{M}} = \langle a_i I + J_i \rangle$, toutes deux diagonales en p blocs de dimensions $n_1 \ n_2 \ \dots \ n_p$, sont semblables.

Comme la "première réduction" (5°) a montré que \overline{M} était semblable à M , il en résulte que toute matrice carrée M peut se réduire par similitude à la forme $\overline{\overline{M}} = \langle a_i I + J_i \rangle$, appelée aussi forme canonique ultime ou *forme de Jordan* de la matrice M .

5 - POLYNOME MINIMAL ; PUISSANCES SUCCESSIVES

5.1 - Blocs diagonaux "fins" ; blocs "principaux"

Dans la forme réduite ultime M , considérons le bloc $a_i J + J_i$, relatif à la valeur spectrale a_i ; nous utiliserons, pour un moment, l'écriture plus légère (sans indice i) $aI + J$.

La matrice J , on l'a vu (II, 6°), est à son tour diagonale par blocs :

$$J = \langle U_{\alpha_1} \ U_{\alpha_2} \ \dots \ U_{\alpha_h} \rangle$$

$\alpha_1 \ \alpha_2 \ \dots \ \alpha_h$ est une suite d'entiers positifs non-croissante, de somme n (en réalité ici n_i), il n'y a donc pas de bloc de plus grand format que le premier $U_{\alpha_1} = U_k$. La matrice nilpotente J a pour ordre de nilpotence $\alpha_1 = k$ (en réalité k_i , qui est $\leq n_1$).

Il en résulte que :

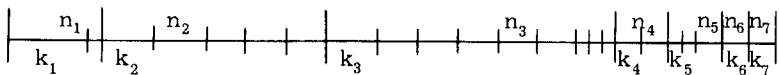
$$(15) \quad aI + J = \langle aI + U_{\alpha_1} \ aI + U_{\alpha_2} \ \dots \ aI + U_{\alpha_h} \rangle ;$$

d'où, en récrivant en indice le numéro de valeur propre i ,

$$a_i I + J_i = \langle a_i I + U_{k_i} \ \dots \rangle ;$$

les blocs non écrits sont tous du type $a_i I + U_{\alpha}$, l'entier positif α , qui est la dimension du bloc correspondant, ne dépassant jamais k_i .

La figure ci-dessous donne un exemple numérique de la façon dont les blocs diagonaux $a_i I + J_i$ peuvent se subdiviser en blocs diagonaux plus "fins" :



Dans cet exemple, on a $m = 38$, $p = 7$ (7 valeurs spectrales distinctes, d'ordres de multiplicité respectifs $n_i = 6, 10, 13, 4, 3, 1, 1$) ; il y a 7 "gros blocs", subdivisés en 22 "blocs fins", dont 7 "principaux" (premiers de chaque gros bloc), de dimensions respectives $k_i = 5, 3, 3, 2, 1, 1, 1$, $\sum k_i = 16$.

5.2 - Définition du polynôme minimal (ou "typique") de M

Il est intéressant, pour le calcul des puissances successives de M, de disposer d'un polynôme $g(\lambda)$, de degré aussi petit que possible, tel que $g(M) = 0$. Le polynôme caractéristique $f(\lambda)$, de degré m , possède bien la propriété voulue, mais rien ne garantit en général qu'il n'en existe pas d'autres, de degré $< m$.

Tout polynôme $\varphi(\lambda)$ tel que $\varphi(M) = 0$ est d'ailleurs divisible par $g(\lambda)$. Car sinon on a l'identité de la division (16) $\varphi(\lambda) = g(\lambda) \cdot Q(\lambda) + R(\lambda)$ avec (degré de R) $<$ (degré de g), d'où $\varphi(M) = g(M) \cdot Q(M) + R(M)$.

Si $\varphi(M) = 0$ et puisque $g(M) = 0$, on en déduit $R(M) = 0$; il y a alors un polynôme de degré inférieur à celui de $g(\lambda)$ qui est annulé par M, ce qui est incompatible avec la définition de $g(\lambda)$. La division (16) se fait donc sans reste, ce qui montre que $\varphi(\lambda)$ est multiple de $g(\lambda)$.

$g(\lambda)$ est appelé le polynôme *minimal* (ou *typique*) de M ; il est en particulier diviseur du polynôme caractéristique $f(\lambda)$, donc ne peut s'annuler pour d'autres valeurs que les valeurs propres $a_1 a_2 \dots a_p$ de M.

5.3 - Nature du polynôme minimal

Pour qu'un polynôme $\varphi(\lambda)$ soit annulé par M il faut (et il suffit) qu'il le soit par \overline{M} , puisque $\varphi(M)$ et $\varphi(\overline{M})$ sont semblables.

Mais pour que $\varphi(\lambda)$ soit annulé par M, il faut qu'il le soit par chacun des blocs diagonaux de M, en particulier par chacun des blocs diagonaux "fins", et plus particulièrement par les blocs "principaux" $a_i I + U_{k_i} = D_i$.

Or tout polynôme en D_i est un polynôme de même degré en $D_i - a_i I$, c'est-à-dire en U_{k_i} . La forme même des puissances successives de U_{k_i} (cf. II, 2°) montre qu'aucun polynôme de degré $< k_i$ en U_{k_i} ne peut être nul, et que $(U_{k_i})^{k_i} = 0 = (D_i - a_i I)^{k_i}$.

Le polynôme de plus bas degré annulé par D_i est donc $g_i(\lambda) = (\lambda - a_i)^{k_i}$. Il est clair que ce même polynôme $g_i(\lambda)$ est annulé par tout autre "bloc fin" appartenant au même "gros bloc" (de

la même valeur propre a_i), puisque k_i est au moins égal à la dimension de ce bloc fin. Et tout polynôme $\varphi(\lambda)$ annulé par D_i est multiple de $(\lambda - a_i)^{k_i}$.

On voit ainsi que le polynôme minimal $g(\lambda)$ de M est un multiple commun, et en fait simplement le produit, des $g_i(\lambda)$ pour $i = 1, 2, \dots, p$:

$$g(\lambda) = (\lambda - a_1)^{k_1} (\lambda - a_2)^{k_2} \dots (\lambda - a_p)^{k_p}$$

Il a les mêmes zéros que le polynôme caractéristique, mais avec des ordres de multiplicité qui peuvent être moindres ($1 \leq k_i \leq n_i$).

5.4 - Puissances successives de M

En effectuant la "première réduction" (diagonalisation par "gros" blocs, III, 5°, (c)), on a écrit l'égalité

$$(11) \quad M = PMP^{-1},$$

dans laquelle P était une écriture condensée de

$$P = \boxed{P_1 \ P_2 \ \dots \ P_p}$$

Si l'on adopte pour P^{-1} l'écriture par blocs

$$P^{-1} = \begin{array}{|c} P_1' \\ P_2' \\ \vdots \\ P_p' \end{array}$$

on voit que, pour tout entier non-négatif s , $M^s = \overline{PM}^s P^{-1}$ peut s'écrire :

$$(17) \quad M^s = P_1(A_1)^s P_1' + P_2(A_2)^s P_2' + \dots + P_p(A_p)^s P_p' \quad (\text{format } m \times m).$$

Or de l'égalité (12) (III, 6°, (b)) il résulte que

$$(A_i)^s = Q_i(a_i I + J_i)^s Q_i^{-1} \quad (\text{format } n \times n)$$

ou, sans indices pour alléger,

$$A^s = Q(aI + J)^s Q^{-1}.$$

Comme précédemment on peut écrire Q en h blocs de colonnes

$$\boxed{R_1 \ R_2 \ \dots \ R_h}, \quad Q^{-1} \text{ en } h \text{ blocs de lignes } \begin{bmatrix} R_1' \\ R_2' \\ \vdots \\ R_h' \end{bmatrix}, \text{ d'où, en tenant compte}$$

de (15)

$$(18) \ A^s = R_1(aI + U_{a_1})^s R_1' + R_2(aI + U_{a_2})^s R_2' + \dots + R_h(aI + U_{a_h})^s R_h'$$

Mais, par développement binomial, on a :

$$(19) \ (aI + U_a)^s = a^s I + \binom{s}{1} a^{s-1} U_a + \binom{s}{2} a^{s-2} (U_a)^2 + \dots \\ \dots + \binom{s}{k-1} a^{s-k+1} (U_a)^{k-1}, \quad (a \leq k)$$

l'addition s'arrêtant au *plus tard* au bout des k termes explicités, en appelant $k(k_i)$ la dimension du "bloc fin principal" relatif à la valeur propre $a(a_i)$.

En combinant (17), (18) et (19), on voit que M^s peut s'écrire comme combinaison linéaire de $k_1 + k_2 + \dots + k_p$ matrices *fixes* (indépendantes de s), les coefficients scalaires de ces matrices faisant seuls intervenir s ; plus précisément :

$$M^s = a_1^s M_{1,1} + \binom{s}{1} a_1^{s-1} M_{1,2} + \dots + \binom{s}{k_1-1} a_1^{s-k_1+1} M_{1,k_1} \\ \left[\begin{array}{l} \text{par convention} \\ s < x \implies \binom{s}{x} = 0 \end{array} \right] \\ + a_2^s M_{2,1} + \binom{s}{1} a_2^{s-1} M_{2,2} + \dots + \binom{s}{k_2-1} a_2^{s-k_2+1} M_{2,k_2} \\ + \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ + a_p^s M_{p,1} + \binom{s}{1} a_p^{s-1} M_{p,2} + \dots + \binom{s}{k_p-1} a_p^{s-k_p+1} M_{p,k_p} \quad (20)$$

5.5 - Equivalence à un calcul par récurrence

Le nombre de coefficients matriciels fixes ci-dessus,

$$k_1 + k_2 + \dots + k_p = \mu \leq m$$

est précisément le degré du polynôme minimal $g(\lambda)$.

Si l'on développe ce polynôme sous la forme :

$$g(\lambda) = \lambda^\mu + g_1 \lambda^{\mu-1} + \dots + g_\mu,$$

on a

$$M^\mu + g_1 M^{\mu-1} + \dots + g_\mu I = 0$$

d'où, en multipliant par $M^{s-\mu}$

$$M^s + g_1 M^{s-1} + \dots + g_\mu M^{s-\mu} = 0$$

Si l'on appelle u_s le terme de M^s qui occupe un emplacement de case *déterminé* (indépendant de s), on a donc l'équation linéaire de récurrence du μ ^{ième} ordre

$$(21) \quad u_s + g_1 u_{s-1} + \dots + g_\mu u_{s-\mu} = 0$$

La résolution de cette équation conduit à combiner linéairement λ solutions indépendantes, que l'on trouve après avoir résolu l'équation dite caractéristique*, laquelle n'est autre que $g(\lambda) = 0$. Ces λ solutions indépendantes sont précisément les λ coefficients scalaires du second membre de (20), sur lequel on retombe ainsi.

5.6 - Cas particuliers

1/ Une valeur propre nulle

La formule (20) reste valable même si une valeur propre, par exemple a_1 , est nulle, à condition d'adopter la convention $0^0 = 1$. Si cette valeur propre nulle est la seule ($p = 1$), la matrice M est nilpotente, d'ordre de nilpotence $k = k_1$; la formule, réduite à sa

* Attention aux confusions : le mot "caractéristique", qui est traditionnel dans la théorie des équations de récurrence, s'applique ici à $g(\lambda) = 0$ et non à $f(\lambda) = 0$ obtenue en annulant le "polynôme caractéristique" de M .

première ligne, demeure valable mais perd son intérêt (les k_i coefficients matriciels ne sont autres que I, M, \dots, M^{k-1} , et M^s est nul dès que $s \geq k$).

2/ Cas "diagonalisable"

C'est celui où $k_1 = k_2 = \dots = k_p = 1$; le second membre de (20) se réduit à sa première colonne

$$M^s = a_1^s M_1 + a_2^s M_2 + \dots + a_p^s M_p$$

M est, dans ce cas, semblable à une matrice \overline{M} diagonale ; on la dit, pour cette raison, *diagonalisable*. Mais rien n'empêche, bien entendu, que p soit strictement inférieur à m . La diagonale de \overline{M} , lue en descendant, présente n_1 fois la valeur propre a_1, \dots, n_p fois la valeur propre a_p , avec toujours $n_1 + n_2 + \dots + n_p = m$; la surdiagonale, elle, est entièrement nulle. Le polynôme caractéristique est toujours $(\lambda - a_1)^{n_1} (\lambda - a_2)^{n_2} \dots (\lambda - a_p)^{n_p}$, mais le polynôme minimal est simplement $(\lambda - a_1) (\lambda - a_2) \dots (\lambda - a_p)$.

Si en outre $p = 1$, (valeur propre unique $a_1 = a$), le polynôme minimal se réduit à $\lambda - a$; $M - aI = 0$, donc $M = aI$ et $M^s = a^s I$.

3/ Cas "régulier" (cf. III, 3°)

C'est un cas particulier du cas "diagonalisable", à savoir celui où $p = m$ et $n_1 = n_2 = \dots = n_m = 1$. Polynôme caractéristique et minimal sont alors tous deux égaux à $(\lambda - a_1) (\lambda - a_2) \dots (\lambda - a_m)$. La première réduction fournit directement la forme canonique ultime.

La formule (17) donne alors :

$$M^s = a_1^s P_1 P_1' + a_2^s P_2 P_2' + \dots + a_m^s P_m P_m'$$

Les matrices fixes $P_i \times P_i'$ sont des produits colonne \times ligne, et sont donc de rang 1.