

Réduction d'une matrice

Ce cours vous est proposé par Odile Brandière, Université de Paris Sud 11, UFR Jean Monnet et AUNEGe, l'Université Numérique en Économie Gestion.

Table des matières

Introduction	2
Réduction d'une matrice	3
Diagonalisation	3
Matrices semblables.....	3
Valeurs propres et vecteurs propres	4
Critères de diagonalisation	6
Forme réduite de Jordan d'une matrice non diagonalisable	8
Matrice triangulaire semblable à une matrice donnée	8
Forme réduite de Jordan	8
Définitions.....	8
Propriétés des blocs de Jordan	10
Application : évolutions séquentielles linéaires	11
Écriture matricielle et résolution du problème	11
Systèmes d'équations de récurrence linéaires avec second membre	13
Étude du comportement d'un système d'équations récurrentes linéaires quand t tend vers l'infini	14
Cas d'un système homogène (ou sans second membre).....	14
Cas d'un système avec second membre	14
Références	15

Introduction

Objectif de la leçon : Dans cette leçon nous essayons de simplifier certains problèmes faisant intervenir des matrices. L'idée principale est de chercher une base dans laquelle une matrice donnée pourra s'écrire le plus simplement possible. On verra que dans certains cas favorables cette forme sera diagonale, on dira alors que la matrice est diagonalisable, dans d'autres elle prendra la forme d'une matrice réduite de Jordan (triangulaire supérieure très simple). La dernière partie du chapitre donne une application fondamentale aux systèmes linéaires récurrents d'ordre 1, généralisation des suites récurrentes abordées en L2 et qui sont un pré requis de ce paragraphe.

Dans tout ce chapitre on considère un espace vectoriel E de dimension finie n sur un corps K ($K = \mathbb{R}$ ou \mathbb{C}). On peut donc considérer que $E = \mathbb{R}^n$ ou $E = \mathbb{C}^n$.

Réduction d'une matrice

Diagonalisation

Les matrices les plus simples sont les matrices diagonales. On est alors amené à se poser le problème suivant : une matrice A quelconque peut-elle être associée simplement à une matrice diagonale. Nous allons tout d'abord préciser le type d'association dont il est question.

Matrices semblables

On considère ici des applications linéaires (souvent notée f) de E dans E , et les matrices associées, E étant muni de la même base au départ et à l'arrivée.

Définition : Deux matrices A et B sont dites semblables si elles représentent la même application linéaire f de E dans E muni de deux bases différentes. Donc A et B sont semblables si et seulement si, il existe une matrice P inversible telle que

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

Si A est la matrice de f dans la base b et si B est celle de f dans la base c , alors P est la matrice de passage de b vers c , c'est à dire la matrice dont les colonnes sont les coordonnées des vecteurs de c dans la base b .

Propriété 1 : Deux matrices semblables ont même trace et même déterminant.

Rappelons que la trace d'une matrice est la somme de ces termes diagonaux.

Démonstration : Supposons A et B semblables. Alors il existe une matrice P inversible telle que $A = PBP^{-1}$. Or dans le cours de Mathématiques 2, nous avons établi que $tr(AB) = tr(BA)$ d'où

$$trA = tr(PBP^{-1}) = tr(BPP^{-1}) = trB$$

Rappelons que $dét(AB) = dét(A) \cdot dét(B)$ et $dét(A) \cdot dét(A^{-1}) = 1$. D'où

$$dét(A) = dét(PBP^{-1}) = dét(P) \cdot dét(B) \cdot dét(P^{-1}) = dét(B).$$

Définition : Une application linéaire f de E dans E est diagonalisable si et seulement si, il existe une base c de E dans laquelle la matrice D de f est diagonale.

La matrice A est diagonalisable si et seulement si A est semblable à une matrice diagonale D .

Si $c = \{v_1, \dots, v_n\}$ et si la matrice diagonale D de f dans c est $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ alors, étant donnée

la définition de la matrice d'une application linéaire dans une base (voir cours de L1), $f(v_1) = \lambda_1 v_1, \dots, f(v_n) = \lambda_n v_n$. Et si A est la matrice de f dans b et V_i celle de v_i pour $i = 1, \dots, n$:

$$AV_i = \lambda_i V_i \text{ et } A = PDP^{-1} \text{ (les colonnes de } P \text{ sont les } V_i).$$

Remarque : Si b est la base canonique et si $v_i = (x_1, \dots, x_n)$, $V_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. On confond souvent les deux notations v_i et V_i dans ce cas.

Valeurs propres et vecteurs propres

On considère dans ce paragraphe une application linéaire f de E dans E et A sa matrice dans une base b de E .

Définition : On appelle valeur propre et vecteur propre associé de l'application linéaire f (respectivement de la matrice A), un élément λ de K (réel ou complexe) et un vecteur v non nul de E (resp. une matrice colonne V non nulle) tels que :

$$f(v) = \lambda v \text{ (resp. } AV = \lambda V)$$

(V est la matrice de v dans b).

On a vu que si f (ou A) est diagonalisable, la base c dans laquelle la matrice de f est la matrice diagonale D , est constituée de vecteurs propres de f (ou de A) et que la diagonale de D est constituée des valeurs propres de f (ou de A).

Réciproquement si $c = \{v_1, \dots, v_n\}$ est une base de vecteurs propres de f , pour tout i , $f(v_i) = \lambda_i v_i$

et dans c la matrice de f est la matrice diagonale $\begin{pmatrix} \lambda_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \lambda_n \end{pmatrix}$.

On obtient la caractérisation suivante :

f (resp. A) est diagonalisable si et seulement si il existe une base c formée de vecteurs propres de f (resp. de A).

D'autre part $\det(A) = \det(D) =$ produit des valeurs propres (distinctes ou non).

Et $\text{tr}(A) = \text{tr}(D) =$ somme des valeurs propres (distinctes ou non).

Notons Id l'application identique de E (pour tout vecteur v de E , $Id(v) = v$).

Soit λ une valeur propre de f .

$v \neq 0$ est un vecteur propre associé à λ si et seulement si $(f - \lambda Id)(v) = 0$.

Puisque $v \neq 0$, $\text{Ker}(f - \lambda Id)$ est un sous-espace vectoriel de dimension ≥ 1 .

Définition : Si λ est une valeur propre de f , le sous-espace vectoriel E_λ de E défini par $E_\lambda = \text{Ker}(f - \lambda Id)$ est appelé **sous espace propre** associé à la valeur propre λ . C'est l'ensemble de tous les vecteurs propres de f associés à λ auquel on a ajouté le vecteur nul (en effet un **vecteur propre n'est jamais nul**!).

On a toujours : $\dim E_\lambda \geq 1$

Soit λ une valeur propre de f , puisque $\dim \text{Ker}(f - \lambda Id) \geq 1$, d'après le théorème aux dimensions $\text{rang}(f - \lambda Id) < n$ et $(f - \lambda Id)$ n'est pas inversible.

Réciproquement si $(f - \lambda Id)$ n'est pas inversible $\dim \text{Ker}(f - \lambda Id) \geq 1$ et il existe $v \neq 0$ tel que $f(v) = \lambda v$ et λ est une valeur propre de f .

On en déduit donc que λ est valeur propre de f si et seulement si $(f - \lambda Id)$ n'est pas inversible, soit $\det(A - \lambda I_n) = 0$.

Propriété 2 et définition : Les valeurs propres de f (resp. de A) sont les racines de $\det(A - \lambda I_n) = 0$.
Le polynôme $P(\lambda) = \det(A - \lambda I_n)$ s'appelle le polynôme caractéristique de f (resp. de A).

Remarque : Si B est la matrice de f dans une autre base, il existe une matrice P inversible telle que $B = PAP^{-1}$ et $PA - \lambda I_n P^{-1} = B - \lambda I_n$. On en déduit alors que $A - \lambda I_n$ et $B - \lambda I_n$ sont semblables et que leurs déterminants sont égaux.

$P(\lambda)$ ne dépend pas de la base dans laquelle on considère f . $P(\lambda)$ ne dépend que de f .

On admettra la propriété suivante qui précise le degré et quelques coefficients de ce polynôme caractéristique :

Propriété 3 : Le polynôme caractéristique $P(\lambda)$ d'une matrice A d'ordre n est de degré n et $P(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \text{tr}(A) \lambda^{n-1} + \dots + \det(A)$.

Dans $\mathbb{C}, P(\lambda)$ a n racines comptées avec leur ordre de multiplicité. Remarquons que si A est réelle, $P(\lambda)$ est à coefficients réels et on a le résultat suivant :

Propriété 4 : Si A est réelle et si λ et V sont une valeur propre de A et un vecteur propre associé, non réels, alors leurs conjugués sont également une valeur propre et un vecteur propre associé.

Démonstration : On a $AV = \lambda V$ donc $\overline{AV} = \overline{\lambda V}$ et $A\overline{V} = \overline{\lambda V}$ puisque $\overline{A} = A$ (A est à coefficients réels). D'où le résultat.

Notons $m(\lambda)$ l'**ordre de multiplicité** de la racine λ dans le polynôme caractéristique $P(\lambda)$. Si $m(\lambda) = 1$, est valeur propre simple, si $m(\lambda) = 2$, est valeur propre double, etc...

Critères de diagonalisation

Théorème 1 : Soit $\lambda_1, \lambda_2, \dots, \lambda_p$ les p valeurs propres de f (resp. A) deux à deux distinctes. Alors la somme $E_{\lambda_1} + E_{\lambda_2} + \dots + E_{\lambda_p}$ des sous-espaces propres de f (resp. A) est **directe** :

$$E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_p}$$

Attention, le théorème ne précise pas que cette somme directe est égale à E .

Démonstration (peut être sautée) : Faisons un raisonnement par récurrence sur p :

Si $p = 2$ nous devons montrer que $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$. Si $v \in E_{\lambda_1} \cap E_{\lambda_2}$, $f(v) = \lambda_1 v$ et $f(v) = \lambda_2 v$, donc $(\lambda_1 - \lambda_2)v = 0$. Or par hypothèse $\lambda_1 \neq \lambda_2$, donc nécessairement $v = 0$ et on a bien $E_{\lambda_1} \cap E_{\lambda_2} = \{0\}$.

Supposons le théorème démontré pour p valeurs propres distinctes, montrons le pour $p + 1$.

Il suffit de montrer que $(\sum_{j=1}^p E_{\lambda_j}) \cap E_{\lambda_{p+1}} = \{0\}$. Soit $v \in (\sum_{i=1}^p E_{\lambda_i}) \cap E_{\lambda_{p+1}}$, on a donc :

$v = v_1 + v_2 + \dots + v_p$ avec $v_i \in E_{\lambda_i}$, et $v \in E_{\lambda_{p+1}}$.

$$\text{Et } f(v) = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_p v_p = \lambda_{p+1} v = \lambda_{p+1}(v_1 + \dots + v_p)$$

D'où $(\lambda_1 - \lambda_{p+1})v_1 + \dots + (\lambda_p - \lambda_{p+1})v_p = 0$. Or par hypothèse de récurrence, la somme $(\sum_{j=1}^p E_{\lambda_j})$ est directe donc la décomposition de 0 est unique et : pour $i = 1, 2, \dots, p$: $(\lambda_i - \lambda_{p+1})v_i = 0$ et puisque $\lambda_i \neq \lambda_{p+1}$, $v_i = 0$. D'où le résultat.

Puisque le polynôme caractéristique de f est de degré n il admet dans \mathbb{C}^n racines comptées avec leur ordre de multiplicité. Si ces racines sont toutes distinctes, il y en a n , auxquelles correspondent n vecteurs propres, chacun base d'un sous-espace propre. Les sous-espaces propres étant en somme directe, la réunion de ces bases est une base de cette somme directe, cette base contient n vecteurs, c'est une base E . Et la matrice de f dans cette base est diagonale. D'où le théorème :

Propriété 5 : Si une application linéaire f de E dans E (resp. une matrice A d'ordre n) admet n valeurs propres distinctes, alors f (resp. A) est diagonalisable.

Nous allons maintenant énoncer et démontrer un théorème plus général qui va préciser dans quelle mesure une application linéaire est diagonalisable lorsque certaines de ses valeurs propres sont multiples. On se place dans le cas complexe, ou dans le cas réel où le polynôme caractéristique a n racines distinctes ou non.

Théorème 2 : Une application linéaire f de E dans E est diagonalisable si et seulement si :

- 1) $P(\lambda)$ a toutes ses racines dans K ,
- 2) $\dim E_{\lambda_i} = m(\lambda_i)$ pour toutes les valeurs propres λ_i de f .

Remarque : Si les valeurs propres ne sont pas toutes réelles, on ne peut pas espérer pouvoir diagonaliser dans \mathbb{R} , en effet D fait apparaître les valeurs propres sur sa diagonale et ne saurait alors être réelle. La première condition est donc indispensable si on veut obtenir une diagonalisation dans \mathbb{R} , par contre, elle n'a pas d'objet si on diagonalise dans \mathbb{C} . Dans la plupart des cas cette condition sera réalisée. Si elle ne l'est pas, on travaillera dans \mathbb{C} , en utilisant la propriété 4 qui fera gagner du temps.

Démonstration (peut être sautée): Supposons que $P(\lambda)$ ait p racines λ_i distinctes, si 2) est vérifié on a donc

$$\sum_{i=1}^p m(\lambda_i) = n.$$

Et si \mathbf{b}_i est une base de E_{λ_i} ; $\cup_{i=1}^p \mathbf{b}_i = \mathbf{b}$ est une base de $E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_p}$.

Or $\text{card} \cup_{i=1}^p \mathbf{b}_i = \sum_{i=1}^p \text{card} \mathbf{b}_i = \sum_{i=1}^p \dim E_{\lambda_i} = \sum_{i=1}^p m(\lambda_i) = n$ d'après 2). Donc \mathbf{b} est une base de E et c'est une base de vecteurs propres de f par construction. f est donc diagonalisable.

Remarquons qu'alors $E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_p}$.

Supposons que $\dim(E_{\lambda_i}) \neq m(\lambda_i)$. Montrons d'abord que de façon générale :

$$\dim E_{\lambda_i} \leq m(\lambda_i)$$

Supposons que l'on puisse associer à λ_i , $r = m(\lambda_i) + 1$ vecteurs propres linéairement indépendants, v_1, \dots, v_r . Soit \mathbf{b} une base de E dont les r premiers vecteurs sont v_1, \dots, v_r .

Si on note \mathbf{B} la matrice de f dans \mathbf{b} , les r premières colonnes de \mathbf{B} sont
$$\begin{pmatrix} \lambda_i & 0 & \dots & 0 \\ 0 & \lambda_i & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_i \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}$$
. Dans le calcul de

$\det(\mathbf{B} - \lambda I_n)$, on peut mettre en facteur $(\lambda_i - \lambda)^r$ (en développant suivant successivement suivant les r premières colonnes) et λ_i apparaît comme une valeur propre de \mathbf{B} d'ordre de multiplicité $\geq m(\lambda_i) + 1$. Ce qui n'est pas possible puisque le polynôme caractéristique de f ne dépend pas de la base utilisée.

Si $\dim E_{\lambda_i} < m(\lambda_i)$, on ne pourra pas trouver une base de vecteurs propres puisque $\sum_{i=1}^p \dim E_{\lambda_i} < n$. Dans ce cas, il n'y a pas assez de vecteurs propres. f n'est pas diagonalisable.

On verra dans le prochain paragraphe que l'on peut trigonaliser f , c'est à dire trouver une base de E dans laquelle la matrice de f est triangulaire, les valeurs propres figurant sur la diagonale avec leur ordre de multiplicité.

Remarque : le théorème 2 peut s'écrire f est diagonalisable si et seulement si

$$\begin{cases} P(\lambda) \text{ a ses racines dans } K, \\ E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots \oplus E_{\lambda_p} \text{ (avec } \lambda_1, \dots, \lambda_p \text{ valeurs propres distinctes de } f) \end{cases}$$

Forme réduite de Jordan d'une matrice non diagonalisable

Il n'est pas toujours possible de diagonaliser une matrice, mais nous allons voir dans ce paragraphe qu'il est toujours possible de trouver une matrice triangulaire supérieure semblable à une matrice donnée. Nous allons même montrer que cette matrice peut prendre une forme particulièrement simple appelée forme réduite de Jordan. De telles formes rendent possibles certains calculs et facilitent les démonstrations.

Matrice triangulaire semblable à une matrice donnée

Nous nous plaçons sur le corps \mathbb{C} et dans le cas où une application linéaire f et sa matrice A dans la base canonique ne sont pas diagonalisables. A a donc au moins une valeur propre multiple dont le sous-espace propre associé est de dimension strictement inférieure à l'ordre de multiplicité de cette valeur propre.

Proposition : Toute matrice triangulaire semblable à une matrice donnée a pour éléments diagonaux les valeurs propres de cette matrice. Chaque valeur propre apparaît sur la diagonale avec son ordre de multiplicité.

Démonstration : Si $T = (t_i)$ est une matrice triangulaire semblable à A ($A = PTP^{-1}$), elle a donc même polynôme caractéristique et mêmes valeurs propres que A .

Or si T est triangulaire, $T - \lambda I_n$ l'est aussi et $\det T - \lambda I_n = (t_{11} - \lambda)(t_{22} - \lambda) \dots (t_{nn} - \lambda)$.

Ainsi les valeurs propres de T donc de A sont bien égales aux éléments diagonaux de T .

Remarquons que si A est réelle, T et P ne le sont donc pas forcément (en effet les valeurs propres de A ne sont pas forcément réelles, et elles figurent sur la diagonale de T).

Forme réduite de Jordan

Définitions

On appelle bloc de Jordan des matrices de la forme : $J(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ \cdot & \lambda & 1 & \cdot \\ 0 & \cdot & \lambda & 1 \\ 0 & \cdot & \cdot & \lambda \end{pmatrix}$

$J(\lambda)$ est une matrice triangulaire supérieure avec λ sur la diagonale principale, 1 sur la diagonale supérieure et des 0 ailleurs.

Si A est une matrice carrée à éléments réels ou complexes, on appelle forme réduite de Jordan de A , une matrice J semblable à A ($A = PJP^{-1}$), formée de blocs diagonaux qui sont des blocs de

$$\text{Jordan : } J = \begin{pmatrix} \mathbf{J}(\lambda_1) & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \mathbf{J}(\lambda_2) & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \mathbf{J}(\lambda_p) \end{pmatrix}$$

Les λ_i sont alors les valeurs propres de A , elles ne sont pas forcément distinctes deux à deux et apparaissent sur la diagonale de J répétées avec leur ordre de multiplicité.

J est triangulaire supérieure avec des zéros partout sauf sur la diagonale principale (qui est formée des valeurs propres répétées avec leur ordre de multiplicité) et sur la diagonale supérieure qui est formée de 1 ou de 0 (si cette diagonale ne comporte pas de 1, J est diagonale).

Exemple : si $J = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}$ J est une forme réduite de Jordan comportant les 3

blocs de Jordan suivant : (λ_1) , $\begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}$ et $\begin{pmatrix} \lambda_2 & 1 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}$.

Supposons que la matrice J de l'exemple précédent soit la forme réduite de Jordan d'une matrice A . Il existe donc une matrice inversible P telle que $A = PJP^{-1}$.

Interprétons A comme la matrice d'une application linéaire f dans la base canonique de \mathbb{R}^6 . Soit v_i le vecteur ayant pour matrice de coordonnées V_i ($i^{\text{ème}}$ colonne de P) dans la base canonique. J est alors la matrice de f dans $\{v_1, v_2, \dots, v_6\}$. Et

$$\begin{cases} f(v_1) = \lambda_1 v_1 \\ f(v_2) = \lambda_1 v_2 \\ f(v_3) = v_2 + \lambda_1 v_3 \\ f(v_4) = \lambda_2 v_4 \\ f(v_5) = v_4 + \lambda_2 v_5 \\ f(v_6) = v_5 + \lambda_2 v_6 \end{cases} \quad \text{ce qui s'écrit matriciellement} \quad \begin{cases} \mathbf{A}V_1 = \lambda_1 V_1 \\ \mathbf{A}V_2 = \lambda_1 V_2 \\ \mathbf{A}V_3 = V_2 + \lambda_1 V_3 \\ \mathbf{A}V_4 = \lambda_2 V_4 \\ \mathbf{A}V_5 = V_4 + \lambda_2 V_5 \\ \mathbf{A}V_6 = V_5 + \lambda_2 V_6 \end{cases}$$

Remarque : De façon générale à chaque bloc de Jordan correspond un vecteur propre (relatif à sa première colonne où il n'y a pas de 1), et s'il y a k blocs de Jordan ayant la même valeur propre λ sur la diagonale, il y a k vecteurs propres (dont les matrices colonnes des coordonnées sont dans P) indépendants et associés à λ et $\dim E_\lambda = k$. Les autres vecteurs sont appelés vecteurs de Jordan associés à λ .

Ici v_1 et v_2 sont deux vecteurs propres indépendants relatifs à λ_1 , par contre il n'y a qu'un vecteur propre (défini à un coefficient de proportionnalité près) relatif à λ_2 .

De façon pratique, après avoir déterminé les valeurs propres de A , on détermine les sousespaces propres dont les bases sont constituées de vecteurs propres indépendants dont les matrices des coordonnées donnent des colonnes de P .

Les autres colonnes de P (ici P_3, P_5 et P_6) sont les matrices colonnes des coordonnées des vecteurs de Jordan (ici v_3, v_5 et v_6) et se déterminent alors simplement à l'aide d'équations analogues aux précédentes.

Propriétés des blocs de Jordan

Si $J(\lambda)$ est un bloc de Jordan on peut écrire : $J(\lambda) = \lambda I + J(0)$ où I est la matrice unité de même ordre que $J(\lambda)$ et où $J(0)$ est un bloc de Jordan ayant une diagonale nulle.

Exemple : à l'ordre 1, $J(0) = (0)$, à l'ordre 2, $J(0) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, à l'ordre 3, $J(0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, à l'ordre 4,

$$J(0) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

On a alors les propriétés suivantes :

1-Si $J(0)$ est d'ordre r , alors $J(0)^r = 0$.

2-Si $J(0)$ est d'ordre r : $J(0) \begin{pmatrix} x_1 \\ \cdot \\ x_{r-1} \\ x_r \end{pmatrix} = \begin{pmatrix} x_2 \\ \cdot \\ x_r \\ 0 \end{pmatrix}$

Enfin nous admettrons le théorème de Jordan, qui prouve l'existence d'une réduite de Jordan quelle que soit la matrice A :

Théorème : Pour toute matrice carrée à éléments dans \mathbb{C} il existe une forme réduite de Jordan qui lui est semblable.

Une application de ce théorème est le calcul de A^t dans le cas où A n'est pas diagonalisable. En effet si J est la forme réduite de Jordan de A :

$$A^t = P J^t P^{-1}$$

Et le calcul de A^t se ramène à celui de J^t .

Or J est de la forme
$$\begin{pmatrix} J(\lambda_1) & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & J(\lambda_2) & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & J(\lambda_p) \end{pmatrix} \text{ et } J^t = \begin{pmatrix} J(\lambda_1)^t & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & J(\lambda_2)^t & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & \cdot & J(\lambda_p)^t \end{pmatrix}.$$

Les calculs de $J(\lambda)^t$ se font en utilisant la décomposition $J(\lambda) = \lambda I + J(0)$, la propriété de $J(0)$ énoncée ci-dessus et le fait que les matrices λI et $J(0)$ commutent.

De façon pratique pour **réduire une matrice** on a les étapes suivantes :

1. Déterminer les valeurs propres, racines du polynôme caractéristique $P_A(\lambda)$, et leur ordre de multiplicité.
2. Pour chaque valeur propre λ déterminer le sous-espace propre E_λ , en donner une base et sa dimension.
3. Donner la réduite de Jordan de A (éventuellement diagonale)
4. Pour chaque λ , si $\dim E_\lambda = m(\lambda)$ il n'y a pas de vecteur de Jordan. Sinon il faut déterminer $(m(\lambda) - \dim E_\lambda)$ vecteur de Jordan indépendants.
5. Expliciter P , formée des bases des E_λ et des vecteurs de Jordan) telle que $A = PJP^{-1}$. (Attention à l'ordre des colonnes de P et à l'ordre des valeurs propres dans J)

Application : évolutions séquentielles linéaires

Ecriture matricielle et résolution du problème

On considère n suites $x_t^1, x_t^2, \dots, x_t^n$ indicées par le temps $t \in \mathbb{N}$ (fonctions du temps), prenant des valeurs réelles ou complexes et qui à l'instant $t + 1$ dépendent linéairement des valeurs prises à l'instant t :

$$\begin{cases} x_{t+1}^1 = a_{11}x_t^1 + a_{12}x_t^2 + \dots + a_{1n}x_t^n \\ \cdot \\ \cdot \\ x_{t+1}^n = a_{n1}x_t^1 + a_{n2}x_t^2 + \dots + a_{nn}x_t^n \end{cases}$$

Posons $A = (a_{ij})$ et $X_t = \begin{pmatrix} x_t^1 \\ \cdot \\ \cdot \\ x_t^n \end{pmatrix}$

Le précédent système s'écrit alors $X_{t+1} = AX_t$ ($t \in \mathbb{N}$)

On en déduit donc par récurrence que $X_t = A^t X_0$

Or si J est la forme réduite de Jordan de A , $A = PJP^{-1}$ et $A^t = PJ^tP^{-1}$.

Si J est diagonale, le calcul de J^t est aisé, sinon on calculera J^t en utilisant la méthode décrite dans le paragraphe précédent.

Remarque : 1. Cette méthode de résolution implique le calcul de P^{-1} .

2. Si A est à coefficients réels et si certaines valeurs propres sont complexes (elles sont alors conjuguées deux à deux ainsi que les vecteurs propres associés), les calculs se font dans \mathbb{C} , mais le résultat final est réel et les parties imaginaires s'éliminent.

Une autre forme de résolution conseillée

Remarquons que $X_{t+1} = AX_t$ s'écrit $X_{t+1} = PJP^{-1}X_t$ avec J forme réduite de Jordan de A . Et en multipliant à gauche par P^{-1} on a : $(P^{-1}X_{t+1}) = J(P^{-1}X_t)$.

1. Changement de variable :

Si on pose $Z_t = P^{-1}X_t$, l'équation précédente s'écrit : $Z_{t+1} = JZ_t$

2. Calcul de Z_t :

Bien sûr on a $Z_t = (J^t)Z_0$. Mais le calcul de J^t n'est pas nécessaire.

Si J est diagonale et si $Z_0 = \begin{pmatrix} z_0^1 \\ z_0^2 \\ \vdots \\ z_0^n \end{pmatrix}$ alors $Z_t = \begin{pmatrix} (\lambda_1)^t z_0^1 \\ (\lambda_2)^t z_0^2 \\ \vdots \\ (\lambda_n)^t z_0^n \end{pmatrix}$

Si J n'est pas diagonale :

$$\begin{cases} z_{t+1}^1 = \lambda_1 z_t^1 + c_{21} z_t^2 \\ z_{t+1}^2 = \lambda_2 z_t^2 + c_{32} z_t^3 \\ \vdots \\ z_{t+1}^n = \lambda_n z_t^n \end{cases} \text{ avec } c_{ij} = 1 \text{ ou } 0$$

Ce système peut être résolu facilement en "remontant" à partir de la dernière équation dont la solution est $z_t^n = (\lambda_n^t) z_0^n$.

En reportant dans l'équation précédente on a : $z_{t+1}^{n-1} = \lambda_{n-1} z_t^{n-1} + c_{n(n-1)} (\lambda_n)^t z_0^n$ et on est ramené à une suite linéaire récurrente du premier ordre (dans le cas où $c_{n(n-1)}$ est non nul $\lambda_n = \lambda_{n-1}$ et il y a résonance comme cela a été vu dans le cours de L_2).

A la fin de cette résolution Z_t dépend des n paramètres $z_0^1, z_0^2, \dots, z_0^n$.

3. Calcul de X_t :

Or $X_t = PZ_t$

Et $P = (P_1 P_2 \dots P_n)$ avec P_i vecteur propre ou vecteur de Jordan.

Remarque : Si J est diagonale : $X_t = (\lambda_1)^t Z_0^1 P_1 + (\lambda_2)^t Z_0^2 P_2 + \dots + (\lambda_n)^t Z_0^n P_n$

Ainsi x_t^1, \dots, x_t^n (coordonnées de X_t) sont des combinaisons linéaires des (λ_i^t) .

Si J n'est pas diagonale les coordonnées de X_t sont des combinaisons linéaires de termes de la forme $t^k \lambda_j^t$ (avec $0 \leq k < m(\lambda_j)$).

A la fin de cette résolution X_t dépend des n paramètres $z_0^1, z_0^2, \dots, z_0^n$.

4. Calcul des paramètres, conditions initiales :

On utilise les conditions initiales X_0 pour calculer les n paramètres intervenus lors du calcul de Z_t .

Remarquons que cette méthode a l'avantage de ne pas avoir à calculer P^{-1} .

Systèmes d'équations de récurrence linéaires avec second membre

On considère donc des systèmes de la forme

$$(1) X_{t+1} = AX_t + F_t$$

où A est une matrice à coefficients réels constants et F_t une matrice colonne dont les coefficients sont des fonctions de t .

1. Première méthode

Cette méthode suppose l'existence d'une solution particulière X_t^* de (1). On remarque alors que $Y_t = X_t - X_t^*$ est une solution de l'équation sans second membre (ou homogène) associée :

$$(2) Y_{t+1} = AY_t$$

(En effet $Y_{t+1} = X_{t+1} - X_{t+1}^* = (AX_t + F_t) - (AX_t^* + F_t) = A(X_t - X_t^*) = AY_t$)

On se ramène alors à la résolution du paragraphe précédent pour calculer Y_t . Et

$$X_t = Y_t + X_t^*$$

On utilise peu cette méthode car souvent il est très difficile de trouver X_t^* .

2. Deuxième méthode

Là encore, on utilisera J , la forme réduite de Jordan de A (J est diagonale ou non) et le changement de variables précédent.

Si $A = PJP^{-1}$, on posera $Z_t = P^{-1}X_t$ et on aura :

$$(X_{t+1} = AX_t + F_t?) \Leftrightarrow (Z_{t+1} = JZ_t + P^{-1}F_t)$$

(par multiplication à gauche par P^{-1})

Ici il est donc nécessaire de calculer P^{-1} . Mais la dernière égalité permet de calculer chaque composante de Z_t qui se présente comme solution d'une équation récurrente d'ordre 1. Enfin on calcule X_t en utilisant : $X_t = PZ_t$.

Etude du comportement d'un système d'équations récurrentes linéaires quand t tend vers l'infini

Cas d'un système homogène (ou sans second membre)

On considère donc un système d'équations de la forme : $X_{t+1} = AX_t$ où A est une matrice réelle à coefficients constants. D'après ce qui a été vu précédemment, les coordonnées de X_t sont des combinaisons linéaires de termes de la forme $t^k \lambda_j^t$ (avec $0 \leq k < m(\lambda_j)$) et le comportement de X_t quand t tend vers $+\infty$ dépend donc du module des valeurs propres λ_j .

Ainsi si **toutes** les valeurs propres de A sont de module strictement inférieur à un, alors

$$X_t \rightarrow 0 \text{ quand } t \rightarrow +\infty$$

Par contre s'il y a des valeurs propres dont le module est strictement supérieur à 1 alors

$$\|X_t\| \rightarrow +\infty \text{ quand } t \rightarrow +\infty$$

(sauf pour certaines conditions initiales).

Définition : On dit alors que le système séquentiel défini par $X_{t+1} = AX_t$ est stable si la suite (X_t) converge vers 0 lorsque $t \rightarrow +\infty$ quel que soit le vecteur initial X_0 .

D'après ce que nous avons vu précédemment, le système est donc stable dès que les modules de **toutes les valeurs propres sont strictement inférieurs à 1**.

On dit alors que le système est **séquentiellement stable**. (On verra dans la leçon suivante une autre notion de stabilité).

Cas d'un système avec second membre

Soit donc le système : $X_{t+1} = AX_t + F_t$. On sait qu'alors $X_t = X_t^* + Y_t$ où X_t^* est une solution particulière et Y_t une la solution générale de l'équation sans second membre.

Si Y_t est séquentiellement stable, $Y_t \rightarrow 0$ quand $t \rightarrow +\infty$ et $X_t \rightarrow X_t^*$. Y_t représente les écarts par rapport à la solution tendancielle ou tendance que représente X_t^* . Si Y_t n'est pas séquentiellement stable, X_t peut ne pas converger.

Le calcul des valeurs propres d'une matrice n'est pas toujours aisé, aussi voici quelques critères permettant de détecter des matrices réelles séquentiellement stables sans calculer leurs valeurs propres.

Critère 1 (conditions nécessaires de stabilité lorsque 1 n'est pas valeur propre) : Pour qu'une matrice A d'ordre n soit séquentiellement stable, il faut nécessairement que :

$$-1 < \det A < 1 \text{ et } -n < \operatorname{tr}(A) < n$$

En effet $\det A = \prod_{j=1}^n \lambda_j$ et $\operatorname{tr}(A) = \sum_{j=1}^n \lambda_j$ et si $|\lambda_j| < 1$ pour tout j , alors $|\det(A)| < 1$ et $|\operatorname{tr}(A)| < n$.

Critère 2 : (condition suffisante de stabilité) :

Si $\sum_{j=1}^n |a_{ij}| < 1$ pour $i = 1, 2, \dots, n$, alors la matrice $A = (a_{ij})$ est séquentiellement stable.

Remarque : $\sum_{j=1}^n |a_{ij}|$ représente la somme des valeurs absolues des coefficients de la $i^{\text{ème}}$ ligne de $A = (a_{ij})$.

Références

Comment citer ce cours ?

Mathématiques 3, Odile Brandière, AUNEGe (<http://auneg.fr>), CC – BY NC ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Cette œuvre est mise à disposition dans le respect de la législation française protégeant le droit d'auteur, selon les termes du contrat de licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). En cas de conflit entre la législation française et les termes de ce contrat de licence, la clause non conforme à la législation française est réputée non écrite. Si la clause constitue un élément déterminant de l'engagement des parties ou de l'une d'elles, sa nullité emporte celle du contrat de licence tout entier.