UE MAT234

Notes de cours sur l'algèbre linéaire

Matrices - Systèmes linéaires - Déterminants - Diagonalisation

Dans tout ce document, K désigne indifféremment le corps des nombres réels $\mathbb{R}$, ou celui des nombres complexes $\mathbb{C}$. E et F sont des espaces vectoriels, de dimensions respectives n et p.

1 Matrices

1.1 Rappels : espaces vectoriels

Définitions

Un ensemble E est un espace vectoriel sur K s'il est muni de 2 lois :

- une addition interne d'éléments de E :
$E \times E \to E, \quad (u, v) \mapsto u + v$

- une multiplication externe d'éléments de E par des scalaires de K :
$K \times E \to E, \quad (\lambda, u) \mapsto \lambda u$

telles que

- (E,+) soit un groupe commutatif : l'addition interne d'éléments de E est associative, commutative, possède un élément neutre ($\exists e \in E, \forall u \in E, u + e = u = e + u$), et tout élément de E a un opposé pour l'addition ($\forall u \in E, \exists u' \in E, u + u' = u' + u = e$).

- la multiplication externe vérifie :

$\lambda(\mu v) = (\lambda\mu)v, \qquad \forall \lambda, \mu \in K,\ \forall v \in E,$

$1v = v, \qquad \forall v \in E,$

$(\lambda + \mu)v = \lambda v + \mu v, \qquad \forall \lambda, \mu \in K,\ \forall v \in E,$

$\lambda(u + v) = \lambda u + \lambda v, \qquad \forall \lambda \in K,\ \forall u, v \in E.$

L'élément neutre e est noté 0_E et l'opposé de u est noté $-u$.

Exemples

L'espace $\mathbb{R}^n$; L'espace des fonctions continues sur $\mathbb{R}$; ...

Combinaison linéaire

Famille libre, liée, génératrice

- La famille $(u_1, ..., u_q)$ de vecteurs de E est dite libre (ou linéairement indépendante) si aucun de ces vecteurs ne peut s'écrire comme combinaison linéaire des autres. Plus précisément si
- Sinon cette famille est dite liée. [Rmq : toute famille contenant 0_E est]
- La famille $(u_1, ..., u_q)$ est dite génératrice dans E si

Base

- **Définition :**
- **Dimension finie :**
- **Propriétés :** si E est de dimension finie n, alors :
  - toutes les bases de E ont ...
  - toute famille libre de n éléments est ...
  - toute famille génératrice de n éléments est ...

Sous espace vectoriel

Sous espace engendré

Somme de sous espaces vectoriels

- **Définition :**
- **Somme directe :**
- **Sous espaces supplémentaires :**

1.2 Matrices : définitions

Déf On appelle **matrice de taille** $\mathbf{p} \times \mathbf{n}$ un tableau à double entrée, comportant p lignes et n colonnes, et dont chaque élément (encore appelé coefficient) appartient à K. L'élément a_{ij} appartenant à la i-ème ligne et à la j-ème colonne est appelé **élément d'indice** $(\mathbf{i}, \mathbf{j})$ de la matrice $A = (a_{ij})$.

Déf L'ensemble des matrices de taille $p \times n$ à éléments dans K est noté $\mathcal{M}_{p,n}(K)$.

Déf L'ensemble des matrices de taille $n \times n$ à éléments dans K, dites **matrices carrées**, est noté $\mathcal{M}_n(K)$.

$$A \text{ appartenant à } \mathcal{M}_{p,n}(K) \text{ est notée } A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le p \\ 1 \le j \le n}}$$

Prop $\mathcal{M}_{p,n}(K)$ est un espace vectoriel sur K, de dimension pn. Sa base canonique est $\{E_{ij}, 1 \le i \le p, 1 \le j \le n\}$ où E_{ij} est la matrice dont tous les éléments sont nuls, sauf celui d'indice (i, j) qui vaut 1. Ainsi toute matrice A de $\mathcal{M}_{p,n}(K)$ s'écrit de manière unique $A = \displaystyle\sum_{i=1}^{p} \sum_{j=1}^{n} a_{ij}\, E_{ij}$.

1.2.1 Quelques matrices carrées particulières

Déf Matrice identité de taille $n \times n$: $I_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ $(I_n(i,j) = 1 \text{ si } i = j,\ 0 \text{ sinon})$.

Déf Matrice diagonale : $D = \begin{pmatrix} d_{11} & & 0 \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$ $(D(i,j) = 0 \text{ si } i \neq j)$.

Déf Matrice triangulaire supérieure : $T = \begin{pmatrix} t_{11} & \cdots & t_{1n} \\ & \ddots & \vdots \\ 0 & & t_{nn} \end{pmatrix}$ $(T(i,j) = 0 \text{ si } i > j)$.

Déf Matrice triangulaire inférieure : $T = \begin{pmatrix} t_{11} & & 0 \\ \vdots & \ddots & \\ t_{n1} & \cdots & t_{nn} \end{pmatrix}$ $(T(i,j) = 0 \text{ si } i < j)$.

Déf Matrice triangulaire inférieure (ou supérieure) stricte : Matrice triangulaire inférieure (ou supérieure) dont la diagonale est nulle.

Déf Matrice bande de largeur $2m+1$: $A(i,j) = 0$ si $|i - j| > m$. Pour $m = 1$ on parle de matrice **tridiagonale**.

1.2.2 Transposition

Déf Soit $A \in \mathcal{M}_{pn}(K)$. On appelle **transposée de A**, notée A^t, la matrice de $\mathcal{M}_{np}(K)$ définie par $A^t(i,j) = A(j,i)$.

Déf $A \in \mathcal{M}_n(K)$ (donc carrée) est **symétrique** ssi $A^t = A$.

Déf $A \in \mathcal{M}_n(K)$ est **antisymétrique** ssi $A^t = -A$. (donc de diagonale nulle).

Prop $A \in \mathcal{M}_n(K)$ peut être décomposée de façon unique en somme d'une matrice symétrique et d'une matrice antisymétrique : $A = \dfrac{1}{2}(A + A^t) + \dfrac{1}{2}(A - A^t)$. On en déduit que $\mathcal{M}_n(K)$ est la somme directe du sous-espace vectoriel des matrices symétriques de taille n et du sous-espace vectoriel des matrices antisymétriques de taille n.

1.2.3 Produit de matrices

Déf Soient $A \in \mathcal{M}_{pn}(K)$ et $B \in \mathcal{M}_{nq}(K)$. On définit alors la matrice $AB \in \mathcal{M}_{pq}(K)$ par $(AB)(i,j) = \displaystyle\sum_{k=1}^{n} A(i,k)\, B(k,j)$. <u>Attention :</u> on voit que le produit AB n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B.

Rmq Disposition pour les calculs

Rmq Action d'une matrice sur un vecteur colonne

Soit $A \in \mathcal{M}_{pn}(K)$ et $x \in K^n$ représenté par le vecteur colonne $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

La décomposition de la matrice A en colonnes permet d'écrire $A = (A_1, ..., A_n)$.
Ainsi, $A\,x = x_1 A_1 + x_2 A_2 + \cdots + x_n A_n$ est donc une CL des colonnes de A.

Prop Le produit de matrices est associatif : $A(BC) = (AB)C$, et distributif par rapport à l'addition : $A(B + C) = AB + AC$.

Par contre il n'est pas commutatif : $AB \neq BA$ en général, même si $A, B \in \mathcal{M}_n(K)$.

Ainsi $(A + B)^2 = A^2 + AB + BA + B^2 \neq A^2 + 2AB + B^2$ en général.

Prop $(AB)^t = B^t A^t$

Prop Le produit de deux matrices triangulaires supérieures (resp. inférieures) est une matrice triangulaire supérieure (resp. inférieure).

Prop Le produit de deux matrices diagonales est une matrice diagonale.

Déf Soit $A \in \mathcal{M}_n(K)$. On appelle **matrice inverse** de A, notée A^{-1}, la matrice de $\mathcal{M}_n(K)$ vérifiant $AA^{-1} = A^{-1}A = I_n$, si elle existe.

Prop Si A inversible, son inverse A^{-1} est unique.

Prop Pour A et B inversibles, on a : $(AB)^{-1} = B^{-1}A^{-1}$.

Prop Pour A inversible, on a : $(A^{-1})^t = (A^t)^{-1}$.

Rmq $AB = AC$ ne conduit pas à $B = C$.

1.2.4 Manipulation par blocs des matrices

On peut, pour simplifier dans certains cas les calculs matriciels, adopter une écriture par blocs. Ceci a de l'intérêt si certains blocs sont particulièrement simples (par exemple nuls, ou égaux à une matrice identité). Les manipulations sont identiques aux calculs matriciels usuels. Il faut simplement s'assurer de la compatibilité des tailles de blocs lors des opérations (*cf* exemples dans les fiches de TD).

Attention : la forme réelle des matrices n'est pas équivalente à la forme par blocs. Par exemple une matrice triangulaire par blocs n'est pas nécessairement triangulaire. De même l'écriture par blocs d'une matrice triangulaire n'est pas nécessairement triangulaire par blocs. Par contre, toute matrice triangulaire peut être écrite sous forme triangulaire par blocs. Et de même pour des matrices, diagonales, bandes, etc.

1.3 Représentation matricielle d'une application linéaire

1.3.1 Rappels : application linéaire

Déf Application linéaire de E dans F

Prop φ linéaire de E dans F (E de dimension finie) est complétement déterminée par la connaissance de l'image d'une base.

Déf $\varphi : E \to F$ linéaire
Noyau de φ : $\text{Ker}\varphi =$
Image de φ : $\text{Im}\varphi =$

Déf Rang d'une application linéaire ($\dim E$ finie) : $\text{rang}(\varphi) = \dim(\text{Im}(\varphi))$.
Puisque $\text{Im}(\varphi)$ est engendré par les vecteurs $\varphi(e_1), ..., \varphi(e_n)$, pour une base $(e_1, ..., e_n)$ de E, le rang de φ est le nombre maximum de vecteurs $\varphi(e_j)$ linéairement indépendants.

Prop Théorème du rang (E de dimension finie)

1.3.2 Représentation matricielle

Soit φ une application linéaire de E vers F. Soient $\mathcal{B}_E = \{e_1, \ldots, e_n\}$ une base de E et $\mathcal{B}_F = \{f_1, \ldots, f_p\}$ une base de F. Pour chaque e_j, $\varphi(e_j) \in F$ et peut donc être décomposé de façon unique sur la base $\{f_1, \ldots, f_p\}$: $\varphi(e_j) = \sum_{i=1}^{p} a_{ij}\, f_i$

On pose : $A = (a_{ij})_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$

Déf A est la **matrice représentative** de l'application linéaire φ dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$. On remarque que chaque colonne j de A contient les coefficients de $\varphi(e_j)$.

Soit $x \in E$. Notons $y = \varphi(x)$. On peut décomposer x dans la base $\mathcal{B}_E$: $x = \sum_{j=1}^{n} x_j\, e_j$ et y dans la base $\mathcal{B}_F$: $y = \sum_{i=1}^{p} y_i\, f_i$. Or $y = \varphi(x) = \sum_{j=1}^{n} x_j\, \varphi(e_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{p} a_{ij}\, f_i = \sum_{i=1}^{p} \left(\sum_{j=1}^{n} a_{ij}\, x_j \right) f_i$.

En posant $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$, on a donc $Y = AX$.

Prop $y = \varphi(x)$ est équivalent à l'écriture matricielle $Y = AX$, où X et Y sont les coordonnées de x et y dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$, et où A est la matrice représentative de φ dans les bases $\mathcal{B}_E$ et $\mathcal{B}_F$.

$$\begin{aligned}
\varphi: \quad E_{(e_1,\ldots,e_n)} &\longrightarrow F_{(f_1,\ldots,f_p)} && (1)\\
x &\mapsto \varphi(x) = y && (2)\\
X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &\mapsto AX = Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} && (3)
\end{aligned}$$

Prop Matrice de l'identité $I_d : E \longrightarrow E, \quad x \mapsto I_d(x) = x,$ avec E muni de la base $\mathcal{B}_E = (e_1, e_2, ..., e_n)$, au départ et à l'arrivée. Alors, $\mathrm{Mat}_{\mathcal{B}_E}(I_d) = I_n$

Rmq Attention - si les bases au départ et à l'arrivée pour l'espace vectoriel E sont différentes, alors l'identité n'est pas représentée par I_n.

Prop Composée d'applications linéaires.
Considérons les applications linéaires suivantes où E, F, G sont de dimension finie.

$$\varphi : E \longrightarrow F, \quad \psi : F \longrightarrow G$$

Alors,

$$\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_G}(\psi \circ \varphi) = \mathrm{Mat}_{\mathcal{B}_F,\mathcal{B}_G}(\psi)\,\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(\varphi)$$

Prop Si φ est bijective de E dans F (de sorte que dim E=dim F), alors $\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(\varphi)$ est inversible et

$$\Big(\mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(\varphi)\Big)^{-1} = \mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(\varphi^{-1})$$

Déf Rang d'une matrice.

1) Si $A = \mathrm{Mat}_{\mathcal{B}_E,\mathcal{B}_F}(\varphi)$, alors, $\mathrm{rang}(A) := \mathrm{rang}(\varphi) = \dim(\mathrm{Im}(\varphi))$.

2) Pour une matrice quelconque $A \in \mathcal{M}_{pn}(K)$, on remarque que A définit une application linéaire de K^n dans K^p : $A : K^n \to K^p, \quad X \mapsto AX$, de sorte que le **rang** de A est la dimension du sous-espace vectoriel engendré par les vecteurs colonnes de A. Autrement dit, c'est le nombre maximum de vecteurs colonnes de A linéairement indépendants les uns des autres.

Prop rang(A)= = rang(A^t)

1.4 Changement de base

1.4.1 Matrice de passage

Soit $\mathcal{B}'_E = \{e'_1, \dots, e'_n\}$ une autre base de E. Chaque e'_j peut être décomposé dans la base $\mathcal{B}_E$: $e'_j = \sum_{i=1}^{n} p_{ij}\, e_i$. Notons $P = (p_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$.

Déf P est la **matrice de passage** de la base $\mathcal{B}_E$ à la base $\mathcal{B}'_E$

Pour $x \in E$, on note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$ ses coordonnées dans $\mathcal{B}_E$ et $\mathcal{B}'_E$.

On a alors : $\mathbf{X} = \mathbf{P}\mathbf{X}'$.

1.4.2 Changement de base pour une application linéaire

Soit $\mathcal{B}'_F = \{f'_1, \ldots, f'_p\}$ une autre base de F. On note Q la matrice de passage de $\mathcal{B}_F$ à $\mathcal{B}'_F$. Pour φ une application linéaire de E vers F, on note A sa matrice représentative dans $\mathcal{B}_E$ et $\mathcal{B}_F$, et B sa matrice représentative dans $\mathcal{B}'_E$ et $\mathcal{B}'_F$.

Soit $x \in E$ et $y = \varphi(x)$. On note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $X' = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$, $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ et $Y' = \begin{pmatrix} y'_1 \\ \vdots \\ y'_p \end{pmatrix}$ les coordonnées de x et y dans les différentes bases. On a : $Y = QY' = QBX'$ et $Y = AX = APX'$, et ceci pour tous x et y. D'où la relation: $\mathbf{B} = \mathbf{Q^{-1}AP}$.

Dans le cas particulier où $E = F$ (c.a.d. où φ est un endomorphisme), la relation précédente devient $\mathbf{B} = \mathbf{P^{-1}AP}$.

2 Chapitre 2 : Systèmes linéaires

En l'absence de précisions, l'espace vectoriel K^m, $(m > 0)$, sera supposé muni de sa base canonique.

• Soit $A \in \mathcal{M}_{pn}(K)$ et $b \in K^p$. On s'intéresse à la résolution du système linéaire $AX = b$, c'est à dire à trouver $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ vérifiant :

$$\begin{cases} a_{11}\,x_1 + \cdots + a_{1n}\,x_n &= b_1 \\ \vdots & \vdots \\ a_{p1}\,x_1 + \cdots + a_{pn}\,x_n &= b_p \end{cases} \qquad (S)$$

• A définit une application linéaire de K^n dans K^p :

$$A : K^n \longrightarrow K^p, \quad X \mapsto AX\,,$$

ce qui permet de définir le noyau $\mathrm{Ker}A$ (sous espace vectoriel de K^n) et l'image $\mathrm{Im}A$ (sous espace vectoriel de K^p).

2.1 Existence et unicité de solutions

Déf Au système (S), on associe le **système homogène** (S_0) : $AX = 0$. Le vecteur nul est évidemment une solution de (S_0).

Prop L'ensemble $\mathcal{S}_0$ des solutions de (S_0) est un sous-espace vectoriel de $\mathbb{R}^p$.
Précisément : $\mathcal{S}_0 = \mathrm{Ker}A$.

Prop L'ensemble $\mathcal{S}$ des solutions de (S) est obtenu en ajoutant à une solution particulière quelconque de (S) l'ensemble des solutions de (S_0).
Précisément :
si $b \notin \mathrm{Im}A$, alors $\mathcal{S} = \emptyset$.
si $b \in \mathrm{Im}A$, alors $\mathcal{S} = \tilde{X} + \mathcal{S}_0$, où $\tilde{X}$ est une solution quelconque de (S).

Preuve - *Si X est une solution quelconque de (S), alors $AX - A\tilde{X} = A(X - \tilde{X}) = b - b = 0_{K^p}$, de sorte que $X - \tilde{X} \in \mathrm{Ker}A$.*

Prop Si rang $A = n$ (nombre de colonnes), (S) admet au plus une solution.

Preuve - *On décompose la matrice $A = (A_1, ..., A_n)$ en n colonnes. Alors $\mathrm{Im}(A) = \mathrm{Vect}(A_1, ..., A_n)$ et puisque $\mathrm{rang}A = n$, les vecteurs colonnes de A sont linéairement indépendants et forment donc une base de $\mathrm{Im}(A)$. Ensuite :*
Si $b \notin ImA$, alors $\mathcal{S} = \emptyset$.
Si $b \in ImA$, alors $\exists X \in K^n$ tel que $AX = b$, c-à-d, tel que $(A_1, ..., A_n)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = b$, c-à-d, tel que $\sum_{j=1}^{n} x_j A_j = b$. Finalement, puisque $b \in \mathrm{Im}(A)$ s'écrit de manière unique dans la base des A_j, la solution X est unique.

Prop Si A est de taille $n \times n$ et si rang $A = n$, alors (S) admet une et une seule solution.

Preuve - *Théorème du rang $\Rightarrow A$ inversible.*

2.2 Résolution d'un système triangulaire

Soit à résoudre $AX = b$ avec $A \in \mathcal{M}_n(K)$ triangulaire inférieure. (S) s'écrit dans ce cas:

$$\left\{\begin{array}{lcl} a_{11}\, x_1 & = & b_1 \\ a_{21}\, x_1 + a_{22}\, x_2 & = & b_2 \\ \vdots \qquad\qquad \ddots & & \vdots \\ a_{n1}\, x_1 + a_{n2}\, x_2 + \cdots + a_{nn}\, x_n & = & b_n \end{array}\right. \qquad (S_T)$$

Prop A est de rang n ssi $a_{ii} \neq 0, \quad \forall i = 1, \ldots, n$

Preuve - *A est de rang $n \Leftrightarrow A$ est inversible $\Leftrightarrow AX = b$ admet une unique solution $\forall b \in K^n$ $\Leftrightarrow$ les a_{ii} sont tous non nuls d'après la résolution décrite ci-dessous.*

Résolution de $(\mathbf{S_T})$ On suppose que les éléments diagonaux a_{ii} sont tous non-nuls. (S_T) admet donc une solution unique, qu'on obtient par descente du système et substitution :

$$\left\{\begin{array}{lcl} x_1 & = & b_1/a_{11} \\ x_2 & = & (b_2 - a_{21}\, x_1)/a_{22} \\ & \vdots & \\ x_n & = & (b_n - a_{n1}\, x_1 - \cdots - a_{n,n-1}\, x_{n-1})/a_{nn} \end{array}\right.$$

Le **coût de cette résolution**, c.a.d. le nombre d'opérations nécessaires pour cet algorithme, est d'environ $n^2/2$ additions/soustractions, et $n^2/2$ multiplications/divisions. (on sépare ainsi ces deux types d'opérations, car les multiplications et divisions coûtent en général nettement plus cher en temps de calcul que les additions ou soustractions).

La méthode est évidemment identique pour un système triangulaire supérieur, en commençant par la dernière ligne et en remontant le système.

2.3 Méthode de Gauss

On considère le système (S) pour une matrice A carrée de rang n. Notons L_k la k-ème ligne de (S). La méthode d'élimination de Gauss consiste à réaliser les étapes suivantes :

- **1ère étape** Si $a_{11} \neq 0$, on élimine x_1 dans les lignes 2 à n par la transformation $L_i \to L_i^{(2)} = L_i - \dfrac{a_{i1}}{a_{11}}\, L_1$ pour $i = 2, \ldots, n$. On obtient donc le système :

$$\left\{\begin{array}{rcl} a_{11}^{(2)}\, x_1 + a_{12}^{(2)}\, x_2 + \cdots + a_{1n}^{(2)}\, x_n & = & b_1^{(2)} \\ a_{22}^{(2)}\, x_2 + \cdots + a_{2n}^{(2)}\, x_n & = & b_2^{(2)} \\ & \vdots & \\ a_{n2}^{(2)}\, x_2 + \cdots + a_{nn}^{(2)}\, x_n & = & b_n^{(2)} \end{array}\right.$$

et on recommence sur le sous-sytème formé des lignes 2 à n.

- **Etape k** Au début de l'étape k, le système est de la forme :

$$\left\{\begin{array}{rcl} a_{11}^{(k)}\,x_1+\cdots+a_{1k}^{(k)}\,x_k+\cdots+a_{1n}^{(k)}\,x_n &=& b_1^{(k)} \\ \ddots\qquad\qquad\qquad && \vdots \\ a_{kk}^{(k)}\,x_k+\cdots+a_{kn}^{(k)}\,x_n &=& b_k^{(k)} \\ && \vdots \\ a_{nk}^{(k)}\,x_k+\cdots+a_{nn}^{(k)}\,x_n &=& b_n^{(k)} \end{array}\right.$$

  Si $a_{kk}^{(k)} \neq 0$, on élimine alors x_k dans les lignes $k+1$ à n par la transformation $L_i^{(k)} \to L_i^{(k+1)} = L_i^{(k)} - \dfrac{a_{ik}^{(k)}}{a_{kk}^{(k)}}\,L_k^{(k)}$ pour $i = k+1,\ldots,n$. On fait donc :

$$a_{ij}^{(k+1)} = \begin{cases} a_{ij}^{(k)} & \text{si } i \leq k \\ a_{ij}^{(k)} - \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}\,a_{kj}^{(k)} & \text{si } i > k \end{cases} \qquad \text{et} \qquad b_i^{(k+1)} = \begin{cases} b_i^{(k)} & \text{si } i \leq k \\ b_i^{(k)} - \frac{b_i^{(k)}}{a_{kk}^{(k)}}\,b_k^{(k)} & \text{si } i > k \end{cases}$$

- **Après $\mathbf{n-1}$ étapes**, on a un système triangulaire supérieur, que l'on peut résoudre aisément.

Remarque : Les $a_{kk}^{(k)}$, qui sont les diviseurs apparaissant dans cet algorithme, sont appelés les **pivots**. Si l'un d'eux est nul, on peut alors permuter les lignes k à n jusqu'à trouver un pivot non-nul (toujours possible car rang $A = n$).

Par ailleurs, lorsqu'un pivot est petit (en valeur absolue), la précision des calculs sur ordinateur peut s'en ressentir. On a plutôt intérêt à adopter la **stratégie du pivot partiel**, qui consiste à intervertir deux lignes ($L_k^{(k)}$ et l'une des $L_i^{(k)}$ pour $i > k$) de façon à placer en $a_{kk}^{(k)}$ le maximum (en valeur absolue) de la colonne k.

Coût de cet algorithme : environ $n^3/3$ additions/soustractions, et $n^3/3$ multiplications/divisions.

2.4 Décomposition LU

La première étape de la méthode de Gauss s'écrit matriciellement :

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{n1}}{a_{11}} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix},$$

soit : $N_1\, A = A^{(2)}$, où la matrice N_1 est triangulaire inférieure, avec des 1 sur la diagonale, tandis que $A^{(2)}$ est la matrice déduite de A à l'issue de la première étape de la méthode de Gauss.

L'étape k s'écrit matriciellement de manière analogue :

$$\begin{pmatrix} 1 & 0 & \cdots & & \cdots & \cdots & 0 \\ 0 & \ddots & & & & & 0 \\ \vdots & & 1 & & & & \vdots \\ \vdots & & -\frac{a_{k+1,k}^{(k)}}{a_{kk}^{(k)}} & 1 & & & \vdots \\ \vdots & & \vdots & & & \ddots & \vdots \\ 0 & \cdots & -\frac{a_{n,k}^{(k)}}{a_{kk}^{(k)}} & 0 & & \cdots & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & \cdots & \cdots & a_{1n} \\ 0 & a_{22}^{(2)} & \cdots & & \cdots & \cdots & a_{2n}^{(2)} \\ \vdots & & \ddots & & & & \vdots \\ \vdots & & & a_{k-1,k-1}^{(k-1)} & \cdots & \cdots & a_{k-1,n}^{(k-1)} \\ \vdots & & & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & a_{nk}^{(k)} & \cdots & a_{nn}^{(k)} \end{pmatrix} = A^{(k+1)}$$

soit : $N_k\, A^{(k)} = A^{(k+1)}$.

Après $n-1$ étapes, on obtient donc :

$$\big(N_{n-1} \cdots N_2 N_1\big) A = A^{(n)} = U$$

où U est une matrice triangulaire supérieure. Les matrices N_k sont triangulaires inférieures avec des 1 sur la diagonale, de sorte que la matrice $L = \big(N_{n-1} \cdots N_2 N_1\big)^{-1}$ existe et est triangulaire inférieure.
Finalement,

$$L^{-1} A = U \;\Rightarrow\; A = LU.$$

• La méthode de Gauss fait donc implicitement la décomposition $A = LU$, où L est triangulaire inférieure avec des 1 sur la diagonale et U est triangulaire supérieure (U est en fait la matrice triangulaire obtenue en fin d'algorithme d'élimination). Connaître explicitement cette décomposition peut réduire énormément les coûts de calcul si l'on a à résoudre plusieurs systèmes $AX = b$ avec différents seconds membres b. En effet, résoudre $LUX = b$ revient simplement à résoudre $LY = b$ puis $UX = Y$, c'est à dire deux systèmes triangulaires. Ainsi, si l'on connait la décomposition LU, le coût de la résolution passe à seulement $2 \times n^2/2 = n^2$, au lieu de $n^3/3$ pour la méthode de Gauss.

Prop Lorsqu'elle existe, la décomposition LU est unique.

• L'obtention de la décomposition LU se fait par exemple par l'algorithme de **réduction de Crout**, qui consiste à écrire formellement l'égalité $A = LU$, et à identifier successivement les élements de la première ligne de A (ce qui donne la première ligne de U), puis ceux de la première colonne de A (ce qui donne la première colonne de L), puis ceux de la deuxième ligne de A (ce qui donne la deuxième ligne de U), etc.

Coût de cet algorithme : on retrouve, ce qui est logique, le coût de la méthode de Gauss, c'est à dire environ $n^3/3$ additions/soustractions, et $n^3/3$ multiplications/divisions. Autrement dit, la décomposition LU ne réduit pas le coût de calcul si l'on n'a qu'une seule résolution de système à effectuer.

Exemple d'utilisation : inversion d'une matrice Soit $A \in \mathcal{M}_n(K)$ de rang n. Si l'on note e_i le i-ème vecteur de la base canonique (c.a.d. le vecteur avec un 1 en i-ème ligne et des 0 ailleurs), et C_i la i-ème colonne de A^{-1}, l'égalité $AA^{-1} = I_n$ est équivalente à $AC_i = e_i$ ($i = 1, \ldots, n$). On a donc à résoudre n systèmes linéaires, ayant tous la même matrice A et des seconds membres différents. Plutôt que mettre en oeuvre n algorithmes de Gauss (ce qui coûterait $n^4/3$), on a intérêt à faire la décomposition LU de A (coût : $n^3/3$) puis à résoudre les n systèmes $LU\,C_i = e_i$ (coût : $n \times n^2$), soit un coût total de $4n^3/3$.

3 Déterminants

3.1 Formes n-linéaires alternées

Soit E un espace vectoriel de dimension n. Soit f définie de E^n vers K (une fonction à image dans K est appelée une **forme**). Autrement dit, f associe à n vecteurs $v_1, \ldots, v_n$ de E un scalaire $f(v_1, \ldots, v_n)$.

Déf f est dite **n-linéaire** ssi elle est linéaire par rapport à chaque variable, c.a.d.

$$f(v_1, \ldots, v_i + v_i', \ldots, v_n) = f(v_1, \ldots, v_i, \ldots, v_n) + f(v_1, \ldots, v_i', \ldots, v_n) \qquad \forall v_1, \ldots, v_n,\ \forall i$$
$$f(v_1, \ldots, \lambda v_i, \ldots, v_n) = \lambda\, f(v_1, \ldots, v_i, \ldots, v_n) \qquad \forall v_1, \ldots, v_n,\ \forall \lambda,\ \forall i$$

Déf Soit f une forme de E^n vers K. On dit que f est **alternée** ssi $f(v_1, \ldots, v_n) = 0$ dès que les v_i ne sont pas tous distincts.

Prop Les formes n-linéaires alternées forment un espace vectoriel de dimension 1. Autrement dit, deux formes n-linéaires alternées sont forcément multiples l'une de l'autre.

3.2 Déterminant d'une matrice carrée

Soit $A \in \mathcal{M}_n(K)$.

Déf On appelle **mineur d'indice $(\mathbf{i}, \mathbf{j})$**, notée M_{ij}, la matrice de $\mathcal{M}_{n-1}(K)$ obtenue en enlevant la i-ème ligne et la j-ème colonne de A.

Déf On appelle **déterminant** l'application qui à toute matrice $A \in \mathcal{M}_n(K)$ associe la valeur définie par récurrence de la façon suivante :

- si $n = 1$, $\det A = a_{11}$

- si $n > 1$, $\det A = \displaystyle\sum_{i=1}^{n} (-1)^{i+j}\, a_{ij} \det M_{ij}$ (développement par rapport à la j-ème colonne),

ou encore $\det A = \displaystyle\sum_{j=1}^{n} (-1)^{i+j}\, a_{ij} \det M_{ij}$ (développement par rapport à la i-ème ligne).

Pour montrer que toutes ces définitions par développement par rapport à une ligne ou à une colonne quelconque sont équivalentes, on montre tout d'abord qu'on définit ainsi des formes n-linéaires alternées. Elles sont donc égales entre elles à un facteur multiplicatif près. De plus, elles prennent toutes la même valeur sur la matrice identité I_n. Donc elles sont égales.

3.3 Quelques propriétés

Soit $A \in \mathcal{M}_n(K)$. On note L_k ses lignes et C_k ses colonnes.

- Le fait de remplacer L_k par $L_k + \displaystyle\sum_{i \neq k} \alpha_i\, L_i$, ou C_k par $C_k + \displaystyle\sum_{i \neq k} \alpha_i\, C_i$ ne change pas la valeur du déterminant.

- Si B est obtenue en permutant deux lignes ou deux colonnes de A, det$B$=-det$A$.
- Si B est obtenue en multipliant une ligne ou une colonne de A par $\lambda \in K$, alors det$B$=λ det$A$.
- detA^t=det$A$
- $\det(\lambda A) = \lambda^n \det A$
- En général, $\det(A+B) \neq \det A + \det B$
- detAB=det$A$.det$B$
- A est inversible ssi $\det A \neq 0$. Dans ce cas, detA^{-1}=1/det$A$
- Le système linéaire $AX = b$ a une solution unique ssi $\det A \neq 0$.
- Si A est triangulaire, det$A$=$\prod_{i=1}^{n} a_{ii}$
- Si A est triangulaire par blocs $A = \begin{pmatrix} A_{11} & \cdots & \cdots & A_{1n} \\ 0 & A_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & A_{nn} \end{pmatrix}$ avec chaque bloc A_{ii} carré, alors det$A$=$\prod_{i=1}^{n} \det A_{ii}$
- Coût de calcul de det$A$: environ $n!$ additions et multiplications. C'est un coût énorme. A titre d'exemple, 20! opérations prendraient des dizaines d'années de calcul sur un ordinateur à plusieurs gigaflops (c.a.d. plusieurs milliards d'opérations par seconde). On doit donc faire apparaitre un maximum de zéros et/ou de symétries dans le déterminant afin de réduire les calculs.

4 Réduction des endomorphismes et des matrices carrées

On rappelle qu'une application linéaire de E dans E est appelée **endomorphisme** de E. A toute matrice $A \in \mathcal{M}_n(K)$ correspond un endomorphisme f de E et réciproquement.

4.1 Eléments propres

Déf $\lambda \in K$ est **valeur propre** de f ssi $(\exists x \in E, x \neq 0 \ / \ f(x) = \lambda x)$

Déf $\lambda \in K$ est **valeur propre** de A ssi $(\exists X \in K^n, X \neq 0 \ / \ AX = \lambda X)$

Déf $x \in E$ est **vecteur propre** de f ssi ($x \neq 0$ et $\exists \lambda \in K \ / \ f(x) = \lambda x$)

Déf $X \in K^n$ est **vecteur propre** de A ssi ($X \neq 0$ et $\exists \lambda \in K \ / \ AX = \lambda X$)

Déf Soient $x \in E$, $x \neq 0$, et $\lambda \in K$ tels que $f(x) = \lambda x$. On dit que x est **vecteur propre de f associé à la valeur propre λ**.

Déf Soient $X \in K^n$, $X \neq 0$, et $\lambda \in K$ tels que $AX = \lambda X$. On dit que X est **vecteur propre de A associé à la valeur propre λ**.

Déf L'ensemble des valeurs propres de f (resp. de A) est appelé **spectre** de f (resp. de A).

Prop Soit $D \in \mathcal{M}_n(K)$ une matrice diagonale : $D = \begin{pmatrix} d_{11} & & 0 \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}$. Chaque d_{ii} est valeur propre de D associée au vecteur propre e_i (i-ème vecteur de la base canonique).

Prop L'ensemble des vecteurs propres associés à une valeur propre λ forme un sous-espace vectoriel, appelé **sous-espace vectoriel propre**, noté E_λ. $x \in E_\lambda$ ssi $f(x) = \lambda x$, c.a.d. ssi $(f - \lambda Id)(x) = 0$. Donc $E_\lambda = \ker(f - \lambda Id)$.

Prop Les sous-espaces vectoriels propres d'un endomorphisme sont en somme directe.

4.2 Polynôme caractéristique

λ est valeur propre de A ssi $\exists X \neq 0 \ / \ AX = \lambda X$, c.a.d. ssi $\exists X \neq 0 \ / \ (A - \lambda Id)X = 0$, c.a.d. ssi $\det(A - \lambda Id) = 0$.

Déf $\det(A - \lambda Id)$ est un polynôme de degré n en λ, appelé **polynôme caractéristique** de A, noté $P_A(\lambda)$. L'ensemble des valeurs propres de A est l'ensemble des racines de $P_A(\lambda)$. (on voit donc qu'il y a au plus n valeurs propres).

Prop Deux matrices semblables ont même polynôme caractéristique. Autrement dit, P_λ est invariant par changement de base.

Déf On appelle **ordre de multiplicité** d'une valeur propre son ordre de multiplicité en tant que racine de $P_A(\lambda)$.

Prop $P_A(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \operatorname{tr}(A) \lambda^{n-1} + \cdots + \det A$.

Prop La dimension du sev propre E_λ est inférieure ou égale à l'ordre de multiplicité de λ.

4.3 Diagonalisation

Déf On dit que A est **diagonalisable** ssi il existe une matrice diagonale D telle que A soit semblable à D. On a donc alors : $D = P^{-1}AP$, où P est une matrice de passage.

Déf On dit que f est **diagonalisable** ssi il existe une base $\mathcal{B}$ dans laquelle la matrice représentative de f soit diagonale.

Prop Les propriétés suivantes sont équivalentes :
(i) f est diagonalisable
(ii) Il existe une base formée de vecteurs propres de f
(iii) La somme des sev propres est égale à E
(iv) La somme des dimensions des sev propres est égale à n.

Déf On dit qu'un polynôme est **scindé** ssi il peut être factorisé par des monômes, c.a.d. être mis sous une forme $P(X) = \prod_{i=1}^{r} (X - x_i)^{m_i}$. Sur $\mathbb{C}$, tout polynôme est scindé.

Prop *Condition nécessaire et suffisante de diagonalisation* : f est diagonalisable ssi P_f est scindé ($P_f(\lambda) = \prod_{i=1}^{r} (\lambda - \lambda_i)^{m_i}$, où les λ_i sont distincts) et chaque sev propre E_{λ_i} a pour dimension l'ordre de multiplicité m_i de la valeur propre λ_i.

Prop *Condition suffisante de diagonalisation* : Si f admet n valeurs propres distinctes, alors f est diagonalisable.

Prop Toute matrice symétrique réelle est diagonalisable (et ses sous-espaces propres sont orthogonaux entre eux pour le produit scalaire usuel, ce qui signifie que deux vecteurs propres correspondant à des valeurs propres distinctes sont orthogonaux).

Bibliographie

[1] J. Lelong-Ferrand et J.-M. Arnaudiès: *Cours de mathématiques - Tome 1*, Dunod 2003.
[2] J.-M. Monier: *Algèbre et géométrie MP*, Dunod 2004.
[3] F. Pham et H. Dillinger: *Algèbre linéaire*, Diderot 1996.