

Algèbre linéaire en dimension finie

Christophe Antonini¹, Olivier Teytaud², Pierre Borgnat³, Annie Chateau⁴, and
Edouard Lebeau⁵

¹Enseignant en CPGE, Institut Stanislas, Cannes

²Chargé de recherche INRIA, Université d'Orsay, Orsay

³Chargé de recherche CNRS, ENS Lyon, Lyon

⁴Maitre de conférence, Université Montpellier-2, Montpellier

⁵Enseignant en CPGE, Lycée Henri Poincaré, Nancy

27 octobre 2022



Algèbre linéaire en dimension finie et calcul matriciel.

1 Algèbre linéaire en dimension finie

On présentera quelques généralités sur l'algèbre linéaire en dimension finie en section 1.1 et sur la dualité en dimension finie en section 1.2.

On consacrera ensuite plusieurs sections aux matrices : le calcul matriciel en section 1.3, les opérations sur les lignes et les colonnes en section 1.3.12 ; enfin, deux dernières sections seront consacrées aux matrices avec des exercices et de la zoologie (sections 1.3.13 et 1.4 respectivement).

On passera ensuite à des sections plus originales ; on reviendra un peu sur la dualité en dimension finie avec un peu de zoologie (section 1.5).

Les espaces vectoriels ont bien sûr une interprétation physique immédiate, en particulier pour la mécanique. Ils sont aussi à la base de structures algébriques comme les espaces de Hilbert, qui auront beaucoup d'applications en analyse fonctionnelle (par exemple pour l'analyse de Fourier). Ils servent aussi à caractériser des solutions d'équations différentielles (l'ensemble des solutions étant en général un espace affine, de dimension évaluable a priori selon les caractéristiques de l'équation). L'analyse matricielle sera à la base de l'optimisation via la matrice hessienne ou les matrices représentant les contraintes en programmation linéaire.

1.1 Généralités

DÉFINITION 0.1 de dimension finie

Un espace vectoriel est dit **de dimension finie** lorsqu'il admet une base finie. Dans le cas contraire il est dit **de dimension infinie**.

Dans un espace fini le cardinal d'une base est appelé **dimension** de l'espace (pour la cohérence de la définition il faudra voir que toutes les bases ont alors même cardinal). Un sous-espace vectoriel d'un espace vectoriel E est dit de **codimension finie** si la dimension de l'espace quotient est finie. On appelle alors **codimension** de cet espace la dimension de l'espace quotient. Sinon il est dit de **codimension infinie**. La notion de base a été définie dans la partie ??.

Dans la suite, E désigne un \mathbb{K} -espace vectoriel de dimension finie.

THÉORÈME 0.1 Théorème de la base incomplète

Si I est une famille libre et K une famille génératrice finie, avec $I \subset K$, alors il existe J avec $I \subset J \subset K$ tel que J soit une base.

Application On verra des applications avec (entre autres) les théorèmes 0.6 (existence de supplémentaire) et 0.18 (à propos d'orthogonalité dans le dual), la proposition 0.30 (rang et matrices extraites).

Démonstration On considère J libre maximale au sens du cardinal vérifiant $I \subset J \subset K$. Si toute famille plus grande incluse dans K est liée, alors tout élément de K est combinaison linéaire d'éléments de J ; tout élément de E s'écrit comme combinaison d'éléments de K , qui eux-mêmes s'écrivent comme combinaisons linéaires d'éléments de J . Donc la famille J est génératrice.

LEMME 0.2 Lemme de Steinitz

Si E est non réduit à $\{0\}$, E admettant une famille génératrice I de cardinal n , toute famille de $n + 1$ vecteurs (ou plus) est liée.

Démonstration

- Le cas $n = 1$ est immédiat. On procède alors par récurrence.
- Supposons la propriété vraie pour $1, 2, \dots, n - 1$; supposons E admettant une famille génératrice b_1, \dots, b_n de cardinal n . Notons H le sous-espace vectoriel de E engendré par b_1, \dots, b_{n-1} .
- Donnons-nous une famille e_1, \dots, e_{n+1} de E . Notre but est de montrer que cette famille est liée.
- Soit, pour $i \in \llbracket 1, n + 1 \rrbracket$, $h_i \in H$ et $\lambda_i \in \mathbb{K}$ tel que $e_i = h_i + \lambda_i b_n$.
- Si tous les λ_i sont nuls, alors l'hypothèse de récurrence appliquée à H permet de conclure (tous les e_i sont alors dans H , et donc ils sont liés).
- Supposons alors (sans perte de généralité) $\lambda_1 \neq 0$; dans ce cas, $b_n = \frac{1}{\lambda_1}(e_1 - h_1)$.
- On peut alors écrire :

$$e_i - h_i = \frac{\lambda_i}{\lambda_1}(e_1 - h_1)e'_i = e_i - \frac{\lambda_i}{\lambda_1}e_1 = \underbrace{h_i}_{\in H} - \underbrace{\frac{\lambda_i}{\lambda_1}h_1}_{\in H}$$

Les e'_i pour $i \in \llbracket 2, n + 1 \rrbracket$ appartiennent donc à H ; ils sont n , et appartiennent à un espace généré par $n - 1$ points, et donc par hypothèse de récurrence ils sont liés. On a donc une combinaison linéaire nulle

des e'_i pour $i \in [[2, n + 1]]$, à coefficients non tous nuls. On a donc une combinaison linéaire nulle des e_i pour $i \in [[1, n + 1]]$, dont au moins un coefficient est non-nul; ceci conclut la preuve.

THÉORÈME 0.3

Dans un espace de dimension finie, toutes les bases ont le même cardinal.

Démonstration Conséquence immédiate du lemme 0.2.

THÉORÈME 0.4

Deux espaces vectoriels de dimension finie sur le même corps sont isomorphes si et seulement si ils ont même dimension.

Démonstration S'ils sont isomorphes, l'image d'une base de l'un par un isomorphisme est base de l'autre, donc ils ont même dimension. S'ils ont même dimension, alors avec (e_1, \dots, e_n) une base de l'un, et (f_1, \dots, f_n) une base de l'autre, l'application $\sum \lambda_i e_i \mapsto \sum \lambda_i f_i$ pour $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ est un isomorphisme, comme on le vérifie facilement (la fonction est bien définie car (e_1, \dots, e_n) est une base, linéarité évidente, injectivité immédiate car (f_1, \dots, f_n) est libre, surjectivité immédiate car (f_1, \dots, f_n) est génératrice).

THÉORÈME 0.5

F , sous-espace vectoriel de E , est égal à E si et seulement si $\dim E = \dim F$.

Démonstration Il est clair que si F et E sont égaux, alors ils ont même dimension. Réciproquement, s'ils ont même dimension, alors une base de F est une famille libre de même cardinal qu'une base de E , donc c'est une base de E (voir le lemme de Steinitz ci-dessus).

THÉORÈME 0.6

Soit E espace vectoriel non nul. Tout sous-espace vectoriel de E admet un supplémentaire, et si $E = F \oplus G$, alors $\dim E = \dim F + \dim G$, et une base de E s'obtient en réunissant une base de F et une base de G .

Démonstration Conséquence du théorème de la base incomplète.

Maintenant quelques théorèmes faciles, sans démonstration, mais fort pratiques néanmoins :

THÉORÈME 0.7

Etant donnés F et G des sous-espaces vectoriels de E , on a $\dim (F + G) + \dim (F \cap G) = \dim F + \dim G$.

On consultera avec profit la section ?? quant à la définition d'une topologie sur les sous-espaces vectoriels d'un espace vectoriel .

THÉORÈME 0.8 Caractérisations de supplémentaires

F et G sont supplémentaires dans E si et seulement si leur intersection est nulle et si la somme de leurs dimensions est la dimension de E .

F et G sont supplémentaires dans E si et seulement si leur intersection est nulle et si leur somme est égale à E .

F et G sont supplémentaires dans E si et seulement si leur somme est égale à E et si la somme de leurs dimensions est égale à la dimension de E .

THÉORÈME 0.9

Etant donné F sous-espace vectoriel de E , la dimension de E/F est égale à la dimension de E moins la dimension de F . D'ailleurs, E/F est isomorphe à tout supplémentaire de F dans E .

THÉORÈME 0.10

Si les E_i sont de dimension finie, alors $\dim \prod_{i=1}^n E_i = \sum_{i=1}^n \dim E_i$.

THÉORÈME 0.11

Si E et F sont de dimension finie, alors $\mathcal{L}(E, F)$ est de dimension finie et $\dim \mathcal{L}(E, F) = \dim E \cdot \dim F$.

Démonstration On considère une base de E et une base de F , notées respectivement (e_i) et (f_i) ; alors les $\varphi_{i,j}$ définies par $\varphi_{i,j}(e_i) = f_j$ et $\varphi_{i,j}(e_l) = 0$ pour $l \neq i$ forment une base de $\mathcal{L}(E, F)$.

DÉFINITION 0.2 rang

On appelle **rang** d'une famille finie de vecteurs la dimension de l'espace vectoriel que cette famille engendre.

On appelle **rang** d'une application linéaire la dimension de son image lorsque celle-ci est finie. C'est en fait le rang de l'image d'une base lorsque le domaine est un espace vectoriel de dimension finie

THÉORÈME 0.12 Théorème du rang

Si $f \in \mathcal{L}(E, F)$ et E de dimension finie, alors $\text{Im } f$ et $\text{Ker } f$ sont de dimension finie, et $\dim E = \dim \text{Im } f + \dim \text{Ker } f$.

Démonstration On considère un supplémentaire du noyau, et on montre qu'il est isomorphe à l'image de f (voir théorème ??).

COROLLAIRE 0.13

Soit f une application d'un espace vectoriel E vers un espace vectoriel F , avec E et F de même dimension finie, alors f est un isomorphisme si et seulement si l'une des propriétés suivantes est vérifiée :

- f a un noyau réduit au singleton $\{0\}$.
- f est injective.
- f est surjective.
- f est bijective.
- Le rang de f est égal à la dimension de E .

PROPOSITION 0.14

Quelques propriétés du rang :

- Le rang d'une famille finie de vecteurs est invariant par permutations
- Multiplier un vecteur d'une famille par un scalaire non nul ne change pas le rang de la famille
- On ne change pas le rang d'une famille de vecteurs en additionnant à un vecteur une combinaison linéaire des autres vecteurs
- Le rang d'une famille de vecteurs est le même si l'on supprime un vecteur nul

1.2 Dualité en dimension finie

La dualité existe dans plusieurs cadres : géométrie, algèbre, topologie. On s'intéresse ici à la dualité au sens algébrique. Le lecteur est invité à différencier les différentes dualités par un rapide tour d'horizon par l'entrée "dualité" de l'index. De même, il convient de ne pas confondre l'orthogonal au sens géométrique, et l'orthogonal dans le dual.

1.2.1 Dualité simple

E est un espace vectoriel normé de dimension finie n .

DÉFINITION 0.3 formes coordonnées associées

Etant donnée une base $(e_i)_{i \in \llbracket 1, n \rrbracket}$ de E , on appelle **formes coordonnées associées** les n formes linéaires e_j^* définies par $e_j^*(e_i) = \delta_{i,j}$.

On note que $x = \sum_{j=1}^n e_j^*(x) \cdot e_j$.

THÉORÈME 0.15

Quelques propriétés éliminatoires de la dualité *en dimension finie* :

- $\dim E^* = \dim E$.
- La famille des (e_j^*) est une base de E^* ; on l'appelle **la base duale** de la base (e_i) .
- Tout f appartenant à E^* s'écrit $f = \sum_i f(e_i) \cdot e_i^*$
- Pour toute base (f_i) de E^* , il existe une base de E dont la base duale est la base des (f_i) .

Application Le dernier point servira par exemple pour relier la dimension d'un sous-espace vectoriel et celle de son orthogonal dans le dual (théorème 0.19).

Démonstration La dimension de E^* est égale à la dimension de E , car $\dim \mathcal{L}(E, \mathbb{K}) = \dim E \cdot \dim \mathbb{K} = \dim E$ (cf théorème 0.11); il suffit donc de montrer que la famille des (e_i^*) est libre. On se donne une combinaison linéaire nulle, dont un coefficient au moins est non nul, des e_j^* ; en considérant l'image de e_i par cette combinaison linéaire nulle, on voit que le coefficient de e_i^* est nul pour tout i .

Donnons-nous ensuite une base f_1, \dots, f_n de E^* . On se donne une application ϕ qui à x associe $(f_1(x), \dots, f_n(x))$. On montre sans difficultés que ϕ est un isomorphisme de E dans \mathbb{K}^n . Donc il existe une base (e_i) telle que $\phi(e_i) = (\delta_{1,i}, \dots, \delta_{i,i}, \dots, \delta_{n,i})$; cette base convient.

On donne sans démonstration le théorème aisé suivant (utilisant la transposition, définie en section ??) :

THÉORÈME 0.16

Soient E et F des \mathbb{K} -espaces vectoriels de dimensions finies (non nécessairement égales). Alors l'application transposition qui à $f \in \mathcal{L}(E, F)$ associe ${}^t f \in \mathcal{L}(F^*, E^*)$ est un isomorphisme. En outre le rang de ${}^t f$ est égal au rang de f et

$$\text{Ker } {}^t f = (\text{Im } f)^\perp$$

1.2.2 Bidual

E est un espace vectoriel de dimension finie n .

THÉORÈME 0.17 Bidual

L'application

$$x \mapsto (\phi \mapsto \phi(x))$$

c'est-à-dire $x \mapsto x^{**}$ avec $x^{**} = \phi \mapsto \phi(x)$, de E dans E^{**} qui à x associe la fonction qui à ϕ dans E^* associe $\phi(x)$ est un isomorphisme; on l'appelle **isomorphisme canonique** de E dans E^{**} . On peut donc identifier, via cet isomorphisme, E et E^{**} .

Démonstration Les dimensions de E et E^{**} étant égales (on est toujours dans le cadre de la dimension finie) il suffit de voir que cette fonction est un morphisme injectif. Supposons x d'image nulle; alors quel que soit f , $f(x) = 0$. Si x est non nul, alors on peut l'appeler e_1 et le compléter en une base e_i ; alors on constate que $e_i^*(x) = 0$ pour tout i , ce qui est contradictoire.

1.2.3 Orthogonalité

Ne confondons pas l'orthogonal au sens géométrique (dans un espace muni d'une forme quadratique, cf section ??, dont les exemples les plus intuitifs sont les espaces de Hilbert, en particulier euclidiens ou hermitiens) et l'orthogonal au sens de la dualité.

E est un espace vectoriel normé de dimension finie n .

THÉORÈME 0.18

Quel que soit F sous-espace vectoriel de E , on a :

- $\dim E = \dim F + \dim F^\perp$
- $F^{\perp\circ} = F$

Démonstration Si $F = \{0\}$ alors le résultat est clair. Sinon, on considère une base $(f_i)_{i \in [1, r]}$ de F , et on la complète en une base $(f_i)_{i \in [1, n]}$ de E . On considère la base duale (f_i^*) ; il est clair alors que les $(f_i)_{i > r}$ forment une famille libre génératrice de F^\perp , d'où la première assertion. On constate bien que $F^{\perp\circ}$ est inclus dans F ; la dimension permet de conclure.

THÉORÈME 0.19

Quel que soit F sous-espace vectoriel de E^* , on a :

- $\dim E = \dim F + \dim F^\circ$
- $F^{\circ\perp} = F$

Démonstration La méthode est la même que précédemment ; il faut juste se souvenir que toute base du dual est la base duale d'une certaine base (cf théorème 0.15).
Donnons enfin sans démonstration le théorème aisé suivant :

THÉORÈME 0.20

Avec ϕ l'isomorphisme canonique de E dans E^{**} , pour tout F sous-espace vectoriel de E , on a $F^{\perp\perp} = \phi(F^{\perp\circ}) = \phi(F)$.

1.3 Calcul matriciel

Le calcul matriciel sert partout. Un grand nombre de problèmes industriels se ramènent à des problèmes linéaires ou quadratiques, donc de représentation matricielle. On parle en général de problème "linéaire" lorsqu'il s'agit de chercher des minima de fonctions de la forme $x \mapsto \sum_i \lambda_i x_i$ pour x tel que $Ax \geq c$ avec A une matrice et c un vecteur. On parle de problèmes quadratiques en général pour la recherche de minima de fonctions de la forme $x \mapsto Q(x)$, avec Q une forme quadratique, pour x tel que $Ax \geq c$ (Notons que parfois on parle de problèmes quadratiques avec des contraintes $B(x) \leq c$ avec B une forme quadratique définie positive). Les problèmes non-linéaires sont en général la recherche de minima de $x \mapsto f(x)$ parmi les x tels que $\forall i \in I, g_i(x) \geq 0$, lorsque le problème ne peut se ramener ni au linéaire, ni au quadratique. De manière amusante, le non-linéaire n'est donc pas simplement ce qui n'est pas linéaire ; le quadratique n'est ni linéaire ni non-linéaire. Même les problèmes non-linéaires se ramènent fréquemment soit à une résolution successive d'un grand nombre de problèmes linéaires, soit à une résolution successive d'un grand nombre de problèmes quadratiques. Les problèmes non-linéaires que l'on ne résout pas par retour à des problèmes quadratiques ou linéaires utilisent néanmoins des matrices et en particulier la matrice dite matrice hessienne : avec f une application suffisamment dérivable, la matrice hessienne de f en x est

$$M_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Les propriétés de cette matrice sont en particulier liées aux extrêma de f ; voir le chapitre ?? . On appelle **hessien** le déterminant de la matrice hessienne. **Attention** Ne pas dire que la hessienne est définie positive dès que le hessien est positif! La hessienne, matrice symétrique par le théorème de Schwartz, est définie positive si pour tout vecteur x non nul, $x'Mx > 0$ avec x' le transposé de x (ou, de manière équivalente, si toutes ses valeurs propres sont positives strictement, ou, de manière encore équivalente, si toute sous-matrice principale a un déterminant positif).

On parlera ici de matrices finies, i.e. avec un nombre fini de lignes et de colonnes. On peut aussi considérer des matrices avec une quantité dénombrables de lignes et de colonnes. Cela sera en particulier utile pour l'analyse du recuit simulé ou des Monte-Carlo par chaînes de Markov. Ces techniques, stochastiques, permettent l'optimisation dans des espaces non-continus (pour le recuit simulé, mais aussi les algorithmes dits évolutionnaires ou les algorithmes dits génétiques) ou le calcul d'intégrales dans des espaces délicats (pour Monte-Carlo par chaînes de Markov). La généralisation à des quantités non dénombrables est usuellement qualifiée de "noyau de transition", i.e. une application $k(x, y)$ (pour x et y dans un même domaine non-dénombrable) telle que $\int k(x, \cdot) = 1$ pour tout x .

Les représentations canoniques de matrices sont capitales pour représenter des problèmes de manière simple. Il est indispensable de bien maîtriser ce que signifie que deux matrices sont équivalentes, ou le fait qu'elles sont semblables, et surtout ne pas croire que c'est pareil.

1.3.1 Bases sur les matrices

DÉFINITION 0.4 Définitions de base sur les matrices

On appelle **matrice de type (n, p) sur un corps \mathbb{K}** toute application de $\llbracket 1, n \rrbracket \times \llbracket 1, p \rrbracket$ (intervalles de \mathbb{N}) dans \mathbb{K} . On la représente généralement comme suit :

$$\begin{pmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,p} \\ m_{2,1} & m_{2,2} & \dots & m_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,p} \end{pmatrix}$$

On note $M_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n, p) sur le corps \mathbb{K} .

On appelle **matrice ligne** une matrice de type $(1, p)$, et **matrice colonne** une matrice de type $(n, 1)$.

On appelle **matrice extraite** d'une matrice de type (n, p) la restriction de cette matrice à $I \times J$, avec $I \subset \llbracket 1, n \rrbracket$ et $J \subset \llbracket 1, p \rrbracket$.

On appelle **i -ième vecteur-ligne** de la matrice M de type (n, p) la restriction de cette matrice à $\{i\} \times \llbracket 1, p \rrbracket$. On peut identifier un vecteur-ligne à un élément de \mathbb{K}^p .

On appelle **j -ième vecteur-colonne** de la matrice M de type (n, p) la restriction de cette matrice à $\llbracket 1, n \rrbracket \times \{j\}$. On peut identifier un vecteur-colonne à un élément de \mathbb{K}^n .

On appelle **matrice associée** à un morphisme f de l'espace vectoriel E de dimension p dans l'espace vectoriel F de dimension n et aux bases $B = (e_i)$ et $B' = (f_i)$ de E et F respectivement la matrice M de type (n, p) telle que $M_{i,j} = f_i^*(e_j)$. On la note $Mat_{B,B'}(f)$.

Inversement, on appelle **application linéaire canoniquement associée à la matrice M** le morphisme de \mathbb{K}^p dans \mathbb{K}^n dont la matrice associée est M .

Application L'introduction de ce chapitre fournit une liste d'applications des matrices. On peut citer aussi les applications de différentes formes spécialisées de matrices à des domaines directement opérationnels :

- matrices de Hadamard, par exemple pour la discrépance (cf e.g. [1]) ;
- matrices stochastiques pour l'étude de chaînes de Markov à espaces d'états finis ou dénombrables (voir [2]) ;
- déterminants pour l'analyse de plans d'expérience ;

et on en trouvera d'autres en zoologie 1.4.

Commençons par deux propositions aisées :

PROPOSITION 0.21

Une matrice de type n, p admet $C_n^{n'}.C_p^{p'}$ matrices extraites de type (n', p') .

PROPOSITION 0.22

E et F étant de dimension respectives p et n sur le corps \mathbb{K} , on a

$$\mathcal{L}(E, F) \simeq \mathcal{M}_{n,p}(\mathbb{K})$$

via l'isomorphisme $f \mapsto \text{Mat}_{B,B'}(f)$.

Avec $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$, et B et B' les bases canoniques, on a alors un isomorphisme canonique entre $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ et $\mathcal{M}_{n,p}(\mathbb{K})$.

PROPOSITION 0.23 Matrice canonique d'une application linéaire en dimension finie

Soit f une application linéaire entre E , espace vectoriel de dimension p , et F , espace vectoriel de dimension n ; soit r le rang de f . Alors il existe une base B de E et une base B' de F telles que

$$\text{Mat}_{B,B'}(f) = M$$

$$\text{avec } M_{i,j} = 1 \text{ si } i = j \leq r \quad M_{i,j} = 0 \text{ sinon}$$

On appelle cette matrice **matrice canonique de f** .

Démonstration Considérer une base B_1 d'un supplémentaire du noyau de f ; considérer $B'_1 = f(B_1)$; considérer B_2 une base du noyau de f , et compléter B'_1 en une base B' . Il reste à considérer $B = B_1 \cup B_2$.

Remarque 0.1

La matrice d'une forme linéaire est une matrice-ligne.

DÉFINITION 0.5 produit matriciel

On appelle **matrice produit** des matrices A et B de types respectifs (n, q) et (q, p) la matrice C de type (n, p) telle que

$$C_{i,j} = \sum_{k=1}^q A_{i,k} \cdot B_{k,j}$$

On note $C = A \times B$ ou $C = A.B$.

Enfin des propriétés faciles, mais capitales :

PROPOSITION 0.24

- Le produit matriciel défini ci-dessus est associatif et bilinéaire.
- Avec $A = \text{Mat}_{B,B'}(f)$ et $B = \text{Mat}_{B',B''}(g)$, $B \times A = \text{Mat}_{B,B''}(g \circ f)$
- Avec $x \in E$, X le vecteur des coordonnées de x dans la base B , alors les coordonnées de $y = f(x)$ dans la base B' sont celles de $Y = M \times X$, et $M = \text{Mat}_{B,B'}(f)$.

1.3.2 Espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$

E et F \mathbb{K} -espace vectoriel de dimensions respectives p et n .

DÉFINITION 0.6 **matrices élémentaires de type (n, p)**

On appelle **matrices élémentaires de type (n, p)** les matrices $E_{i,j}$ avec $E_{i,j}(a, b) = \delta_{a,i} \cdot \delta_{b,j}$; c'est-à-dire les matrices de type (n, p) ne comportant qu'un 1 (en position (i, j)) et des 0 partout ailleurs.

PROPOSITION 0.25

$\mathcal{M}_{n,p}(\mathbb{K})$ est un \mathbb{K} -espace vectoriel pour l'addition terme à terme et pour la multiplication terme à terme par un scalaire.

L'application de $\mathcal{L}(E, F)$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ qui à une application f associe $\text{Mat}_{B,B'}(f)$ est un isomorphisme.

Les matrices élémentaires forment une base de cet espace vectoriel qui est donc de dimension $n.p$.

On retrouve ainsi le fait que $\dim \mathcal{L}(E, F) = \dim E \cdot \dim F$.

1.3.3 Transposition de matrices

On définit ici la transposition de matrices et son lien avec la transposition (plus abstraite) de morphismes d'espaces vectoriels (cf section ??).

DÉFINITION 0.7 **Transposée d'une matrice**

Etant donnée une matrice M on appelle **matrice transposée** de M la matrice N avec $N_{i,j} = M_{j,i}$. Si M est de type (n,p) , alors N est de type (p,n) .
On note $N = {}^tM$.

PROPOSITION 0.26

- L'application $M \mapsto {}^tM$ est un isomorphisme entre $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{M}_{p,n}(\mathbb{K})$ en tant que \mathbb{K} -espaces vectoriels .
- ${}^t(A \times B) = {}^tB \times {}^tA$
- $Mat_{B'^*, B^*}({}^t f) = {}^t Mat_{B, B'}(f)$

1.3.4 Le cas des matrices carrées : la \mathbb{K} -algèbre $\mathcal{M}_n(\mathbb{K})$

DÉFINITION 0.8 **matrice carrée**

On appelle **matrice carrée** une matrice de type (n,n) pour un certain n . On note $\mathcal{M}_n(\mathbb{K}) = \mathcal{M}_{n,n}(\mathbb{K})$.

On appelle **matrice d'un endomorphisme** f associée à une base B (finie) la matrice $Mat_{B,B}(f)$; on la note aussi $Mat_B(f)$.

On appelle **diagonale** d'une matrice carrée M de type (n,n) le vecteur $(M_{1,1}, \dots, M_{i,i}, \dots, M_{n,n})$.

On appelle **trace** d'une matrice carrée M la somme $\sum_{i=1}^n M_{i,i}$. On la note $tr(M)$. L'application $M \rightarrow tr(M)$ est une application linéaire.

On appelle **matrice unité d'ordre** n la matrice M avec $M_{i,j} = \delta_{i,j}$. C'est la matrice dans toute base de l'endomorphisme identité.

On appelle **matrice scalaire** une matrice égale à $\lambda.I$ avec λ un scalaire et I une matrice unité. C'est la matrice dans toute base de l'homothétie de rapport λ .

On appelle **matrice diagonale associée à un n -uplet** m la matrice M de type (n,n) définie par $M_{i,i} = m_i$ et $M_{i,j} = 0$ si $i \neq j$. On note $M = diag(m)$.

Une matrice est dite **symétrique** si elle est égale à sa transposée.

Une matrice M est dite **antisymétrique** si elle est égale à l'opposée de sa transposée, c'est-à-dire si ${}^tM = -M$.

Une matrice carrée est dite **triangulaire supérieure** si $j < i \rightarrow M_{i,j} = 0$

On note \mathcal{T}_n^s l'ensemble des matrices carrées triangulaires supérieures d'ordre n .

Une matrice carrée est dite **triangulaire inférieure** si $j > i \rightarrow M_{i,j} = 0$

On note \mathcal{T}_n^i l'ensemble des matrices carrées triangulaires inférieures d'ordre n .

PROPOSITION 0.27

Quelques propriétés des matrices carrées :

- $\mathcal{M}_n(\mathbb{K})$ est une \mathbb{K} -algèbre, isomorphe à la \mathbb{K} -algèbre des endomorphismes de \mathbb{K}^n (ou de E , pour tout espace vectoriel E de dimension n).
- Si M est inversible, alors sa transposée aussi et $({}^tM)^{-1} = {}^t(M^{-1})$.
- $tr(AB) = tr(BA)$ (que A et B soient carrées ou non, pourvu que le produit soit carré)
- Une matrice est scalaire si et seulement si c'est la matrice associée à un endomorphisme de la forme $x \mapsto \lambda.x$, i.e. à une homothétie.
- L'ensemble des matrices diagonales de $\mathcal{M}_n(\mathbb{K})$ est une sous-algèbre commutative de $\mathcal{M}_n(\mathbb{K})$.
- L'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{K})$ et l'ensemble des matrices antisymétriques de $\mathcal{M}_n(\mathbb{K})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{K})$; si \mathbb{K} n'est pas de caractéristique 2, ils sont supplémentaires; ils sont alors de dimensions respectives $\frac{n.(n+1)}{2}$ et $\frac{n.(n-1)}{2}$, engendrés respectivement par les $E_{i,j} + E_{j,i}$ pour $i \leq j$ et par les $E_{i,j} - E_{j,i}$ pour $i < j$. Toute matrice carrée M s'écrit $A + S$, avec A antisymétrique et S symétrique, avec $A = \frac{1}{2}(M - {}^tM)$ et $S = \frac{1}{2}(M + {}^tM)$
- Le produit de deux matrices symétriques est symétrique si et seulement si les deux matrices commutent.
- L'ensemble des matrices triangulaires supérieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $n.(n+1)/2$.
- L'ensemble des matrices triangulaires inférieures est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$, de dimension $n.(n+1)/2$.
- L'ensemble des matrices triangulaires inférieures est isomorphe à l'ensemble des matrices triangulaires supérieures. La transposition réalise un isomorphisme entre ces espaces vectoriels.
- Les éléments inversibles de l'ensemble des matrices triangulaires supérieures (resp. inférieures) sont les matrices triangulaires dont tous les coefficients diagonaux sont non nuls; ils forment un sous-groupe multiplicatif de $\mathcal{M}_n(\mathbb{K})$.
- Etant donnée une matrice A de $\mathcal{M}_n(\mathbb{K})$, on appelle **commutant de A** le sous-ensemble de $\mathcal{M}_n(\mathbb{K})$ des matrices commutant avec A .
- Etant donnée A une matrice de $\mathcal{M}_n(\mathbb{K})$, on définit $A^0 = I$ avec I la matrice unité d'ordre n , et $A^{i+1} = A.A^i$ (au sens du produit matriciel). On peut ainsi définir étant donné un polynôme l'image de A par ce polynôme, en utilisant la multiplication matricielle et la multiplication par un scalaire.
- Etant donnée A une matrice de $\mathcal{M}_n(\mathbb{K})$, l'application de $\mathbb{K}[X] \rightarrow \mathcal{M}_n(\mathbb{K})$ qui à P associe $P(A)$ est un morphisme d'algèbres. Son image est une sous-algèbre de $\mathcal{M}_n(\mathbb{K})$; on la note $\mathbb{K}(A)$, elle est commutative.
- $\mathbb{K}(A)$ est de dimension finie, au plus n .

Démonstration Seul le dernier point pose difficulté. Il nécessitera le théorème de Cayley-Hamilton, qui montre que $A^n \in Vect(A^0, \dots, A^{n-1})$. La suite de la preuve est claire.

1.3.5 Changement de bases

Soit E espace vectoriel de dimension n , et (e_i) et (f_j) des bases de E .

DÉFINITION 0.9 **matrice de passage**

On appelle **matrice de passage** de la base (e_i) à la base (f_j) la matrice P de type (n, n) définie par $P_{i,j} = e_i^*(f_j)$; on la note $P_{(e_i),(f_j)}$. Il s'agit en fait de la matrice $Mat_{(f_j),(e_i)}(Id)$.

PROPOSITION 0.28

Quelques propriétés des matrices de passage :

- Le produit de la matrice de passage de B à B' par la matrice de passage de B' à B'' est la matrice de passage de B à B'' .
- $P_{B,B'}^{-1} = P_{B',B}$
- Etant donné X le vecteur des coordonnées de x dans une base B , les coordonnées de x dans la base B' sont données par X' avec $X' = P_{B',B}X$.
- $Mat_{B',C'}(f) = P_{C',C} \cdot Mat_{B,C}(f) \cdot P_{B,B'}$
- Dans le cas d'un endomorphisme $Mat_{B'}(f) = P_{B',B} \cdot Mat_B(f) \cdot P_{B,B'}$

Attention La matrice de passage de B à B' donne les coordonnées dans B en fonction des coordonnées dans B' et pas le contraire... La terminologie vaut ce qu'elle vaut ! On peut le retenir en considérant que la matrice de passage de la base B à la base B' est la matrice dans B de l'endomorphisme dont l'image de B est B' .

1.3.6 Groupe linéaire et groupe spécial linéaire

Voir ??.

1.3.7 Groupe orthogonal réel et groupe spécial orthogonal réel

Voir ??.

Ne pas confondre avec groupe linéaire et groupe spécial linéaire.

1.3.8 Rang d'une matrice

DÉFINITION 0.10 Matrice associée à une famille finie de vecteurs

Etant donnée une famille de vecteurs (x_1, \dots, x_p) d'un espace vectoriel de dimension finie n et une base (v_1, \dots, v_n) de E , on appelle **matrice associée à la famille des x_i et à la base des v_i** la matrice M définie par $M_{i,j} = v_i^*(x_j)$.

On appelle **rang d'une matrice** M le rang de la famille de ses vecteurs colonnes. On le note $rg(M)$.

Une matrice N extraite de M , avec N carrée inversible d'ordre le rang de M , est appelée **matrice principale** de M .

Appliation Les matrices principales permettent de caractériser commodément les matrices définies positives, qui servent beaucoup en optimisation; voir la section ??.

La preuve des propriétés qui suivent est simple et permet de jongler un peu entre matrices et morphismes :

PROPOSITION 0.29

Quelques propriétés des rangs de matrices : •Le rang d'une matrice associée à un système de vecteurs et à une base est indépendant de cette base.

- $rg(Mat_{B,B'}(f)) = rg(f)$
- $rg({}^t M) = rg(M) \leq \min(n, p)$, avec M de type (n, p) .
- $rg(M) = n \iff f$ surjective
- $rg(M) = p \iff f$ injective
- $rg(MM') \leq \min(rg(M), rg(M'))$

PROPOSITION 0.30

Le rang d'une matrice M non nulle est égal à n maximal tel qu'il existe une matrice extraite inversible de type (n, n) de M .

Démonstration Il est clair que le rang de M est supérieur au rang de toute matrice extraite inversible (considérer une combinaison linéaire nulle des vecteurs correspondants de M).

Etant donné r le rang de M il existe une famille libre de r vecteurs parmi la famille des vecteurs colonnes de M (par le théorème de la base incomplète).

On peut alors considérer la transposée de cette matrice son rang est le même, et donc on peut se restreindre au même nombre de colonnes.

1.3.9 Matrices équivalentes, matrices semblables

Nous allons introduire et étudier les matrices semblables et les matrices équivalentes ; il faut bien noter que deux matrices sont équivalentes si elles représentent le même morphisme dans des bases différentes, alors que deux matrices sont semblables si elles représentent le même endomorphisme dans des bases différentes. Pour l'endomorphisme, on n'a qu'une base à changer ; pour un morphisme, on peut changer la base de départ et la base d'arrivée. Beaucoup de matrices de même type sont équivalentes : il suffit qu'elles aient même rang. La relation "être semblable" est beaucoup plus rare.

On va parler ici de matrices inversibles, mais on n'a pas de préoccupation vis à vis d'une éventuelle forme quadratique équipant l'espace vectoriel .

DÉFINITION 0.11 Matrices équivalentes

Deux matrices A et B de même type (n, p) sont dites **équivalentes** si il existe P et Q des matrices inversibles de types respectifs (p, p) et (n, n) telles que

$$B = Q.A.P$$

PROPOSITION 0.31

Quelques propriétés immédiates :

- Il s'agit d'une relation d'équivalence.
- Deux matrices de même type sont équivalentes si et seulement si elles représentent un même morphisme dans des bases différentes.
- Deux matrices de même type sont équivalentes si et seulement si elles ont le même rang.

☐ **Matrices équivalentes** **Attention** Dans la deuxième propriété, il faut bien voir qu'il faut changer éventuellement à la fois la base de départ et la base d'arrivée. On pourra faire le lien avec la proposition 0.23 quant aux matrices canoniques.

DÉFINITION 0.12 **Matrices semblables**

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

$$A = P.B.P^{-1}$$

PROPOSITION 0.32

- Il s'agit d'une relation d'équivalence
- Deux matrices sont semblables si elles représentent un même endomorphisme dans deux bases différentes
- Deux matrices semblables ont même rang
- Deux matrices semblables ont même trace

☐ **Matrices semblables**

Démonstration Facile pour les trois premiers points ; pour le quatrième il suffit de se rappeler que la trace de AB est égale à la trace de BA , cf le troisième point de la proposition 0.27. C'est grâce à ce quatrième point que l'on peut définir la trace d'un endomorphisme, indépendamment du choix d'une base.

Attention Cette fois-ci, contrairement au cas des matrices équivalentes, deux matrices de même rang ne sont pas nécessairement semblables.

Attention La deuxième caractérisation fait appel à un endomorphisme et pas une application linéaire quelconque ; la base de départ est la même que celle d'arrivée.

1.3.10 Cofacteurs et inversion

On va introduire différentes définitions dont un aboutissement sera la détermination analytique de l'inverse d'une matrice. Quoiqu'informatiquement inefficace (beaucoup trop lente), cette inversion est utile pour les propriétés analytiques qu'elle entraîne.

DÉFINITION 0.13 Mineur et cofacteur

Le déterminant de la matrice M à laquelle on ôte la j -ième colonne et la i -ième ligne est appelé **mineur** (i, j) de M . On le note généralement $\Delta_{i,j}$.

Le déterminant de la matrice M à laquelle on ôte la j -ième colonne pour la remplacer par le i -ième vecteur de la base est appelé **cofacteur** (i, j) de M . On le note généralement $\gamma_{i,j}(M)$.

La matrice γ ainsi définie est appelée **comatrice** de M . On la note généralement $com(M)$.

La matrice ${}^t\gamma$ est appelée **matrice complémentaire** de M . On la note généralement \tilde{M} .

Pour y voir plus clair, le mineur (i, j) de M est :

$$\begin{vmatrix} M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & M_{1,j+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & M_{2,j+1} & \dots & M_{2,n} \\ M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & M_{3,j+1} & \dots & M_{3,n} \\ M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & M_{4,j+1} & \dots & M_{4,n} \\ \vdots & \vdots \\ M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & M_{i-1,j+1} & \dots & M_{i-1,n} \\ M_{i+1,1} & M_{i+1,2} & M_{i+1,3} & \dots & M_{i+1,j-1} & M_{i+1,j+1} & \dots & M_{i+1,n} \\ \vdots & \vdots \\ M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & M_{n,j+1} & \dots & M_{n,n} \end{vmatrix}$$

et le cofacteur (i, j) de M est :

$$\begin{vmatrix} M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & 0 & M_{1,j+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & 0 & M_{2,j+1} & \dots & M_{2,n} \\ M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & 0 & M_{3,j+1} & \dots & M_{3,n} \\ M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & 0 & M_{4,j+1} & \dots & M_{4,n} \\ \vdots & \vdots \\ M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\ M_{i,1} & M_{i,2} & M_{i,3} & \dots & M_{i,j-1} & 1 & M_{i,j+1} & \dots & M_{i,n} \\ M_{i+1,1} & M_{i+1,2} & M_{i+1,3} & \dots & M_{i+1,j-1} & 0 & M_{i+1,j+1} & \dots & M_{i+1,n} \\ \vdots & \vdots \\ M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & 0 & M_{n,j+1} & \dots & M_{n,n} \end{vmatrix}$$

qui est d'ailleurs égal à

$$\begin{vmatrix} M_{1,1} & M_{1,2} & M_{1,3} & \dots & M_{1,j-1} & 0 & M_{1,j+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & M_{2,3} & \dots & M_{2,j-1} & 0 & M_{2,j+1} & \dots & M_{2,n} \\ M_{3,1} & M_{3,2} & M_{3,3} & \dots & M_{3,j-1} & 0 & M_{3,j+1} & \dots & M_{3,n} \\ M_{4,1} & M_{4,2} & M_{4,3} & \dots & M_{4,j-1} & 0 & M_{4,j+1} & \dots & M_{4,n} \\ \vdots & \vdots \\ M_{i-1,1} & M_{i-1,2} & M_{i-1,3} & \dots & M_{i-1,j-1} & 0 & M_{i-1,j+1} & \dots & M_{i-1,n} \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ M_{i+1,1} & M_{i+1,2} & M_{i+1,3} & \dots & M_{i+1,j-1} & 0 & M_{i+1,j+1} & \dots & M_{i+1,n} \\ \vdots & \vdots \\ M_{n,1} & M_{n,2} & M_{n,3} & \dots & M_{n,j-1} & 0 & M_{n,j+1} & \dots & M_{n,n} \end{vmatrix}$$

PROPOSITION 0.33

$$\gamma_{i,j} = (-1)^{i+j} \cdot \Delta_{i,j}$$

Démonstration Nous donnons ici un rapide plan de preuve. Considérer les permutations σ telles que $\sigma(i) = j$. Considérer alors leurs restrictions à $I = [[1, n]] \setminus \{i\}$ et $J = [[1, n]] \setminus \{j\}$ (on restreint à la fois le domaine et le codomaine). Considérer alors pour chaque σ ainsi restreint $\sigma' \circ \sigma \circ \sigma''$ avec σ' la seule bijection croissante de J vers $[[1, n-1]]$ et σ'' la seule bijection croissante de $[[1, n-1]]$ vers I . On a alors $\epsilon(\sigma) = \epsilon(\sigma' \circ \sigma \circ \sigma'') \times (-1)^{i+j}$.

PROPOSITION 0.34

La comatrice de la transposée de M est la transposée de la comatrice de M .

Il est nécessaire pour la suite d'avoir lu la partie ??.

THÉORÈME 0.35

$$\tilde{M} \cdot M = \det M \cdot Id$$

et $M \cdot \tilde{M} = \det M \cdot Id$ en particulier, si M est inversible, $M^{-1} = \frac{1}{\det M} \tilde{M}$

Démonstration Considérons le terme (i, j) de la matrice $\tilde{M} \cdot M$. Il s'agit de $\sum_{k=1}^n \gamma_{k,i} \cdot M_{k,j}$.
Considérons le terme (i, j) de la matrice $\det M \cdot Id$.

Il s'agit de $\delta_{i,j} \cdot \det M$.

On cherche donc à montrer que $\delta_{i,j} \cdot \det M = \sum_{k=1}^n \gamma_{k,i} \cdot M_{k,j}$.

On distingue deux cas :

- $i \neq j$

On considère alors la matrice M , sur laquelle on remplace la colonne i par la colonne j (on ne les permute pas, on supprime la colonne i , et on copie la colonne j à la place).

Le déterminant de la matrice obtenue est nul, puisqu'on a deux fois la même colonne. On développe par rapport à la colonne i . On obtient

$$\sum_{k=1}^n M_{k,j} \cdot \gamma_{k,i} = 0$$

et donc le résultat dans ce cas est bien montré.

• $i = j$

Dans ce cas on considère le développement de M par rapport à la i -ième colonne et on obtient

$$\det M = \sum_{k=1}^n \gamma_{k,i} \cdot M_{k,i} = \sum_{k=1}^n \gamma_{k,i} \cdot M_{k,j}$$

d'où le résultat souhaité.

Le second résultat se déduit du premier en considérant la transposée de chacun des deux produits.

1.3.11 Matrices par blocs

Les matrices par blocs sont fort pratiques pour les produits, les déterminants, les inversions. En particulier, les ordinateurs fonctionnent en général avec des caches, c'est-à-dire des mémoires très rapides intermédiaires entre le processeur et la mémoire vive : ces "caches" ont une taille très limitée, et dimensionner les blocs d'un calcul par blocs de manière à ce que le cache ne soit pas débordé (ce qui entraîne des communications trop fréquentes avec la beaucoup plus lente mémoire vive) change tout à la performance d'un algorithme. On peut en fait déterminer la taille du cache en examinant la vitesse d'un calcul matriciel en fonction de la taille des blocs. Notons que des découpages par bandes existent aussi.

Les méthodes expliquées ci-dessous dans le cas de quatre blocs, se généralisent naturellement au cas d'un nombre quelconque de blocs.

Produit par blocs Etant données les matrices A, B, C, D, A', B', C' et D' , avec

$$\text{largeur}(A) = \text{largeur}(C) = \text{hauteur}(A') = \text{hauteur}(B')$$

$$\text{largeur}(B) = \text{largeur}(D) = \text{hauteur}(C') = \text{hauteur}(D')$$

on considère les matrices M et M' définies comme suit :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad M' = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

Alors

$$M.M' = \begin{pmatrix} A.A' + B.C' & A.B' + B.D' \\ C.A' + D.C' & C.B' + D.D' \end{pmatrix}$$

Inverse par blocs Si $M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$, avec A et B inversibles, alors M est inversible et

$$M^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}.C.B^{-1} \\ 0 & B^{-1} \end{pmatrix}.$$

Déterminant par blocs Si $M = \begin{pmatrix} A & C & B \end{pmatrix}$, alors $\det M = \det A \cdot \det B$.

1.3.12 Opérations sur les lignes et les colonnes

On trouvera des manipulations proches de celles décrites ici dans la partie sur les déterminants ???. La proposition ??? illustre aussi des opérations sur les lignes et les colonnes.

DÉFINITION 0.14 système d'équations linéaires

On appelle **système d'équations linéaires** une équation de la forme $MX = Y$, où M (matrice) et Y (vecteur) sont donnés et où X est l'inconnue.

Les **opérations sur les lignes et les colonnes** d'une matrice ou d'un système linéaire sont par définition :

- l'addition d'une ligne (resp. colonne) i à une ligne (resp. colonne) $j \neq i$
- la multiplication d'une ligne (resp. colonne) i par un scalaire $\lambda \neq 0$
- la permutation de deux lignes (resp. colonnes) i et $j \neq i$

Ces opérations seront notées respectivement :

- $L_i \leftarrow L_i + L_j$ (resp. $C_i \leftarrow C_i + C_j$)
- $L_i \leftarrow \lambda \cdot L_i$ (resp. $C_i \leftarrow \lambda \cdot C_i$)
- $L_i \leftrightarrow L_j$ (resp. $C_i \leftrightarrow C_j$)

On pourra éventuellement ajouter à une ligne (resp. une colonne) une autre ligne (resp. colonne) multipliée par un scalaire λ ; cela se notera $L_i \leftarrow L_i + \lambda L_j$ (resp. $C_i \leftarrow C_i + \lambda C_j$) : il s'agit d'une composition des deux premiers types d'opérations.

Exemple 0.1 • Examinons un système sur lequel on "bricole" sur les lignes ou colonnes. Avant :

$$\begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 1 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Après $L_1 \leftarrow L_1 + 2 L_2$:

$$\begin{pmatrix} 1 & 4 & -1 \\ 0 & 1 & 1 \\ -1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 3 \end{pmatrix}$$

• Sur une matrice maintenant. Avant :

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Après $C_1 \leftrightarrow C_2$

$$\begin{pmatrix} 2 & 1 & 3 \\ 5 & 4 & 6 \\ 8 & 7 & 9 \end{pmatrix}$$

PROPOSITION 0.36

Les opérations sur les lignes correspondent à des multiplications à gauche par des matrices inversibles ; les opérations sur les colonnes correspondent à des multiplications à droite par des matrices inversibles. L'inverse d'une opération sur les lignes (resp. les colonnes) est une opération sur les lignes (resp. les colonnes).

Démonstration $L_i \leftarrow L_i + L_j$ correspond à la multiplication à gauche par $I + E_{i,j}$;
 $L_i \leftarrow \lambda L_i$ correspond à la multiplication à gauche par $\text{diag}(1, \dots, 1, \lambda, 1, \dots, 1)$, i.e. par la matrice avec des 1 sur la diagonale, sauf λ en place (i, i) ;
 $L_i \leftrightarrow L_j$ correspond à la multiplication à gauche par $I + E_{i,j} + E_{j,i} - E_{i,i} - E_{j,j}$.
 Pour les opérations sur les colonnes, on se ramène au même type d'opérations sur les lignes par transposition.

PROPOSITION 0.37

- Le déterminant d'une matrice n'est pas modifié en ajoutant à une ligne une combinaison linéaire des autres lignes.
 - Multiplier une ligne par λ multiplie le déterminant par λ .
 - Permuter des lignes multiplie le déterminant par -1 .
- Les mêmes résultats sont valables pour les colonnes.

Démonstration Cela découle immédiatement des propriétés du déterminant, et du fait que le déterminant d'une matrice est égal au déterminant de sa transposée. Pour des rappels, on peut consulter la partie ??.

THÉORÈME 0.38 **Formules de Cramer**

Considérons le système d'équations linéaire $MX = Y$, avec M de type (n, n) :

$$M = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,n} \end{pmatrix}, \quad Y = {}^t(y_1, \dots, y_n) = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

On suppose en outre que M est inversible.
 Alors X est solution, avec

$$x_i = \frac{\begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & Y_1 & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & Y_2 & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & Y_n & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}}{\det M}$$

Démonstration La solution X est clairement unique, car par inversibilité de M $X = M^{-1}Y$.

On a alors $Y = \sum_k X_k C_k$ avec C_k la k -ième colonne de M .

Donc, quel que soit i , on a alors $\det M_k^{(i)} = \sum X_k \det M'_k{}^{(i)}$ avec

$$M_k^{(i)} = \begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & Y_1 & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & Y_2 & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & Y_n & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}$$

et

$$M'_k{}^{(i)} = \begin{vmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,i-1} & M_{1,k} & M_{1,i+1} & \dots & M_{1,n} \\ M_{2,1} & M_{2,2} & \dots & M_{2,i-1} & M_{2,k} & M_{2,i+1} & \dots & M_{2,n} \\ \vdots & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,i-1} & M_{n,k} & M_{n,i+1} & \dots & M_{n,n} \end{vmatrix}$$

On en déduit donc, en supprimant de la somme $\sum X_k C_k$ les termes nuls, $X_i \det M = M_i^{(i)}$, ce qui est précisément le résultat désiré.

THÉORÈME 0.39 Méthode du pivot de Gauss

La **méthode de Gauss** consiste à :

- 1) permuter les lignes pour avoir un coefficient non nul en haut à gauche de la matrice ; ce coefficient est appelé **pivot**
- 2) soustraire la première ligne multipliée par un coefficient adéquat à chacune des autres lignes de manière à avoir des zéros sur toute la première colonne en dehors du premier coefficient
- 3) Procéder de même sur la matrice extraite, simplement dépourvue de sa première ligne et sa première colonne.

Le point 1) pourra toujours être réalisé si on trouve toujours un coefficient non nul à échanger ; pour peu que la matrice soit inversible, cette condition sera toujours vérifiée. Si elle ne l'est pas, on peut travailler sur la matrice extraite par suppression de la première colonne.

En répétant cette méthode, on arrive à obtenir une matrice triangulaire supérieure. En fait la matrice obtenue est de la forme illustrée sur la figure 1, du moins après permutation des colonnes.

La matrice ainsi obtenue est donc beaucoup plus maniable : le calcul du déterminant (si la matrice est carrée) se résume à la multiplication des éléments diagonaux, la résolution de systèmes linéaires est plus aisée (il convient de noter pour cela que les opérations sur les colonnes sont de simples permutations sur les inconnues), le calcul du rang est immédiat (nombre d'éléments avant le dernier élément non nul de la dernière colonne).

Remarque Afin de minimiser les pertes de précision d'un calcul numérique, il est préférable de choisir un pivot grand en valeur absolue.

La méthode du **pivot total** consiste à chercher le pivot non pas seulement sur la colonne en cours, mais de permuter éventuellement les colonnes pour avoir un pivot plus grand. Par opposition au pivot total, la méthode ci-dessus est dite **pivot partiel**.

THÉORÈME 0.40 Décomposition $A = LU$

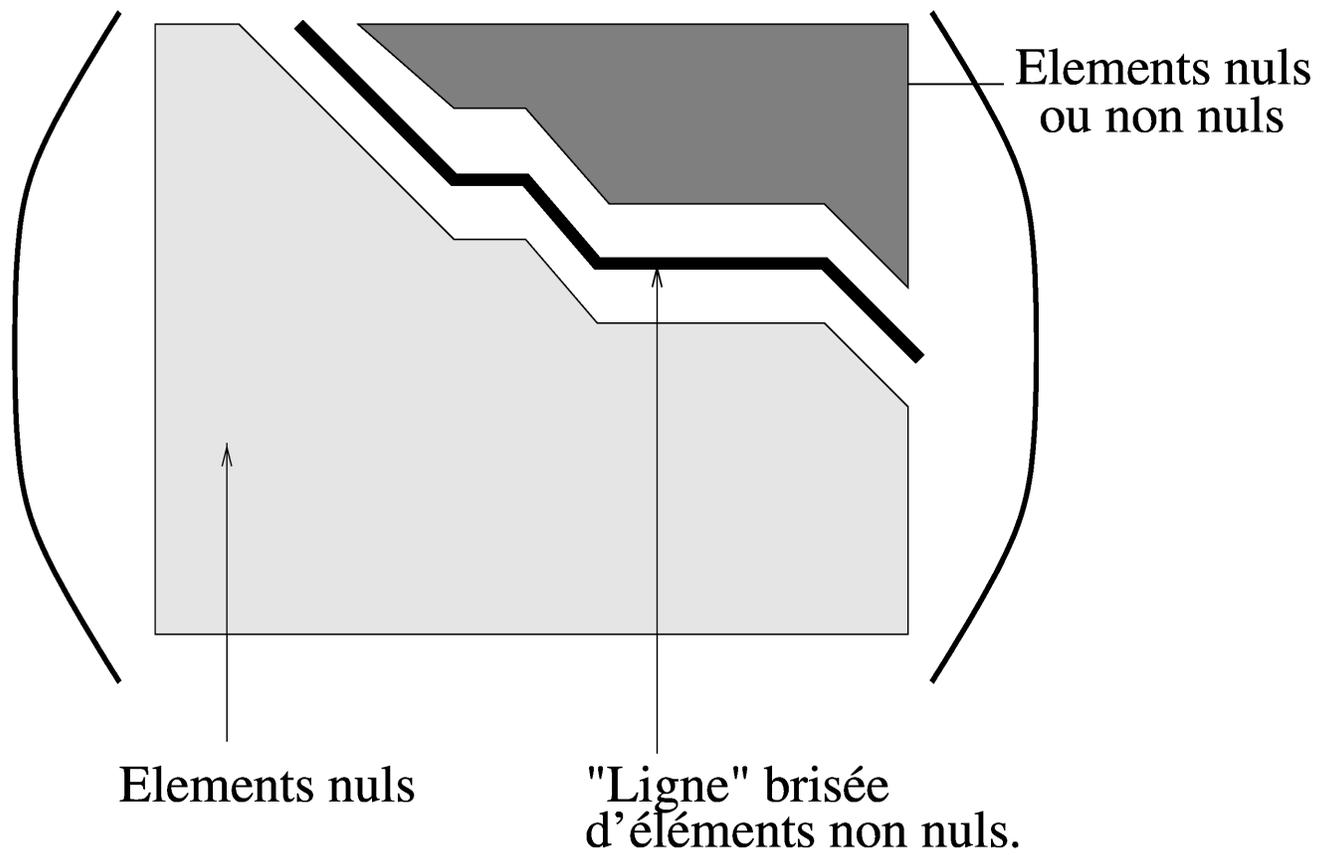


FIGURE 1 – Matrice obtenue après pivot de Gauss. La "ligne brisée" évoquée comporte soit des sauts à droite, soit des sauts en diagonale en bas à droite. Le premier élément de la ligne brisée se trouve quelque part sur la première ligne (éventuellement en haut à gauche).

Etant donnée une matrice A supposée inversible, on définit $a_k = |(A_{i,j})_{i,j \leq k}|$, déterminant de la matrice obtenue en se restreignant aux k premières lignes et k premières colonnes.

On appelle **décomposition** $A = LU$ un produit du type $A = LU$ avec L matrice triangulaire inférieure ne comportant que des 1 sur la diagonale, U matrice triangulaire supérieure.

Alors, il existe une décomposition $A = LU$ si et seulement si les a_k sont non nuls pour tout k dans $\llbracket 1, n \rrbracket$.

Démonstration Il suffit d'utiliser la méthode de Gauss, en considérant les matrices correspondant aux opérations sur les lignes et les colonnes. C'est-à-dire que l'on obtient un produit $\prod_{i=1}^n M_{n-i} A = U$, avec M_i la matrice correspondant à l'opération sur les lignes et les colonnes effectué à la i -ième étape. Mais l'inverse d'une opération sur les lignes ou les colonnes est une opération sur les lignes ou les colonnes; donc on peut aussi écrire $A = \prod_{i=1}^n N_i U$, avec N_i l'inverse de M_i . Le produit des N_i est bien de la forme désirée, triangulaire supérieure à diagonale unité, comme on s'en convaincra en considérant la stabilité de l'ensemble des matrices triangulaires supérieures à diagonale unité par multiplication par les matrices des opérations sur les lignes et les colonnes utilisées dans le pivot de Gauss.

La condition sur les déterminants ne couvre pas toutes les matrices A inversibles possibles. Cela va être résolu avec la proposition qui suit :

PROPOSITION 0.41

Si A est inversible, il existe une matrice de permutation P telle que $PA = LU$.

Démonstration Si a_k est nul, il existe nécessairement une permutation de lignes qui arrange ça, sinon A ne pourrait pas être de rang plein - il suffit donc de multiplier les différentes permutations de lignes nécessaires pour obtenir une matrice vérifiant les conditions demandées.

THÉORÈME 0.42 Décomposition de Cholesky

On appelle **décomposition de Cholesky** ou **décomposition** $A = {}^t R R$, un produit de la forme $A = {}^t R R$, avec R triangulaire inférieure inversible.

A admet une décomposition de Cholesky si et seulement si A est symétrique définie positive.

Démonstration On pourra consulter le livre [3] pour cette preuve.

1.3.13 Exercices sur les matrices

Exemple 0.2 Une forme linéaire f sur $\mathcal{M}_n(\mathbb{K})$ telle que $f(XY) = f(YX)$ est proportionnelle à l'application $M \mapsto \text{tr}(M)$.

Démonstration Il suffit de considérer X et Y des matrices élémentaires, et de généraliser à toute matrice par linéarité de f .

Exemple 0.3 Une matrice de $\mathcal{M}_n(\mathbb{K})$ commute avec toutes les matrices de

$\mathcal{M}_n(\mathbb{K})$ si et seulement si elle est scalaire, i.e. de la forme λId avec Id la matrice identité et $\lambda \in \mathbb{K}$.

Démonstration Ecrire la condition de commutativité avec les matrices élémentaires suffit.

1.4 Zoologie sur les matrices et leurs déterminants

On réutilise la notion de matrice circulante vue en section ?? :

PROPOSITION 0.43

L'ensemble des matrices circulantes de type (n, n) est une K -algèbre engendrée par la matrice M_0 suivante :

$$M_0 = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ 1 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Démonstration Il convient de vérifier tout d'abord qu'il s'agit bien d'une K -algèbre. Pour ensuite voir qu'elle est engendrée par M_0 , il suffit de voir que la matrice circulante associée à (x_1, \dots, x_n) est la matrice

$$x_1.M_0^0 + x_2.M_0^1 + \dots + x_n.M_0^{n-1}$$

et le résultat est acquis.

1.5 Zoologie de la dualité en dimension finie

On va présenter ici les polynômes de Lagrange (avec leurs propriétés d'interpolation), et la réécriture de sous-espaces vectoriels en tant qu'intersection d'hyperplans.

1.5.1 Polynômes de Lagrange

On considère l'espace $E = \mathbb{R}_n[X]$ des polynômes à une indéterminée et de degré au plus n . Soient a_0, \dots, a_n des réels deux à deux distincts. On définit $n + 1$ formes linéaires sur E par $f_i(P) = P(a_i)$.

PROPOSITION 0.44

Les (f_i) forment une base de E^* .

Démonstration Puisque la dimension de E est égale à la dimension de E^* , il suffit de voir que la famille est de rang $n + 1$, ce qui est vérifié si et seulement si l'espace vectoriel dual est de dimension 0. Supposons qu'un certain polynôme P appartienne à cet orthogonal, alors il s'annule en a_0, \dots, a_n ; donc il est nul, puisqu'il est de degré au plus n .

PROPOSITION 0.45 Polynômes de Lagrange

(f_i) est la base duale des P_i , avec

$$P_i(x) = \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}$$

Les P_i sont appelés polynômes de Lagrange. L'intérêt de ces polynômes va être la facilité avec laquelle on va interpoler avec ces polynômes ; il suffit de connaître leurs valeurs en les a_i . Malheureusement, cette technique est une interpolation, c'est-à-dire que le polynôme obtenu par interpolation (corollaire ci-dessous) incorpore tous les bruits et converge en pratique très mal vers une vraie fonction connue seulement par ses valeurs en un nombre fini de points, dès que cette fonction n'est pas "vraiment" un polynôme. Ce constat est à la base de la théorie de la régularisation (e.g., régularisation de Tychonov).

Démonstration Facile, il suffit de vérifier que $P_i(a_j) = \delta_{i,j}$.

COROLLAIRE 0.46

On en déduit notamment que tout polynôme de degré n s'écrit comme combinaison linéaire des P_i , les coefficients étant donnés par les f_i , c'est-à-dire que tout P de degré $\leq n$ s'écrit

$$P = \sum_{i=1}^n f_i(P)P_i = \sum_{i=1}^n P_i^*(P)P_i$$

Un exemple d'utilisation Maple :

Exemple Maple

`> interp([0, 1, 2, 3], [exp(0), exp(1), exp(2), exp(3)], x);`

$$\frac{1}{6}e^3x^3 - \frac{1}{2}e^3x^2 + \frac{1}{3}e^3x - \frac{1}{2}e^2x^3 + 2e^2x^2 - \frac{3}{2}e^2x + \frac{1}{2}ex^3 - \frac{5}{2}ex^2 + 3ex - \frac{1}{6}x^3 + x^2 - \frac{11}{6}x + 1$$

Il est intéressant de tracer ensuite la courbe exponentielle et les graphes des interpolations à différents ordres superposées.

1.5.2 Définition d'un sous-espace vectoriel par une famille d'équations

PROPOSITION 0.47

Soit F un sous-espace vectoriel de E de dimension p , E étant un espace vectoriel de dimension finie n . Alors il existe $n - p$ formes linéaires linéairement indépendantes f_i telles que $F = \{x ; \forall i f_i(x) = 0\}$ c'est-à-dire que F s'exprime comme intersection de $n - p$ hyperplans. Il est en outre impossible de définir F comme intersection de moins de $n - p$ hyperplans.

Démonstration Pour voir qu'une telle famille existe, il suffit de considérer une base de F et sa base duale. Pour vérifier qu'on ne peut faire moins, il suffit de considérer que pour tout G sous-espace vectoriel de E et tout H hyperplan de E on a $\dim(G \cap H) \geq \dim G - 1$ (par la formule $\dim G + \dim H = \dim(G + H) + \dim(G \cap H)$).

Cette proposition permet de voir certaines choses clairement : ainsi, tout sous-espace vectoriel de \mathbb{R}^n peut être écrit sous formes d'un système d'équations linéaires ; toute droite de \mathbb{R}^3 est une intersection de deux hyperplans.

Références

- [1] B. Chazelle, *The Discrepancy Method*, Cambridge University Press, 2000.
- [2] H. Brézis, *Analyse fonctionnelle*, Masson, 1983.
- [3] P.G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Dunod, 1998.