Chapitre 16

Matrices

16.1 Introduction historique

> *Il nous faut commencer, non par un carré, mais par un arrangement oblong de termes [...]. Cela ne représentera pas en soi un déterminant, mais c'est pour ainsi dire une matrice, dont nous pouvons tirer divers systèmes de déterminants.*
>
> James Joseph Sylvester - *Additions to the articles « On a new class of theorems » and « On Pascal's theorem »*, Philosophical Magazine, 1850

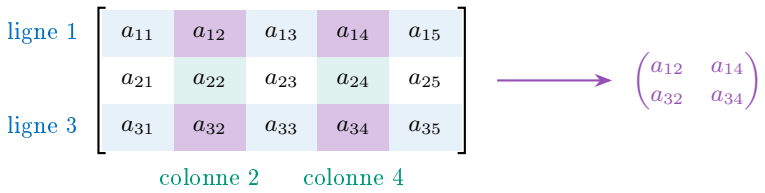
16.1.1 Un mot venu du latin

Les tableaux de nombres sont aussi anciens que les systèmes linéaires : le chapitre 3 a rappelé que les *Neuf chapitres sur l'art du calcul* disposaient déjà les coefficients en tableau pour les combiner. Pendant deux millénaires, pourtant, le tableau n'est qu'un support de calcul, et l'objet étudié est le système, puis, à partir du XVIII^{e} siècle, le nombre qu'on lui associe et que **Cauchy** appellera *déterminant* (chapitre 17). En 1850, **James Joseph Sylvester** s'intéresse aux déterminants que l'on peut extraire d'un tableau rectangulaire en choisissant autant de lignes que de colonnes ; il appelle ce tableau une *matrice*, mot latin qui désigne l'utérus, le lieu où quelque chose prend naissance. La matrice est alors le « moule » d'où naissent les déterminants, et le mot dit encore qu'elle n'est pas un nombre.

16.1.2 Cayley et la naissance d'une algèbre

Le pas décisif est franchi en 1858 par **Arthur Cayley**, ami et collègue de Sylvester, dans son *Memoir on the theory of matrices*. Cayley considère les matrices comme des objets qu'on additionne et qu'on multiplie, et il définit le produit de façon que le produit de deux matrices représente la composée des deux substitutions linéaires qu'elles décrivent : c'est exactement le point de vue de la section 16.4, où le produit matriciel apparaîtra comme la traduction de la composition des applications linéaires. Cayley observe que ce produit n'est pas commutatif, ce qui fait des matrices, après les quaternions de Hamilton (1843), l'un des premiers anneaux non commutatifs de l'histoire. Le chapitre se termine en revenant à l'idée de Sylvester : le rang d'une matrice se lit sur les tableaux carrés qu'on en extrait.

Dans le tableau oblong ci-dessous, de trois lignes et cinq colonnes, le choix de deux lignes, en bleu, et de deux colonnes, en vert, fait « naître » un tableau carré, en violet : c'est l'image que Sylvester avait en tête.



16.2 Calcul matriciel

Dans tout le chapitre, $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$, et n, p, q, r des entiers naturels non nuls.

16.2.1 L'espace des matrices

Définition : Une **matrice** à n lignes et p colonnes à coefficients dans $\mathbb{K}$ est une famille $A = (a_{ij})_{1\leqslant i\leqslant n,\, 1\leqslant j\leqslant p}$ d'éléments de $\mathbb{K}$, indexée par $[\![1\,,n]\!]\times[\![1\,,p]\!]$. Le scalaire a_{ij} est le **coefficient** d'indice (i,j) : le premier indice repère la **ligne**, le second la **colonne**. On la représente par le tableau

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}.$$

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble de ces matrices, et $\mathcal{M}_n(\mathbb{K}) = \mathcal{M}_{n,n}(\mathbb{K})$ celui des **matrices carrées** d'ordre n. Une matrice de $\mathcal{M}_{n,1}(\mathbb{K})$ est une **matrice colonne**, une matrice de $\mathcal{M}_{1,p}(\mathbb{K})$ une **matrice ligne**.

Remarques :

- Deux matrices sont égales si elles ont la même taille et les mêmes coefficients : une matrice est une famille, au sens du chapitre 14, et l'égalité de deux familles est celle de leurs termes. Les matrices $\begin{pmatrix}1 & 2\end{pmatrix}$ et $\begin{pmatrix}1\\2\end{pmatrix}$ sont donc distinctes.
- On identifiera $\mathcal{M}_{n,1}(\mathbb{K})$ et $\mathbb{K}^n$, en écrivant en colonne les n-uplets ; ce sera la convention de tout le chapitre à partir de la section 16.4.

L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ n'est autre que $\mathbb{K}^{[\![1\,,n]\!]\times[\![1\,,p]\!]}$, espace des familles indexées par un ensemble fini de np éléments : on le munit des opérations du chapitre 14, terme à terme. Pour $A = (a_{ij})$, $B = (b_{ij})$ et $\lambda \in \mathbb{K}$, on pose

$$A + B = (a_{ij} + b_{ij}) \qquad \text{et} \qquad \lambda A = (\lambda a_{ij}).$$

Définition : Pour $(k,\ell) \in [\![1\,,n]\!]\times[\![1\,,p]\!]$, on note $E_{k\ell}$ la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont tous les coefficients sont nuls, sauf celui d'indice (k,ℓ), qui vaut 1. Ce sont les **matrices élémentaires**. À l'aide du **symbole de Kronecker**

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases} \qquad \text{on a} \qquad E_{k\ell} = (\delta_{ik}\,\delta_{j\ell})_{i,j}.$$

Proposition 1 : L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$, muni de ces opérations, est un $\mathbb{K}$-espace vectoriel. La famille $(E_{k\ell})_{1\leqslant k\leqslant n,\, 1\leqslant \ell\leqslant p}$ en est une base, dite **base canonique**, et

$$\dim \mathcal{M}_{n,p}(\mathbb{K}) = np, \qquad A = \sum_{k=1}^{n}\sum_{\ell=1}^{p} a_{k\ell}\,E_{k\ell}.$$

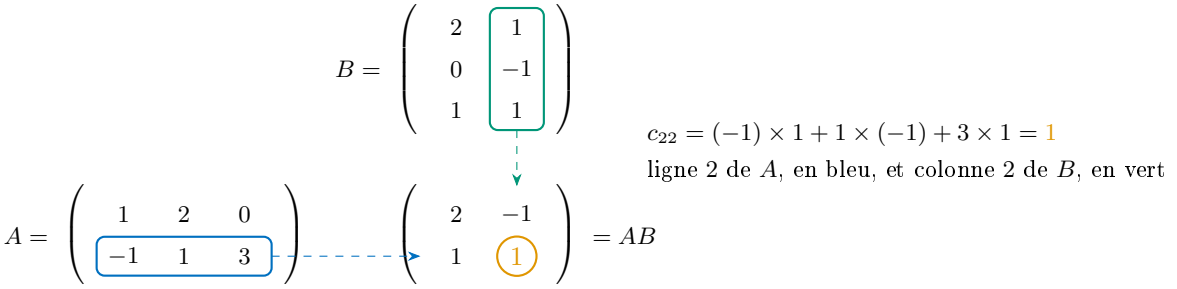
∘ Démonstration : C'est l'espace $\mathbb{K}^I$ des familles indexées par $I = [\![1\,,n]\!]\times[\![1\,,p]\!]$, dont on a vu au chapitre 14 qu'il est un espace vectoriel. Le coefficient d'indice (i,j) de $\sum a_{k\ell}E_{k\ell}$ est $\sum_{k,\ell} a_{k\ell}\,\delta_{ik}\,\delta_{j\ell} = a_{ij}$: toute matrice s'écrit de façon unique comme combinaison des $E_{k\ell}$, qui forment donc une base, de cardinal np. □

16.2.2 Produit matriciel

Définition : Soit $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{jk}) \in \mathcal{M}_{p,q}(\mathbb{K})$. Le **produit** AB est la matrice $C = (c_{ik}) \in \mathcal{M}_{n,q}(\mathbb{K})$ définie par

$$c_{ik} = \sum_{j=1}^{p} a_{ij}\,b_{jk} \qquad \text{pour } 1 \leqslant i \leqslant n \text{ et } 1 \leqslant k \leqslant q.$$

Le coefficient c_{ik} s'obtient en parcourant la ligne i de A et la colonne k de B, et en sommant les produits des coefficients qui se correspondent. Le produit n'est défini que si le nombre de colonnes de A est égal au nombre de lignes de B ; le résultat a le nombre de lignes de A et le nombre de colonnes de B.



On place B au-dessus et à droite de A : chaque coefficient du produit se trouve à la croisée de la ligne de A et de la colonne de B qui servent à le calculer.

Remarques :

- Le produit BA est ici aussi défini, mais il appartient à $\mathcal{M}_3(\mathbb{K})$:

$$BA = \begin{pmatrix} 1 & 5 & 3 \\ 1 & -1 & -3 \\ 0 & 3 & 3 \end{pmatrix} \neq AB.$$

  L'erreur la plus répandue est de croire que AB et BA sont comparables ; ils n'ont même pas, en général, la même taille.

- La définition peut sembler arbitraire. Elle est en réalité imposée par la composition des applications linéaires, comme le montrera le théorème 2 : c'est la définition de Cayley.

Proposition 2 : Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$, de colonnes $C_1, \dots, C_p \in \mathcal{M}_{n,1}(\mathbb{K})$, et $X \in \mathcal{M}_{p,1}(\mathbb{K})$ de coefficients $x_1, \dots, x_p$. Alors

$$AX = x_1 C_1 + \cdots + x_p C_p.$$

Plus généralement, la k-ième colonne de AB est A multiplié par la k-ième colonne de B, et la i-ième ligne de AB est la i-ième ligne de A multipliée par B.

∘ Démonstration : Le i-ième coefficient de AX est $\sum_j a_{ij} x_j$, qui est aussi celui de $\sum_j x_j C_j$. Les deux autres affirmations se lisent sur la formule définissant c_{ik}, à k fixé, puis à i fixé. □

Remarque : Le produit AX est donc une *combinaison linéaire des colonnes* de A, de coefficients les x_j. Cette lecture, plus que la formule, est celle qu'il faut retenir : elle fera le lien avec les familles de vecteurs du chapitre 15, et c'est elle qui expliquera que l'image d'une matrice est engendrée par ses colonnes.

Proposition 3 : Lorsque les produits sont définis, on a :

(i) *bilinéarité :* $A(\lambda B + B') = \lambda AB + AB'$ et $(\lambda A + A')B = \lambda AB + A'B$;

(ii) *associativité :* $(AB)C = A(BC)$ pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$ et $C \in \mathcal{M}_{q,r}(\mathbb{K})$;

(iii) *élément neutre :* en notant $I_n = (\delta_{ij}) \in \mathcal{M}_n(\mathbb{K})$ la **matrice identité**, $I_n A = A = A I_p$ pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$.

∘ Démonstration :

▷ *(i)* Le coefficient $\sum_j a_{ij}(\lambda b_{jk} + b'_{jk})$ vaut $\lambda \sum_j a_{ij} b_{jk} + \sum_j a_{ij} b'_{jk}$, par linéarité de la somme (chapitre 3). Le second point est identique.

▷ *(ii)* Le coefficient d'indice (i, ℓ) de $(AB)C$ et celui de $A(BC)$ sont

$$\sum_{k=1}^{q} \Big(\sum_{j=1}^{p} a_{ij} b_{jk} \Big) c_{k\ell} \qquad \text{et} \qquad \sum_{j=1}^{p} a_{ij} \Big(\sum_{k=1}^{q} b_{jk} c_{k\ell} \Big),$$

qui sont tous deux égaux à la somme double $\sum_{(j,k)} a_{ij} b_{jk} c_{k\ell}$ sur $[\![1, p]\!] \times [\![1, q]\!]$, par interversion des sommes (proposition 6 du chapitre 3).

▷ *(iii)* Le coefficient d'indice (i,j) de $I_n A$ est $\sum_k \delta_{ik} a_{kj} = a_{ij}$, puisque seul le terme $k = i$ est non nul. De même à droite. □

Proposition 4 : Les matrices élémentaires $E_{ij} \in \mathcal{M}_{n,p}(\mathbb{K})$ et $E_{k\ell} \in \mathcal{M}_{p,q}(\mathbb{K})$ vérifient

$$E_{ij}\, E_{k\ell} = \delta_{jk}\, E_{i\ell}.$$

∘ Démonstration : Le coefficient d'indice (s,t) du produit est $\sum_m \delta_{si}\delta_{mj}\,\delta_{mk}\delta_{t\ell}$. Seul le terme $m = j$ peut être non nul, et il vaut $\delta_{si}\,\delta_{jk}\,\delta_{t\ell}$, qui est le coefficient d'indice (s,t) de $\delta_{jk}E_{i\ell}$. □

Remarque : Le produit $E_{ij}E_{k\ell}$ est nul dès que $j \neq k$: on dispose ainsi de produits nuls de deux matrices non nulles. Par exemple, dans $\mathcal{M}_2(\mathbb{K})$, $E_{12}E_{12} = 0$. La section suivante exploite ce phénomène, qui n'a pas d'analogue dans $\mathbb{K}$.

16.2.3 Transposition

Définition : La **transposée** de $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ est la matrice $A^T = (a_{ji})_{1\leqslant i\leqslant p,\, 1\leqslant j\leqslant n} \in \mathcal{M}_{p,n}(\mathbb{K})$: les lignes de A^T sont les colonnes de A.

Proposition 5 : L'application $A \mapsto A^T$ est un isomorphisme de $\mathcal{M}_{n,p}(\mathbb{K})$ sur $\mathcal{M}_{p,n}(\mathbb{K})$, et $\left(A^T\right)^T = A$. Pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$,

$$(AB)^T = B^T\, A^T.$$

∘ Démonstration : La linéarité et l'involutivité se lisent sur les coefficients ; une application linéaire involutive est bijective, sa réciproque étant elle-même. Le coefficient d'indice (k,i) de $(AB)^T$ est celui d'indice (i,k) de AB, soit $\sum_j a_{ij}b_{jk}$; c'est aussi $\sum_j b_{jk}\, a_{ij}$, coefficient d'indice (k,i) de $B^T\, A^T$, la ligne k de B^T étant la colonne k de B. □

Remarque : L'ordre des facteurs s'inverse : $(AB)^T = A^T\, B^T$ n'a d'ailleurs en général pas de sens, A^T ayant n colonnes et B^T ayant p lignes. On retrouvera la même inversion pour l'inverse d'un produit, et pour la même raison qu'au chapitre 11, où $(gh)^{-1} = h^{-1}g^{-1}$ dans un groupe.

16.3 Matrices carrées

16.3.1 L'anneau des matrices carrées

Le produit de deux matrices carrées d'ordre n est une matrice carrée d'ordre n : la multiplication est une loi de composition interne sur $\mathcal{M}_n(\mathbb{K})$.

Théorème 1 : Anneau des matrices carrées :
$(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau, d'élément neutre I_n. Pour $n \geqslant 2$, il n'est ni commutatif ni intègre : il existe des matrices A, B telles que $AB \neq BA$, et des matrices non nulles dont le produit est nul.

∘ Démonstration : Le groupe $(\mathcal{M}_n(\mathbb{K}), +)$ est abélien par la proposition 1. La multiplication est associative, distributive par rapport à l'addition et admet I_n pour neutre, par la proposition 3. Pour $n \geqslant 2$, la proposition 4 donne

$$E_{12}E_{21} = E_{11} \neq E_{22} = E_{21}E_{12} \qquad \text{et} \qquad E_{12}E_{12} = 0.$$

□

Exemple : Dans $\mathcal{M}_2(\mathbb{K})$, avec $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, on trouve

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A^2 = 0.$$

La matrice A est non nulle, mais son carré est nul : on dit qu'elle est **nilpotente**, ce qui signifie qu'une de ses puissances est nulle.

Remarques :

- Les règles de calcul d'un anneau quelconque (chapitre 11) s'appliquent, et seulement elles. L'erreur la plus répandue est d'écrire $(A+B)^2 = A^2 + 2AB + B^2$: le développement correct est $A^2 + AB + BA + B^2$, et les deux coïncident si et seulement si $AB = BA$.
- De même, $AB = AC$ n'entraîne pas $B = C$, même si $A \neq 0$: avec la matrice A de l'exemple, $A \times A = A \times 0$. On ne peut simplifier que par une matrice inversible, notion étudiée plus bas.

16.3.2 Matrices remarquables

Définition : Soit $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$. Les coefficients $a_{11}, \dots, a_{nn}$ forment la **diagonale** de A. On dit que A est :

- **diagonale** si $a_{ij} = 0$ pour $i \neq j$; on note alors $A = \mathrm{diag}(a_{11}, \dots, a_{nn})$;
- **scalaire** si elle est de la forme λI_n, avec $\lambda \in \mathbb{K}$;
- **triangulaire supérieure** si $a_{ij} = 0$ pour $i > j$, et **triangulaire inférieure** si $a_{ij} = 0$ pour $i < j$.

Proposition 6 : Les matrices diagonales, les matrices triangulaires supérieures et les matrices triangulaires inférieures forment trois sous-anneaux de $\mathcal{M}_n(\mathbb{K})$, qui sont aussi des sous-espaces vectoriels. Le produit de deux matrices triangulaires supérieures A et B est triangulaire supérieure, de coefficients diagonaux les produits $a_{ii}b_{ii}$; en particulier,

$$\mathrm{diag}(\alpha_1, \dots, \alpha_n)\,\mathrm{diag}(\beta_1, \dots, \beta_n) = \mathrm{diag}(\alpha_1\beta_1, \dots, \alpha_n\beta_n).$$

∘ Démonstration : La stabilité par combinaison linéaire est claire, et I_n est diagonale. Soit A et B triangulaires supérieures. Dans la somme $\sum_j a_{ij}b_{jk}$, un terme ne peut être non nul que si $i \leqslant j$ et $j \leqslant k$. Si $i > k$, il n'y a aucun tel j, et le coefficient est nul ; si $i = k$, seul $j = i$ convient, et le coefficient vaut $a_{ii}b_{ii}$. Le cas inférieur s'obtient par transposition (proposition 5), et le cas diagonal en combinant les deux. □

Remarque : Les matrices diagonales commutent entre elles, et les matrices scalaires commutent avec toutes les matrices. La réciproque de ce dernier point est vraie : une matrice qui commute avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$ est scalaire. C'est la traduction matricielle de l'exercice 39 de la feuille du chapitre 15, sur les endomorphismes qui commutent avec tous les autres.

Définition : Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est **symétrique** si $A^T = A$, et **antisymétrique** si $A^T = -A$. On note $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ leurs ensembles.

Proposition 7 : $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont deux sous-espaces supplémentaires de $\mathcal{M}_n(\mathbb{K})$, de dimensions respectives

$$\dim \mathcal{S}_n(\mathbb{K}) = \frac{n(n+1)}{2} \qquad \text{et} \qquad \dim \mathcal{A}_n(\mathbb{K}) = \frac{n(n-1)}{2}.$$

∘ Démonstration : Ce sont les noyaux des endomorphismes $A \mapsto A^T - A$ et $A \mapsto A^T + A$. Toute matrice s'écrit

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2},$$

somme d'une matrice symétrique et d'une matrice antisymétrique, et une matrice à la fois symétrique et antisymétrique vérifie $A = -A$, donc est nulle. Pour les dimensions, les matrices E_{ii} et $E_{ij} + E_{ji}$, pour $i < j$, forment une base de $\mathcal{S}_n(\mathbb{K})$, et les $E_{ij} - E_{ji}$, pour $i < j$, une base de $\mathcal{A}_n(\mathbb{K})$: il y en a $n + \frac{n(n-1)}{2}$ et $\frac{n(n-1)}{2}$. □

La décomposition est celle d'une fonction en partie paire et partie impaire (chapitre 14) : l'application $s : A \mapsto A^T$ est une symétrie de $\mathcal{M}_n(\mathbb{K})$, au sens du théorème 6 du chapitre 14, et $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$ sont les sous-espaces $\mathrm{Ker}(s - \mathrm{id})$ et $\mathrm{Ker}(s + \mathrm{id})$ qui lui sont associés. On vérifie que les dimensions ont pour somme n^2, comme le veut la proposition 3 du chapitre 15.

16.3.3 Puissances d'une matrice

Pour $A \in \mathcal{M}_n(\mathbb{K})$, on pose $A^0 = I_n$ et $A^{k+1} = A^k A$. Comme dans tout anneau (chapitre 11), la formule du binôme vaut pour deux matrices *qui commutent* : si $AB = BA$, alors pour tout $m \in \mathbb{N}$,

$$(A + B)^m = \sum_{k=0}^{m} \binom{m}{k} A^k B^{m-k}.$$

♡ **Méthode : Pour calculer les puissances d'une matrice A :**

- ▶ On cherche à écrire $A = \lambda I_n + N$, où N est nilpotente, c'est-à-dire telle que $N^s = 0$ pour un certain s.
- ▶ Les matrices λI_n et N commutent, puisqu'une matrice scalaire commute avec toute matrice.
- ▶ La formule du binôme ne comporte alors que s termes non nuls, ceux d'indice $k < s$.

Exemple : Soit $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = 2I_3 + N$, avec $N = E_{12} + E_{23}$. La proposition 4 donne $N^2 = E_{13}$ et $N^3 = 0$. Pour $m \geqslant 2$,

$$A^m = 2^m I_3 + m\,2^{m-1} N + \binom{m}{2} 2^{m-2} N^2 = \begin{pmatrix} 2^m & m\,2^{m-1} & m(m-1)\,2^{m-3} \\ 0 & 2^m & m\,2^{m-1} \\ 0 & 0 & 2^m \end{pmatrix},$$

formule qui reste vraie pour $m = 0$ et $m = 1$. Pour $m = 5$, on trouve les coefficients 32, 80 et 80, ce que confirme un calcul direct. La diagonale croît comme 2^m, mais les coefficients au-dessus croissent plus vite, d'un facteur polynomial en m : c'est l'effet de la partie nilpotente.

16.3.4 Matrices inversibles

Définition : Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est **inversible** s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$. La matrice B, unique, est l'**inverse** de A, notée A^{-1}. On note $\mathrm{GL}_n(\mathbb{K})$ l'ensemble des matrices inversibles d'ordre n.

Proposition 8 : $\mathrm{GL}_n(\mathbb{K})$ est un groupe pour la multiplication, appelé **groupe linéaire**. Pour A, $B \in \mathrm{GL}_n(\mathbb{K})$,

$$(AB)^{-1} = B^{-1}A^{-1} \qquad \text{et} \qquad A^T \in \mathrm{GL}_n(\mathbb{K}), \quad (A^T)^{-1} = (A^{-1})^T.$$

∘ Démonstration : C'est le groupe des inversibles de l'anneau $\mathcal{M}_n(\mathbb{K})$ (chapitre 11), et la formule $(AB)^{-1} = B^{-1}A^{-1}$ est celle de tout groupe. En transposant $AA^{-1} = A^{-1}A = I_n$, la proposition 5 donne $(A^{-1})^T A^T = A^T (A^{-1})^T = {I_n}^T = I_n$. □

Proposition 9 : La matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible si et seulement si $ad - bc \neq 0$, et alors

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

∘ Démonstration : Posons $B = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$; un calcul direct donne $AB = BA = (ad - bc)I_2$. Si $ad - bc \neq 0$, la matrice $\frac{1}{ad-bc}B$ est donc un inverse de A. Si $ad - bc = 0$ et si A était inversible, on aurait $B = A^{-1}(AB) = 0$, donc $a = b = c = d = 0$ et $A = 0$, qui n'est pas inversible : c'est absurde. □

Remarque : Le nombre $ad - bc$ est le *déterminant* de A ; le chapitre 17 le généralisera aux matrices de tout ordre, avec le même rôle : une matrice carrée est inversible si et seulement si son déterminant est non nul.

♡ **Méthode : Pour inverser une matrice $A \in \mathcal{M}_n(\mathbb{K})$ par un système :**

- ▶ On résout le système $AX = Y$, d'inconnue $X \in \mathcal{M}_{n,1}(\mathbb{K})$, pour un second membre Y quelconque.
- ▶ Si, pour tout Y, il admet une unique solution, qui s'écrit $X = BY$, alors A est inversible et $A^{-1} = B$.
- ▶ Si, pour un Y, il n'admet pas de solution, ou en admet plusieurs, A n'est pas inversible.

Justifions la méthode. Si $X = BY$ est l'unique solution de $AX = Y$, alors $A(BY) = Y$ pour tout Y ; avec pour Y les colonnes de I_n, la proposition 2 donne $AB = I_n$. Pour tout X, le vecteur X est l'unique solution du système de second membre $Y = AX$, donc $X = BAX$, et de même $BA = I_n$. Réciproquement, si A est inversible, $AX = Y$ équivaut à $X = A^{-1}Y$.

Exemple : Soit $C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$. Le système $CX = Y$ s'écrit $x_1 + x_2 = y_1$, $x_2 + x_3 = y_2$ et $x_1 + x_3 = y_3$. En ajoutant les trois équations, on obtient $x_1 + x_2 + x_3 = \frac{1}{2}(y_1 + y_2 + y_3)$, puis, en retranchant chacune d'elles,

$$C^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

On vérifie que $CC^{-1} = I_3$. L'inverse d'une matrice à coefficients entiers n'est donc pas toujours à coefficients entiers ; on verra au chapitre 17 que le facteur $\frac{1}{2}$ vient du déterminant de C, qui vaut 2.

Remarque : La définition exige $AB = I_n$ *et* $BA = I_n$. On montrera au théorème 3 que, pour des matrices carrées, l'une des deux égalités suffit ; ce résultat, qui n'a rien d'évident sur les coefficients, viendra de la dimension finie, par le théorème 10 du chapitre 15.

16.4 Matrices et applications linéaires

Dans cette section, E, F, G sont des $\mathbb{K}$-espaces de dimensions finies p, n, m, munis de bases $\mathcal{B} = (e_1, \dots, e_p)$, $\mathcal{C} = (f_1, \dots, f_n)$ et $\mathcal{D}$. On identifie $\mathcal{M}_{n,1}(\mathbb{K})$ à $\mathbb{K}^n$.

16.4.1 Matrice d'une application linéaire

Définition :

- La **matrice d'un vecteur** $x = \sum_j x_j e_j$ dans la base $\mathcal{B}$ est la colonne $\mathrm{Mat}_{\mathcal{B}}(x) \in \mathcal{M}_{p,1}(\mathbb{K})$ de ses coordonnées $x_1, \dots, x_p$. La matrice d'une famille $(x_1, \dots, x_q)$ de vecteurs de E est la matrice de $\mathcal{M}_{p,q}(\mathbb{K})$ dont la k-ième colonne est $\mathrm{Mat}_{\mathcal{B}}(x_k)$.
- La **matrice de** $u \in \mathcal{L}(E,F)$ dans les bases $\mathcal{B}$ et $\mathcal{C}$ est la matrice de la famille $\big(u(e_1), \dots, u(e_p)\big)$ dans la base $\mathcal{C}$:

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K}), \qquad \text{où} \qquad u(e_j) = \sum_{i=1}^{n} a_{ij} f_i.$$

- Si $u \in \mathcal{L}(E)$, on note $\mathrm{Mat}_{\mathcal{B}}(u) = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(u) \in \mathcal{M}_p(\mathbb{K})$ la matrice de l'endomorphisme u dans la base $\mathcal{B}$.

La j-ième colonne de la matrice contient donc les coordonnées de l'image du j-ième vecteur de la base de départ. La matrice a autant de colonnes que la dimension de l'espace de départ et autant de lignes que la dimension de l'espace d'arrivée.

♡ **Méthode : Pour écrire la matrice d'une application linéaire u dans des bases $\mathcal{B}$ et $\mathcal{C}$:**

- ▶ On calcule les images $u(e_1), \dots, u(e_p)$ des vecteurs de la base de départ.
- ▶ On décompose chacune d'elles dans la base d'arrivée $\mathcal{C}$.
- ▶ On écrit les coordonnées obtenues en colonnes, dans l'ordre des e_j.

Exemples :

- La dérivation D de $\mathbb{K}_3[X]$, dans la base $(1, X, X^2, X^3)$, envoie ces vecteurs sur 0, 1, $2X$ et $3X^2$:

$$\mathrm{Mat}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

  C'est une matrice triangulaire supérieure à diagonale nulle ; on vérifie que sa puissance quatrième est nulle, ce qui traduit le fait que la dérivée quatrième d'un polynôme de degré au plus 3 est nulle.

- Voyons $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel de dimension 2, de base $(1, \mathrm{i})$, et soit $u : z \mapsto (a+\mathrm{i}b)z$, avec a, b réels. Alors $u(1) = a + \mathrm{i}b$ et $u(\mathrm{i}) = -b + \mathrm{i}a$, d'où

$$\mathrm{Mat}_{(1,\mathrm{i})}(u) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

  Pour $a + \mathrm{i}b = \mathrm{e}^{\mathrm{i}\theta}$, on retrouve la matrice de la rotation d'angle θ : la multiplication par un complexe, vue au chapitre 4 comme une similitude, est un endomorphisme du plan réel.

- L'application $u : P \mapsto P(X+1)$ de $\mathbb{K}_2[X]$ envoie 1, X et X^2 sur 1, $1+X$ et $1+2X+X^2$:

$$\mathrm{Mat}_{(1,X,X^2)}(u) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Proposition 10 : Soit $u \in \mathcal{L}(E,F)$, $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ et $x \in E$. Les coordonnées $X = \mathrm{Mat}_{\mathcal{B}}(x)$ et $Y = \mathrm{Mat}_{\mathcal{C}}\big(u(x)\big)$ vérifient

$$Y = AX.$$

◦ Démonstration : Si $x = \sum_j x_j e_j$, la linéarité de u donne

$$u(x) = \sum_{j=1}^{p} x_j u(e_j) = \sum_{j=1}^{p} x_j \sum_{i=1}^{n} a_{ij} f_i = \sum_{i=1}^{n} \Big(\sum_{j=1}^{p} a_{ij} x_j \Big) f_i,$$

et la i-ième coordonnée de $u(x)$ est le i-ième coefficient de AX. □

16.4.2 Le dictionnaire

Théorème 2 : Matrices et applications linéaires :

(i) L'application $u \mapsto \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ est un isomorphisme d'espaces vectoriels de $\mathcal{L}(E,F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$.

(ii) Pour $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$, $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(v \circ u) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(v)\ \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$.

(iii) L'application $u \mapsto \mathrm{Mat}_{\mathcal{B}}(u)$ est un isomorphisme d'anneaux de $\mathcal{L}(E)$ sur $\mathcal{M}_p(\mathbb{K})$.

(iv) Si $\dim E = \dim F$, u est un isomorphisme si et seulement si $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ est inversible, et alors

$$\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u^{-1}) = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)^{-1}.$$

◦ Démonstration :

▷ *(i)* La linéarité vient de ce que les coordonnées de $(\lambda u + v)(e_j)$ sont λ fois celles de $u(e_j)$ plus celles de $v(e_j)$. La bijectivité est le théorème 4 du chapitre 14 : une application linéaire est déterminée par les images des e_j, qui peuvent être choisies arbitrairement, c'est-à-dire par des colonnes arbitraires.

▷ *(ii)* La j-ième colonne de $\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(v \circ u)$ est $\mathrm{Mat}_{\mathcal{D}}\big(v(u(e_j))\big)$. Par la proposition 10 appliquée à v, c'est $\mathrm{Mat}_{\mathcal{C},\mathcal{D}}(v)$ multiplié par $\mathrm{Mat}_{\mathcal{C}}\big(u(e_j)\big)$, qui est la j-ième colonne de $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$. La proposition 2 conclut.

▷ *(iii)* C'est (i) et (ii) pour $E = F = G$, avec $\mathrm{Mat}_{\mathcal{B}}(\mathrm{id}_E) = I_p$.

▷ *(iv)* Si u est un isomorphisme, (ii) appliqué à $u^{-1} \circ u = \mathrm{id}_E$ et à $u \circ u^{-1} = \mathrm{id}_F$ montre que $\mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u^{-1})$ est un inverse de $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$. Réciproquement, si $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ est inversible, soit $v \in \mathcal{L}(F,E)$ de matrice A^{-1}, qui existe par (i) ; alors $v \circ u$ et $u \circ v$ ont pour matrices I_p, donc sont les identités. □

$$\begin{array}{ccccc} E & \xrightarrow{u} & F & \xrightarrow{v} & G \\ \downarrow{\scriptstyle \mathrm{Mat}_{\mathcal{B}}}\simeq & & \downarrow{\scriptstyle \mathrm{Mat}_{\mathcal{C}}}\simeq & & \downarrow{\scriptstyle \mathrm{Mat}_{\mathcal{D}}}\simeq \\ \mathbb{K}^p & \xrightarrow[X \mapsto AX]{} & \mathbb{K}^n & \xrightarrow[Y \mapsto BY]{} & \mathbb{K}^m \end{array}$$

Avec $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ et $B = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(v)$, on peut passer de E à $\mathbb{K}^m$ par le haut, puis descendre, ou descendre d'abord, puis multiplier par A, puis par B : le résultat est le même. Composer u puis v revient à multiplier par A puis par B, c'est-à-dire par BA.

Remarques :

- Le théorème justifie la définition du produit matriciel, que Cayley avait choisie pour cette raison : c'est la seule qui fasse correspondre le produit à la composition. On retrouve aussi la proposition 8 du chapitre 15, $\dim \mathcal{L}(E,F) = np$, puisque $\mathcal{L}(E,F)$ est isomorphe à $\mathcal{M}_{n,p}(\mathbb{K})$.
- Le dictionnaire dépend des bases choisies : une même application linéaire a autant de matrices que de couples de bases. La section 16.5 étudie comment la matrice change avec les bases ; l'erreur la plus répandue est d'écrire « la matrice de u » sans préciser dans quelles bases.
- Pour $u : P \mapsto P(X+1)$, la réciproque est $P \mapsto P(X-1)$, dont la matrice dans la base $(1, X, X^2)$ est

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} :$$

c'est l'inverse de la matrice de u, ce que confirme le produit des deux, et que le point (iv) prédisait sans calcul.

16.4.3 Application linéaire canoniquement associée à une matrice

Définition : Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. L'**application linéaire canoniquement associée** à A est

$$u_A : \left| \begin{array}{ccc} \mathbb{K}^p & \longrightarrow & \mathbb{K}^n \\ X & \longmapsto & AX \end{array} \right.$$

C'est l'unique application linéaire dont la matrice dans les bases canoniques est A. On appelle **noyau**, **image** et **rang** de A ceux de u_A :

$$\operatorname{Ker} A = \{X \in \mathbb{K}^p \mid AX = 0\}, \qquad \operatorname{Im} A = \{AX \mid X \in \mathbb{K}^p\}, \qquad \operatorname{rg}(A) = \dim \operatorname{Im} A.$$

Proposition 11 : Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$.

(i) L'image de A est engendrée par ses colonnes, et le rang de A est le rang de la famille de ses colonnes.

(ii) Le noyau de A est l'ensemble des solutions du système homogène dont les équations sont données par les lignes de A.

(iii) Si $A = \operatorname{Mat}_{\mathcal{B},\mathcal{C}}(u)$, alors $\operatorname{rg}(A) = \operatorname{rg}(u)$.

∘ Démonstration :

▷ *(i)* Par la proposition 2, $AX = \sum_j x_j C_j$: l'image de A est l'ensemble des combinaisons linéaires des colonnes C_j, c'est-à-dire $\operatorname{Vect}(C_1, \dots, C_p)$.

▷ *(ii)* Le i-ième coefficient de AX est $a_{i1}x_1 + \cdots + a_{ip}x_p$, et $AX = 0$ est le système formé de ces n équations.

▷ *(iii)* Notons $\varphi : x \mapsto \operatorname{Mat}_{\mathcal{B}}(x)$ et $\psi : y \mapsto \operatorname{Mat}_{\mathcal{C}}(y)$, qui sont des isomorphismes. La proposition 10 s'écrit $\psi \circ u = u_A \circ \varphi$, soit $u = \psi^{-1} \circ u_A \circ \varphi$, et la composition par des isomorphismes ne change pas le rang (proposition 7 du chapitre 15). □

Exemple : La matrice $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 0 & 1 \end{pmatrix}$ a deux premières colonnes non colinéaires, et sa troisième colonne est la somme des deux premières : $\operatorname{rg}(A) = 2$. Par le théorème du rang, $\operatorname{Ker} A$ est une droite, et le vecteur $(1, 1, -1)$, qui exprime la relation $C_1 + C_2 - C_3 = 0$, l'engendre. La deuxième ligne, double de la première, n'apporte aucune équation nouvelle au noyau.

Théorème 3 : Caractérisations des matrices inversibles :
Pour $A \in \mathcal{M}_n(\mathbb{K})$, les assertions suivantes sont équivalentes :

(i) A est inversible ;
(ii) $\operatorname{Ker} A = \{0\}$;
(iii) les colonnes de A engendrent $\mathbb{K}^n$;
(iv) $\operatorname{rg}(A) = n$;
(v) il existe B telle que $BA = I_n$;
(vi) il existe B telle que $AB = I_n$.

Dans les deux derniers cas, $B = A^{-1}$.

∘ Démonstration : La matrice A est celle de u_A dans la base canonique de $\mathbb{K}^n$; par le théorème 2 (iv), (i) équivaut à la bijectivité de u_A. Le théorème 10 du chapitre 15 montre que celle-ci équivaut à l'injectivité de u_A, soit (ii), et à sa surjectivité, soit (iii) par la proposition 11, ou encore (iv). Si $BA = I_n$, alors $u_B \circ u_A = \mathrm{id}$ par le théorème 2 (ii), donc u_A est injective et A est inversible ; en multipliant à droite par A^{-1}, on obtient $B = A^{-1}$. Le cas $AB = I_n$ est symétrique, u_A étant alors surjective. Les implications (i) ⇒ (v) et (i) ⇒ (vi) sont claires. □

Remarque : Ce théorème est propre aux matrices *carrées*. La matrice ligne $A = \begin{pmatrix}1 & 0\end{pmatrix}$ et la matrice colonne $B = \begin{pmatrix}1\\0\end{pmatrix}$ vérifient $AB = I_1$, alors que $BA = E_{11} \neq I_2$: une matrice non carrée peut avoir un inverse d'un côté, jamais des deux, puisque u_A irait alors d'un espace dans un autre de dimension différente.

Proposition 12 : Une matrice triangulaire supérieure $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si ses coefficients diagonaux sont tous non nuls. Son inverse est alors triangulaire supérieure, de coefficients diagonaux $\frac{1}{a_{11}}, \dots, \frac{1}{a_{nn}}$.

∘ Démonstration : Notons $(\varepsilon_1, \dots, \varepsilon_n)$ la base canonique de $\mathbb{K}^n$ et $F_k = \text{Vect}(\varepsilon_1, \dots, \varepsilon_k)$. La matrice étant triangulaire supérieure, sa j-ième colonne $A\varepsilon_j$ appartient à F_j.

▷ *Si un coefficient diagonal est nul.* Soit k tel que $a_{kk} = 0$. Les k premières colonnes appartiennent alors toutes à F_{k-1}, de dimension $k-1$: elles sont liées (théorème 1 du chapitre 15), donc $\text{Ker}\, A \neq \{0\}$, et A n'est pas inversible.

▷ *Si les coefficients diagonaux sont non nuls.* Le système $AX = 0$ se résout de bas en haut : la dernière équation $a_{nn}x_n = 0$ donne $x_n = 0$, puis l'avant-dernière $x_{n-1} = 0$, et ainsi de suite. Donc $\text{Ker}\, A = \{0\}$ et A est inversible (théorème 3). Pour chaque k, $u_A(F_k) \subset F_k$, et u_A est injective, donc $u_A(F_k) = F_k$ par égalité des dimensions ; ainsi $A^{-1}\varepsilon_k \in F_k$, ce qui signifie que A^{-1} est triangulaire supérieure. Les coefficients diagonaux de $AA^{-1} = I_n$ sont alors les produits des coefficients diagonaux, par la proposition 6. □

16.5 Changements de bases, équivalence et similitude

16.5.1 Matrices de passage

Définition : Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases d'un espace E de dimension n. La **matrice de passage** de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$:

$$P_{\mathcal{B}}^{\mathcal{B}'} = \text{Mat}_{\mathcal{B}}(\mathcal{B}') \in \mathcal{M}_n(\mathbb{K}).$$

Sa j-ième colonne contient les coordonnées du j-ième vecteur de la *nouvelle* base $\mathcal{B}'$ dans l'*ancienne* base $\mathcal{B}$.

Proposition 13 : Soit $\mathcal{B}$, $\mathcal{B}'$ et $\mathcal{B}''$ trois bases de E.

(i) $P_{\mathcal{B}}^{\mathcal{B}'} = \text{Mat}_{\mathcal{B}',\mathcal{B}}(\text{id}_E)$, et $P_{\mathcal{B}}^{\mathcal{B}'} P_{\mathcal{B}'}^{\mathcal{B}''} = P_{\mathcal{B}}^{\mathcal{B}''}$.

(ii) $P_{\mathcal{B}}^{\mathcal{B}'}$ est inversible, et $\left(P_{\mathcal{B}}^{\mathcal{B}'}\right)^{-1} = P_{\mathcal{B}'}^{\mathcal{B}}$.

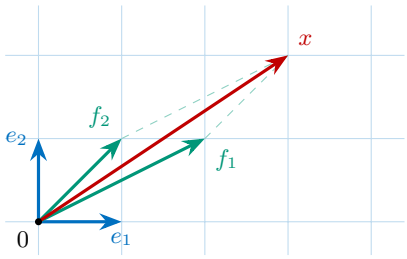
(iii) Pour $x \in E$, les coordonnées $X = \text{Mat}_{\mathcal{B}}(x)$ et $X' = \text{Mat}_{\mathcal{B}'}(x)$ vérifient $X = P_{\mathcal{B}}^{\mathcal{B}'} X'$.

∘ Démonstration :

▷ *(i)* La j-ième colonne de $\text{Mat}_{\mathcal{B}',\mathcal{B}}(\text{id}_E)$ contient les coordonnées dans $\mathcal{B}$ de $\text{id}_E(e'_j) = e'_j$. La relation vient du théorème 2 (ii), appliqué à $\text{id}_E \circ \text{id}_E$.

▷ *(ii)* Avec $\mathcal{B}'' = \mathcal{B}$, (i) donne $P_{\mathcal{B}}^{\mathcal{B}'} P_{\mathcal{B}'}^{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{B}} = I_n$, et le théorème 3 conclut.

▷ *(iii)* C'est la proposition 10 appliquée à id_E, de matrice $P_{\mathcal{B}}^{\mathcal{B}'}$ dans les bases $\mathcal{B}'$ au départ et $\mathcal{B}$ à l'arrivée. □



dans $\mathcal{B} = (e_1, e_2)$: $X = \begin{pmatrix}3\\2\end{pmatrix}$

dans $\mathcal{B}' = (f_1, f_2)$: $X' = \begin{pmatrix}1\\1\end{pmatrix}$

$P_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix}2 & 1\\1 & 1\end{pmatrix}$, $X = P_{\mathcal{B}}^{\mathcal{B}'} X'$

Le vecteur x, en rouge, a pour coordonnées $(3,2)$ dans la base canonique, en bleu, et $(1,1)$ dans la base (f_1,f_2), en vert, puisque $x = f_1 + f_2$: le parallélogramme en pointillés le montre. Le produit $P_{\mathcal{B}}^{\mathcal{B}'}X'$ redonne bien $(3,2)$.

Remarque : La matrice $P_{\mathcal{B}}^{\mathcal{B}'}$ exprime la nouvelle base dans l'ancienne, mais elle transforme les *nouvelles* coordonnées en *anciennes*. L'erreur la plus répandue est d'écrire $X' = P_{\mathcal{B}}^{\mathcal{B}'}X$; pour obtenir les nouvelles coordonnées, il faut inverser : $X' = \left(P_{\mathcal{B}}^{\mathcal{B}'}\right)^{-1}X$. Sur la figure, $\left(P_{\mathcal{B}}^{\mathcal{B}'}\right)^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$ envoie bien $(3,2)$ sur $(1,1)$.

Théorème 4 : Formules de changement de bases :
Soit $u \in \mathcal{L}(E,F)$, $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, $\mathcal{C}$ et $\mathcal{C}'$ deux bases de F. Avec $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$, $A' = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(u)$, $P = P_{\mathcal{B}}^{\mathcal{B}'}$ et $Q = P_{\mathcal{C}}^{\mathcal{C}'}$,

$$A' = Q^{-1}AP.$$

En particulier, si $u \in \mathcal{L}(E)$, $A = \mathrm{Mat}_{\mathcal{B}}(u)$ et $A' = \mathrm{Mat}_{\mathcal{B}'}(u)$, alors $A' = P^{-1}AP$.

∘ *Démonstration :* On écrit $u = \mathrm{id}_F \circ u \circ \mathrm{id}_E$, en munissant les espaces successifs des bases $\mathcal{B}'$, $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{C}'$. Le théorème 2 (ii) et la proposition 13 donnent

$$\mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(u) = \mathrm{Mat}_{\mathcal{C},\mathcal{C}'}(\mathrm{id}_F)\,\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)\,\mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E) = P_{\mathcal{C}'}^{\mathcal{C}}\,A\,P_{\mathcal{B}}^{\mathcal{B}'} = Q^{-1}AP.$$

□

♡ **Méthode : Pour simplifier la matrice d'un endomorphisme u par un changement de base :**

- On cherche des vecteurs dont l'image par u s'exprime simplement, par exemple des vecteurs v non nuls tels que $u(v)$ soit colinéaire à v, ou des sous-espaces que u laisse stables.
- On vérifie qu'ils forment une base $\mathcal{B}'$, et l'on écrit directement $\mathrm{Mat}_{\mathcal{B}'}(u)$.
- La matrice de passage P donne alors $\mathrm{Mat}_{\mathcal{B}}(u) = P\,\mathrm{Mat}_{\mathcal{B}'}(u)\,P^{-1}$, ce qui permet par exemple de calculer les puissances.

Exemple : Soit u l'endomorphisme de $\mathbb{R}^2$ de matrice $A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}$ dans la base canonique. Avec les vecteurs $f_1 = (2,1)$ et $f_2 = (1,1)$ de la figure, on calcule $Af_1 = (6,3) = 3f_1$ et $Af_2 = (2,2) = 2f_2$. Dans la base $\mathcal{B}' = (f_1,f_2)$, la matrice de u est donc $D = \mathrm{diag}(3,2)$, et $A = PDP^{-1}$, d'où pour tout $m \in \mathbb{N}$,

$$A^m = PD^mP^{-1} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3^m & 0 \\ 0 & 2^m \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2\cdot 3^m - 2^m & 2^{m+1} - 2\cdot 3^m \\ 3^m - 2^m & 2^{m+1} - 3^m \end{pmatrix}.$$

Pour $m = 1$, on retrouve A. Géométriquement, u dilate d'un facteur 3 dans la direction de f_1 et d'un facteur 2 dans celle de f_2 ; c'est ce que la base canonique cachait.

16.5.2 Matrices équivalentes et rang

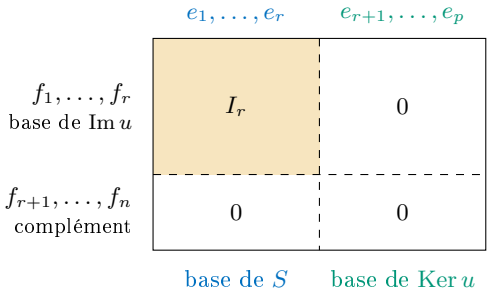
Pour $r \leqslant \min(n,p)$, on note $J_r \in \mathcal{M}_{n,p}(\mathbb{K})$ la matrice dont tous les coefficients sont nuls, sauf les r premiers coefficients diagonaux, égaux à 1 :

$$J_r = \sum_{i=1}^{r} E_{ii} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Théorème 5 : Forme normale d'une application linéaire :
Soit $u \in \mathcal{L}(E,F)$ de rang r, avec $\dim E = p$ et $\dim F = n$. Il existe une base $\mathcal{B}$ de E et une base $\mathcal{C}$ de F telles que

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) = J_r.$$

∘ Démonstration : Soit S un supplémentaire de $\operatorname{Ker} u$ dans E, de dimension r par le théorème du rang, et $(e_1, \dots, e_r)$ une base de S, complétée par une base $(e_{r+1}, \dots, e_p)$ de $\operatorname{Ker} u$: on obtient une base $\mathcal{B}$ de E (proposition 3 du chapitre 15). Par le théorème 8 du chapitre 15, u induit un isomorphisme de S sur $\operatorname{Im} u$, donc les vecteurs $f_i = u(e_i)$, pour $i \leqslant r$, forment une base de $\operatorname{Im} u$ (proposition 6 du chapitre 15) ; on la complète en une base $\mathcal{C} = (f_1, \dots, f_n)$ de F. Alors $u(e_j) = f_j$ pour $j \leqslant r$, et $u(e_j) = 0$ pour $j > r$: c'est exactement dire que $\operatorname{Mat}_{\mathcal{B},\mathcal{C}}(u) = J_r$. □



les r premiers vecteurs de $\mathcal{B}$ sont envoyés sur les r premiers vecteurs de $\mathcal{C}$, les $p - r$ derniers sur 0

Les colonnes de J_r correspondent aux vecteurs de la base de départ, ses lignes à ceux de la base d'arrivée : le bloc I_r, en orange, traduit l'isomorphisme de S sur $\operatorname{Im} u$, et les colonnes nulles le noyau. Toute l'information sur u, à changement de bases près, tient dans le seul entier r.

Définition : Deux matrices A, $A' \in \mathcal{M}_{n,p}(\mathbb{K})$ sont **équivalentes** s'il existe $Q \in \mathrm{GL}_n(\mathbb{K})$ et $P \in \mathrm{GL}_p(\mathbb{K})$ telles que $A' = Q^{-1}AP$.

Cette relation est une relation d'équivalence sur $\mathcal{M}_{n,p}(\mathbb{K})$: elle est réflexive avec $P = I_p$ et $Q = I_n$, symétrique puisque $A = QA'P^{-1}$, et transitive puisque $Q'^{-1}(Q^{-1}AP)P' = (QQ')^{-1}A(PP')$. Par le théorème 4, deux matrices qui représentent la même application linéaire dans deux couples de bases sont équivalentes ; réciproquement, toute matrice inversible est une matrice de passage, puisque ses colonnes forment une base (théorème 3), de sorte que deux matrices équivalentes représentent la même application linéaire dans deux couples de bases.

Théorème 6 : Classification par le rang :

(i) Une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est de rang r si et seulement si elle est équivalente à J_r.

(ii) Deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$ sont équivalentes si et seulement si elles ont le même rang.

(iii) Pour toute matrice A, $\operatorname{rg}(A^T) = \operatorname{rg}(A)$: le rang des lignes est égal au rang des colonnes.

∘ Démonstration :

▷ *(i)* Si $\operatorname{rg}(A) = r$, le théorème 5 appliqué à u_A donne des bases dans lesquelles u_A a pour matrice J_r, et le théorème 4 montre que A et J_r sont équivalentes. Réciproquement, multiplier une matrice à droite ou à gauche par une matrice inversible ne change pas son rang, par la proposition 7 du chapitre 15 appliquée aux applications canoniquement associées ; et $\operatorname{rg}(J_r) = r$, ses colonnes non nulles étant les r premiers vecteurs de la base canonique.

▷ *(ii)* Par (i), deux matrices de même rang r sont équivalentes à J_r, donc entre elles par transitivité ; et deux matrices équivalentes ont le même rang par l'argument précédent.

▷ *(iii)* Si $A = QJ_rP^{-1}$, la proposition 5 donne $A^T = (P^{-1})^T {J_r}^T Q^T$, où ${J_r}^T$ est la matrice J_r de $\mathcal{M}_{p,n}(\mathbb{K})$ et où les matrices extrêmes sont inversibles (proposition 8). Donc A^T est équivalente à J_r, et elle est de rang r par (i). □

Remarque : Le point (iii) est remarquable : rien ne laisse prévoir, sur une matrice de trois lignes et cinq colonnes, que le nombre maximal de lignes indépendantes, vecteurs de $\mathbb{K}^5$, soit égal au nombre maximal de colonnes indépendantes, vecteurs de $\mathbb{K}^3$. On en déduit que $\operatorname{rg}(A) \leqslant \min(n, p)$, et que les lignes d'une matrice carrée inversible sont libres.

Définition : Une **matrice extraite** de $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est une matrice obtenue en ne conservant que les coefficients situés sur certaines lignes $i_1 < \cdots < i_s$ et certaines colonnes $j_1 < \cdots < j_t$ de A.

Théorème 7 : Rang et matrices extraites :
Le rang d'une matrice extraite de A est au plus $\mathrm{rg}(A)$. Le rang de A est la taille maximale d'une matrice carrée inversible extraite de A.

∘ Démonstration :

▷ *Première affirmation.* Ne garder que certaines colonnes revient à passer à une sous-famille des colonnes, ce qui ne peut pas augmenter le rang. Ne garder que certaines lignes revient, après transposition, à ne garder que certaines colonnes de A^T, qui a le même rang que A (théorème 6). Une matrice extraite s'obtient en faisant les deux.

▷ *Seconde affirmation.* Une matrice carrée inversible extraite de taille s est de rang s, donc $s \leqslant \mathrm{rg}(A)$. Réciproquement, soit $r = \mathrm{rg}(A)$. On choisit r colonnes libres de A : elles forment une matrice $B \in \mathcal{M}_{n,r}(\mathbb{K})$ de rang r. Par le théorème 6, B^T est aussi de rang r, et l'on peut en choisir r colonnes libres, c'est-à-dire r lignes de B formant une matrice carrée de rang r. C'est une matrice carrée extraite de A, inversible par le théorème 3. □

Remarque : On retrouve l'idée de Sylvester rappelée dans l'introduction : le rang se lit sur les matrices carrées qui « naissent » de A. Combiné au chapitre 17, qui caractérisera l'inversibilité d'une matrice carrée par un déterminant non nul, ce théorème permettra de calculer le rang en examinant des déterminants extraits.

16.5.3 Matrices semblables et trace

Définition : Deux matrices A, $A' \in \mathcal{M}_n(\mathbb{K})$ sont **semblables** s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $A' = P^{-1}AP$.

C'est encore une relation d'équivalence, et par le théorème 4, deux matrices sont semblables si et seulement si elles représentent le même endomorphisme dans deux bases. Le diagramme suivant résume la situation.

$$\begin{array}{ccc} \mathbb{K}^n & \xrightarrow{A'=P^{-1}AP} & \mathbb{K}^n \\ {\scriptstyle P}\downarrow & & \downarrow{\scriptstyle P} \\ \mathbb{K}^n & \xrightarrow[A]{} & \mathbb{K}^n \end{array}$$

En haut, on lit les coordonnées dans la nouvelle base ; la matrice de passage P, en violet, les convertit en anciennes coordonnées, sur lesquelles on applique A. Aller par la droite puis descendre, ou descendre puis aller par la droite, donne le même résultat : $PA' = AP$.

Remarques :

- Deux matrices semblables sont équivalentes, donc ont le même rang ; la réciproque est fausse. Les matrices E_{12} et E_{11} de $\mathcal{M}_2(\mathbb{K})$ sont toutes deux de rang 1, mais elles ne sont pas semblables : si $E_{11} = P^{-1}E_{12}P$, alors $E_{11} = E_{11}^2 = P^{-1}E_{12}^2P = 0$, ce qui est absurde.
- Une matrice scalaire n'est semblable qu'à elle-même, puisque pour toute matrice P inversible,
$$P^{-1}(\lambda I_n)P = \lambda I_n :$$
l'homothétie de rapport λ a la même matrice dans toutes les bases.

Définition : La **trace** d'une matrice carrée $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ est la somme de ses coefficients diagonaux :
$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Proposition 14 : La trace est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$. Pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$,
$$\mathrm{tr}(AB) = \mathrm{tr}(BA).$$
En particulier, deux matrices semblables ont la même trace.

∘ Démonstration : La linéarité est claire. On calcule, par interversion des sommes,

$$\mathrm{tr}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{ij}b_{ji} = \sum_{j=1}^{p}\sum_{i=1}^{n} b_{ji}a_{ij} = \mathrm{tr}(BA).$$

Si $A' = P^{-1}AP$, alors $\mathrm{tr}(A') = \mathrm{tr}\left((P^{-1}A)P\right) = \mathrm{tr}\left(P(P^{-1}A)\right) = \mathrm{tr}(A)$. □

Remarques :

- La relation ne permet pas de permuter librement : $\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$, mais $\mathrm{tr}(ABC)$ diffère en général de $\mathrm{tr}(BAC)$. Avec $A = E_{12}$, $B = E_{21}$ et $C = E_{11}$, on trouve $\mathrm{tr}(ABC) = \mathrm{tr}(E_{11}) = 1$ et $\mathrm{tr}(BAC) = \mathrm{tr}(E_{22}E_{11}) = 0$.
- Il n'existe pas de matrices A, $B \in \mathcal{M}_n(\mathbb{K})$ telles que $AB - BA = I_n$: la trace du membre de gauche est nulle, celle du membre de droite vaut n. La trace interdit ainsi, en une ligne, une égalité que l'on aurait bien du mal à réfuter sur les coefficients.

Définition : Soit E de dimension finie et $u \in \mathcal{L}(E)$. La **trace** de u est $\mathrm{tr}(u) = \mathrm{tr}\left(\mathrm{Mat}_{\mathcal{B}}(u)\right)$, où $\mathcal{B}$ est une base quelconque de E.

La définition a un sens : les matrices de u dans deux bases sont semblables (théorème 4), donc ont la même trace (proposition 14).

Proposition 15 : La trace est une forme linéaire sur $\mathcal{L}(E)$, et $\mathrm{tr}(u \circ v) = \mathrm{tr}(v \circ u)$ pour u, $v \in \mathcal{L}(E)$. La trace d'un projecteur est égale à son rang.

∘ Démonstration : Les deux premiers points découlent de la proposition 14 par le théorème 2. Si p est un projecteur, $E = \mathrm{Im}\, p \oplus \mathrm{Ker}\, p$ (théorème 5 du chapitre 14) ; dans une base adaptée à cette décomposition, p a pour matrice $\mathrm{diag}(1, \dots, 1, 0, \dots, 0)$, avec $\mathrm{rg}(p)$ coefficients égaux à 1, et sa trace vaut $\mathrm{rg}(p)$. □

Exemple : La projection p de $\mathbb{R}^2$ sur la droite $\mathbb{R}(2,1)$ parallèlement à $\mathbb{R}(1,1)$ a pour matrice $\mathrm{diag}(1,0)$ dans la base $\big((2,1),(1,1)\big)$, donc, par le théorème 4, pour matrice dans la base canonique

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}.$$

Sa trace vaut $2 - 1 = 1$, son rang aussi ; on vérifie que cette matrice est égale à son carré. La trace d'un projecteur est donc un entier, ce que rien n'impose a priori à la somme $2 + (-1)$.

16.6 Opérations élémentaires et systèmes linéaires

16.6.1 Opérations élémentaires et produit matriciel

On reprend les trois opérations élémentaires du chapitre 3, appliquées aux lignes $L_1, \dots, L_n$ ou aux colonnes $C_1, \dots, C_p$ d'une matrice : l'échange de deux lignes, la multiplication d'une ligne par un scalaire non nul, l'ajout à une ligne d'un multiple d'une autre ligne, et de même pour les colonnes.

Définition : Soit $i \neq j$ dans $[\![1, n]\!]$, $\lambda \in \mathbb{K}$ et $\alpha \in \mathbb{K}^*$. On définit dans $\mathcal{M}_n(\mathbb{K})$:

- la **matrice de transvection** $T_{ij}(\lambda) = I_n + \lambda E_{ij}$;
- la **matrice de dilatation** $D_i(\alpha) = I_n + (\alpha - 1)E_{ii}$;
- la **matrice de transposition** $P_{ij} = I_n - E_{ii} - E_{jj} + E_{ij} + E_{ji}$, obtenue en échangeant les lignes i et j de I_n.

Proposition 16 : Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Multiplier A à gauche par une matrice de $\mathcal{M}_n(\mathbb{K})$ agit sur ses lignes :

$$T_{ij}(\lambda)A : L_i \leftarrow L_i + \lambda L_j, \qquad D_i(\alpha)A : L_i \leftarrow \alpha L_i, \qquad P_{ij}A : L_i \leftrightarrow L_j.$$

Multiplier A à droite par les matrices analogues de $\mathcal{M}_p(\mathbb{K})$ agit sur ses colonnes : $AT_{ij}(\lambda)$ effectue $C_j \leftarrow C_j + \lambda C_i$, $AD_j(\alpha)$ effectue $C_j \leftarrow \alpha C_j$, et AP_{ij} échange C_i et C_j. Ces matrices sont inversibles :

$$T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda), \qquad D_i(\alpha)^{-1} = D_i(\alpha^{-1}), \qquad P_{ij}^{-1} = P_{ij}.$$

∘ Démonstration : Par la proposition 4, la matrice $E_{ij}A$ a pour i-ième ligne la j-ième ligne de A, et ses autres lignes sont nulles. Donc $T_{ij}(\lambda)A = A+\lambda E_{ij}A$ se déduit de A en ajoutant λL_j à L_i ; les deux autres cas sont analogues. De même, AE_{ij} a pour j-ième colonne la i-ième colonne de A, ses autres colonnes étant nulles, d'où l'effet sur les colonnes. Enfin, l'opération inverse de chaque opération élémentaire est une opération élémentaire du même type ; appliquée à la matrice I_n, elle fournit l'inverse annoncé. □

Remarque : L'erreur la plus répandue est d'inverser les rôles de i et j pour les colonnes : à gauche, $T_{ij}(\lambda)$ ajoute à la ligne i un multiple de la ligne j, mais à droite, elle ajoute à la colonne j un multiple de la colonne i. Il suffit de le vérifier sur I_n, dont $T_{ij}(\lambda)$ est à la fois l'image par l'une et par l'autre opération.

Proposition 17 : Les opérations élémentaires sur les lignes conservent le noyau d'une matrice, les opérations élémentaires sur les colonnes conservent son image, et toutes conservent son rang.

∘ Démonstration : Une suite d'opérations sur les lignes remplace A par GA, où G est un produit de matrices élémentaires, donc inversible ; et $GAX = 0$ équivaut à $AX = 0$. Une suite d'opérations sur les colonnes remplace A par AG, avec $G \in \mathrm{GL}_p(\mathbb{K})$; comme $X \mapsto GX$ est surjective, $\mathrm{Im}(AG) = \mathrm{Im}\, A$. Dans les deux cas, le rang est conservé, par le théorème 6 (ii). □

16.6.2 L'algorithme de Gauss-Jordan

Une matrice est **échelonnée en lignes** si chaque ligne non nulle commence par strictement plus de zéros que la précédente, les lignes nulles étant placées en dernier ; le premier coefficient non nul d'une ligne est son **pivot**. Elle est **échelonnée réduite** si, de plus, chaque pivot vaut 1 et est le seul coefficient non nul de sa colonne.

Théorème 8 : Échelonnement :
Toute matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ peut être transformée, par une suite d'opérations élémentaires sur les lignes, en une matrice échelonnée réduite. Le rang de A est le nombre de pivots de cette matrice.

∘ Démonstration :

▷ *Existence.* C'est l'algorithme du pivot du chapitre 3, que l'on prolonge. Si la première colonne de A est nulle, on passe à la suivante. Sinon, on amène par un échange une ligne de premier coefficient non nul en première position, on la divise par ce coefficient, et l'on annule les autres coefficients de la colonne par des transvections $L_i \leftarrow L_i - a_{i1}L_1$. On recommence sur les lignes 2 à n et les colonnes suivantes, ce qui produit une matrice échelonnée en un nombre fini d'étapes, par récurrence sur le nombre de colonnes. On annule enfin les coefficients situés au-dessus de chaque pivot, en remontant de la dernière ligne non nulle à la première, par des transvections.

▷ *Rang.* Les lignes non nulles d'une matrice échelonnée sont libres : dans une combinaison nulle, la colonne du premier pivot ne fait intervenir que la première ligne, dont le coefficient est donc nul, puis la colonne du deuxième pivot donne la nullité du deuxième coefficient, et ainsi de suite. Le rang des lignes est donc le nombre de pivots, et il est égal au rang de la matrice (théorème 6 (iii)), qui est celui de A (proposition 17). □

♡ **Méthode : Pour calculer le rang d'une matrice :**
- ▶ On l'échelonne par des opérations sur les lignes, ou sur les colonnes, ou sur les deux, puisque toutes conservent le rang.
- ▶ On compte les pivots de la matrice échelonnée obtenue.

Il est inutile d'aller jusqu'à la forme réduite : une forme échelonnée suffit.

Exemple : Pour la matrice A de la proposition 11, les opérations $L_2 \leftarrow L_2 - 2L_1$ et $L_3 \leftarrow L_3 - L_1$, puis l'échange $L_2 \leftrightarrow L_3$, donnent

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 0 & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & -2 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Il y a deux pivots : $\mathrm{rg}(A) = 2$, comme on l'avait vu en cherchant une relation entre les colonnes. Le noyau, conservé par ces opérations, se lit sur la forme échelonnée : $x_2 = -x_3$ et $x_1 = -2x_2 - 3x_3 = -x_3$, ce qui redonne la droite $\mathbb{R}(1, 1, -1)$.

♡ **Méthode : Pour inverser une matrice $A \in \mathcal{M}_n(\mathbb{K})$ par la méthode de Gauss-Jordan :**

- On écrit côte à côte A et I_n, sous la forme d'un tableau $(A \mid I_n)$ à n lignes et $2n$ colonnes.
- On effectue sur les lignes de ce tableau les opérations qui transforment A en une matrice échelonnée réduite.
- Si l'on obtient I_n à gauche, A est inversible et la partie droite est A^{-1} ; si une ligne nulle apparaît à gauche, A n'est pas inversible.

En effet, la suite d'opérations revient à multiplier le tableau à gauche par une matrice inversible G, ce qui donne $(GA \mid G)$. Si $GA = I_n$, alors $G = A^{-1}$ par le théorème 3, et la partie droite est A^{-1}. Si GA a une ligne nulle, elle est de rang strictement inférieur à n, et A, de même rang, n'est pas inversible.

Exemple : Inversons $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. L'opération $L_3 \leftarrow L_3 - L_1$, puis $L_3 \leftarrow L_3 + L_2$, donne

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right).$$

La partie gauche est triangulaire supérieure à pivots égaux à 1 : A est inversible. On remonte par $L_2 \leftarrow L_2 - L_3$ et $L_1 \leftarrow L_1 - L_3$, puis $L_1 \leftarrow L_1 - 2L_2$:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 2 & -1 & -1 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{array}\right).$$

La partie gauche est devenue I_3, et la partie droite est l'inverse cherché :

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

On vérifie que $AA^{-1} = I_3$. Ici l'inverse est à coefficients entiers, contrairement à celui de la matrice C de la section 16.3 ; le chapitre 17 l'expliquera par le déterminant de A, qui vaut 1.

Proposition 18 : Toute matrice inversible est un produit de matrices de transvection, de dilatation et de transposition.

◦ Démonstration : Si A est inversible, la méthode de Gauss-Jordan fournit des matrices élémentaires $G_1, \ldots, G_k$ telles que $G_k \cdots G_1 A = I_n$. Alors $A = G_1^{-1} \cdots G_k^{-1}$, et chaque G_i^{-1} est une matrice élémentaire du même type (proposition 16). □

16.6.3 Systèmes linéaires

Un système de n équations linéaires à p inconnues, au sens du chapitre 3, s'écrit $AX = B$, où $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est la matrice de ses coefficients, $X \in \mathbb{K}^p$ la colonne des inconnues et $B \in \mathbb{K}^n$ celle des seconds membres. Le **système homogène associé** est $AX = 0$; le système est **compatible** s'il admet au moins une solution. Le **rang** du système est celui de A.

Théorème 9 : Structure des solutions d'un système linéaire :
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ de rang r et $B \in \mathbb{K}^n$.

(i) L'ensemble des solutions du système homogène $AX = 0$ est $\operatorname{Ker} A$, sous-espace de $\mathbb{K}^p$ de dimension $p - r$.

(ii) Le système $AX = B$ est compatible si et seulement si $B \in \operatorname{Im} A$, c'est-à-dire si B est combinaison linéaire des colonnes de A.

(iii) S'il est compatible et si X_0 en est une solution, ses solutions sont les $X_0 + Y$, où Y parcourt $\operatorname{Ker} A$.

(iv) Si $n = p$, le système admet une unique solution pour tout B si et seulement si A est inversible ; cette solution est $X = A^{-1}B$, et l'on dit que le système est **de Cramer**.

∘ Démonstration :

▷ *(i)* C'est la définition du noyau et le théorème du rang appliqué à u_A, ce qui retrouve la proposition 9 du chapitre 15.

▷ *(ii)* Par définition, $\operatorname{Im} A$ est l'ensemble des AX, et la proposition 11 l'identifie à l'ensemble des combinaisons linéaires des colonnes.

▷ *(iii)* Si $AX_0 = B$, alors $AX = B$ équivaut à $A(X - X_0) = 0$. C'est le cas particulier, pour l'application u_A, de la structure des solutions d'une équation linéaire (chapitre 14).

▷ *(iv)* Si A est inversible, $AX = B$ équivaut à $X = A^{-1}B$. Réciproquement, si le système admet une unique solution pour un second membre, (iii) montre que $\operatorname{Ker} A = \{0\}$, et A est inversible par le théorème 3. □

Géométriquement, chaque équation non nulle $a_{i1}x_1 + \cdots + a_{ip}x_p = b_i$ définit un hyperplan affine de $\mathbb{K}^p$, et l'ensemble des solutions en est l'intersection : il est vide, ou c'est un sous-espace affine de direction $\operatorname{Ker} A$, de dimension $p - r$. Le rang r est le nombre d'équations « vraiment indépendantes », comme au chapitre 15.

Exemples :

- Le système $x + 2y - z = 2$, $2x + y + z = 7$, $-x + y + 2z = 3$ du chapitre 3 admet l'unique solution $(2, 1, 2)$. Par le point (iv), sa matrice est inversible, et le système obtenu en changeant les seconds membres admet encore une unique solution, quels qu'ils soient : le calcul fait au chapitre 3 pour un seul second membre suffisait à le prévoir.
- La matrice $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix}$ est de rang 2, sa troisième ligne étant la somme des deux premières, et son noyau est la droite $\mathbb{R}(1, -2, 1)$. Pour $B = (3, 6, 9)$, le chapitre 3 avait trouvé les solutions $(t, 3 - 2t, t) = (0, 3, 0) + t(1, -2, 1)$: c'est une solution particulière plus le noyau, une droite affine de $\mathbb{R}^3$, de dimension $3 - 2 = 1$.
- Avec la matrice $A' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 3 \end{pmatrix}$, de rang 2 puisque sa troisième ligne est la somme des deux premières, le système de second membre $(1, 2, 4)$ est incompatible, alors qu'il devient compatible pour le second membre $(1, 2, 3)$: l'image de A' est le plan des (b_1, b_2, b_3) vérifiant $b_3 = b_1 + b_2$, et seul le second des deux seconds membres y appartient.

Remarque : Un système carré n'est donc pas « en général » de Cramer sans vérification : l'erreur la plus répandue est de conclure à l'unicité de la solution parce que le système a autant d'équations que d'inconnues. C'est l'inversibilité de A, c'est-à-dire son rang, qui décide ; le chapitre 17 fournira avec le déterminant un critère calculable directement.

16.7 Bilan

Le chapitre a construit un dictionnaire entre deux langages : celui, géométrique, des applications linéaires des chapitres 14 et 15, et celui, numérique, des tableaux de nombres. Chaque notion de l'un a sa traduction dans l'autre, et l'on passe de l'un à l'autre en choisissant des bases.

Applications linéaires	**Matrices**	**Référence**
$y = u(x)$	$Y = AX$	proposition 10
composition $v \circ u$	produit BA	théorème 2
isomorphisme	matrice inversible	théorèmes 2 et 3
$\operatorname{Im} u$, $\operatorname{Ker} u$, $\operatorname{rg}(u)$	colonnes, lignes, $\operatorname{rg}(A)$	proposition 11
changement de bases	$A' = Q^{-1}AP$	théorème 4
u de rang r	A équivalente à J_r	théorèmes 5 et 6
endomorphisme dans deux bases	matrices semblables, $A' = P^{-1}AP$	théorème 4
$\operatorname{tr}(u)$, $\operatorname{tr}(u \circ v) = \operatorname{tr}(v \circ u)$	$\operatorname{tr}(A)$, $\operatorname{tr}(AB) = \operatorname{tr}(BA)$	propositions 14 et 15
équation linéaire $u(x) = b$	système $AX = B$	théorème 9

Trois idées sont à retenir. Le produit matriciel n'est pas une convention arbitraire, mais la traduction de la composition. Le rang est le seul invariant d'une matrice à équivalence près, et il se calcule par l'algorithme du pivot. La similitude, plus fine que l'équivalence, ne se réduit pas au rang ni à la trace : décider si deux matrices sont semblables, et trouver une base où la matrice d'un endomorphisme est la plus simple possible, est le problème de la *réduction*, qui occupera une large part de l'année suivante. Le chapitre 17 apporte un outil décisif, le déterminant, qui caractérise l'inversibilité par le calcul d'un seul nombre.