Chapitre 31

Matrix Reloaded

Dans ce chapitre, on note comme d'habitude $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et on se donne trois $\mathbb{K}$-espaces vectoriels E, F et G de dimension finie non nulle. Sauf indication contraire (dans la partie III), on suppose que $\dim(E) = p$, $\dim(F) = n$ et $\dim(G) = q$ et on munit ces espaces des bases $B_E = (e_1, \dots, e_p)$, $B_F = (f_1, \dots, f_n)$ et $B_G = (g_1, \dots, g_q)$.

Comme dans les chapitres précédents, on se limite à $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ car c'est le cadre du programme mais les résultats de ce chapitre sont encore vrais sur un corps quelconque.

I Représentation matricielle

I.1 Matrices de vecteurs

On rappelle le résultat suivant (cf. chapitre 28) : B_E étant une base de E, tout vecteur $x \in E$ s'écrit de manière unique comme combinaison linéaire des e_i, $1 \leqslant i \leqslant p$. c'est-à-dire que pour tout $x \in E$, il existe $(\alpha_1, \dots, \alpha_p) \in \mathbb{K}^p$ unique tel que $x = \alpha_1 e_1 + \cdots + \alpha_p e_p$. Par conséquent, pour caractériser totalement un élément de E, il suffit de donner ses coordonnées dans la base B_E choisie.

Les α_i, $1 \leqslant i \leqslant p$, sont appelées coordonnées de x dans la base B_E et dépendent de la base. Par exemple, les coordonnées de $x = (1,2)$ dans la base canonique de $\mathbb{K}^2$ sont 1 et 2 mais ses coordonnées dans la base (c'en est une car formée de deux vecteurs non colinéaires en dimension 2) $((1,2), (-217, 1789))$ sont 1 et 0.

Définition. Soit $x \in E$. Soient $x_1, \dots, x_p$ les coordonnées de E dans la base B_E. On définit la matrice colonne des coordonnées de x dans la base B_E, noté $\mathrm{Mat}_{B_E}(x)$, X_{B_E} ou X si aucune confusion n'est possible, par :

$$X_{B_E} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

Proposition. L'application

$$\varphi : \left\{ \begin{array}{ccc} E & \longrightarrow & \mathscr{M}_{p,1}(\mathbb{K}) \\ x & \longmapsto & X \end{array} \right.$$

est un isomorphisme.

Démonstration.

- Montrons tout d'abord que φ est linéaire. Soient $(x,y) \in E^2$ et $(\lambda, \mu) \in \mathbb{K}^2$. Notons $x = x_1 e_1 + \cdots + x_p e_p$ et $y = y_1 e_1 + \cdots + y_p e_p$ et donc

$$\lambda x + \mu y = (\lambda x_1 + \mu y_1) e_1 + \cdots + (\lambda x_p + \mu y_p) e_p$$

Dès lors :

$$\begin{aligned} \varphi(\lambda x + \mu y) &= \begin{pmatrix} \lambda x_1 + \mu y_1 \\ \vdots \\ \lambda x_p + \mu y_p \end{pmatrix} \\ &= \lambda \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} + \mu \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \\ &= \lambda \varphi(x) + \mu \varphi(y) \end{aligned}$$ $\square$

- Soit $x \in \ker(\varphi)$. Alors $X = 0$ donc les coordonnées de x sont nulles donc $x = 0.e_1 + \cdots + 0.e_p = 0$: on en déduit que $\ker(\varphi) = \{0\}$ donc φ est injective.

- Enfin, soit $X = (x_1 \dots x_p)^\top \in \mathscr{M}_{p,1}(\mathbb{K})$ et soit $x = x_1e_1 + \cdots + x_pe_p$: alors $\varphi(x) = X$, φ est surjective.

Remarque : On pourra donc (et on ne se privera pas de le faire) identifier x et X. On dira indifféremment que $X \in \mathbb{K}^p$ ou que $X \in \mathscr{M}_{p,1}(\mathbb{K})$, ce qui n'a aucune importance puisque ces deux espaces vectoriels sont isomorphes car ils ont la même dimension (et donc on peut les identifier).

On peut généraliser cette définition à une famille de plusieurs vecteurs en mettant « côte à côte » les matrices colonnes associées :

Définition. Soit $(x_1, \dots, x_n) \in E^n$. On définit la matrice des coordonnées des x_j dans la base B_E par :

$$\text{Mat}_{B_E}(x_1, \dots, x_n) = (X_1|X_2|\dots|X_n)$$

c'est-à-dire la matrice obtenue en mettant côte à côte les matrices colonnes des coordonnées de $x_1, \dots, x_n$.

Ici, X_1 désigne évidemment la matrice colonne des coordonnées de x_1 dans la base B_E définie précédemment, c'est-à-dire $\text{Mat}_{B_E}(x_1)$, et idem pour les autres.

Définition (matrice canoniquement associée). Si E admet une base dite canonique (donc si E est du type $\mathbb{K}^n$, $\mathbb{K}_n[X]$ ou $\mathscr{M}_{n,p}(\mathbb{K})$) et si B_E est la base canonique de E, on dit que $\text{Mat}_{B_E}(x_1, \dots, x_n)$ est la matrice canoniquement associée aux x_j.

Remarques :

- C'est une matrice de taille $p \times n$: le nombre de colonnes est égal au nombre de vecteurs (n, dans notre définition), et le nombre de lignes à la dimension de l'espace (i.e. le nombre de coordonnées de chaque vecteur, p dans notre définition).
- En d'autres termes, pour tout $j \in [\![1\,;\, p]\!]$, la j-ème colonne de cette matrice est formée des coordonnées de x_j dans la base B_E, c'est-à-dire que $\text{Mat}_{B_E}(x_1, \dots, x_p) = (a_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$ où, pour tout $j \in [\![1\,;\, p]\!]$:

$$x_j = \sum_{i=1}^{n} a_{i,j} f_i$$

- Encore en d'autres termes, pour tous i et j, $a_{i,j}$ est la i-ème coordonnée de x_j.
- Méthode :
  - ⋆ Tracer des parenthèses ouvrante et fermante.
  - ⋆ Écrire au-dessus $x_1, \dots, x_p$.
  - ⋆ Écrire à droite $e_1, \dots, e_n$.
  - ⋆ Constater que les parenthèses n'ont pas la bonne taille et rectifier...
  - ⋆ Écrire dans chaque colonne les coordonnées du vecteur correspondant.

Exemples :

- Donner la matrice canoniquement associée aux vecteurs $x_1 = (2,3), x_2 = (0,1)$ et $x_3 = (4,5)$.

  Ici, c'est facile, puisque les coordonnées dans la base canonique de $\mathbb{K}^2$ sont les coordonnées des vecteurs. Si on note $B = (e_1, e_2)$ la base canonique de $\mathbb{K}^2$:

$$\text{Mat}_B(x_1, x_2, x_3) = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{pmatrix} 2 & 0 & 4 \\ 3 & 1 & 5 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \end{array}$$

- Donner la matrice canoniquement associée aux polynômes $(1, 2X, 3X^2, \dots, nX^{n-1})$ dans $E = \mathbb{K}_n[X]$.

  Notons $B = (1, X, \dots, X^n)$ la base canonique de $\mathbb{K}_n[X]$. Alors la matrice canoniquement associée à $(1, 2X, \dots, nX^{n-1})$ est la matrice de taille $n \times n+1$ suivante :

Ce n'est pas une matrice carrée !

$$\mathrm{Mat}_B(1, 2X, \dots, nX^{n-1}) = \begin{array}{cc} \begin{array}{ccccc} 1 & 2X & \cdots & \cdots & nX^{n-1} \end{array} & \\ \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 2 & & & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & n \\ 0 & 0 & \cdots & & 0 \end{pmatrix} & \begin{array}{c} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{array} \end{array}$$

  Cependant, si on prend comme base $B' = (1, 2X, 3X^2, \dots, (n+1)X^n)$, la matrice associée est :

C'est bien une base : en effet, c'est une famille libre, car échelonnée en degré, à $n+1$ éléments dans un espace de dimension $n+1$.

$$\mathrm{Mat}_{B'}(1, 2X, \dots, nX^{n-1}) = \begin{array}{cc} \begin{array}{ccccc} 1 & 2X & \cdots & \cdots & nX^{n-1} \end{array} & \\ \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & & & 0 \\ \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & & 0 \end{pmatrix} & \begin{array}{c} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{array} \end{array}$$

- Donner la matrice canoniquement associée à

$$M = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 4 \end{pmatrix} \in \mathscr{M}_{2,3}(\mathbb{K})$$

  Attention, ici, on a beau avoir une matrice, ce n'est rien d'autre qu'un seul vecteur de l'espace vectoriel $\mathscr{M}_{2,3}(\mathbb{K})$! Et puisque la base canonique de $\mathscr{M}_{2,3}(\mathbb{K})$ est $B = (E_{1,1}, E_{1,2}, E_{1,3}, E_{2,1}, E_{2,2}, E_{2,3})$, la matrice canoniquement associée à M est donnée ci-contre.

$$\mathrm{Mat}_B(M) = \begin{array}{cc} M & \\ \begin{pmatrix} 1 \\ -1 \\ 2 \\ 3 \\ 1 \\ 4 \end{pmatrix} & \begin{array}{c} E_{1,1} \\ E_{1,2} \\ E_{1,3} \\ E_{2,1} \\ E_{2,2} \\ E_{2,3} \end{array} \end{array}$$

Il n'y a pas d'ordre naturel pour cette base canonique : on pourrait très bien l'ordonner différemment, par exemple écrire $B' = (E_{1,1}, E_{2,1}, E_{1,2}, E_{2,2}, E_{1,3}, E_{2,3})$ et alors cela modifierait la matrice. De même, dans l'exemple précédemment, on aurait pu écrire la base canonique $B' = (X^n, \dots, 1)$ ce qui aurait changé la matrice. Il y a évidemment un lien entre les matrices d'une même famille dans deux bases différentes : cf. paragraphe VI.2.

I.2 Matrices et applications linéaires

I.2.a Introduction

Le résultat rappelé au début du paragraphe précédent étant vrai dans n'importe quel espace vectoriel, il est également vrai sur F, c'est-à-dire que, puisque B_F est une base de F, tout vecteur $y \in F$ s'écrit de manière unique comme combinaison linéaire des f_i, $1 \leqslant i \leqslant n$, c'est-à-dire que pour tout $y \in F$, il existe $(\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n$ (appelés encore une fois les coordonnées de y dans la base B_F et dépendant de la base choisie) uniques tels que $y = \alpha_1 f_1 + \cdots + \alpha_n f_n$.

On rappelle également (cf. chapitre 29) qu'une application linéaire de E dans F est entièrement déterminée par l'image des vecteurs d'une base.

On peut même aller plus loin en combinant les deux résultats précédents : une application linéaire u de E dans F est entièrement déterminée par les coordonnées des $u(e_j)$, $1 \leqslant j \leqslant p$, dans la base B_F.

Exemples :

- On prend $E = F = \mathbb{K}^2$ muni de sa base canonique (e_1, e_2). Soit $u : \mathbb{K}^2 \longrightarrow \mathbb{K}^2$ l'unique application linéaire vérifiant $u(e_1) = (1,0)$ et $u(e_2) = (3,1)$. Alors pour tout $(x,y) \in \mathbb{K}^2$:

$$\begin{aligned} u(x,y) &= u(xe_1 + ye_2) \\ &= xu(e_1) + yu(e_2) \\ &= x(1,0) + y(3,1) \\ &= (x+3y, y) \end{aligned}$$

- On prend à présent $E = \mathbb{K}^2, F = \mathbb{K}_3[X]$ munis de leurs bases canoniques. Soit $u : \mathbb{K}^2 \longrightarrow \mathbb{K}_3[X]$ l'unique application linéaire vérifiant $u(e_1) = (1,0,-1,2)$ et $u(e_2) = (3,1,1,0)$ (en notant les coordonnées dans la base canonique $(1, X, X^2, X^3)$ de F). Alors, pour tout $(x,y) \in \mathbb{K}^2$:

$$\begin{aligned} u(x,y) &= xu(e_1) + yu(e_2) \\ &= x(1 - X^2 + 2X^3) + y(3 + X + X^2) \\ &= 2xX^3 + (y-x)X^2 + yX + 3y + x \end{aligned}$$

I.2.b Matrice d'une application linéaire dans un couple de bases

Définition. Soit $u \in \mathscr{L}(E,F)$. On appelle matrice de u dans les bases B_E et B_F la matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$ dont la j-ème colonne est formée des coordonnées de $u(e_j)$ dans la base B_F. On la note $\mathrm{Mat}_{B_F,B_E}(u)$ ou $\mathrm{Mat}(u)$ quand aucune confusion n'est possible.

Dans la définition ci-contre, quand on dit « la j-ème colonne », il est sous-entendu : pour tout $j \in [\![1\,;\,p]\!]$ (puisque, par définition, la matrice a p colonnes). On ne l'écrit pas pour ne pas alourdir la définition.

Remarques :

- En d'autres termes, $A = (A_{i,j})_{1 \leqslant i \leqslant n, 1 \leqslant j \leqslant p}$ où, pour tout $j \in [\![1\,;\,p]\!]$,

$$u(e_j) = \sum_{i=1}^{n} A_{i,j} f_i$$

c'est-à-dire que pour tous i et j, $A_{i,j}$ est la i-ème coordonnée de $u(e_j)$. On note :

$$\mathrm{Mat}_{B_F,B_E}(u) = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & \dots & u(e_p) \end{array} \\ \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,p} \\ A_{2,1} & A_{2,2} & \dots & A_{2,p} \\ \vdots & \vdots & & \vdots \\ A_{n,1} & A_{n,2} & \dots & A_{n,p} \end{pmatrix} \end{array} \begin{array}{c} f_1 \\ f_2 \\ \vdots \\ f_n \end{array}$$

- Moyen mnémotechnique pour le placement des bases (Mat_{B_F,B_E} et non pas Mat_{B_E,B_F}) : dans l'égalité ci-dessus, B_E et B_E sont côte à côte, ainsi que B_F et B_F (en faisant le tour de la feuille...).
- Si A est la matrice de u, on dit que u est représentée par A.
- Le nombre de lignes de $A = \mathrm{Mat}(u)$ est égal à la dimension de l'espace d'arrivée, et le nombre de colonnes à la dimension de l'espace de départ. Par exemple, si u est une forme linéaire, alors $A = \mathrm{Mat}(u)$ est une matrice ligne.
- La matrice A n'est rien d'autre que la matrice des coordonnées des vecteurs $(u(e_1), \dots, u(e_p))$ dans la base B_F. On l'obtient donc de la même façon que dans le paragraphe I.1 :

- ⋆ On trace des parenthèses ouvrante et fermante.
- ⋆ On écrit au-dessus $u(e_1), \dots, u(e_p)$ (et on retrouve le fait que le nombre de colonnes est la dimension de l'espace de départ).
- ⋆ On écrit à droite $f_1, \dots, f_n$ (et on retrouve le fait que le nombre de lignes est la dimension de l'espace d'arrivée).
- ⋆ On constate que les parenthèses n'ont pas la bonne taille et on rectifie...
- ⋆ On écrit dans chaque colonne les coordonnées du vecteur correspondant.

- Là aussi, pour certains espaces vectoriels, on peut définir la matrice canoniquement associée à une application linéaire :

Définition (matrice canoniquement associée). Si E et F admettent des bases dites canoniques (donc si E et F sont du type $\mathbb{K}^n$, $\mathbb{K}_n[X]$ ou $\mathscr{M}_{n,p}(\mathbb{K})$) et si B_E et B_F sont les bases canoniques de E et F, on dit que $\text{Mat}_{B_F,B_E}(u)$ est la matrice canoniquement associée à u.

Exemples :

- Les matrices canoniquement associées aux applications linéaires du paragraphe précédent sont respectivement :

$$A_1 = \begin{array}{cc} \begin{array}{cc} u(e_1) & u(e_2) \end{array} & \\ \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \end{array} \end{array} \qquad \text{et} \qquad A_2 = \begin{array}{cc} \begin{array}{cc} u(e_1) & u(e_2) \end{array} & \\ \begin{pmatrix} 1 & 3 \\ 0 & 1 \\ -1 & 1 \\ 2 & 0 \end{pmatrix} & \begin{array}{c} 1 \\ X \\ X^2 \\ X^3 \end{array} \end{array}$$

- Considérons

$$D : \left\{ \begin{array}{ccc} D_n[X] & \longrightarrow & \mathbb{K}_n[X] \\ P & \longmapsto & P' \end{array} \right.$$

La matrice canoniquement associée (carrée de taille $n+1$) à D est :

$$A = \begin{array}{cc} \begin{array}{cccccc} D(1) & D(X) & D(X^2) & \cdots & \cdots & D(X^n) \end{array} & \\ \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & & & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & n \\ 0 & 0 & 0 & \cdots & & 0 \end{pmatrix} & \begin{array}{c} 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{array} \end{array}$$

Cependant, si on prend comme base de départ $B_E = (1, X, \dots, X^n)$ (c'est-à-dire la base canonique) et comme base d'arrivée $B_F = (1, 2X, 3X^2, \dots, (n+1)X^n)$, la matrice associée est :

$$\widetilde{A} = \begin{array}{cc} \begin{array}{cccccc} D(1) & D(X) & D(X^2) & \cdots & \cdots & D(X^n) \end{array} & \\ \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & & & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & 0 & \cdots & & 0 \end{pmatrix} & \begin{array}{c} 1 \\ 2X \\ 3X^2 \\ \vdots \\ nX^{n-1} \\ (n+1)X^n \end{array} \end{array}$$

On peut remarquer que la matrice dépend des bases B_E et B_F choisies, c'est-à-dire que si on prend des bases différentes, on aura (pour une même application linéaire !) des matrices différentes. Y a-t-il un lien entre les matrices associées à une même application linéaire ? Réponse au paragraphe VI.

- Considérons

$$u : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x,y,z) & \longmapsto & (x+2y, z-x-y) \end{array} \right.$$

La matrice canoniquement associée à u est :

$$A = \begin{array}{c} \begin{array}{ccc} u(e_1) & u(e_2) & u(e_3) \end{array} \\ \begin{pmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \end{array}$$

Si on ne fait pas attention, il peut y avoir une ambiguïté ici concernant les notations e_1 et e_2. En effet, quand on se place dans $\mathbb{K}^n$, la base canonique est habituellement notée $(e_1, \dots, e_n)$. Or, ici, on prend à la fois la base canonique de $\mathbb{K}^3$, notée (e_1, e_2, e_3), et celle de $\mathbb{K}^2$, notée (e_1, e_2), ce qui peut prêter à confusion : dans le premier cas, $e_1 = (1,0,0)$ (c'est un vecteur de $\mathbb{K}^3$), tandis que dans le second cas, $e_1 = (1,0)$ (c'est un vecteur de $\mathbb{K}^2$). On pourrait prendre des notations différentes pour empêcher toute ambiguïté (par exemple noter (e_1, e_2, e_3) la base canonique de $\mathbb{K}^3$ et (f_1, f_2) celle de $\mathbb{K}^2$), mais bon... prendre des notations différentes serait couper les cheveux en quatre dans le sens de la longueur pour quelque chose qui, somme toute, est assez anecdotique et n'est pas du tout ambigu si on réfléchit tant soit peu à l'espace dans lequel on se trouve. De plus, ce genre d'ambiguïté ne se produira pas à l'écrit.

- Soit

$$u : \left\{ \begin{array}{ccc} \mathbb{R}^3 & \longrightarrow & \mathbb{R}_4[X] \\ (x,y,z) & \longmapsto & (3y+x)X^4 - yX^2 + (z+x)X + y \end{array} \right.$$

La matrice canoniquement associée à u est :

$$A = \begin{array}{c} \begin{array}{ccc} u(e_1) & u(e_2) & u(e_3) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 3 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ X^2 \\ X^3 \\ X^4 \end{array}$$

- Soit $M \in \mathscr{M}_{n,p}(\mathbb{K})$ et soit

$$u : \left\{ \begin{array}{ccc} \mathscr{M}_{p,1}(\mathbb{K}) & \longrightarrow & \mathscr{M}_{n,1}(\mathbb{K}) \\ X & \longmapsto & MX \end{array} \right.$$

Donnons sa matrice canoniquement associée. Notons $E_{1,1}, \dots, E_{p,1}$ la base canonique de $\mathscr{M}_{p,1}(\mathbb{K})$ (ce sont des vecteurs colonnes avec p lignes). Soit $j \in [\![1\,;\,p]\!]$. Alors :

Comme ci-dessus, nous notons de la même façon la base canonique de $\mathscr{M}_{p,1}(\mathbb{K})$ et celle de $\mathscr{M}_{n,1}(\mathbb{K})$ car il n'y a aucun risque de confusion dès qu'on s'intéresse à la taille des matrices.

$$ME_{j,1} = \begin{pmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,p} \\ M_{2,1} & M_{2,2} & \dots & M_{2,p} \\ \vdots & \vdots & & \vdots \\ M_{n,1} & M_{n,2} & \dots & M_{n,p} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{c} 1 \\ \vdots \\ j \\ \vdots \\ p \end{array} = \begin{pmatrix} M_{1,j} \\ M_{2,j} \\ \vdots \\ M_{n,j} \end{pmatrix} = \sum_{i=1}^{n} M_{i,j} E_{i1}$$

où l'on a noté $E_{1,1}, \dots, E_{n,1}$ la base canonique de $\mathscr{M}_{n,1}(\mathbb{K})$ (ce sont, eux, des vecteurs colonnes à n lignes). Or, par définition, la j-ième colonne de $\mathrm{Mat}(u)$ est formée des coordonnées de $ME_{j,1}$, c'est-à-dire que la j-ième colonne de $\mathrm{Mat}(u)$ est la j-ième colonne de M. Par conséquent, la matrice canoniquement associée à u est M.

- Soit $z_0 = a + ib \in \mathbb{C}$. L'application

$$s : \begin{cases} \mathbb{C} \longrightarrow \mathbb{C} \\ z \longmapsto z \times z_0 \end{cases}$$

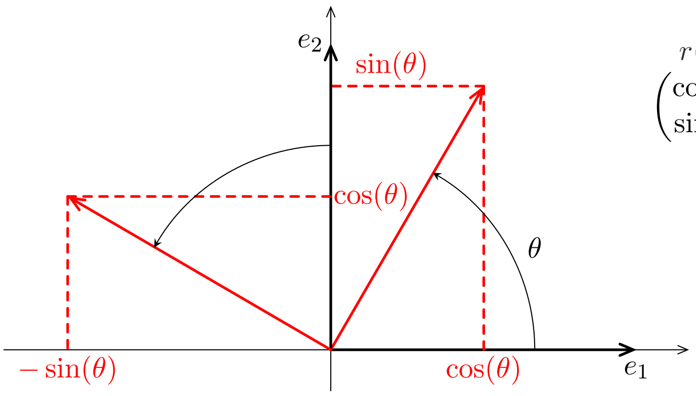
est $\mathbb{R}$-linéaire (exo). Donnons sa matrice dans la base $(1, i)$ de $\mathbb{C}$ vu comme $\mathbb{R}$-espace vectoriel : on sait que $s(1) = a + ib$ et $s(i) = -b + ia$ si bien que la matrice recherchée est

$$\begin{array}{cc} \begin{array}{cc} s(1) & s(i) \end{array} & \\ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} & \begin{array}{c} 1 \\ i \end{array} \end{array}$$

Par exemple, la matrice associée à la multiplication par $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ est

$$\begin{array}{cc} \begin{array}{cc} s(1) & s(i) \end{array} & \\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} & \begin{array}{c} 1 \\ i \end{array} \end{array}$$

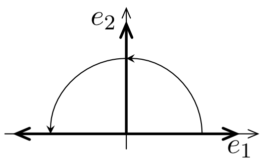
et on sait (cf. chapitre 7) que multiplier par $e^{i\theta}$ revient à effectuer une rotation d'angle θ. En conclusion, la matrice ci-dessus est la matrice associée à la rotation d'angle θ, ce qu'on aurait aussi pu trouver en travaillant sur $\mathbb{R}^2$ muni de la base canonique :



$$\begin{array}{cc} \begin{array}{cc} r(e_1) & r(e_2) \end{array} & \\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \end{array} \end{array}$$

- Par exemple, la rotation d'angle $\pi/2$, c'est-à-dire la multiplication par i, est représentée par la matrice

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



Il est capital de bien se représenter cette application linéaire géométriquement. Elle est utilisée de nombreuses fois comme exemple ou contre-exemple en algèbre linéaire.

- Donnons un dernier exemple, plus difficile (mais fréquent dans les exercices, cf. exercices 10, 15, 36, 64). Donnons la matrice canoniquement associée à

$$u : \begin{cases} \mathbb{K}_n[X] \longrightarrow \mathbb{K}_n[X] \\ P \longmapsto P(X+2) \end{cases}$$

On cherche donc :

$$A = \begin{array}{c} \begin{array}{cccccc} u(1) & u(X) & u(X)^2 & u(X)^3 & \dots & u(X)^n \end{array} \\ \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} & & \cdots & \\ A_{2,1} & A_{2,2} & A_{2,3} & & \cdots & \\ & & & & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ & & & & \ddots & \\ & & & & \cdots & \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{array}$$

On a donc une matrice de taille $(n+1)\times(n+1)$. Soit donc $(i,j) \in [\![1\,;\, n+1]\!]^2$. Le coefficient $A_{i,j}$ en ligne i et j est la coordonnée selon X^{j-1} de $u(X)^{i-1}$: il y a en effet un décalage puisque les lignes et colonnes des matrices sont numérotées à partir du rang 1 mais les polynômes à partir de la puissance 0. Or,

On voit par exemple que la troisième colonne correspond à $u(X)^2$ et que la quatrième ligne correspond à X^3.

$$u(X)^{j-1} = (X+2)^{j-1} = \sum_{k=0}^{j-1} \binom{j-1}{k} X^k \times 2^{j-1-k}$$

En prenant $k = i-1$, le coefficient de X^{i-1} est $\binom{j-1}{i-1}(2)^{j-i}$ et c'est donc la valeur de $A_{i,j}$ (et il est nul si $i > j$ donc la matrice est triangulaire supérieure), si bien que :

$$A = \begin{pmatrix} 2^0\binom{0}{0} & 2^1\binom{1}{0} & 2^2\binom{2}{0} & \cdots & 2^n\binom{n}{0} \\ 0 & 2^0\binom{1}{1} & 2^1\binom{2}{1} & \cdots & 2^{n-1}\binom{n}{1} \\ 0 & 0 & 2^0\binom{2}{2} & \cdots & 2^{n-2}\binom{n}{2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 2^0\binom{n}{n} \end{pmatrix}$$

I.3 Écriture matricielle de $y = u(x)$

Rappel : Soit $x \in E$ qu'on écrit $x = x_1 e_1 + \cdots + x_p e_p$. On définit le vecteur colonne des coordonnées de x dans la base B_E, noté $\mathrm{Mat}_{B_E}(x)$, X_{B_E} ou X si aucune confusion n'est possible, par :

$$X_{B_E} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

De plus, les applications suivantes sont des isomorphismes :

Si $y = y_1 f_1 + \cdots + y_n f_n$, on définit de même Y_{B_F} (ou Y si aucune confusion n'est possible) par

$$Y_{B_F} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

On dira indifféremment que $X \in \mathbb{K}^p$ ou que $X \in \mathscr{M}_{p,1}(\mathbb{K})$, ce qui n'a aucune importance puisque ces deux espaces vectoriels sont isomorphes (et donc on peut les identifier), et même chose pour Y.

$$\left\{ \begin{array}{ccc} E & \longrightarrow & \mathscr{M}_{p,1}(\mathbb{K}) \\ x & \longmapsto & X \end{array} \right. \qquad \text{et} \qquad \left\{ \begin{array}{ccc} F & \longrightarrow & \mathscr{M}_{n,1}(\mathbb{K}) \\ y & \longmapsto & Y \end{array} \right.$$

On pourra donc (et on ne se privera pas de le faire) identifier x et X, ainsi que y et Y.

Proposition. Soient $x \in E$, $u \in \mathscr{L}(E,F)$, $A = \mathrm{Mat}(u)$ et $y = u(x)$. Alors $Y = AX$.

Remarque : En d'autres termes, l'évaluation par une application linéaire u se traduit matriciellement par un produit. Nous verrons dans le paragraphe II.3 que la composition d'applications linéaires se traduit elle aussi matriciellement par un produit.

Démonstration. Notons $A = (A_{i,j})$. Par définition, pour tout $j \in [\![1\,;\, p]\!]$,

$$u(e_j) = \sum_{i=1}^{n} A_{i,j} f_i$$

Par conséquent :

$$
\begin{aligned}
y &= u(x) \\
&= u\left(\sum_{j=1}^{p} x_j e_j\right) \\
&= \sum_{j=1}^{p} x_j u(e_j) \\
&= \sum_{j=1}^{p}\sum_{i=1}^{n} x_j A_{i,j} f_i \\
&= \sum_{i=1}^{n}\left(\sum_{j=1}^{p} x_j A_{i,j}\right) f_i \\
&= \sum_{i=1}^{n}\left(\sum_{j=1}^{p} A_{i,j} x_j\right) f_i
\end{aligned}
$$

Par linéarité de u

Par unicité des coordonnées sur une base :

$$\forall i \in [\![1\,;\, n]\!], y_i = \sum_{j=1}^{p} A_{i,j} x_j$$

□

c'est-à-dire que le terme général de Y est le terme général du produit AX. En conclusion, $Y = AX$.

II Matrices/AL : même combat

II.1 Isomorphismes d'espaces vectoriels

Théorème. Avec les notations de la définition,

$$\varphi : \begin{cases} \mathscr{L}(E,F) \longrightarrow \mathscr{M}_{n,p}(\mathbb{K}) \\ u \longmapsto \mathrm{Mat}_{B_F,B_E}(u) \end{cases}$$

est un isomorphisme d'espaces vectoriels.

On pourra ainsi identifier les deux espaces vectoriels $\mathscr{L}(E,F)$ et $\mathscr{M}_{n,p}(\mathbb{K})$ puisqu'ils sont isomorphes.

Remarque : Puisque φ est bijective, on peut donc faire le cheminement inverse du paragraphe précédent, c'est-à-dire qu'une fois les bases B_E et B_F fixées, on pourra parler de l'application linéaire $u \in \mathscr{L}(E,F)$ représentée par une matrice $A \in \mathscr{M}_{n,p}(\mathbb{K})$, et si B_E et B_F sont les bases canoniques de E et F, on pourra parler de l'application linéaire canoniquement associée à A. En particulier, on pourra partir d'une matrice pour donner l'application linéaire associée.

Plus précisément, on pourra utiliser l'article défini « le/la » au lieu de l'article indéfini « un/une »

Exemple : L'endomorphisme de $\mathbb{K}_2[X]$ canoniquement associé à

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 4 \\ 0 & 4 & 5 \end{pmatrix}$$

est l'endomorphisme $u : \mathbb{K}_2[X] \longrightarrow \mathbb{K}_2[X]$ vérifiant $u(1) = 3X + 1$, $u(X) = 4X^2 + X$ et $u(X^2) = 5X^2 + 4X + 2$. On en déduit facilement (cf. paragraphe I.2.a) l'image d'un polynôme quelconque $P = aX^2 + bX + c \in \mathbb{K}_2[X]$.

Démonstration. • Montrons tout d'abord que φ est linéaire. Soient $(u,v) \in \mathscr{L}(E,F)^2$ et $(\lambda,\mu) \in \mathbb{K}^2$. Notons $\mathrm{Mat}(u) = (a_{i,j})$ et $\mathrm{Mat}(v) = (b_{i,j})$ (on ne précise plus les bases ni les tailles des matrices). Soit $j \in [\![1\,;\, p]\!]$. Par définition,

$$u(e_j) = a_{1,j} f_1 + \cdots a_{n,j} f_n \qquad \text{et} \qquad v(e_j) = b_{1,j} f_1 + \cdots b_{n,j} f_n$$

Alors

$$(\lambda u + \mu v)(e_j) = (\lambda a_{1,j} + \mu b_{1,j}) f_1 + \cdots + (\lambda a_{1,j} + \mu b_{1,j}) f_n$$

Ainsi, si on note $\text{Mat}(\lambda u + \mu v) = (c_{i,j})$, on a, pour tous i et j, $c_{i,j} = \lambda a_{i,j} + \mu b_{i,j}$ donc, par définition d'une somme de matrices, et de la multiplication d'une matrice par un scalaire, $\text{Mat}(\lambda u + \mu v) = \lambda \,\text{Mat}(u) + \mu \,\text{Mat}(v)$ c'est-à-dire que $\varphi(\lambda u + \mu v) = \lambda \varphi(u) + \mu \varphi(v)$: φ est linéaire.

- Une fois n'est pas coutume, nous allons montrer la bijectivité en une seule étape. Soit $A = (a_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$. Il existe une unique application linéaire $u \in \mathscr{L}(E,F)$ vérifiant :

$$\forall j \in [\![1\,;\, p]\!], \qquad u(e_j) = \sum_{i=1}^{n} a_{i,j} f_i$$

□

c'est-à-dire $\varphi(u) = A$: A admet un unique antécédent par φ : φ est bijective.

Remarque : Cet isomorphisme est utile à plus d'un titre :

- Puisque φ est bijective, on pourra « identifier » les deux espaces (voir plus haut). C'est l'idée générale de tout le chapitre : « matrices/applications linéaires : même combat ».
- Puisque φ est surjective, si $A \in \mathscr{M}_{n,p}(\mathbb{K})$, on pourra dire : « Soit $u \in \mathscr{L}(E,F)$ telle que $A = \text{Mat}_{B_F, B_E}(u)$ » (u existe bien et est unique par bijectivité). Les entiers n et p étant quelconques, c'est bien sûr aussi valable avec une matrice $B \in \mathscr{L}(F,E)$ (on dira alors : « Soit $v \in \mathscr{L}(F,E)$ telle que $B = \text{Mat}_{B_E, B_F}(v)$ »).
- Enfin, puisque φ est injective (et n et p quelconques), on pourra écrire (par exemple) : « $\text{Mat}_{B_E}(u \circ v) = \text{Mat}_{B_E}(\text{Id}_E)$ donc $u \circ v = \text{Id}_E$ ». cf. partie III.3. On pourra aussi dire : $\text{Mat}_{B_F, B_E}(u) = 0$ donc $u = 0$.

La notation Mat_{B_E} sera vue dans le paragraphe III.1.

Corollaire. $\mathscr{L}(E,F)$ est de dimension finie égale à $\dim(E) \times \dim(F)$. En particulier :

- $\mathscr{L}(E)$, l'ensemble des endomorphismes de E, est de dimension $(\dim(E))^2$.
- $\mathscr{L}(E,\mathbb{K})$, l'ensemble des formes linéaires de E dans $\mathbb{K}$, est de même dimension que E.

Bien sûr, on rappelle que E et F sont supposés de dimension finie.

II.2 Noyau, image d'une matrice

II.2.a Application des matrices à l'étude de $\text{Im}(u)$ et de $\ker(u)$

Si $u \in \mathscr{L}(\mathbb{K}^p, \mathbb{K}^n)$ et si A est la matrice canoniquement associée à u, alors on peut déduire de A des informations sur $\text{Im}(u)$ et $\text{Ker}(u)$. Plus précisément :

- $\text{Im}(u)$ est engendré par les vecteurs colonnes de A. En effet, ceux-ci sont égaux à $u(e_1), \dots, u(e_n)$ et on sait que l'image d'une base est une famille génératrice de l'image.
- Si A a une colonne nulle, disons la j-ième, alors le vecteur e_j est dans le noyau.
- Si une colonne est combinaison linéaire d'autres, on peut en déduire que la différence est dans le noyau.

Exemple : Soit $u \in \mathscr{L}(\mathbb{K}^3)$, on suppose que la matrice canoniquement associée à u est

$$\begin{array}{cc} & \begin{array}{ccc} u(e_1) & u(e_2) & u(e_3) \end{array} \\ A = & \begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 4 \\ 3 & 0 & 6 \end{pmatrix} \begin{array}{c} e_1 \\ e_2 \\ e_3 \end{array} \end{array}$$

Alors on a :

- $e_2 = (0, 1, 0) \in \operatorname{Ker}(u)$.
- $u(e_3) = 2u(e_1)$ donc $u(e_3 - 2e_1) = 0$ si bien que $e_3 - 2e_1 \in \operatorname{Ker}(u)$.
- Puisque l'image d'une famille génératrice est une famille génératrice de l'image, $\operatorname{Im}(u) = \operatorname{Vect}(u(e_1), u(e_2), u(e_3))$. Or, on sait qu'on peut « supprimer les vecteurs superflus » (c'est-à-dire ceux qui sont CL des autres), et puisque $u(e_2) = 0$ et $u(e_3) = 2u(e_1)$, alors $\operatorname{Im}(u) = \operatorname{Vect}(u(e_1)) = \operatorname{Vect}((1, 2, 3))$.

On en déduit que :

- $(1, 2, 3)$ est une base de $\operatorname{Im}(u)$.
- $\operatorname{rg}(u) = 1$ et donc $\dim \operatorname{Ker}(u) = 2$ (d'après le théorème du rang).
- $e_2 = (1, 0, 0)$ et $e_3 - 2e_1 = (-2, 0, 1)$ forment une base de $\operatorname{Ker}(u)$ car sont libres dans un espace de dimension 2.

Remarque : Dans le cas d'une application linéaire $u \in \mathscr{L}(E, F)$ quelconque avec des bases pas forcément canoniques, le raisonnement est tout-à-fait analogue, si ce n'est qu'il faut remplacer les vecteurs colonnes par les éléments de F dont ce sont les coordonnées dans la base B_F. Par exemple, dans le cas d'une matrice à 3 lignes, il faut remplacer le vecteur $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ par $1.f_1 + 2.f_2 + 3.f_3$. Mais cela se produit assez peu en pratique.

Pour le noyau : Soient $u \in \mathscr{L}(\mathbb{K}^p, \mathbb{K}^n)$ et A est la matrice canoniquement associée à u. Puisque l'écriture matricielle de $y = u(x)$ est $Y = AX$, c'est-à-dire que l'évualuation se traduit par un produit, alors on peut en déduire des informations supplémentaires sur $\operatorname{Ker} u$ à l'aide de A. Plus précisément :

$$\begin{aligned} x \in \operatorname{Ker} u & \iff u(x) = 0 \\ & \iff y = 0 \\ & \iff Y = 0 \\ & \iff AX = 0 \\ & \iff \left\{ \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,p}x_p & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{n,1}x_1 & + & a_{n,2}x_2 & + & \cdots & + & a_{n,p}x_p & = & 0 \end{array} \right. \end{aligned}$$

Nous déduirons de l'équivalence $u(x) = 0 \iff AX = 0$ que $\ker(u)$ et $\ker(A)$ sont isomorphes, cf. paragraphe II.2.b.

En d'autres termes, les lignes donnent un système d'équations du noyau. Par exemple, le noyau de $u \in \mathscr{L}\left(\mathbb{K}^3\right)$ canoniquement associé à $A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$ est l'ensemble des $(x_1, x_2, x_3) \in \mathbb{K}^3$ vérifiant

$$\left\{ \begin{array}{ccccccc} x_1 & + & 2x_2 & + & x_3 & = & 0 \\ x_1 & + & 2x_2 & + & 2x_3 & = & 0 \\ x_1 & + & 2x_2 & + & 3x_3 & = & 0 \end{array} \right.$$

II.2.b Noyau, image d'une matrice

On peut s'affranchir des applications linéaires et définir directement l'image et le noyau d'une matrice de la façon suivante :

Définition. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. On définit l'image de A par :

$$\mathrm{Im}(A) = \{AX \mid X \in \mathscr{M}_{p,1}(\mathbb{K})\}$$

Remarques :

- $\mathrm{Im}(A)$ est incluse dans $\mathscr{M}_{n,1}(\mathbb{K})/\mathbb{K}^n$ (rappelons qu'on identifie $\mathscr{M}_{n,1}(\mathbb{K})$ et $\mathbb{K}^n$). Plus précisément, on montre aisément que $\mathrm{Im}(A)$ est un sev de $\mathscr{M}_{n,1}(\mathbb{K})/\mathbb{K}^n$.
- Puisqu'on identifie $\mathscr{M}_{p,1}(\mathbb{K})$ à $\mathbb{K}^p$, on peut aussi définir l'image de A par :

$$\mathrm{Im}(A) = \{AX \mid X \in \mathbb{K}^p\}$$

- Cette définition est tout à fait analogue à l'image d'une application linéaire : c'est l'ensemble des vecteur/matrices colonnes qui s'écrivent sous la forme AX donc qui sont « images » de A.
- En effet, on a vu que les images des AL s'écrivaient à l'aide d'un produit matriciel : si u est représentée par A, l'opération $y = u(x)$ s'écrit matriciellement $Y = AX$. Puisque les $u(x)$ forment l'image de u, on définit l'image de A par l'ensemble des AX quand X décrit $\mathbb{K}^p/\mathscr{M}_{p,1}(\mathbb{K})$. De façon plus précises, « Matrices/AL : même combat » :

Proposition. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et soit $u \in \mathscr{L}(E,F)$ représentée par A. Alors $\mathrm{Im}(A)$ et $\mathrm{Im}(u)$ sont isomorphes.

Démonstration. Soit

$$\varphi : \begin{cases} \mathrm{Im}(u) \longrightarrow \mathrm{Im}(A) \\ \quad y \quad \longmapsto \quad Y \end{cases}$$ □

où on a noté Y la matrice colonne Y_{B_F} de y dans la base B_F.

- φ est évidemment linéaire.
- φ est bien à valeurs dans $\mathrm{Im}(A)$: en effet, si $y \in \mathrm{Im}(u)$, alors il existe $x \in E$ tel que $y = u(x)$ si bien que (cf. paragraphe I.3) $Y = AX \in \mathrm{Im}(A)$ (on a évidemment noté X la matrice colonne X_{B_E}).
- On sait déjà (cf. paragraphe I.3) que $y \mapsto Y$ est injective.
- Enfin, si $T \in \mathrm{Im}(A)$, alors il existe $X \in \mathscr{M}_{p,1}(\mathbb{K})/\mathbb{K}^p$ tel que $T = AX$. Par conséquent, si on note x le vecteur de E associé à X (on sait que l'application $x \mapsto X$ est bijective, cf. paragraphe I.3) et si on note $y = u(x)$, alors $Y = AX$ donc $Y = T$ c'est-à-dire que $T = \varphi(y)$: φ est surjective.

Remarques :

- En particulier ces deux espaces ont même dimension : on pourra donc identifier image de u et image de A. Nous utiliserons ce résultat dans le paragraphe V.1 pour définir le rang d'une matrice.
- Par conséquent, tous les résultats concernant $\mathrm{Im}(u)$ sont encore valables pour $\mathrm{Im}(A)$ en adaptant les termes : par exemple, l'image de A est engendrée par les vecteurs colonnes de A. Cela vient du fait que les vecteurs colonnes de A sont les $u(e_j)$ qui engendrent $\mathrm{Im}(u)$.
- On peut définir de la même façon le noyau d'une matrice :

Définition. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. On définit le noyau de A par :

$$\mathrm{Ker}(A) = \{X \in \mathscr{M}_{p,1}(\mathbb{K}) \mid AX = 0\}$$

Remarques :

- $\ker(A)$ est inclus dans $\mathscr{M}_{p,1}(\mathbb{K})/\mathbb{K}^p$ (rappelons qu'on identifie $\mathscr{M}_{p,1}(\mathbb{K})$ et $\mathbb{K}^p$). Plus précisément, on montre aisément que $\ker(A)$ est un sev de $\mathscr{M}_{p,1}(\mathbb{K})/\mathbb{K}^p$.
- On peut aussi définir le noyau de A par :

$$\mathrm{Ker}(A) = \{X \in K^p \,|\, AX = 0\}$$

- On prouve comme ci-dessus que l'application $x \mapsto X$ est un isomorphisme de $\ker(u)$ dans $\ker(A)$. On pourra donc identifier ces deux espaces, et on en déduit en particulier que ces deux espaces ont même dimension.
- Il en découle par exemple par exemple que, là aussi, les lignes fournissent un système d'équations du noyau, et que le théorème du rang est encore valable pour les matrices :

Théorème (Théorème du rang). Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Alors :

$$\dim(\ker(A)) + \dim(\mathrm{Im}(A)) = p$$

Remarques :

- Là aussi, la dimension du noyau + la dimension de l'image est égale à la dimension de **l'espace de départ**.
- La somme des dimension est égale à p, pas $n \times p$ (la taille de la matrice) ! En particulier, pour une matrice carrée, la somme de la dimension de l'image et de celle du noyau est égale à n et non pas n^2 !

II.3 Matrice d'une composée

Théorème. Soient $u \in \mathscr{L}(E, F)$ et $v \in \mathscr{L}(F, G)$.

$$\begin{tikzcd} E{,}B_E \arrow[r, "u"] \arrow[rr, bend right, "v \circ u"'] & F{,}B_F \arrow[r, "v"] & G{,}B_G \end{tikzcd}$$

Alors : $\mathrm{Mat}_{B_G,B_E}(v \circ u) = \mathrm{Mat}_{B_G,B_F}(v) \times \mathrm{Mat}_{B_F,B_E}(u)$.

En d'autres termes, quand elle est bien définie, et avec les bonnes bases, la matrice d'une composée est le produit des matrices. C'est d'ailleurs pour cela qu'on a défini le produit de matrices comme on l'a fait.

Démonstration. On note

- $A = \mathrm{Mat}(u) = (A_{i,j}) \in \mathscr{M}_{n,p}(\mathbb{K})$.
- $B = \mathrm{Mat}(v) = (B_{i,j}) \in \mathscr{M}_{q,n}(\mathbb{K})$.
- $C = \mathrm{Mat}(v \circ u) = (C_{i,j}) \in \mathscr{M}_{q,p}(\mathbb{K})$.

On n'écrit pas que $1 \leqslant i \leqslant n$ pour A, que $1 \leqslant i \leqslant q$ pour B etc. En effet, on précise la taille des matrices donc il n'y a aucune ambiguïté.

Par définition, pour tout $j \in [\![1 \,;\, p]\!]$,

$$u(e_j) = \sum_{k=1}^{n} A_{k,j} f_k \qquad \text{et} \qquad v \circ u(e_j) = \sum_{i=1}^{q} C_{i,j} g_i$$

De plus, pour tout $k \in [\![1 \,;\, n]\!]$, $v(f_k) = \displaystyle\sum_{i=1}^{q} B_{i,k} g_i$. Ainsi, on a également (par linéarité de v) :

$$\begin{aligned} v \circ u(e_j) &= v\left(\sum_{k=1}^{n} A_{k,j} f_k\right) \\ &= \sum_{k=1}^{n} A_{k,j} v(f_k) \\ &= \sum_{k=1}^{n} \sum_{i=1}^{q} A_{k,j} B_{i,k} g_i \end{aligned}$$

Finalement, $v \circ u(e_j) = \displaystyle\sum_{i=1}^{q} \left(\sum_{k=1}^{n} B_{i,k} A_{k,j} \right) g_i$. Par unicité des coordonnées sur une base :

$$\forall (i,j), \qquad C_{i,j} = \sum_{k=1}^{n} B_{i,k} A_{k,j}$$

c'est-à-dire que le terme général de C est égal au terme général de BA. En conclusion, $C = BA$. $\square$

III Cas particulier des matrices carrées

Dans cette partie, on suppose que $\dim(E) = \dim(F) = n$ et dans le paragraphe III.1, on s'intéressera exclusivement aux endomorphismes de E.

III.1 Cas particulier des endomorphismes

Un endomorphisme étant une application linéaire d'un espace vectoriel dans lui-même, les résultats des paragraphes précédent s'appliquent (d'ailleurs, on ne s'est pas privé de donner les matrices canoniquement associées à des endomorphismes). Les endomorphismes jouent cependant un rôle à part dans ce chapitre car on utilise la plupart du temps la même base au départ et à l'arrivée.

Définition. Soit $u \in \mathscr{L}(E)$. La matrice (carrée) $\text{Mat}_{B_E,B_E}(u)$ est appelée matrice associée à u dans la base B_E et est notée plus simplement $\text{Mat}_{B_E}(u)$.

Remarque : Quand nous parlerons d'endomorphismes (par exemple dans le paragraphe VI.3 quand nous parlerons de matrice semblables), il sera sous-entendu que nous prendrons la même base au départ et à l'arrivée. Cependant, parfois (par exemple lorsque nous parlerons de matrices équivalentes dans le paragraphe VI.4), nous ne prendrons pas la même base de E au départ qu'à l'arrivée, et alors, même lorsque nous manipulerons des endomorphismes, nous parlerons plutôt d'applications linéaires de E dans E (sous-entendu, donc, avec des bases pas forcément égales au départ et à l'arrivée).

Nous avons par exemple donné la matrice canoniquement associée à la dérivation sur $\mathbb{K}_n[X]$ (donc en la considérant comme un endomorphisme de $\mathbb{K}_n[X]$) mais nous avons aussi donné sa matrice dans les bases $(1, X, \dots, X^n)$ et $(1, 2X, \dots, (n+1)X^n)$ (donc on la considérant plutôt comme une AL de $\mathbb{K}_n[X]$ dans lui-même). Donnons un autre exemple d'application de ce principe.

Exemple : Si on se replace dans le cas général, c'est-à-dire qu'on prend un espace E quelconque et on considère l'identité comme un endomorphisme de E, et si on note $B_E = (e_1, \dots, e_n)$ une base de E, de cardinal n (puisque $\dim(E) = \dim(F) = n$) alors, pour tout $\lambda \in \mathbb{K}$:

$$\text{Mat}_{B_E}(\lambda \,\text{Id}_E) = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & \cdots & u(e_n) \end{array} \\ \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ \vdots \\ e_n \end{array} = \lambda \text{I}_n$$

En particulier, $\text{Mat}_{B_E}(\lambda \,\text{Id}_E) = \lambda \text{I}_n$ et ce, quelle que soit la base choisie ! Cependant, si on prend des bases de E différentes au départ et à l'arrivée, donc si on considère plutôt l'identité comme une application linéaire de E dans E, ce n'est plus forcément le cas. Par exemple, si on note toujours $B_E = (e_1, \dots, e_n)$ une base, alors $B_{E'} = (e_1, 2e_2, \dots, ne_n)$ est également une base de E (exo) et alors :

$$\mathrm{Mat}_{B_{E'},B_E}(\lambda\,\mathrm{Id}_E) = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & \cdots & u(e_n) \end{array} \\ \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda/2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda/n \end{pmatrix} \end{array} \begin{array}{c} e_1 \\ 2e_2 \\ \vdots \\ ne_n \end{array} \neq \lambda \mathrm{I}_n$$

Bon, ce cas est tout de même rare en pratique... Quand on manipulera des endomorphismes, on les traitera en général... comme des endomorphismes, c'est-à-dire qu'on prendra la même base au départ et à l'arrivée. On pourra en particulier retenir le résultat suivant (utile dans le paragraphe III.3 et plus particulièrement l'année prochaine) :

Corollaire. Si $\lambda \in \mathbb{K}$, alors la matrice de $\lambda\,\mathrm{Id}_E$ est λI_n, peu importe la base choisie. En particulier, les matrices des homothéties sont les mêmes dans toutes les bases.

Les résultats du paragraphe II.3 s'applique en particulier aux endomorphismes :

Corollaire. Avec les notations précédentes,

$$\varphi : \begin{cases} \mathscr{L}(E) & \longrightarrow & \mathscr{M}_{n,p}(\mathbb{K}) \\ u & \longmapsto & \mathrm{Mat}_{B_E}(u) \end{cases}$$

est un isomorphisme d'anneaux.

Rappelons que $\mathscr{L}(E)$ est un anneau pour les lois $+$ et $\circ$, et que $\mathscr{M}_n(\mathbb{K})$ est un anneau pour les lois $+$ et $\times$.

Démonstration. On sait déjà que c'est une bijection linéaire (cf. paragraphe II.1) donc, pour tous endomorphismes u et v, $\varphi(u+v) = \varphi(u)+\varphi(v)$. De plus, d'après ce qui précède, $\varphi(\mathrm{Id}_E) = I_n$: φ envoie le neutre de la seconde loi sur le neutre de la seconde loi. Enfin, d'après le paragraphe II.3 :

$$\begin{aligned} \varphi(u\circ v) &= \mathrm{Mat}_{B_E}(u\circ v) \\ &= \mathrm{Mat}_{B_E}(u) \times \mathrm{Mat}_{B_E}(v) \\ &= \varphi(u)\times\varphi(v) \end{aligned}$$

□

φ est donc un morphisme d'anneaux, et puisqu'il est bijectif, c'est un isomorphisme d'anneaux.

Matrices/AL : même combat !

Corollaire. Soient $u \in \mathscr{L}(E)$ et $A = \mathrm{Mat}_{B_E}(u)$. Alors $A^2 = \mathrm{Mat}_{B_E}\left(u^2\right)$, $A^3 = \mathrm{Mat}_{B_E}\left(u^3\right)$ et plus généralement, pour tout $k \in \mathbb{N}$, $A^k = \mathrm{Mat}_{B_E}\left(u^k\right)$.

Démonstration. ⇝ Exercice.

Corollaire. Soit $u \in \mathscr{L}(E)$ et soit $A = \mathrm{Mat}_{B_E}(u)$. Alors A est nilpotente si et seulement si u est nilpotent, et alors A et u ont même indice de nilpotence.

Démonstration. Supposons A d'indice de nilpotence p, c'est-à-dire que $A^p = 0$ et $A^{p-1} \neq 0$. D'après ce qui précède, $\mathrm{Mat}_{B_E}(u^p) = 0$ et $\mathrm{Mat}_{B_E}(u^{p-1}) \neq 0$. On en déduit que $u^p = 0 \neq u^{p-1}$ donc u est nilpotent d'indice p. La réciproque analogue.

III.2 De l'art de bien choisir ses bases

On a vu que la matrice d'un endomorphisme changeait lorsqu'on changeait les bases (même si, pour certains endomorphismes, les homothéties, on avait toujours la même matrice). Pour beaucoup de raisons, il peut être pratique de rechercher une base dans laquelle la matrice est « la plus simple possible » (même si c'est parfois plus facile à dire qu'à faire)'.

Nous verrons comment passer d'une matrice à une autre dans le paragraphe VI.1.

Exemples :

- Reprenons l'exemple du chapitre 29 :

$$u : \left\{ \begin{array}{ccc} \mathbb{K}^3 & \longrightarrow & \mathbb{K}^3 \\ (a,b,c) & \longmapsto & \left(\dfrac{a-b-c}{2}, 0, \dfrac{-a+b+c}{2} \right) \end{array} \right.$$

est le projecteur sur $E_1 = \text{Vect}((-1,0,1))$ parallèlement à $E_2 = \{(x,y,z) \in \mathbb{K}^3 \,|\, x - y - z = 0\}$. Sa matrice canoniquement associée est :

$$A = \begin{array}{c} \begin{array}{ccc} u(e_1) & u(e_2) & u(e_3) \end{array} \\ \begin{pmatrix} 1/2 & -1/2 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 1/2 & 1/2 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \end{array}$$

On peut trouver une base dans laquelle la matrice de u est plus simple : il est immédiat que $\varepsilon_1 = (-1,0,1)$ est une base de E_1, et on trouve avec la méthode vue dans le chapitre 29 (exo) que $\varepsilon_2 = (1,0,1)$ et $\varepsilon_3 = (0,1,-1)$ forment une base de E_2 donc, d'après la formule de concaténation des bases, $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ est une base de E. On trouve que $u(\varepsilon_1) = \varepsilon_1$ et que $u(\varepsilon_2) = u(\varepsilon_3) = 0$ donc la matrice de u dans cette base est :

$$A = \begin{array}{c} \begin{array}{ccc} u(\varepsilon_1) & u(\varepsilon_2) & u(\varepsilon_3) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{array}$$

- La forme de la matrice ci-dessus n'est pas si surprenante et peut être généralisée à n'importe quel projecteur. Rappelons que si E_1 et E_2 sont supplémentaires alors, pour tout $x \in E$, il existe $x_1 \in E_1$ et $x_2 \in E_2$ uniques tels que $x = x_1 + x_2$. L'application

$$p : \left\{ \begin{array}{ccc} E & \to & E \\ x & \mapsto & x_1 \end{array} \right.$$

est appelée projecteur sur E_1 parallèlement à E_2. De plus, $E_2 = \ker(p)$ donc, pour tout $x \in E_2, p(x) = 0$ et $\text{Im}(p) = E_1$ donc, pour tout $x \in E_1$, il existe t tel que $x = p(t)$ si bien que $p(x) = p^2(t) = p(t) = x$. Si on note $d = \dim(E_1)$, alors $\dim(E_2) = n - d$ (rappelons qu'ils sont supplémentaires). Soient $(e_1, \dots, e_d)$ une base de E_1 et $(e_{d+1}, \dots, e_n)$ une base de B_2, si bien que $B = (e_1, \dots, e_n)$ est une base de E d'après le théorème de concaténation des bases, et la matrice de p dans cette base est :

$$A = \begin{array}{cc} \begin{array}{ccccccc} u(e_1) & u(e_2) & \dots & u(e_d) & u(e_{d+1}) & \dots & u(e_n) \end{array} & \\ \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \dots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ \vdots & \vdots & & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \\ \vdots \\ e_d \\ e_{d+1} \\ \vdots \\ e_n \end{array} \end{array}$$

En d'autres termes, tout projecteur peut s'écrire (dans une base adaptée à la décomposition $E = E_1 \oplus E_2$) comme une matrice diagonale avec des 1 et des 0 sur la diagonale.

On en déduira dans le paragraphe VII.3.c que la trace d'un projecteur est égale à son rang.

- On peut donner un résultat analogue pour les symétries : en notant encore E_1 et E_2 deux sous-espaces supplémentaires, $(e_1, \dots, e_d)$ une base de E_1 et $(e_{d+1}, \dots, e_n)$ une base de E_2, alors la matrice de la symétrie par rapport à E_1 parallèlement à E_2 est :

Rappelons que $s(x) = x$ si $x \in E_1$ et $s(x) = -x$ si $x \in E_2$.

$$A = \begin{array}{cc} \begin{array}{ccccccc} u(e_1) & u(e_2) & \dots & u(e_d) & u(e_{d+1}) & \dots & u(e_n) \end{array} & \\ \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \dots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ \vdots & \vdots & & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & -1 \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \\ \vdots \\ e_r \\ e_{d+1} \\ \vdots \\ e_n \end{array} \end{array}$$

- Par exemple (cf. chapitre 29), rappelons que

$$s \colon \begin{cases} \mathbb{K}^3 & \longrightarrow \qquad\qquad \mathbb{K}^3 \\ (a,b,c) & \longmapsto \left(\dfrac{a-b-c}{2}, 0, \dfrac{-a+b+c}{2}\right) - \left(\dfrac{a+b+c}{2}, b, \dfrac{a-b+c}{2}\right) = (-b-c, -b, -a+b) \end{cases}$$

est la symétrie par rapport à E_1 parallèlement à E_2 (avec les mêmes que précédemment) donc la matrice de s dans la même base que ci-dessus :

$$A = \begin{array}{cc} \begin{array}{ccc} u(\varepsilon_1) & u(\varepsilon_2) & u(\varepsilon_3) \end{array} & \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} & \begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{array} \end{array}$$

ce qui est tout de même plus simple que sa matrice dans la base canonique :

$$A = \begin{array}{cc} \begin{array}{ccc} u(e_1) & u(e_2) & u(e_3) \end{array} & \\ \begin{pmatrix} 0 & -1 & -1 \\ 0 & -1 & 0 \\ -1 & 1 & 0 \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \\ e_3 \end{array} \end{array}$$

Remarque : ⚠ Attention, malgré les exemples précédents et tous les exemples faits au chapitre 21, il ne faut surtout pas croire que, pour tout endomorphisme u, il existe une base B telle que $\mathrm{Mat}_B(u)$ soit diagonale (un tel endomorphisme u est dit diagonalisable : l'étude des endomorphismes diagonalisables est une part importante du programme de deuxième année). Montrons par exemple que si u est un endomorphisme nilpotent non nul, il n'existe aucune base B telle que $\mathrm{Mat}_B(u)$ soit diagonale.

Soit donc $u \in \mathscr{L}(E)$ nilpotent non nul. Supposons qu'il existe une base $B = (e_1, \dots, e_n)$ telle que $\mathrm{Mat}_B(u)$ soit diagonale, c'est-à-dire qu'il existe $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ tel que

Avec le vocabulaire de deuxième année : un endomorphisme nilpotent non nul n'est pas diagonalisable.

$$\mathrm{Mat}_B(u) = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & \cdots & u(e_n) \end{array} \\ \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ \vdots \\ e_n \end{array}$$

Soit $i \in [\![1\,;n]\!]$. Alors $u(e_i) = \lambda_i e_i$. Par une récurrence immédiate, en utilisant la linéarité de u, pour tout k, $u^k(e_i) = {\lambda_i}^k e_i$. Or, u est nilpotente donc il existe p tel que $u^p = 0$ si bien que ${\lambda_i}^p e_i = 0$ et puisque e_i est non nul (c'est un élément d'une base donc d'une famille libre), $\lambda_i = 0$. i étant quelconque, $u(e_i) = 0$ pour tout i et une AL est entièrement déterminée par l'image d'une base donc $u = 0$ (ou alors on dit que $\mathrm{Mat}_B(u) = 0$ mais la matrice nulle est la même dans toutes les bases donc $u = 0$) ce qui est absurde.

Ainsi, il n'est pas possible d'écrire la matrice d'un endomorphisme nilpotent de la même façon que ci-dessus. La classification des endomorphismes nilpotents n'est pas quelque-chose de facile et nécessite un outil sophistiqué appelé tableau de Young (nous le verrons peut-être en DM). Contentons-nous du cas de la dimension 3.

Exemple : Supposons que $\dim(E) = 3$. Soit $u \in \mathscr{L}(E)$ telle que $u^3 = 0$. Donnons la matrice de u sous la forme la plus simple possible.

Rappelons (cf. chapitre 30) que si u est nilpotent en dimension n alors $u^n = 0$.

Rappelons que l'indice de nilpotence est la plus petite puissance p telle que $u^p = 0$, et on a donc $p \leqslant 3$. Séparons les cas selon la valeur de p.

- Si $p = 1$ alors $u = 0$ donc sa matrice est la matrice nulle dans toutes les bases.
- Supposons que $p = 3$ donc que $u^2 \neq 0$. Il existe alors x tel que $u^2(x) \neq 0$. Montrons que $(x, u(x), u^2(x))$ est une base de $\mathbb{K}^3$. Puisqu'on a une famille à trois éléments en dimension 3, il suffit de prouver que c'est une famille libre. Soient donc $\lambda_0, \lambda_1, \lambda_2$ tels que $\lambda_0 x + \lambda_1 u(x) + \lambda_2 u^2(x) = 0$. En composant par u^2, par linéarité, il vient : $\lambda_0 u^2(x) + \lambda_1 u^3(x) + \lambda_2 u^4(x) = 0$. Or, $u^2(x) \neq 0$ et $u^3(x) = u^4(x) = 0$ donc $\lambda_0 u^2(x) = 0$ donc $\lambda_0 = 0$. L'égalité initiale devient $\lambda_1 u(x) + \lambda_2 u^2(x) = 0$ et en composant par u, on obtient de même $\lambda_1 = 0$ et donc $\lambda_2 u^2(x) = 0$ si bien que $\lambda_2 = 0$: la famille est libre donc forme une base de E. Notons $B = (u^2(x), u(x), x)$ qui est donc aussi une base de E, la matrice de u dans cette base est donc :

Raisonnement classique ! cf. chapitre 30.

On préfère changer l'ordre car on préfère travailler avec des matrices triangulaires supérieures que des matrices triangulaires inférieures.

$$\mathrm{Mat}_B(u) = \begin{array}{c} \begin{array}{ccc} u(u^2(x)) & u(u(x)) & u(x) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ u^2(x) \\ u(x) \\ x \end{array}$$

- Supposons enfin que $p = 2$ i.e. $u^2 = 0$ mais $u \neq 0$. Le raisonnement précédent n'est plus valide puisqu'il n'existe pas de vecteur x tel que $u^2(x) \neq 0$. On peut cependant prendre $x \in E$ tel que $u(x) \neq 0$. Puisque $u(u(x)) = 0$ alors $u(x) \in \ker(u)$. Prouvons que $\ker(u)$ est de dimension 2. On sait que $\dim \ker(u) \in [\![0\,;3]\!]$.
  - ⋆ Si $\dim \ker(u) = 3$ alors $\ker(u) = E$ donc $u = 0$ ce qui est exclu.
  - ⋆ Si $\dim \ker(u) = 0$ alors $\ker(u) = \{0\}$ donc u est injectif. Or, une composition d'injections est une injection donc u^2 est une injection ce qui est absurde puisque $u^2 = 0$ (plus généralement, on montre de même qu'un endomorphisme nilpotent n'est jamais injectif).

- ⋆ Si $\dim(\ker(u)) = 1$ alors, d'après le théorème du rang, $\dim \operatorname{Im}(u) = 2$. Or, $u^2 = 0$ donc, pour tout x, $u(u(x)) = 0$ si bien que $u(x) \in \ker(u)$: en d'autres termes, $\operatorname{Im}(u) \subset \ker(u)$ ce qui est absurde avec les dimensions.

On en déduit que $\ker(u)$ est de dimension 2. Puisque $u(x) \neq 0$, on peut le compléter en base de $\ker(u)$ à l'aide du théorème de la base incomplète. On se donne donc y tel que $(u(x), y)$ soit une base de $\ker(u)$. Montrons que $(x, u(x), y)$ est une base de E. Il suffit de prouver que c'est une famille libre puisqu'on a une famille à trois éléments en dimension 3. Soient donc $\lambda_1, \lambda_2, \lambda_3$ tels que $\lambda_1 x + \lambda_2 u(x) + \lambda_3 y = 0$. En composant par u, par linéarité, et puisque $u(x)$ et y sont dans le noyau, il vient $\lambda_1 u(x) = 0$ donc $\lambda_1 = 0$ puisque $u(x) \neq 0$. Ainsi, $\lambda_2 u(x) + \lambda_3 y = 0$ mais $u(x)$ et y sont libres car forment une base du noyau donc $\lambda_2 = \lambda_3 = 0$ ce qui permet de conclure. La matrice de u dans la base $B = (u(x), x, y)$ est alors :

$$\operatorname{Mat}_B(u) = \begin{array}{c c} \begin{array}{ccc} u(u(x)) & u(x) & u(y) \end{array} & \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \begin{array}{c} u(x) \\ x \\ y \end{array} \end{array}$$

Nous dirons dans le paragraphe VII.3 que ces trois matrices (les deux ci-dessus et la matrice nulle) ne sont pas semblables car elles n'ont pas le même rang : par conséquent, tout endomorphisme en dimension 3 est représenté (dans une base bien choisie) par une seule des trois matrices précédentes.

C'est l'exercice 57 du chapitre 30.

Tant que nous sommes dans les matrices nilpotentes, rappelons le résultat admis dans le chapitre 21 :

Proposition.

- Une matrice triangulaire supérieure stricte (i.e. avec une diagonale nulle) est nilpotente d'indice inférieur ou égal à n.
- Si $k \in [\![2 \,;\, n]\!]$, alors la matrice

$$N = \begin{array}{cc} \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \end{pmatrix} & \begin{array}{c} 1 \\ 2 \\ \vdots \\ k-1 \\ \\ \\ \\ [n] \end{array} \end{array}$$

  est nilpotente d'indice k.
- L'indice d'une matrice nilpotente de taille n (pas forcément triangulaire) est inférieur ou égal à n. En d'autres termes, si $N \in \mathscr{M}_n(\mathbb{K})$ est nilpotente, alors $N^n = 0$.

C'est intuitif : les rangées de coefficients vont « monter d'un étage » à chaque fois jusqu'à disparaître. Cependant, attention, la réciproque est fausse ! Une matrice peut être nilpotente sans être triangulaire ! Par exemple, la matrice

$$N = \begin{pmatrix} 1 & j & j^2 \\ j & j^2 & 1 \\ j^2 & 1 & j \end{pmatrix}$$

est nilpotente d'indice 2 (j est ici le complexe $e^{2i\pi/3}$).

DÉMONSTRATION. • Soit $T \in \mathscr{M}_n(\mathbb{K})$ triangulaire supérieure stricte. Soit $u \in \mathscr{L}(E)$ représenté par N dans la base B_E.

$$
\begin{array}{cc}
 & \begin{array}{cccccc} u(e_1) & u(e_2) & u(e_3) & \cdots & \cdots & u(e_n) \end{array} \\
T = & \begin{pmatrix}
0 & T_{1,2} & T_{1,3} & \cdots & \cdots & T_{1,n} \\
0 & 0 & T_{2,3} & & & T_{2,n} \\
\vdots & \vdots & 0 & \ddots & & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \vdots & & \ddots & T_{n-1,n} \\
0 & 0 & 0 & \cdots & & 0
\end{pmatrix}
\begin{array}{c} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{n-1} \\ e_n \end{array}
\end{array}
$$

Par conséquent, pour tout $j \in [\![1\,;\,n]\!]$, $u(e_j) = \displaystyle\sum_{i=1}^{j-1} T_{i,j} e_i \in \operatorname{Vect}(e_1, \ldots, e_{j-1})$.

L'idée est que les coefficients vont remonter d'un rang à chaque nouvelle itération. Montrons cela rigoureusement : montrons par récurrence que pour tout $k \in [\![1\,;\,n-1]\!]$, $u^k(E) \subset \operatorname{Vect}(e_1, \ldots, e_{n-k})$. Le résultat est vrai au rang 1 puisque $\operatorname{Im}(u) \subset \operatorname{Vect}(e_1, \ldots, e_{n-1})$ (les vecteurs colonnes engendrent l'image). Soit $k \in [\![1\,;\,n-2]\!]$. Supposons qu'il soit vrai au rang k et prouvons qu'il soit vrai au rang $k-1$. Soit $x \in E$. $u^{k+1}(x) = u(u^k(x))$. Or, par hypothèse de récurrence, $u^k(x) \in \operatorname{Vect}(e_1, \ldots, e_{n-k})$ donc $u^{k+1}(x) \in \operatorname{Vect}(e_1, \ldots, e_{n-k-1})$ ce qui clôt la récurrence. En particulier, $u^{n-1}(E) \subset \operatorname{Vect}(e_1)$ et puisque $u(e_1) = 0$, $u^n(E) = \{0\}$, $u^n = 0$ donc $T^n = 0$.

L'image perd une dimension à chaque itération.

- Soit $u \in \mathscr{L}(E)$ représenté par N dans la base B_E :

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} u(e_1) & u(e_2) & u(e_3) & \cdots & u(e_k) & u(e_{k+1}) & \cdots & u(e_n) \end{array} \\
N = & \begin{pmatrix}
0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0
\end{pmatrix}
\begin{array}{c} e_1 \\ e_2 \\ \vdots \\ e_{k-1} \\ e_k \\ \\ e_{n-1} \\ e_n \end{array}
\end{array}
$$

En d'autres termes, $u(e_1) = 0$, pour tout $i \in [\![2\,;\,k]\!]$, $u(e_i) = e_{i-1}$ et si $i \geqslant k+1$, alors $u(e_i) = 0$.

On a $u^2(e_2) = u(e_1) = 0$, $u^2(e_3) = u(e_2) = e_1, \ldots, u^2(e_k) = u(e_{k-1}) = e_{k-2}$ donc :

$$
\begin{array}{cc}
 & \begin{array}{cccccccc} u^2(e_1) & u^2(e_2) & u^2(e_3) & \cdots & u^2(e_k) & u^2(e_{k+1}) & \cdots & u^2(e_n) \end{array} \\
N^2 = & \begin{pmatrix}
0 & 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & \cdots & 0 & 0
\end{pmatrix}
\begin{array}{c} e_1 \\ e_2 \\ \vdots \\ e_{k-2} \\ e_{k-1} \\ e_k \\ \\ e_{n-1} \\ e_n \end{array}
\end{array}
$$

De même, $u^3(e_3)=0, u^3(e_4)=e_1,\dots,u^3(e_k)=e_{k-3}$, et ainsi de suite : $u^{k-1}(e_k)=e_1$ donc $u^{k-1}\neq 0$, et pour tout $i\in [\![1\,;\,n]\!]$, $u^k(e_i)=0$: une application linéaire étant entièrement déterminée par l'image d'une base, $u^k=0$ ce qui permet de conclure.

- Soit $u\in \mathscr{L}(E)$ représentée par N. Dès lors, u est un endomorphisme de E nilpotent. On a vu dans le chapitre 30 que $u^n=0$ donc $N^n=0$.

Morale de l'histoire : Il peut être pratique de chercher la matrice la plus simple possible pour représenter un endomorphisme u (c'est parfois plus facile à dire qu'à faire), c'est-à-dire avec beaucoup de coefficients nuls. On cherchera donc une base $B=(e_1,\dots,e_n)$ pour laquelle les $u(e_i)$ s'expriment en fonction des e_i, par exemple une base pour laquelle on a des égalités du type $u(e_i)=\lambda_i e_i$ (pour obtenir une matrice diagonale, on en reparle l'année prochaine) ou $u(e_i)=e_j$ (pour une matrice triangulaire comme ci-dessus, mais en général on vous guide). Vaste programme, comme dirait l'autre...

III.3 Matrices inversibles

On se replace dans le cas évoqué au début de la partie III, c'est-à-dire qu'on s'intéresse aux applications linéaires $u\in\mathscr{L}(E,F)$ avec $\dim(E)=\dim(F)=n$.

Théorème. Soient $u\in\mathscr{L}(E,F)$ et $A=\mathrm{Mat}(u)$. Alors u est bijective si et seulement si A est inversible, et alors $A^{-1}=\mathrm{Mat}(u^{-1})$.

On n'a pas précisé les bases pour A et A^{-1} puisqu'elles sont largement sous-entendues : $A=\mathrm{Mat}_{B_F,B_E}(u)$ et, quand elle existe,

$$A^{-1}=\mathrm{Mat}_{B_E,B_F}(u).$$

Remarque : En d'autres termes, une application linéaire est un isomorphisme si et seulement si sa matrice associée (dans des bases quelconques) est inversible.

Démonstration.
- Supposons A inversible. Alors $AA^{-1}=A^{-1}A=\mathrm{I}_n$. Soit $v\in \mathscr{L}(E,F)$ tel que $\mathrm{Mat}(v)=A^{-1}$ (l'existence et l'unicité d'une telle application linéaire découle du théorème d'isomorphisme, voir la dernière remarque du II.1). Puisque $AA^{-1}=\mathrm{I}_n=\mathrm{Mat}_{B_F}(\mathrm{Id}_F)$ alors

$$\mathrm{Mat}_{B_F,B_E}(u)\times\mathrm{Mat}_{B_E,B_F}(v)=\mathrm{Mat}_{B_F}(\mathrm{Id}_F)$$

donc $\mathrm{Mat}_{B_F}(u\circ v)=\mathrm{Mat}_{B_F}(\mathrm{Id}_F)$ donc $v\circ u=\mathrm{Id}_F$. Puisque E et F sont de même dimension finie, u est bijective et $v=u^{-1}$ donc $A^{-1}=\mathrm{Mat}_{B_F}\left(u^{-1}\right)$.

On utilise le dernier résultat du III du chapitre 30 : si une application linéaire entre deux espaces vectoriels de même dimension finie admet un inverse à gauche ou à droite, alors elle est bijective et l'inverse à gauche ou à droite est sa bijection réciproque.

- Réciproquement, supposons u bijective. Notons $B=\mathrm{Mat}_{B_E,B_F}\left(u^{-1}\right)$. Alors

$$AB=\mathrm{Mat}(u)\times\mathrm{Mat}\left(u^{-1}\right)=\mathrm{Mat}\left(u\circ u^{-1}\right)=\mathrm{Mat}(\mathrm{Id}_F)=\mathrm{I}_n$$

De même, $BA=\mathrm{Mat}(\mathrm{Id}_E)=\mathrm{I}_n$ donc A est inversible et $B=A^{-1}$. □

Remarque : On est à présent en mesure de prouver le résultat admis au chapitre 21, à savoir : « si B est un inverse à droite ou à gauche de A, alors B est un inverse de A » (et c'est pour cela qu'on s'est retenu de l'utiliser dans la démonstration ci-dessus) : supposons donc que B est un inverse à gauche (raisonnement analogue dans l'autre cas) de A. Soit $u\in\mathscr{L}(E,F)$ tel que $A=\mathrm{Mat}(u)$. Soit $v\in\mathscr{L}(F,E)$ tel que $B=\mathrm{Mat}(v)$. Puisque $BA=\mathrm{I}_n$ alors de même que ci-dessus, $v\circ u=\mathrm{Id}_E$ donc, puisqu'on est en dimension finie, u est bijective et $v=u^{-1}$ donc A est inversible et $B=\mathrm{Mat}\left(u^{-1}\right)=\mathrm{Mat}(u)^{-1}=A^{-1}$.

Remarque : Prouvons également le résultat évoqué dans le chapitre 21, à savoir qu'une matrice de $\mathscr{M}_n(\mathbb{K})$ est un diviseur de zéro si et seulement si elle n'est pas inversible (et non nulle). On sait déjà (cf. chapitres 18 et 21) qu'un diviseur de zéro n'est pas inversible.

Rappelons qu'une matrice A est un diviseur de zéro si et seulement s'il existe $B\neq 0$ telle que $AB=0$ ou $BA=0$. En particulier, un diviseur de zéro est non nul par définition.

Prouvons la réciproque : soit $A\in\mathscr{M}_n(\mathbb{K})$ non inversible et prouvons que c'est un diviseur de zéro. Soit $u\in\mathscr{L}(\mathbb{K}^n)$ canoniquement associée à A. Alors u n'est pas bijective donc n'est pas surjective car est linéaire entre deux espaces vectoriels de même dimension finie. A n'étant pas nulle, u n'est pas l'application nulle donc $\mathrm{Im}(u)\neq\{0\}$, et puisque u n'est pas surjective, $\dim \mathrm{Im}(u)\leqslant n-1$. Soit S un supplémentaire de $\mathrm{Im}(u)$ (possible car on est en dimension finie). Soit v l'unique AL nulle sur $\mathrm{Im}(u)$ et égale à l'identité sur S. Alors v n'est pas l'application nulle, pourtant $v\circ u$ est la fonction nulle : en effet, pour tout $x\in\mathbb{K}^n$,

cf. chapitre 29 : une AL est entièrement déterminée par les valeurs prises sur deux sous-espaces vectoriels supplémentaires.

$u(x) \in \mathrm{Im}(u)$ donc $v(u(x)) = 0$. Si on note B la matrice canoniquement associée à v, alors $B \neq 0$ mais

$$\begin{aligned} BA &= \mathrm{Mat}(v) \times \mathrm{Mat}(u) \\ &= \mathrm{Mat}(v \circ u) \\ &= 0 \end{aligned}$$

c'est-à-dire que A est un diviseur de zéro.

Corollaire. Une matrice carrée A est inversible si et seulement si ses vecteurs colonnes forment une base de $\mathbb{K}^n$ si et seulement si ses vecteurs lignes forment une base de $\mathbb{K}^n$.

Démonstration. Soit $u \in \mathscr{L}(\mathbb{K}^n, \mathbb{K}^n)$ canoniquement associée à A.

$$\begin{aligned} A \text{ est inversible} &\iff u \text{ est bijective} \\ &\iff u \text{ envoie une base sur une base} \\ &\iff (u(e_1), \dots, u(e_n)) \text{ est une base de } \mathbb{K}^n \end{aligned}$$

et donc A est inversible si et seulement si ses vecteurs colonnes forment une base de $\mathbb{K}^n$ puisque, par définition de A, ses vecteurs colonnes sont $(u(e_1), \dots, u(e_n))$. Enfin,

$$\begin{aligned} A \text{ est inversible} &\iff A^\top \text{ est inversible} \\ &\iff \text{les colonnes de } A^\top \text{ forment une base de } \mathbb{K}^n, \end{aligned}$$

ce qui permet de conclure puisque les vecteurs colonnes de $A^\top$ sont les vecteurs lignes de A. $\square$

Exemple : Considérons

$$u : \left\{ \begin{array}{ccc} \mathbb{K}_n[X] & \longrightarrow & \mathbb{K}_n[X] \\ P & \longmapsto & P - P' \end{array} \right.$$

L'application u est bien linéaire et à valeurs dans $\mathbb{K}_n[X]$: en effet, si P est constant alors $P' = 0$ donc $P - P' = P \in \mathbb{K}_n[X]$, et si P n'est pas constant, alors $\deg(P') = \deg(P) - 1$, et donc $\deg u(P) = \deg P$ car $u(P)$ est la différence de deux polynômes de degrés distincts, et en particulier $u(P) \in \mathbb{K}_n[X]$. Montrons de deux manières différentes que u est bijective.

- Première méthode : « point de vue application linéaire ». Soit $P \in \mathrm{Ker}(u)$. Alors $P = P'$, ce qui n'est possible que si $P = 0$ (voir ci-contre). Ainsi, $\mathrm{Ker}(u) = \{0\}$, u est une application linéaire injective entre deux espaces vectoriels de même dimension finie donc est bijective.
- Deuxième méthode : « point de vue matriciel ». La matrice canoniquement associée à u est :

$$A = \begin{array}{c} \begin{array}{cccccc} u(1) & u(X) & u(X^2) & \cdots & \cdots & u(X^n) \end{array} \\ \begin{pmatrix} 1 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & -2 & & & 0 \\ \vdots & \vdots & 1 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & -n \\ 0 & 0 & 0 & \cdots & & 1 \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ X^2 \\ \vdots \\ X^{n-1} \\ X^n \end{array}$$

Puisque A est triangulaire supérieure avec des coefficients diagonaux tous non nuls, elle est inversible : u est bijective.

Si P est constant alors $P' = 0$ donc $P = P'$ si et seulement si $P = 0$. Si P n'est pas constant, alors $\deg(P) \neq \deg(P')$ donc $P \neq P'$.

Théorème. Soient $A \in \mathscr{M}_{n,p}(\mathbb{K})$, $B \in \mathrm{GL}_n(\mathbb{K})$ et $C \in \mathrm{GL}_p(\mathbb{K})$.

- $\ker(BA) = \ker(A)$.
- $\mathrm{Im}(AC) = \mathrm{Im}(A)$.

Démonstration. Soient $u \in \mathscr{L}(\mathbb{K}^p, \mathbb{K}^n)$, $g \in \mathscr{L}(\mathbb{K}^n)$ et $h \in \mathscr{L}(\mathbb{K}^p)$ canoniquement associées à A, B et C respectivement. B et C étant inversibles, g et h sont bijectives. D'après le chapitre 30, $\ker(g \circ u) = \ker(u)$ et $\mathrm{Im}(u \circ h) = \mathrm{Im}(u)$. Or, si on note

$$\varphi : \begin{cases} \mathbb{K}^n \longrightarrow \mathscr{M}_{n,1}(\mathbb{K}) \\ \quad y \longmapsto \quad Y \end{cases} \qquad \square$$

alors on sait (cf. paragraphe II.2.b) que $\varphi(\mathrm{Im}(u)) = \mathrm{Im}(A)$ et que $\varphi(\mathrm{Im}(u \circ h)) = \mathrm{Im}(AC)$. D'où l'égalité des images. Le cas des noyaux est analogue.

Corollaire. Les opérations élémentaires sur les lignes (respectivement colonnes) préservent le noyau (respectivement l'image).

Démonstration. Découle du fait (cf. chapitre 21) qu'effectuer une opération élémentaire sur les lignes (respectivement colonnes) se traduit matriciellement par une multiplication à gauche (respectivement à droite) par une matrice inversible. Il suffit ensuite d'appliquer le théorème précédent.

III.4 De l'art de bien choisir ses espaces

Depuis le début du chapitre (sauf quand on a donné des exemples explicites), E et F sont quelconques. En d'autres termes, si $A \in \mathscr{M}_{n,p}(\mathbb{K})$, alors pour tout choix de E et F « qui ont les bonnes dimensions » (et pour tout choix de bases B_E et B_F), il existe une application linéaire $u \in \mathscr{L}(E, F)$ unique telle que $A = \mathrm{Mat}_{B_F, B_E}(u)$. Par exemple, si

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \mathscr{M}_3(\mathbb{K}),$$

alors A est à la fois la matrice canoniquement associée

- à $u \in \mathscr{L}(\mathbb{K}_2[X])$ définie par $u(1) = 1, u(X) = 2 + 3X$ et $u(X^2) = 4 + 5X + 6X^2$. En effet, on peut voir A comme la matrice suivante :

$$A = \begin{array}{c} \begin{array}{ccc} u(1) & u(X) & u(X^2) \end{array} \\ \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 5 \\ 0 & 0 & 6 \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ X^2 \end{array}$$

- à $\tilde{u} \in \mathscr{L}(\mathbb{K}^3)$ définie par $\tilde{u}(e_1) = (1, 0, 0), \tilde{u}(e_2) = (2, 3, 0)$ et $\tilde{u}(e_3) = (4, 5, 6)$. En effet, on peut également voir A comme la matrice suivante :

$$A = \begin{array}{c} \begin{array}{ccc} \tilde{u}(e_1) & \tilde{u}(e_2) & \tilde{u}(e_3) \end{array} \\ \begin{pmatrix} 1 & 2 & 4 \\ 0 & 3 & 5 \\ 0 & 0 & 6 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \end{array}$$

Par conséquent, si on veut montrer qu'une matrice A est inversible, il suffit d'exhiber **une** application linéaire bijective telle que $A = \mathrm{Mat}(u)$ et c'est parfois plus facile dans certains espaces que dans d'autres (typiquement dans des espaces de polynômes car on dispose d'outils puissants pour montrer qu'une famille est libre ou qu'un polynôme est nul). Donnons trois applications de ce principe.

Remarque : Démontrons justement le résultat admis au chapitre 21, c'est-à-dire qu'une matrice triangulaire supérieure est inversible si et seulement si ses coefficients diagonaux sont tous non nuls, et qu'alors son inverse est également triangulaire supérieure. On se donne donc $A \in \mathscr{M}_n(\mathbb{K})$ triangulaire supérieure. Soit $u \in \mathscr{L}(\mathbb{K}_{n-1}[X])$ canoniquement associée à A (voir l'exemple de u ci-contre, même si évidemment ici on est dans le cas général) c'est-à-dire que :

$$\forall j \in [\![1 \,;\, n]\!], \qquad u\left(X^{j-1}\right) = A_{1,j} + A_{2,j}X + \cdots + A_{j,j}X^{j-1}$$

(on s'arrête à X^{j-1} car A est triangulaire supérieure).

- Supposons qu'il existe j tel que $A_{j,j} = 0$ et montrons que A n'est pas inversible. Pour tout $k \in [\![0 \,;\, j-1]\!]$, $\deg(X^k) \leqslant j-2$ donc $u(1), \dots, u(X^{j-1})$ appartiennent à $\mathbb{K}_{j-2}[X]$. Or, j vecteurs dans un espace de dimension $j-1$ forment une famille liée. Ainsi u envoie une famille libre sur une famille liée donc n'est pas injective, donc pas bijective. Ainsi, A n'est pas inversible.
- Réciproquement, si tous les coefficients diagonaux de A sont non nuls, alors la famille de polynômes $u(1), u(X), \dots, u(X^{n-1})$ est échelonnée en degré donc est une famille libre à n éléments dans un espace de dimension n, donc c'est une base. L'application linéaire u envoie une base sur une base donc est bijective, ce qui implique que A est inversible.

Par exemple, si

$$A = \begin{pmatrix} 1 & 2 & 4 & 6 \\ 0 & 3 & 5 & 7 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

alors $j = 3$, $u(1) = 1$, $u(X) = 2+3X$ et $u(X^2) = 4+5X$, tous trois de degré inférieur ou égal à 1.

Supposons enfin u bijective et montrons que A^{-1} est également triangulaire supérieure. Soit P non nul. Notons $P = \displaystyle\sum_{k=0}^{d} c_k X^k \in \mathbb{K}_{n-1}[X]$ avec $c_d \neq 0$ (et donc $d = \deg(P)$). Par linéarité de u,

$$u(P) = \sum_{k=0}^{d} c_k u\left(X^k\right) = c_d u\left(X^d\right) + \sum_{k=0}^{d-1} c_k u(X^k).$$

Or, $c_d \neq 0$ et on a montré ci-dessus que pour tout k, $\deg(u(X^k)) = k$ donc $u\left(X^d\right)$ est de degré d et la somme est une somme de termes de degrés inférieurs ou égaux à $d-1$ donc est de degré inférieur ou égal à $d-1$. En d'autres termes, $\deg(u(P)) = \deg(P)$ (vrai aussi si P est nul par linéarité de u). Dès lors, $\deg\left(u^{-1}(P)\right) = \deg(P)$ (on applique cette relation à $u^{-1}(P)$). En particulier, pour tout $j \in [\![1 \,;\, n]\!]$, $\deg u^{-1}\left(X^{j-1}\right) = j-1$ et cela implique que $(A^{-1})_{i,j} = 0$ dès que $i > j$ donc A^{-1} est triangulaire supérieure.

Exemple : Prouver que la matrice $A \in \mathscr{M}_{n+1}(\mathbb{R})$ de terme général $\dbinom{j-1}{i-1}$ est inversible et donner son inverse.

Comme dans le paragraphe I.2.b, on pense à un endomorphisme polynomial. Plus précisément, A peut être vue comme la matrice canoniquement associée à l'endomorphisme de $\mathbb{R}_n[X]$ défini par $\varphi(P) = P(X+1)$: on a bien, pour tout $j \in [\![1 \,;\, n+1]\!]$,

Les coefficients binomiaux font penser à un binôme de Newton.

$$\varphi(X^{j-1}) = (X+1)^{j-1} = \sum_{k=0}^{j-1} \binom{j-1}{k} X^k$$

Attention au décalage (cf. paragraphe I.2.b) : les puissances commencent en 0, les lignes et les colonnes au rang 1, et donc le coefficient en position (i,j) de la matrice canoniquement associée à φ est le coefficient selon X^{i-1} de $\varphi(X^{j-1})$ c'est-à-dire $\binom{j-1}{i-1}$, et donc cette matrice est bien égale à A. Puisque φ est bijectif de réciproque $\varphi^{-1} : P \mapsto P(X-1)$, alors A est inversible d'inverse la matrice canoniquement associée à φ^{-1}, et on montre comme au paragraphe I.2.b que cette matrice est la matrice de terme général $\binom{j-1}{i-1}(-1)^{i-j}$:

$$A^{-1} = \begin{pmatrix} \binom{0}{0} & -\binom{1}{0} & \binom{2}{0} & \cdots & (-1)^n\binom{n}{0} \\ 0 & \binom{1}{1} & -\binom{2}{1} & \cdots & (-1)^{n-1}\binom{n}{1} \\ 0 & 0 & \binom{2}{2} & \cdots & (-1)^{n-2}\binom{n}{2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \binom{n}{n} \end{pmatrix}$$

Exemple : Inversibilité de la matrice de Vandermonde. Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$. On définit

$$VDM = \begin{pmatrix} 1 & \lambda_1 & {\lambda_1}^2 & \cdots & {\lambda_1}^{n-1} \\ 1 & \lambda_2 & {\lambda_2}^2 & \cdots & {\lambda_2}^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & {\lambda_n}^2 & \cdots & {\lambda_n}^{n-1} \end{pmatrix} \in \mathscr{M}_n(\mathbb{K})$$

Donnons une CNS pour que VDM soit inversible.

- Si deux scalaires parmi les λ_i, $1 \leqslant i \leqslant n$, sont égaux, alors VDM a deux lignes égales donc n'est pas inversible.
- Supposons que les λ_i, $1 \leqslant i \leqslant n$, soient tous distincts. Soit

$$u : \left\{ \begin{array}{ccc} \mathbb{K}_{n-1}[X] & \longrightarrow & \mathbb{K}^n \\ P & \longmapsto & (P(\lambda_1), \dots, P(\lambda_n)) \end{array} \right.$$

L'application u est linéaire. Donnons la matrice canoniquement associée à u.

$$\mathrm{Mat}(u) = \begin{array}{c} \begin{array}{ccccc} u(1) & u(X) & u(X^2) & \cdots & u(X^{n-1}) \end{array} \\ \begin{pmatrix} 1 & \lambda_1 & {\lambda_1}^2 & \cdots & {\lambda_1}^{n-1} \\ 1 & \lambda_2 & {\lambda_2}^2 & \cdots & {\lambda_2}^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \lambda_n & {\lambda_n}^2 & \cdots & {\lambda_n}^{n-1} \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ \vdots \\ \vdots \\ e_n \end{array} = VDM.$$

Soit $P \in \mathrm{Ker}(u)$. Alors $P(\lambda_1) = \cdots = P(\lambda_n) = 0$. P admet donc au moins n racines distinctes. Or, $P \in \mathbb{K}_{n-1}[X]$ donc $\deg(P) \leqslant n-1$. Il en découle que $P = 0$ donc $\mathrm{Ker}(u) = \{0\}$: u est injective. De plus, $\dim(\mathbb{K}_{n-1}[X]) = \dim(\mathbb{K}^n) = n$. Par conséquent u est une application linéaire injective entre deux espaces vectoriels de même dimension finie : u est bijective donc VDM est inversible.

On a déjà prouvé la bijectivité de u dans le chapitre 30. Cela permet (cf. chapitre 30) de prouver l'existence et l'unicité des polynômes de Lagrange. Nous reverrons la matrice VDM dans le chapitre 33.

Conclusion : VDM est inversible si et seulement si les λ_i, $1 \leqslant i \leqslant n$, sont deux à deux distincts.

III.5 Critère du noyau

Théorème (critère du noyau). Une matrice carrée A est inversible si et seulement son noyau est nul.

Encore une fois : « matrices/AL, même combat ».

Remarque : En d'autres termes, une matrice $A \in \mathscr{M}_n(\mathbb{K})$ est inversible si et seulement si le vecteur colonne nul est la seule solution du système $AX = 0$.

Ci-contre, X et 0 sont des vecteurs colonnes à n coordonnées.

Démonstration. Soit $u \in \mathscr{L}(E, F)$ représentée par A. Rappelons (cf. paragraphe II.2.b) que $\ker(u)$ et $\ker(A)$ sont isomorphes. Par conséquent :

$$\begin{aligned} \ker(A) = \{0\} & \iff \ker(u) = \{0\} \\ & \iff u \text{ injective} \\ & \iff u \text{ bijective} \\ & \iff A \text{ inversible} \end{aligned}$$

Car u est une AL entre deux espaces vectoriels de même dimension finie.

□

Activité : critère d'Hadamard. Soit $A = (a_{i,j}) \in \mathscr{M}_n(\mathbb{K})$ telle que :

$$\forall i \in [\![1 \,;\, n]\!], \qquad |a_{i,i}| > \sum_{j \neq i} |a_{i,j}|.$$

Montrons que A est inversible.

Soit $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathscr{M}_{n,1}(\mathbb{K})$ tel que $AX = 0$. Montrons que X est nul.

Soit i_0 tel que $|x_{i_0}| = \max\limits_{1 \leqslant j \leqslant n} |x_j|$ (c'est-à-dire que x_{i_0} est la coordonnée de X dont la valeur absolue est maximale). Puisque $AX = 0$ (vecteur colonne nul), alors, en particulier, la coordonnée en ligne i_0 est nulle, donc $\displaystyle\sum_{j=1}^{n} a_{i_0,j} \times x_j = 0$, d'où $a_{i_0,i_0} x_{i_0} = -\displaystyle\sum_{j \neq i_0} a_{i_0,j} \times x_j$. En prenant la valeur absolue et en appliquant l'inégalité triangulaire, il vient :

$$|a_{i_0,i_0} x_{i_0}| \leqslant \sum_{j \neq i_0} |a_{i_0,j} \times x_j|$$

Or, par définition de i_0, on a $|x_j| \leqslant |x_{i_0}|$ pour tout $j \in [\![1 \,;\, n]\!]$ donc

$$|a_{i_0,i_0} x_{i_0}| \leqslant |x_{i_0}| \sum_{j \neq i_0} |a_{i_0,j}|$$

Si $|x_{i_0}| \neq 0$, on a $|a_{i_0,i_0}| \leqslant \displaystyle\sum_{j \neq i_0} |a_{i_0,j}|$ ce qui est absurde. Ainsi, $|x_{i_0}| = 0$ donc $|x_j| = 0$ pour tout j, donc $X = 0$: A est inversible.

En d'autres termes, chaque coefficient diagonal est en valeur absolue (ou en module si on est sur $\mathbb{C}$) strictement supérieur à la somme des modules des autres termes de la ligne (on dit que A est à diagonale strictement dominante). Par exemple,

$$\begin{pmatrix} -6 & 3 & 2 \\ -13 & -19 & 2 \\ 0 & -1 & 3 \end{pmatrix}$$

est inversible.

IV Bilan

Pour les matrices quelconques (pas forcément carrées) avec $\dim E = p, \dim F = n$ et $\dim G = q$:

Matrices	**Applications linéaires**
$A = \operatorname{Mat}_{B_F,B_E}(u)$	$u \in \mathscr{L}(E,F)$
$B = \operatorname{Mat}_{B_E,B_F}(v)$	$v \in \mathscr{L}(F,E)$
BA	$v \circ u$
$X_{B_E} \in \mathscr{M}_{p,1}(\mathbb{R})$ (ou $\mathbb{R}^p$)	$x \in E$
$Y = AX$	$y = u(x)$
$Y \in \operatorname{Im}(A)$	$y \in \operatorname{Im}(u)$
$X \in \operatorname{Ker} A$ ie $AX = 0$	$x \in \operatorname{Ker} u$

Pour les matrices carrées avec $\dim E = \dim F = n$ (et à partir de la troisième ligne, on suppose que $u \in \mathscr{L}(E)$ et que les bases de départ et d'arrivée sont les mêmes).

Matrices	**Applications linéaires**
$A \in \mathscr{M}_n(\mathbb{R})$ inversible	$u \in \mathscr{L}(E,F)$ bijective
A^{-1} (si A est inversible)	u^{-1} (si u est bijective)
$A = \operatorname{Mat}_{B_E}(u)$	$u \in \mathscr{L}(E)$
A^2	u^2
A^k	u^k
λI_n	$\lambda \operatorname{Id}_E$

V Rang(s)

On se replace dans le cas général, c'est-à-dire $\dim E = p$ et $\dim F = n$.

V.1 Rang d'une matrice

Proposition. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Toutes les applications linéaires représentées par A ont le même rang.

Démonstration. Soit u une application linéaire représentée par A. Alors :

$$\begin{aligned} \mathrm{rg}(u) &= \dim \mathrm{Im}(u) \\ &= \dim \mathrm{Im}(A) \end{aligned}$$

□

puisque $\mathrm{Im}(A)$ et $\mathrm{Im}(u)$ sont isomorphes (cf. paragraphe II.3.b). En d'autres termes, toutes les applications linéaires représentées par A ont un rang égal à $\dim \mathrm{Im}(A)$ et en particulier ont le même rang.

Définition. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. On appelle rang de A, noté $\mathrm{rg}(A)$, le rang de toute application linéaire représentée par A.

Matrices/AL : même combat.

Remarque : En d'autres termes, par définition, $\mathrm{rg}(\mathrm{Mat}(u)) = \mathrm{rg}(u) = \dim \mathrm{Im}(u)$ et cela ne dépend pas des espaces vectoriels E et F ni des bases choisies.

Proposition. Soient $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $u \in \mathscr{L}(E,F)$ tel que $A = \mathrm{Mat}(u)$.

- $\mathrm{rg}(A) = \dim(\mathrm{Im}(A))$
- $\mathrm{rg}(A) + \dim(\ker(A)) = p$ **(Théorème du rang)**
- $\mathrm{rg}(A) = 0 \iff A = 0$
- $\mathrm{rg}(A) \leqslant n$ $(= \dim F)$ avec égalité si et seulement si u est surjective.
- $\mathrm{rg}(A) \leqslant p$ $(= \dim E)$ avec égalité si et seulement si u est injective.

Démonstration. On rappelle que $\mathrm{rg}(A) = \mathrm{rg}(u) = \dim \mathrm{Im}(u)$.

- On sait (cf. paragraphe II.3.b) que $\mathrm{Im}(A)$ et $\mathrm{Im}(u)$ ont même dimension. Or, $\mathrm{rg}(A) = \mathrm{rg}(u) = \dim(\mathrm{Im}(u))$ ce qui permet de conclure.
- Découle du point précédent et du théorème du rang vu au paragraphe II.3.b.
- $\mathrm{rg}(A) = 0 \iff \mathrm{Im}(u) = \{0\} \iff u = 0$
- $\mathrm{rg}(A) = \dim \mathrm{Im}(u) \leqslant \dim F$, car $\mathrm{Im}(u) \subset F$, avec égalité si et seulement si $\mathrm{Im}(u) = F$ c'est-à-dire si u est surjective.
- $\mathrm{rg}(A) = \mathrm{rg}(u) = \dim E - \dim \mathrm{Ker}\, u$ d'après le théorème du rang, ce qui permet de conclure. □

Remarques :

- En particulier, si $A \in \mathscr{M}_{n,p}(\mathbb{K})$ alors $\mathrm{rg}(A) \leqslant \min(n,p)$, c'est-à-dire que le rang d'une matrice est inférieur à son nombre de lignes et à son nombre de colonnes. Par exemple, on peut affirmer directement que $\mathrm{rg}\begin{pmatrix} 1 & -2 & -1 & 3 & 5 \end{pmatrix} = 1$ puisque son rang est inférieur ou égal à 1 et que ce n'est pas la matrice nulle.
- Comme dit dans le paragraphe II.3.b, quand on applique le théorème du rang à une matrice, la somme des dimensions du noyau et de l'image est égale à la dimension de l'espace de départ (i.e. le nombre de colonnes), pas à la taille de la matrice ! En particulier, pour une matrice A carrée de taille n, $\dim(\ker(A)) + \dim(\mathrm{Im}(A)) = n$ et non pas n^2 !

Nous avons vu ce résultat pour les applications linéaires dans le chapitre 30, c'est-à-dire que si $u \in \mathscr{L}(E,F)$, alors $\mathrm{rg}(u) \leqslant \min(\dim(E), \dim(F))$. Encore une fois : Matrices/AL, même combat !

Corollaire. Soit $A \in \mathscr{M}_n(\mathbb{K})$. Alors $\text{rg}(A) \leqslant n$ et, si u est une application linéaire représentée par A :

$$\begin{aligned}\text{rg}(A) = n \quad &\iff \quad u \text{ est bijective}\\ &\iff \quad u \text{ est surjective}\\ &\iff \quad u \text{ est injective}\\ &\iff \quad A \text{ est inversible}\end{aligned}$$

Nous compléterons cette suite d'équivalences dans le paragraphe V.5.

Démonstration. Il suffit d'appliquer la proposition précédente et de se souvenir que u est linéaire entre deux espaces de même dimension finie (puisque A est carrée) donc est injective si et seulement si elle est surjective si et seulement si elle est bijective.

Proposition. On ne change pas le rang d'une matrice en multipliant (à droite ou à gauche) par une matrice inversible.

Encore une fois : Matrices/AL, même combat.

Démonstration. Soient $A \in \mathscr{M}_{n,p}(\mathbb{K})$, $B \in \text{GL}_n(\mathbb{K})$ et $C \in \text{GL}_p(\mathbb{K})$. Soient $u \in \mathscr{L}(\mathbb{K}^p, \mathbb{K}^n)$, $g \in \mathscr{L}(\mathbb{K}^n)$ et $h \in \mathscr{L}(\mathbb{K}^p)$ canoniquement associées à A, B et C respectivement. B et C étant inversibles, g et h sont bijectives. D'après le chapitre 30, $\text{rg}(g \circ u) = \text{rg}(u)$ et $\text{rg}(u \circ h) = \text{rg}(u)$. Or, $A = \text{Mat}(u)$ et la composition se traduit matriciellement par le produit donc $BA = \text{Mat}(g \circ u)$ et $AC = \text{Mat}(u \circ h)$. Dès lors, $\text{rg}(u) = \text{rg}(A)$, $\text{rg}(u \circ h) = \text{rg}(BA)$ et $\text{rg}(u \circ g) = \text{rg}(AC)$ donc $\text{rg}(A) = \text{rg}(AC) = \text{rg}(BA)$.

Corollaire. Les opérations élémentaires (sur les lignes et les colonnes) préservent le rang.

Démonstration. Il suffit de se souvenir (cf. chapitre 21) qu'effectuer une opération élémentaire sur les lignes (respectivement les colonnes) se traduit matriciellement par une multiplication à gauche (respectivement à droite) par une matrice inversible.

Remarque : Comme pour les applications linéaires, dans le cas où on multiplie par une matrice quelconque, on a tout de même le résultat suivant :

Matrices/AL : même combat.

Proposition. Soient $A \in \mathscr{M}_{n,p}(\mathbb{K})$ et $B \in \mathscr{M}_{p,q}(\mathbb{K})$. Alors $\text{rg}(AB) \leqslant \min(\text{rg}(A), \text{rg}(B))$.

Démonstration. Découle du résultat analogue pour les AL. Les détails sont laissés en exercice.

On peut évidemment généraliser à un produit d'un nombre quelconque de matrices : le rang d'un produit est inférieur au minimum des rangs (cf. exercice 19 par exemple).

V.2 Et en pratique ?

On sait que le rang d'une matrice carrée de taille n inversible vaut n, mais on n'a aucune information sur les matrices non inversibles (à part qu'il est inférieur ou égal à $n-1$ si la matrice est carrée, qu'il est nul si et seulement si la matrice est nulle, et qu'il est inférieur à $\min(n,p)$ dans le cas général). Il nous faut donc un moyen de donner le rang explicitement.

En pratique, on donne le rang d'une matrice en se ramenant, à l'aide d'opérations élémentaires, à une matrice dont on connaît le rang : puisque les opérations élémentaires préservent le rang, le rang de la matrice initiale sera le même que celui de la matrice finale. Il nous faut pour cela un type de matrices « de référence » dont on peut donner le rang directement.

Définition. Soit $A \in \mathscr{M}_{n,k}(\mathbb{K})$. On dit que A est échelonnée s'il existe $r \leqslant n$ tel que

- Pour tout $i \in [\![1 ; r]\!]$, $a_{i,i} \neq 0$ et pour tout $j < i$, $a_{i,j} = 0$.
- Pour tout $i \geqslant r+1$, la ligne L_i est nulle.

En d'autres termes, une matrice échelonnée est une matrice du type :

$$\begin{pmatrix} a_{1,1} & * & * & * & \cdots & * \\ 0 & a_{2,2} & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & \cdots & * \\ \vdots & \vdots & & a_{r,r} & \cdots & * \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}$$

avec les $a_{i,i}$, $1 \leqslant i \leqslant r$, non nuls et les coefficients représentés sous forme d'étoiles étant quelconques (nuls ou non).

Exemple : La matrice suivante est échelonnée :

$$\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Pour faire simple, une matrice est échelonnée quand :

- ses coefficients diagonaux sont non nuls jusqu'à un certain point.
- tous les termes « à droite » de la partie non nulle de la diagonale sont quelconques.
- tous les termes « à gauche et sous » la partie non nulle de la diagonale sont nuls.

Théorème. Le rang d'une matrice échelonnée A est égal au nombre de lignes non nulles de A.

On peut aussi dire que le rang d'une matrice échelonnée est égal au nombre de coefficients diagonaux non nuls de cette matrice.

DÉMONSTRATION. On reprend les notations de la définition (c'est-à-dire que A a r lignes non nulles). Soit $u \in \mathscr{L}(\mathbb{K}_{p-1}[X], \mathbb{K}_{n-1}[X])$ canoniquement associée à A. On a donc :

$$A = \begin{array}{c} \begin{array}{cccccc} u(1) & u(X) & \dots & u(X^{r-1}) & \dots & u(X^{p-1}) \end{array} \\ \begin{pmatrix} a_{1,1} & * & * & * & \cdots & * \\ 0 & a_{2,2} & * & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots & \cdots & * \\ \vdots & \vdots & & a_{r,r} & \cdots & * \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ \vdots \\ X^{r-1} \\ X^r \\ \vdots \\ X^{n-1} \end{array}$$

Pour tout $k \leqslant r-1$, $\deg(X^k) = k$ donc $u(1), \dots, u(X^{r-1})$ est échelonnée en degré donc libre, et puisque c'est une famille libre à r éléments de $\mathbb{K}_{r-1}[X]$, c'est une base de $\mathbb{K}_{r-1}[X]$. De plus, si $k \geqslant r$, $u(X^k) \in \text{Vect}(1, X, \dots, X^{r-1}) = \mathbb{K}_{r-1}[X]$ donc

$$\text{Im}(u) = \text{Vect}(u(1), \dots, u(X^{p-1})) = \text{Vect}(u(1), \dots, u(X^{r-1})).$$

en supprimant les vecteurs superflus donc $\text{Im}(u) = \mathbb{K}_{r-1}[X]$, si bien que $\text{rg}(A) = \text{rg}(u) = r$. $\square$

De l'art de bien choisir ses espaces...

Exemple : $\text{rg}\begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix} = 2.$

Conclusion : si on veut donner le rang d'une matrice, on se ramène à une matrice échelonnée à l'aide d'opérations élémentaires sur les lignes ou les colonnes

Ici on a le droit de mélanger les deux, contrairement au calcul de l'inverse, cf. chapitre 21.

Exemple :

$$\begin{aligned} \text{rg}\begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 3 \\ 3 & 1 & 5 \end{pmatrix} &= \text{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & -2 & 2 \end{pmatrix} \quad \begin{matrix} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 3L_1 \end{matrix} \\ &= \text{rg}\begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad L_3 \leftarrow L_3 - 2L_2 \\ &= 2. \end{aligned}$$

Exemple : Soit $\lambda \in \mathbb{K}$. Donnons, en fonction de λ, le rang des matrices

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 3 & 2 & -1 & 3 \\ \lambda & 3 & -2 & 0 \\ -1 & 0 & -4 & 3 \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} 1 & 1 & -1 & 2 \\ \lambda & 1 & 1 & 1 \\ 1 & -1 & 3 & -3 \\ 4 & 2 & 0 & \lambda \end{pmatrix}$$

D'abord A (on commence par échanger C_1 et C_4 car « on cherche à repousser λ le plus loin possible » pour simplifier les calculs) :

$$\begin{aligned}
\operatorname{rg}(A) &= \operatorname{rg}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 3 & 2 & -1 & 3 \\ 0 & 3 & -2 & \lambda \\ 3 & 0 & -4 & -1 \end{pmatrix} && C_1 \leftrightarrow C_4 \\
&= \operatorname{rg}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 3 & -2 & \lambda \\ 0 & -3 & -4 & -4 \end{pmatrix} && \begin{array}{l} L_2 \leftarrow L_2 - 3L_1 \\ L_4 \leftarrow L_4 - 3L_1 \end{array} \\
&= \operatorname{rg}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -5 & \lambda \\ 0 & 0 & -1 & -4 \end{pmatrix} && \begin{array}{l} L_3 \leftarrow L_3 + 3L_2 \\ L_4 \leftarrow L_4 - 3L_2 \end{array} \\
&= \operatorname{rg}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & -4 \\ 0 & 0 & -5 & \lambda \end{pmatrix} && L_3 \leftrightarrow L_4 \\
&= \operatorname{rg}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & -4 \\ 0 & 0 & 0 & \lambda + 20 \end{pmatrix} && L_4 \leftarrow L_4 - 5L_3
\end{aligned}$$

En particulier, la matrice ci-dessus n'est pas inversible (rappelons qu'une matrice carrée de taille n est inversible si et seulement si son rang est égal à n). On a donc à présent une nouvelle CNS pour montrer qu'une matrice est inversible : une matrice (carrée) de taille n est inversible si et seulement si son rang est égal à n. Nous donnerons une autre CNS dans le chapitre 33.

Il y a donc deux cas : si $\lambda = -20$, alors

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & -1 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 3$$

tandis que si $\lambda \neq -20$, alors $\operatorname{rg}(A) = 4$. Passons à présent à B :

En particulier, A est inversible si et seulement si $\lambda \neq -20$. L'année prochaine, vous direz que $\lambda = -20$ est la seule valeur propre de A et que l'espace propre associé est de dimension 1.

$$\begin{aligned}
\operatorname{rg}(B) &= \operatorname{rg}\begin{pmatrix} -1 & 1 & 1 & 2 \\ 1 & 1 & \lambda & 1 \\ 3 & -1 & 1 & -3 \\ 0 & 2 & 4 & \lambda \end{pmatrix} && C_1 \leftrightarrow C_3 \\
&= \operatorname{rg}\begin{pmatrix} -1 & 1 & 1 & 2 \\ 0 & 2 & \lambda+1 & 3 \\ 0 & 2 & 4 & 3 \\ 0 & 2 & 4 & \lambda \end{pmatrix} && \begin{array}{l} L_2 \leftarrow L_2 + L_1 \\ L_3 \leftarrow L_3 + 3L_1 \end{array} \\
&= \operatorname{rg}\begin{pmatrix} -1 & 1 & 1 & 2 \\ 0 & 2 & \lambda+1 & 3 \\ 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & \lambda - 3 \end{pmatrix} && L_4 \leftarrow L_4 - L_3 \\
&= \operatorname{rg}\begin{pmatrix} -1 & 1 & 1 & 2 \\ 0 & 2 & \lambda+1 & 3 \\ 0 & 0 & 3-\lambda & 0 \\ 0 & 0 & 0 & \lambda - 3 \end{pmatrix} && L_3 \leftarrow L_3 - L_2
\end{aligned}$$

Là aussi, il y a deux cas : si $\lambda = 3$ alors

$$\operatorname{rg}(B) \quad = \quad \operatorname{rg}\begin{pmatrix} -1 & 1 & 1 & 2 \\ 0 & 2 & 4 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad = \quad 2$$

tandis que si $\lambda \neq 3$, $\operatorname{rg}(B) = 4$.

V.3 Rang d'une famille de vecteurs

V.3.a Définition et premiers exemples

Définition. Soient $k \in \mathbb{N}^*$ et $(x_1, \dots, x_k) \in E^k$. On appelle rang des vecteurs $x_1, \dots, x_k$, noté $\operatorname{rg}(x_1, \dots, x_k)$, la dimension de $\operatorname{Vect}(x_1, \dots, x_k)$.

En d'autres termes, le rang d'une famille de vecteurs est égal à la dimension de l'espace engendré par ces vecteurs.

Rappel (cf. chapitre 28) : Si y est combinaison linéaire de $(x_1, \dots, x_q)$ alors

$$\operatorname{Vect}(x_1, \dots, x_q, y) = \operatorname{Vect}(x_1, \dots, x_q).$$

Le but du jeu va être d'enlever tous les vecteurs superflus jusqu'à arriver à une famille libre, qui sera donc une base de l'espace engendré, ce qui donnera la dimension et donc le rang.

Exemples : On se place dans $\mathbb{K}^3$ (qu'on identifie à $\mathscr{M}_{3,1}(\mathbb{K})$). Donnons le rang des familles suivantes :

- $x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $x_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$.

  Réponse : $\operatorname{rg}(x_1, x_2, x_3) = \dim \operatorname{Vect}(x_1, x_2, x_3)$. Or, par définition, x_1, x_2, x_3 forment une famille génératrice de $\operatorname{Vect}(x_1, x_2, x_3)$. De plus (x_1, x_2, x_2) est libre (c'est la base canonique de $\mathbb{K}^3$). Ainsi, $\operatorname{Vect}(x_1, x_2, x_3)$ admet une base à trois éléments donc est de dimension 3. Finalement, $\operatorname{rg}(x_1, x_2, x_3) = 3$.

  On aurait aussi pu dire directement que $\operatorname{Vect}(x_1, x_2, x_3) = \mathbb{K}^3$ car (x_1, x_2, x_3) est la base canonique de $\mathbb{R}^3$ donc cet espace est de dimension 3.

- $x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

  Réponse : $\operatorname{rg}(x_1, x_2) = \dim \operatorname{Vect}(x_1, x_2)$. Par définition, x_1 et x_2 forment une famille génératrice de $\operatorname{Vect}(x_1, x_2)$ et ces vecteurs sont non colinéaires donc (x_1, x_2) est libre donc est une base de $\operatorname{Vect}(x_1, x_2)$. Ainsi, $\operatorname{Vect}(x_1, x_2)$ admet une base à deux éléments donc est de dimension 2. Finalement, $\operatorname{rg}(x_1, x_2) = 2$.

- $x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$, $x_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

  Réponse : $\operatorname{rg}(x_1, x_2, x_3) = \dim \operatorname{Vect}(x_1, x_2, x_3)$. Or, par définition, x_1, x_2, x_3 forment une famille génératrice de $\operatorname{Vect}(x_1, x_2, x_3)$. Cependant la famille (x_1, x_2, x_3) n'est pas libre car $x_2 = 2x_1$. On peut donc en déduire que $\operatorname{Vect}(x_1, x_2, x_3) = \operatorname{Vect}(x_1, x_3)$ (cf. rappel ci-dessus), et donc x_1 et x_3 forment une famille génératrice de cet espace. Or x_1 et x_3 sont non colinéaires donc (x_1, x_3) est libre et donc est une base de $\operatorname{Vect}(x_1, x_3)$. Ainsi, $\operatorname{Vect}(x_1, x_2, x_3)$ admet une base à deux éléments donc est de dimension 2. Finalement, $\operatorname{rg}(x_1, x_2, x_3) = 2$.

- $x_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $x_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $x_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

  Réponse : $\operatorname{rg}(x_1, x_2, x_3, x_4) = \dim \operatorname{Vect}(x_1, x_2, x_3, x_4)$. Or, par définition, x_1, x_2, x_3, x_4 forment une famille génératrice de l'espace $\operatorname{Vect}(x_1, x_2, x_3, x_4)$. Cependant la famille (x_1, x_2, x_3, x_4) n'est pas libre car $x_4 = x_1 + x_2 + x_3$. On peut cependant en déduire que $\text{Vect}(x_1, x_2, x_3, x_4) = \text{Vect}(x_1, x_2, x_3)$ (cf. rappel ci-dessus) et on a déjà vu que cet espace est de dimension 3 (cf. premier exemple ci-dessus). Finalement, $\text{rg}(x_1, x_2, x_3, x_4) = 3$.

V.3.b Premières propriétés

Dans la suite du paragraphe on se donne $k \in \mathbb{N}^*$ et $(x_1, \dots, x_k) \in E^k$.

Proposition. Le rang d'une sous-famille de $(x_1, \dots, x_k)$ est inférieur ou égal à $\text{rg}(x_1, \dots, x_k)$, c'est-à-dire que si on retire vecteurs, le rang ne peut que diminuer.

Démonstration. Il suffit de voir que si $(x_i)_{i\in I}$ est une sous-famille de $(x_1, \dots, x_k)$, alors $\text{Vect}(x_i)_{i\in I} \subset \text{Vect}(x_1, \dots, x_k)$ donc a une dimension inférieure (ou égale).

Proposition. $\text{rg}(x_1, \dots, x_k) = 0 \iff x_1 = \cdots = x_k = 0$.

En particulier, si l'un des vecteurs est non nul, le rang de la famille est supérieur ou égal à 1.

Démonstration.

$$\begin{aligned} \text{rg}(x_1, \dots, x_k) = 0 &\iff \dim \text{Vect}(x_1, \dots, x_k) = 0 \\ &\iff \text{Vect}(x_1, \dots, x_k) = \{0\} \\ &\iff x_1 = \cdots = x_k = 0 \end{aligned}$$

□

Proposition.

- $\text{rg}(x_1, \dots, x_k) \leqslant k$ avec égalité si et seulement si $(x_1, \dots, x_k)$ est libre.
- $\text{rg}(x_1, \dots, x_k) \leqslant \dim E$ avec égalité si et seulement si $(x_1, \dots, x_k)$ est génératrice de E.

En d'autres termes, le rang d'une famille de vecteurs est inférieur ou égal au nombre de vecteurs et à la dimension de l'espace « ambiant », avec égalité si et seulement si...

⇝ Exercice.

Plus généralement, on a le résultat fondamental suivant :

Théorème. Le rang de la famille $(x_1, \dots, x_k)$ est le nombre maximal de vecteurs formant une famille libre que l'on peut extraire de $(x_1, \dots, x_k)$.

Ce résultat est à rapprocher de celui reliant le rang avec la taille maximale d'une matrice extraite inversible : cf. paragraphe V.4.a.

Démonstration. Soit $r = \text{rg}(x_1, \dots x_k)$. On a $r = \dim \text{Vect}(x_1, \dots, x_k)$ par définition, donc une famille libre extraite de $x_1, \dots, x_k$ (et donc en particulier constituée d'éléments de $\text{Vect}(x_1, \dots, x_k)$) est de cardinal inférieur ou égal à r (rappelons qu'une famille libre est de cardinal inférieur ou égal à la dimension de l'espace).

Il ne reste plus qu'à montrer qu'on peut effectivement extraire de $(x_1, \dots, x_k)$ une famille libre de cardinal r. D'après le théorème de la base extraite, puisque $(x_1, \dots, x_k)$ est par définition une famille génératrice de $\text{Vect}(x_1, \dots, x_k)$, on peut en extraire une base (de cardinal r car r est la dimension de cet espace), en particulier on peut en extraire une famille libre de cardinal r. □

Corollaire.

- $\text{rg}(x_1, \dots, x_k) = 1$ si et seulement si les x_k sont tous proportionnels (et au moins un est non nul).
- Si E est supposé de dimension p, alors $\text{rg}(x_1, \dots, x_p) = p$ si et seulement si $(x_1, \dots, x_p)$ est une base de E.

Démonstration. $\text{rg}(x_1, \dots, x_k) = 1$ si et seulement si on peut extraire au maximum un vecteur libre de cette famille : cette famille n'est donc pas constituée de vecteurs tous nuls, et si on en prend deux, ils forment une famille liée donc sont proportionnels.

$\text{rg}(x_1, \dots, x_p) = p$ si et seulement si $(x_1, \dots, x_p)$ est une famille libre, si et seulement si c'est une base (puisque c'est une famille à p éléments en dimension p).

Proposition.

1. Le rang d'une famille de vecteurs ne dépend pas de l'ordre des vecteurs.
2. $\forall \lambda \in \mathbb{K}^*$, $\text{rg}(\lambda x_1, x_2, \dots, x_k) = \text{rg}(x_1, \dots, x_k)$.
3. $\forall \lambda \in \mathbb{K}$, $\forall i \in [\![2\,;\,k]\!]$, $\text{rg}(x_1 + \lambda x_i, x_2, \dots, x_k) = \text{rg}(x_1, \dots, x_k)$.
4. $\text{rg}(x_1, \dots, x_k, 0) = \text{rg}(x_1, \dots, x_k)$.
5. Plus généralement, si $y \in \text{Vect}(x_1, \dots, x_k)$, $\text{rg}(x_1, \dots, x_k, y) = \text{rg}(x_1, \dots, x_k)$.

Puisque le rang ne dépend pas de l'ordre des vecteurs, on peut effectuer les opérations 2 et 3 sur tous les vecteurs, et pas forcément sur le premier, et on peut « supprimer » tous les vecteurs superflus (opérations 4 et 5), pas forcément le dernier.

Démonstration.

1. Puisque la somme est commutative sur E, $\text{Vect}(x_1, \dots, x_k)$ (et donc sa dimension) ne dépend pas de l'ordre des vecteurs.
2. Puisque $\lambda \neq 0$, on montre par double inclusion que $\text{Vect}(x_1, x_2, \dots, x_k) = \text{Vect}(\lambda x_1, x_2, \dots, x_k)$ et donc ces espaces ont la même dimension.
3. $\rightsquigarrow$ Exercice.
4. Découle du rappel se trouvant au début du paragraphe.
5. Idem.

Morale de l'histoire : On ne change pas le rang en :

- modifiant l'ordre des vecteurs.
- multipliant l'un des vecteurs par un scalaire non nul.
- ajoutant à un vecteur une combinaison linéaire des autres vecteurs.
- supprimant un vecteur qui est combinaison linéaire des autres.

V.3.c Encore une définition de rang !

En fait, ce sont les mêmes ! Notons $C_1, \dots, C_p$ les vecteurs colonnes de A (qui appartiennent donc à $\mathbb{K}^n$ qu'on identifie à $\mathscr{M}_{n,1}(\mathbb{K})$). Si u est l'application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ canoniquement associée à A, alors $C_1, \dots, C_p$ engendrent l'image de u (cf. paragraphe II.2.b) donc :

$$\begin{aligned} \text{rg}(A) &= \text{rg}(u) \\ &= \dim \text{Im}(u) \\ &= \dim \text{Vect}(C_1, \dots, C_p) \\ &= \text{rg}(C_1, \dots, C_p) \end{aligned}$$

En d'autres termes :

Théorème. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Le rang de A est égal au rang de ses vecteurs colonnes.

On retrouve le fait qu'on ne change pas le rang d'une matrice lorsqu'on effectue des opérations élémentaires sur les colonnes (ce qu'on savait déjà), mais cela nous donne de nouveaux outils pour calculer le rang d'une matrice :

Corollaire. Soit A une matrice.

- On ne modifie pas le rang de A en « supprimant » un vecteur colonne nul ou combinaison linéaire d'autres colonnes.
- Le rang de A est le nombre maximal de vecteurs libres parmi ses vecteurs colonnes.
- Si A est carrée de taille n, alors $\text{rg}(A) = n$ si et seulement si ses vecteurs colonnes forment une base de $\mathbb{K}^n$.
- $\text{rg}(A) = 1$ si et seulement si ses vecteurs colonnes sont tous proportionnels (et au moins un est non nul).

Pas forcément carrée.

On retrouve le fait qu'une matrice carrée ayant deux colonnes proportionnelles n'est pas inversible. Attention, ce n'est pas équivalent : une matrice n'est pas inversible si et seulement si ses vecteurs colonnes ne sont pas libres...

Exemple : Prouvons que

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

est inversible.

Les vecteurs colonnes de M forment une base de $\mathbb{K}^n$: en effet, si on note comme d'habitude $(e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$, alors les vecteurs colonnes de M sont $(e_n, e_1, \dots, e_{n-1})$, et donc c'est encore une base de $\mathbb{K}^n$ (le fait que des vecteurs forment une base ou non ne dépend pas de l'ordre des vecteurs). Il en découle que $\text{rg}(M) = n$ donc que M est inversible.

...donc si et seulement si l'un d'entre eux est CL des autres, mais cela ne se voit pas forcément à l'oeil nu, cela ne veut pas forcément dire que deux d'entre eux sont proportionnels. C'est la même chose avec les vecteurs lignes (voir ci-dessous).

Théorème (admis provisoirement). Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Alors $\text{rg}(A) = \text{rg}\left(A^\top\right)$, c'est-à-dire qu'une matrice et sa transposée ont même rang.

Nous le prouverons dans le paragraphe VI.4.

Corollaire. Le rang d'une matrice est égal au rang de ses vecteurs lignes.

Démonstration. Les lignes de A sont les colonnes de $A^\top$ et $\text{rg}(A) = \text{rg}(A^\top)$. □

Conséquence : Tous les résultats précédents sont donc encore vrais en remplaçant « colonnes » par « lignes » :

- On ne modifie pas le rang d'une matrice en effectuant des opérations élémentaires sur les lignes (ce qu'on savait déja)
- On ne modifie pas le rang de A en « supprimant » un vecteur ligne nul ou combinaison linéaire d'autres lignes.
- Le rang de A est le nombre maximal de vecteurs libres parmi ses vecteurs lignes.
- Si A est carrée de taille n, alors $\text{rg}(A) = n$ si et seulement si ses vecteurs lignes forment une base de $\mathbb{K}^n$.
- $\text{rg}(A) = 1$ si et seulement si ses vecteurs lignes sont tous proportionnels (et au moins un est non nul).

V.3.d Et en pratique ?

Si on veut donner le rang d'une famille de vecteurs, on commence par se ramener à une matrice (en effet, le rang d'une matrice est égal au rang de ses vecteurs colonnes) et on se ramène à la méthode vue au paragraphe V.2. Par exemple, si on veut donner le rang de la famille de vecteurs $\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\3\\5\end{pmatrix}$, on écrit

$$\text{rg}\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\3\\5\end{pmatrix}\right) = \text{rg}\begin{pmatrix}1 & 1 & 1\\ 2 & 1 & 3\\ 3 & 1 & 5\end{pmatrix}$$

et on fait comme dans le paragraphe V.2.

V.4 Rang et matrices extraites

V.4.a Matrices extraites

Définition. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$. Une matrice B est extraite de A s'il existe $I \subset [\![1\,;n]\!]$ et $J \subset [\![1\,;p]\!]$ tels que $B = (A_{i,j})_{i \in I, j \in J}$.

Remarque : Si on note $I = \{i_1;\dots;i_q\}$ et $J = \{j_1;\dots;j_r\}$, la matrice B est la matrice obtenue en ne gardant que les lignes d'indices $i_1,\dots,i_q$ et les colonnes d'indices $j_1,\dots,j_r$. En d'autres termes, une matrice B est extraite d'une matrice A lorsqu'on peut l'obtenir en barrant des lignes et des colonnes de A (pas forcément le même nombre de lignes et le même nombre de colonnes).

Avec les notations ci-contre, B est de taille $q\times r$: B n'a aucune raison d'être carrée !

$$\begin{pmatrix}
a_{1,1} & \cdots & a_{1,j_1} & \cdots & a_{1,j_2} & \cdots & a_{1,j_r} & \cdots & a_{1,p}\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
a_{i_1,1} & \cdots & a_{i_1,j_1} & \cdots & a_{i_1,j_2} & \cdots & a_{i_1,j_r} & \cdots & a_{i_1,p}\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
a_{i_2,1} & \cdots & a_{i_2,j_1} & \cdots & a_{i_2,j_2} & \cdots & a_{i_2,j_r} & \cdots & a_{i_2,p}\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
a_{i_q,1} & \cdots & a_{i_q,j_1} & \cdots & a_{i_q,j_2} & \cdots & a_{i_q,j_r} & \cdots & a_{i_q,p}\\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots\\
a_{n,1} & \cdots & a_{n,j_1} & \cdots & a_{n,j_2} & \cdots & a_{n,j_r} & \cdots & a_{n,p}
\end{pmatrix}
=
\begin{pmatrix}
a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_r}\\
a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_r}\\
\vdots & \vdots & \cdots & \vdots\\
\vdots & \vdots & \cdots & \vdots\\
a_{i_q,j_1} & a_{i_q,j_2} & \cdots & a_{i_q,j_r}
\end{pmatrix}$$

Exemple : La matrice $B = \begin{pmatrix} 4 & 10\\ 6 & 12\end{pmatrix}$ est extraite de $A = \begin{pmatrix} 1 & 4 & 7 & 10 & 13\\ 2 & 5 & 8 & 11 & 14\\ 3 & 6 & 9 & 12 & 15\end{pmatrix}$: on a rayé les lignes $1,3,5$ et la deuxième colonne.

Théorème. Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$.

- Si B est extraite de A alors $\mathrm{rg}(B) \leqslant \mathrm{rg}(A)$.
- Le rang de A est la taille de la plus grande matrice (carrée) inversible qu'on peut extraire de A.

Démonstration.

- On obtient B à partir de A en deux étapes (qu'on peut évidemment effectuer dans l'ordre qu'on veut) : on supprime des colonnes et on supprime des lignes. Notons C la matrice obtenue à partir de A en ne supprimant que les colonnes. Or, $\mathrm{rg}(A)$ est le rang de ses vecteurs colonnes, et on a déjà vu qu'en supprimant des vecteurs, le rang diminue donc $\mathrm{rg}(A) \geqslant \mathrm{rg}(C)$. De plus, on obtient B à partir de C en supprimant des lignes et puisque le rang d'une matrice est aussi le rang de ses vecteurs lignes, on diminue aussi le rang en supprimant des vecteurs lignes, donc $\mathrm{rg}(C) \geqslant \mathrm{rg}(B)$ ce qui permet de conclure.
- Supposons évidemment A non nulle. Notons $r = \mathrm{rg}(A) \in \mathbb{N}^*$. D'après ce qui précède, il suffit de prouver qu'on peut extraire de A une matrice (carrée) inversible de taille r. D'après ce qui précède, A possède r vecteurs colonnes libres (le rang est le nombre maximal de vecteurs libres) : en ne prenant que ces vecteurs (et en rayant donc les autres), cela donne une matrice C à r colonnes de rang r (puisque ses r colonnes sont libres). Le rang d'une matrice étant aussi le rang de ses vecteurs lignes, la matrice C possède r vecteurs lignes libres. Dès lors, en prenant uniquement ces r vecteurs lignes, cela donne une matrice B à r lignes et r colonnes, donc carrée, dont les vecteurs lignes

sont libres donc cette matrice est de rang r donc est inversible (car carrée de taille r). La matrice B est obtenue en rayant des lignes et des colonnes de A donc est extraite de A, ce qui est le résultat voulu.

Exemple : $\mathrm{rg}\begin{pmatrix}2&3&1&0\\1&0&3&1\\3&0&0&2\end{pmatrix} = 3$ car la matrice extraite $\begin{pmatrix}3&1&0\\0&3&1\\0&0&2\end{pmatrix}$ est inversible (pourquoi, au fait ?) et car le rang de la matrice de départ est inférieur à 3.

V.5 Bilan

Si $A \in \mathscr{M}_{n,p}(\mathbb{K})$:

- $\mathrm{rg}(A)$ est égal au rang de toute application linéaire représentée par A.
- $\mathrm{rg}(A) = \dim \mathrm{Im}(A)$.
- $\mathrm{rg}(A) = \mathrm{rg}(A^\top)$.
- $\mathrm{rg}(A) \leqslant \min(n, p)$.
- $\mathrm{rg}(A) = n\,(= \dim(F))$ si et seulement si u est surjective.
- $\mathrm{rg}(A) = p\,(= \dim(E))$ si et seulement si u est injective.
- On ne modifie pas le rang d'une matrice en la multipliant par une matrice inversible.
- Les opérations élémentaires (sur les lignes et les colonnes) laissent le rang invariant.
- On ne modifie pas le rang en supprimant une colonne qui est combinaison linéaire des autres colonnes.
- On ne modifie pas le rang en supprimant une ligne qui est combinaison linéaire des autres lignes.

De plus, le rang de A est égal :

- au rang de ses vecteurs lignes.
- au nombre maximal de vecteurs libres parmi ses vecteurs lignes.
- au rang de ses vecteurs colonnes.
- à la taille de la plus grande matrice inversible qu'on peut extraire de A.
- au nombre maximal de vecteurs libres parmi ses vecteurs colonnes.

En particulier :

- $\mathrm{rg}(A) = 0$ si et seulement si $A = 0$.
- $\mathrm{rg}(A) = 1$ si et seulement si toutes les lignes de A sont proportionnelles si et seulement si toutes les colonnes de A sont proportionnelles.

Et A non nulle !

Enfin, si $A \in \mathscr{M}_n(\mathbb{K})$ (i.e. si $\dim(E) = \dim(F) = n$) et si $u \in \mathscr{L}(E, F)$ est une application linéaire représentée par A : ,

Nous rajouterons, dans le chapitre 33 : si et seulement si $\det(A) \neq 0$, si et seulement si $\det(u) \neq 0$.

$$\begin{aligned}
\mathrm{rg}(A) = n \quad &\iff \quad u \text{ est injective}\\
&\iff \quad u \text{ est surjective}\\
&\iff \quad u \text{ est bijective}\\
&\iff \quad A \text{ est inversible}\\
&\iff \quad \text{Les vecteurs colonnes de } A \text{ sont libres}\\
&\iff \quad \text{Les vecteurs lignes de } A \text{ sont libres}\\
&\iff \quad \text{Les vecteurs lignes de } A \text{ forment une base de } \mathbb{K}^n\\
&\iff \quad \text{Les vecteurs colonnes de } A \text{ forment une base de } \mathbb{K}^n.
\end{aligned}$$

V.6 Rang d'un système linéaire

Rappel (cf. chapitre 21) : Soit

$$(S)\quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p = b_2 \\ \qquad\vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = b_n \end{cases}$$

un système linéaire de n équations à p inconnues où $a_{i,j} \in \mathbb{K}$ et $b_i \in \mathbb{K}$ pour tout $(i,j) \in [\![1\,;\,n]\!] \times [\![1\,;\,p]\!]$. Posons

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p} \\ \vdots & \vdots & & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,p} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}.$$

$A \in \mathscr{M}_{n,p}(\mathbb{K})$, $B \in \mathscr{M}_{n,1}(\mathbb{K})$ et $X \in \mathscr{M}_{p,1}(\mathbb{K})$. Cependant, comme précédemment, nous dirons parfois que $B \in \mathbb{K}^n$ et $X \in \mathbb{K}^p$.

La matrice A est appelée la matrice associée au système (S). Alors S est équivalent à l'équation matricielle $AX = B$.

V.6.a Cas particulier des systèmes homogènes

On suppose dans ce paragraphe que le système est homogène, c'est-à-dire que $B = 0$.

Proposition. L'ensemble des solutions du système homogène $AX = 0$ est le noyau de A.

Démonstration. Rien à prouver : l'ensembles des X tels que $AX = 0$ est par définition le noyau de A. D'ailleurs, on avait déjà dit que les lignes formaient un système d'équations du noyau.

Un noyau étant un espace vectoriel, on retrouve le fait (cf. chapitre 28) que l'ensemble des solutions d'un système linéaire est un espace vectoriel.

Définition. On appelle rang d'un système linéaire le rang de la matrice A associée.

Proposition. L'ensemble des solutions du système homogène $AX = 0$ est un espace vectoriel de dimension égale à $p - \mathrm{rg}(A)$, c'est-à-dire le nombre d'équations moins le rang du système.

Démonstration. On vient de prouver que cet ensemble était égal à $\ker(A)$: il suffit ensuite d'appliquer le théorème du rang.

V.6.b Cas général

On ne suppose plus que $B = 0$. On rappelle qu'un système est dit compatible lorsqu'il admet des solutions.

Proposition. Le système $AX = B$ est compatible si et seulement si $B \in \mathrm{Im}(A)$.

Démonstration. Notons $C_1, \dots, C_n$ les colonnes de A. On sait (cf. chapitre 21) que le système est compatible si et seulement si $B \in \mathrm{Vect}(C_1, \dots, C_n)$, et on sait (cf. paragraphe II.2.b) que $\mathrm{Vect}(C_1, \dots, C_n) = \mathrm{Im}(A)$.

Les vecteurs colonnes engendrent l'image, et les vecteurs lignes donnent un système d'équation du noyau.

Remarque : Rappel : les solutions d'un système compatible $AX = B$ sont les $X_0 + Y$, où X_0 est une solution particulière et où Y parcourt l'ensemble des solutions du système homogène associé. Cela munit l'ensemble des solutions d'une structure d'espace affine (cf. chapitre 36).

VI Matrices de passage et changement de base, cas général

On suppose dans les deux premiers paragraphes que $\dim(E) = n$.

VI.1 Matrices de passage

Définition. Soient $B = (e_1, \dots, e_n)$ et $B' = (\varepsilon_1, \dots, \varepsilon_n)$ deux bases de E. On appelle matrice de passage de B' vers B la matrice $P_{B,B'} \in \mathscr{M}_n(\mathbb{K})$ dont la j-ième colonne est formée des coordonnées de ε_j dans la base B.

On expliquera ce nom dans le paragraphe VI.2, et pourquoi on la note $P_{B,B'}$ avec B en premier, alors qu'on va de B' vers B.

Remarques :

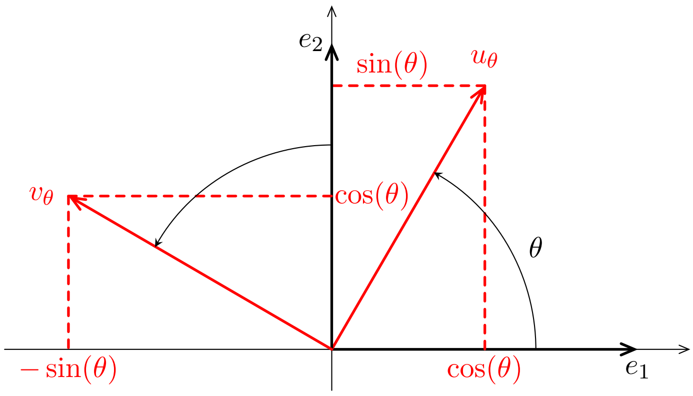
- En d'autres termes, $P_{B,B'} = (p_{i,j})_{1\leqslant i,j\leqslant n}$ où, pour tout $j \in [\![1\,;\, n]\!]$,
$$\varepsilon_j = \sum_{i=1}^{n} p_{i,j} e_i$$
c'est-à-dire que pour tous i et j, $p_{i,j}$ est la i-ème coordonnée de ε_j dans la base $(e_1, \dots, e_n)$. On note :
$$P_{B,B'} = \begin{array}{c} \begin{array}{cccc} \varepsilon_1 & \varepsilon_2 & \dots & \varepsilon_n \end{array} \\ \begin{pmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,n} \\ p_{2,1} & p_{2,2} & \dots & p_{2,n} \\ \vdots & \vdots & & \vdots \\ p_{n,1} & p_{n,2} & \dots & p_{n,n} \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ \vdots \\ e_n \end{array}$$
- En d'autres termes, $P_{B,B'} = \mathrm{Mat}_{B,B'}(\mathrm{Id}_E)$: quand la base de départ et la base d'arrivée ne sont pas les mêmes, la matrice de Id_E n'est pas forcément I_n !
- Par conséquent, si B, B' et B'' sont trois bases de E, $P_{B,B'} \times P_{B',B''} = \mathrm{Mat}_{B,B'}(\mathrm{Id}_E) \times \mathrm{Mat}_{B',B''}(\mathrm{Id}_E)$: il découle du théorème du paragraphe II.3 que $P_{B,B'} \times P_{B',B''} = \mathrm{Mat}_{B,B''}(\mathrm{Id}_E)$ donc que $P_{B,B'} \times P_{B',B''} = P_{B,B''}$.
- Une matrice de passage est forcément carrée (de taille la dimension de l'espace) puisque toutes les bases d'un espace vectoriel (de dimension finie) ont le même cardinal (égal à la dimension de l'espace).

On en déduit le même moyen mnémotechnique que précédemment pour se souvenir quels vecteurs on met au-dessus : dans l'égalité ci-contre, B' est « à côté » des vecteurs $\varepsilon_1, \dots, \varepsilon_n$ (qui forment la base B'), et éloignée des vecteurs $(e_1, \dots, e_n)$ qui forment la base B. Intuitivement (comme ci-dessous) : « B' est à côté d'elle-même ».

Exemples :

- Dans les exercices du chapitre 21, quand je vous demandais d'expliciter « la matrice P dont les vecteurs colonnes sont X, Y, Z dans cet ordre », je vous demandais en fait d'expliciter la matrice de passage $P_{B,B'}$ avec B' la base (X, Y, Z). La matrice $D = PMP^{-1}$ qu'on calculait ensuite était alors $\mathrm{Mat}_{B'}(u)$: cf. paragraphe VII.
- Si on note B la base canonique de $\mathbb{K}_n[X]$ et B' la base $(1, 2X, 3X^2, \dots, (n+1)X^n)$, alors la matrice de passage de B' vers B est la matrice (de taille $n+1$) suivante :
$$P_{B,B'} = \begin{array}{c} \begin{array}{cccc} 1 & 2X & \dots & (n+1)X^n \end{array} \\ \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (n+1) \end{pmatrix} \end{array} \begin{array}{c} \\ 1 \\ X \\ \vdots \\ X^n \end{array}$$
- Dans $\mathbb{K}^2$, notons $\varepsilon_1 = (1, 2)$, $\varepsilon_2 = (3, 4)$, et $B = (e_1, e_2)$ la base canonique de $\mathbb{R}^2$. Les vecteurs ε_1 et ε_2 ne sont pas colinéaires donc forment une famille libre. Ils forment une famille libre à deux éléments dans un espace de dimension 2 donc forment une base. On l'appelle B', et on a :
$$P_{B,B'} = \begin{array}{c} \begin{array}{cc} \varepsilon_1 & \varepsilon_2 \end{array} \\ \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \end{array}$$

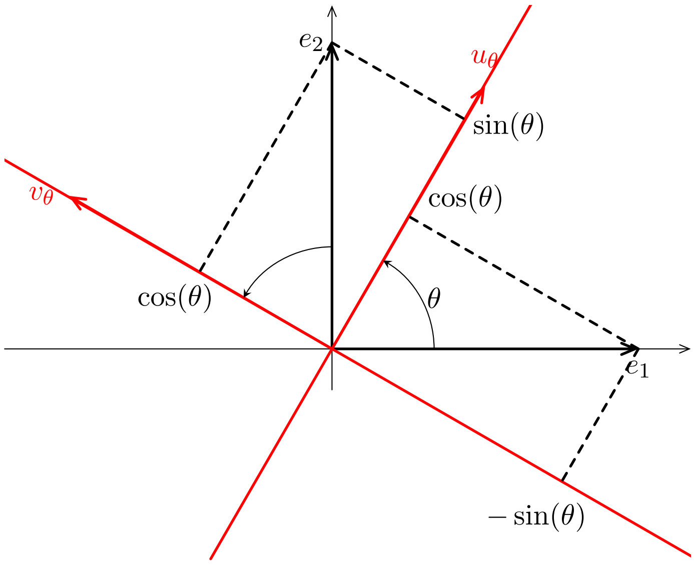
- Notons $B_\theta = (u_\theta, v_\theta)$ la base de $\mathbb{K}^2$ obtenue en appliquant une rotation d'angle θ à (e_1, e_2), la base canonique de $\mathbb{R}^2$:



Dès lors, $P_{B,B_\theta} = \begin{array}{cc} \begin{array}{cc} u_\theta & v_\theta \end{array} & \\ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \end{array} \end{array}$.

- Dans l'autre sens, donnons la matrice de passage de B vers B_θ : on cherche pour cela les coordonnées de (e_1, e_2) dans la base B_θ :

Un peu de trigo ne fait jamais de mal...



Dès lors, $P_{B_\theta,B} = \begin{array}{cc} \begin{array}{cc} e_1 & e_2 \end{array} & \\ \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} & \begin{array}{c} u_\theta \\ v_\theta \end{array} \end{array}$.

Proposition. Avec les mêmes notations que ci-dessus, $P_{B,B'}$ est inversible, et $\left(P_{B,B'}\right)^{-1} = P_{B',B}$.

Démonstration. Découle du fait (voir ci-dessus) que $P_{B,B'} = \mathrm{Mat}_{B,B'}(\mathrm{Id}_E)$: Id_E est bijective donc $P_{B,B'}$ est inversible et son inverse est $\mathrm{Mat}_{B',B}(\mathrm{Id}_E) = P_{B',B}$.

C'est le résultat du paragraphe III.3 avec $E = F$, $u = \mathrm{Id}_E$, $B_E = B$ et $B_F = B'$.

Remarque : C'est intuitif : une matrice de passage va d'une base dans une autre, pour revenir en sens inverse... il suffit d'aller de la seconde base vers la première.

Exemple : Reprenons les notations précédentes, à savoir B la base canonique et B_θ son image par la rotation d'angle θ. Un calcul simple donne bien $P_{B,B_\theta} \times P_{B_\theta,B} = I_2$ donc on a bien $P_{B_\theta,B} = (P_{B,B_\theta})^{-1}$.

Remarque : La réciproque est vraie, c'est-à-dire que toute matrice inversible est une matrice de passage : plus précisément, si P est carrée de taille égale à $\dim(E)$, alors il existe deux bases B_1 et B_2 de E telles que $P = P_{B_2,B_1}$. En effet, si P est une matrice inversible inversible de taille $\dim(E)$, si on note B_2 une base de E et si on note B_1 la famille de vecteurs de E dont les coordonnées dans la base B_2 sont les vecteurs colonnes de P, alors B_1 est une base de E (puisque P est inversible) et, par définition, puisque P est la matrice dont les coordonnées des vecteurs de B_1 dans la base B_2, alors $P = P_{B_2,B_1}$.

VI.2 Changement de coordonnées dans une base (pour les vecteurs)

Théorème. Soit $x \in E$. On note X_B et $X_{B'}$ les matrices colonnes de $\mathscr{M}_{n,1}(\mathbb{K})$ dont les coordonnées sont les coordonnées de x dans, respectivement, les bases B et B' (cf. paragraphe I.4). Alors : $X_B = P_{B,B'}X_{B'}$.

Moyen mnémotechnique : dans l'égalité ci-contre, B et B sont côte à côte, ainsi que B' et B'. Ce théorème est la raison pour laquelle on note cette matrice $P_{B,B'}$, pour pouvoir appliquer « la relation de Chasles ». De plus, le nom de matrice de passage de B' vers B prend tout son sens : on part de $X_{B'}$, donc de la base B', et on arrive à X_B, donc dans la base B.

Démonstration. Analogue à la démonstration du théorème du paragraphe II.3 et laissée en exercice. □

Exemples :

- Reprenons les notations $(e_1, e_2, \varepsilon_1, \varepsilon_2, B, B')$ du paragraphe précédent, c'est-à-dire que $\varepsilon_1 = (1,2), \varepsilon_2 = (3,4)$ etc. Soit $(x,y) \in \mathbb{R}^2$ Puisque $P_{B',B} = (P_{B,B'})^{-1}$, alors

$$P_{B',B} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & -3/2 \\ 1 & 1/2 \end{pmatrix}$$

Les coordonnées de (x,y) dans la base B sont x et y, donc ses coordonnées dans la base B' sont :

$$X_{B'} = P_{B',B}X_B = \begin{pmatrix} -2x - 3y/2 \\ x + y/2 \end{pmatrix}.$$

- Reprenons aussi les notations précédentes et notons $B = (e_1, e_2)$ la base canonique et $B_\theta = (u_\theta, v_\theta)$ son image par la rotation d'angle θ. Soit X un vecteur de $\mathbb{R}^2$ et notons ses coordonnées (x,y) dans la base canonique et (x_θ, v_θ) ses coordonnées dans la base B_θ, si bien qu'on note $X_B = (x,y)$ et $X_{B_\theta} = (x_\theta, y_\theta)$. Dès lors : $X_{B_\theta} = P_{B_\theta,B}X_B$ donc on obtient :

$$x_\theta = x\cos(\theta) + y\sin(\theta) \qquad \text{et} \qquad y_\theta = -x\sin(\theta) + y\cos(\theta)$$

et on trouverait x et y en fonction de x_θ et y_θ en effectuant le produit $P_{B,B_\theta}X_{B_\theta}$.

VI.3 Changement de coordonnées dans une base (pour les AL)

Théorème. Soit $u \in \mathscr{L}(E,F)$. Soient (B_1, B_2) deux bases de E et (C_1, C_2) deux bases de F. Alors :

$$\mathrm{Mat}_{C_2,B_2}(u) = P_{C_2,C_1} \times \mathrm{Mat}_{C_1,B_1}(u) \times P_{B_1,B_2}$$

Moyen mnémotechnique : chaque base est collée à elle-même.

Démonstration. La matrice de droite est égale à

$$\mathrm{Mat}_{C_2,C_1}(\mathrm{Id}_F) \times \mathrm{Mat}_{C_1,B_1}(u) \times \mathrm{Mat}_{B_1,B_2}(\mathrm{Id}_E)$$ □

qui, d'après le paragraphe II.3, est égale à $\mathrm{Mat}_{C_2,B_2}(\mathrm{Id}_F \circ u \circ \mathrm{Id}_E)$ ce qui est le résultat voulu.

Remarques :

- Ce théorème fait un peu peur, mais il est très simple ! Il suffit de bien visualiser le diagramme ci-dessous :

$$\begin{tikzcd} E, B_1 \arrow[d, "u"] & E, B_2 \arrow[l, "\mathrm{Id}_E"'] \arrow[d, "u"] \\ F, C_1 \arrow[r, "\mathrm{Id}_F"'] & F, C_2 \end{tikzcd}$$

  et son équivalent matriciel :

$$\begin{tikzcd} E, B_1 \arrow[d, "\mathrm{Mat}_{C_1,B_1}(u)"'] & E, B_2 \arrow[l, "P_{B_1,B_2}"'] \arrow[d, "\mathrm{Mat}_{C_2,B_2}(u)"] \\ F, C_1 \arrow[r, "P_{C_2,C_1}"'] & F, C_2 \end{tikzcd}$$

  Rappelons que, même si on garde la même AL, si on change les bases, alors la matrice change... C'est d'ailleurs tout l'objet du paragraphe ! Il ne faut pas être surpris qu'on trouve deux matrices différentes à un endroit où on a la même AL puisque les bases de départ et d'arrivée ont changé !

  Le résultat du théorème précédent est très simple : si on va directement de (E, B_2) dans (F, B_2), alors la matrice est $\mathrm{Mat}_{C_2,B_2}(u)$, mais on peut d'abord aller chez (E, B_1) (matrice de passage P_{B_1,B_2}) puis aller chez (F, C_1) (matrice $\mathrm{Mat}_{C_1,B_1}(u)$) et enfin aller chez (F, C_2) (matrice de passage P_{C_2,C_1}). En clair, on applique une matrice de passage à droite (au départ, puisqu'on va de la droite vers la gauche quand on compose) pour changer la base de départ, et on applique une matrice de passage à gauche (à l'arrivée) pour changer la base d'arrivée. Rien que de naturel !
- Il y a deux formules de changements de base : une pour les vecteurs, et une pour les matrices. Merci de ne pas les confondre. De toute façon, il n'y a pas lieu de les confondre : pour un vecteur, on reste dans le même espace, donc on applique une seule matrice de passage, et pour une AL on change de base à l'arrivée et au départ, donc on applique deux fois une matrice de passage (ce sera la même chose avec les endomorphismes, cf. paragraphe VII.2).
- Là aussi, c'est ce que je vous demandais de faire dans le chapitre 21 lorsque je vous demandais de calculer la matrice $P^{-1}AP$: la matrice P étant la matrice formée par les trois vecteurs X, Y, Z, c'était la matrice de passage de l'ancienne base (la base canonique) vers la nouvelle base (la base (X, Y, Z)) et la matrice P^{-1} était la matrice de passage dans le sens inverse donc la matrice $D = P^{-1}AP$ était la matrice de u dans la base (X, Y, Z). Je vous faisais changer de base sans vous le dire !

Nous en reparlerons quand nous étudierons le cas particulier des endomorphismes dans le paragraphe VII.2. Nous verrons des exemples à ce moment là, puisque c'est ce que nous manierons le plus en pratique.

VI.4 Matrices équivalentes

Dans ce paragraphe, on se replace dans le cas général, c'est-à-dire que $\dim(E) = p$ et $\dim(F) = n$.

Définition. Soit $r \in [\![0\,;\, \min(n,p)]\!]$. On note J_r la matrice de taille $n \times p$ dont tous les coefficients sont nuls sauf les r premiers coefficients diagonaux qui valent 1 :

$$J_r = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \cdots & 0 \\ \vdots & \vdots & & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\0\\\vdots\\\vdots\end{array}}\right\} r \\ \\ \\ \\ \end{array} \in \mathscr{M}_{n,p}(\mathbb{K})$$

Remarques :

- La matrice J_r n'est pas forcément carrée. De plus, on la note de la même façon peu importe sa taille, celle-ci dépend du cadre de l'exercice, il n'y aura jamais d'ambiguité.
- $\text{rg}(J_r) = r$ puisque J_r est échelonnée. En fait, J_r est l'archétype de la matrice de rang r :

Proposition. Soit $u \in \mathscr{L}(E, F)$ et soit $r = \text{rg}(u)$. Alors il existe un couple de bases dans lequel u a pour matrice J_r. En d'autres termes, il existe une base B_1 de E et une base C_1 de F telles que $\text{Mat}_{C_1,B_1}(u) = J_r$.

Démonstration. $\text{rg}(u) = r$ donc, d'après le théorème du rang, $\ker(u)$ est de dimension $p-r$. Soit S un supplémentaire de $\ker(u)$, qui est donc de dimension r. Soit $(e_1, \dots, e_r)$ une base de S et soit $(e_{r+1}, \dots, e_n)$ une base de $\ker(u)$. D'après le théorème de concaténation des bases, $(e_1, \dots, e_n)$ est une base de E qu'on note B_1.

Montrons que $(u(e_1), \dots, u(e_r))$ est une famille libre. Soient $\lambda_1, \dots, \lambda_r$ tels que $\lambda_1 u(e_1) + \dots + \lambda_r u(e_r) = 0$. Par linéarité de u, $u(\lambda_1 e_1 + \dots + \lambda_r e_r) = 0$ c'est-à-dire que $\lambda_1 e_1 + \dots + \lambda_r e_r \in \ker(u)$. Or, $\lambda_1 e_1 + \dots + \lambda_r e_r \in S$ donc appartient à $S \cap \ker(u) = \{0\}$ puisque ces espaces sont supplémentaires. On en déduit que $\lambda_1 e_1 + \dots + \lambda_r e_r = 0$ donc que les λ_i sont nuls puisque les e_i sont libres (c'est une sous-famille d'une base donc une famille libre). D'après le théorème de la base incomplète, on peut compléter $(u(e_1), \dots, u(e_r))$ en base avec $n - r$ vecteurs qu'on note $f_{r+1}, \dots, f_n$, si bien que $(u(e_1), \dots, u(e_r), f_{r+1}, \dots, f_n)$ est une base de F qu'on note C_1, et on a finalement :

Rappelons que $e_{r+1}, \dots, e_p$ sont des éléments de $\ker(u)$ donc ont une image nulle.

$$\text{Mat}_{C_1,B_1}(u) = \begin{array}{c} \begin{array}{ccccccc} u(e_1) & u(e_2) & \dots & u(e_r) & u(e_{r+1}) & \dots & u(e_p) \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots & \cdots & 0 \\ \vdots & \vdots & & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ u(e_1) \\ u(e_2) \\ \vdots \\ u(e_r) \\ f_{r+1} \\ \vdots \\ f_n \end{array}$$

Remarque : Si on note B_E et B_F d'autres bases de E et F respectivement, il découle de la formule de changement de base pour les applications linéaires que

$$\text{Mat}_{B_F,B_E}(u) = P_{B_F,C_1} \times J_r \times P_{B_1,B_E}$$

Puisqu'une matrice de passage est inversible, cela justifie la notion de matrices équivalentes définies ci-dessous.

Attention à la taille des matrices P et Q ! Mais cela ne présente aucune difficulté quand on réfléchit puisque ces tailles sont indispensables pour que les produits soient définis.

Définition. Soit $(M, N) \in \mathscr{M}_{n,p}(\mathbb{K})^2$. On dit que M est équivalente à N s'il existe $P \in \text{GL}_n(\mathbb{K})$ et $Q \in \text{GL}_p(\mathbb{K})$ telles que $M = PNQ$.

Proposition. La relation « être équivalente à » est une relation d'équivalence.

DÉMONSTRATION.
- Soit $M \in \mathscr{M}_{n,p}(\mathbb{K})$. Alors $M = I_n \times M \times I_p$ donc M est équivalente à M : cette relation est réflexive.
- Soit $(M, N,) \in \mathscr{M}_{n,p}(\mathbb{K})^2$. Supposons que M soit équivalente à N : il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $Q \in \mathrm{GL}_p(\mathbb{K})$ telles que $M = PNQ$ donc, en multipliant à gauche par P^{-1} et à droite par Q^{-1}, il vient : $N = P^{-1}MQ^{-1}$, et puisque P^{-1} et Q^{-1} sont inversibles, on en déduit que N est équivalente à M : la relation est symétrique.
- Soit $(M_1, M_2, M_3) \in \mathscr{M}_{n,p}(\mathbb{K})^2$. Supposons que M_1 soit équivalente à M_2 et M_2 à M_3 : il existe $P_1 \in \mathrm{GL}_n(\mathbb{K})$ et $Q_1 \in \mathrm{GL}_p(\mathbb{K})$ telles que $M_1 = P_1M_2Q_1$ et P_2 et Q_2 telles que $M_2 = P_2M_3Q_2$ donc $M_1 = (P_1P_2)M_3(Q_2Q_1)$: P_1P_2 et Q_2Q_1 sont inversibles car produit de matrices inversibles donc M_1 est équivalente à M_3 : la relation est transitive, c'est bien une relation d'équivalence.

Remarque : Puisque la relation est symétrique, N et M jouent le même rôle : on dira donc que les matrices sont équivalentes.

Remarque : Il découle de la formule de changement de base vue en VI.2. que deux matrices d'une même application linéaire dans deux bases différentes sont équivalentes. Montrons que la réciproque est vraie, c'est-à-dire que deux matrices équivalentes représentent la même AL dans des bases différentes.

Puisqu'une matrice inversible est une matrice de passage, il existe C_1, C_2 bases de F et B_1, B_2 bases de E telles que $P = P_{C_2,C_1}$ et $Q = P_{B_1,B_2}$. Dès lors : $M = P_{C_2,C_1}NP_{B_1,B_2}$. Si on note $M = \mathrm{Mat}_{C_2,B_2}(u)$ et $N = \mathrm{Mat}_{C_1,B_1}(v)$ (possible d'après le théorème d'isomorphisme vu en II.1.) alors

P est une matrice de passage entre deux bases de E et Q entre deux bases de F puisque P est carrée de taille $p = \dim(E)$ et Q est carrée de taille $n = \dim(F)$.

$$\mathrm{Mat}_{C_2,B_2}(u) = P_{C_2,C_1} \times \mathrm{Mat}_{C_1,B_1}(v) \times P_{B_1,B_2}$$

Or, d'après la formule de changement de base vue au paragraphe VI.2 :

$$\mathrm{Mat}_{C_2,B_2}(u) = P_{C_2,C_1} \times \mathrm{Mat}_{C_1,B_1}(u) \times P_{B_1,B_2}$$

On en déduit que

$$P_{C_2,C_1} \times \mathrm{Mat}_{C_1,B_1}(u) \times P_{B_1,B_2} = P_{C_2,C_1} \times \mathrm{Mat}_{C_1,B_1}(v) \times P_{B_1,B_2}$$

Il en découle (en multipliant à gauche et à droite par les inverses des matrices de passage) que $\mathrm{Mat}_{C_1,B_1}(v) = \mathrm{Mat}_{C_1,B_1}(u)$, c'est-à-dire que $N = \mathrm{Mat}_{C_1,B_1}(u)$. Puisque $M = \mathrm{Mat}_{C_2,B_2}(u)$, en conclusion :

Théorème. Deux matrices sont équivalentes lorsqu'elles représentent la même application linéaires dans des bases différentes.

On devrait plutôt dire : dans des couples de bases différents.

Il existe une caractérisation simple pour les matrices équivalentes :

Proposition. Deux matrices sont équivalentes si et seulement si elles ont le même rang. En d'autres termes, les classes d'équivalence de la relation « être équivalentes » sont entièrement déterminées par le rang.

Pour les matrices semblables (cf. paragraphe VII.3), ce sera une condition nécessaire mais pas suffisante !

DÉMONSTRATION. Soient M et N deux matrices équivalentes : il existe donc P et Q inversibles (de taille p et de taille n respectivement) telles que $M = PNQ$. Puisqu'on ne change pas le rang en multipliant par une matrice inversible, alors deux matrices inversibles ont le même rang.

Réciproquement, supposons que M et N aient le même rang. Soit u l'application linéaire représentée par M dans les bases B_E et B_F c'est-à-dire que $M = \mathrm{Mat}_{B_F,B_E}(u)$. D'après ce qu'on a vu plus haut :

$$M = P_{B_F,C_1} \times J_r \times P_{B_1,B_E} \qquad \square$$

et puisqu'une matrice de passage est inversible, alors M et J_r sont équivalentes. Par symétrie des rôles, N et J_r sont équivalentes, et par transitivité, M et N sont équivalentes.

Remarque : Quand on parle de matrices équivalentes, est sous-entendu que les matrices ont la même taille. De plus, comme on l'a dit, on note de la même façon la matrice J_r peu importe la taille, et donc il est sous-entendu qu'on parle de la matrice J_r de la même taille que les autres matrices. Mais en pratique, il n'y aura aucune ambiguïté.

Corollaire. Une matrice est de rang r si et seulement si elle est équivalente à J_r.

Remarque : Par conséquent, lorsqu'on aura une matrice de rang r, on pourra l'écrire directement sous la forme $M = PJ_rQ$ ce qui permettra de prouver facilement certains résultats puisque J_r est une matrice très simple. Donnons une illustration de ce principe en démontrant le résultat admis au paragraphe V.3.c et prouvons qu'une matrice et sa transposée ont même rang.

Soit donc $M \in \mathscr{M}_{n,p}(\mathbb{K})$ de rang r et montrons que $M^\top$ est aussi de rang r. M est de rang r donc est équivalente à J_r (c'est-à-dire la matrice J_r de taille $n \times p$) : il existe $P \in \mathrm{GL}_p(\mathbb{K})$ et $Q \in \mathrm{GL}_n(\mathbb{K})$ telles que $P = PJ_rQ$ donc $M^\top = Q^\top {J_r}^\top P^\top$. Or :

Ci-contre, nous préciserons les tailles des matrices J_r puisque nous manipulerons deux matrices J_r qui n'ont pas la même taille.

$$J_r{}^\top = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & 0 \\ \vdots & \vdots & & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} 1 \\ 0 \\ \vdots \\ \vdots \end{matrix}}\right\} r \\ \vphantom{\begin{matrix} 0 \\ \vdots \\ \vdots \\ 0 \end{matrix}} \end{array} \in \mathscr{M}_{p,n}(\mathbb{K})$$

c'est-à-dire que la transposée de la matrice J_r de taille $n \times p$ est la matrice J_r de taille $p \times n$, qui est également de rang r (toutes les matrices J_r, peu importe leur taille, sont de rang r). Puisque multiplier une matrice par une matrice inversible ne change pas son rang, on en déduit que

$$\begin{aligned} \mathrm{rg}\left(M^\top\right) &= \mathrm{rg}\left({J_r}^\top\right) \\ &= r \\ &= \mathrm{rg}(M) \end{aligned}$$

c'est-à-dire qu'une matrice et sa transposée ont même rang.

Attention de ne pas croire que toutes les propriétés sont gardées par équivalence : par exemple, deux matrices équivalentes n'ont pas forcément la même trace (voir la suite), une matrice équivalente à une matrice nilpotente n'est pas forcément nilpotente (en effet, si A est non inversible, alors $r = \mathrm{rg}(A) \leqslant n - 1$ donc A est équivalente à K_r qui est nilpotente mais A ne l'est pas forcément, prendre $A = J_r$ par exemple).

Remarque : Si $r \leqslant n-1$, notons K_r la matrice suivante

$$K_r = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & 0 \\ \vdots & \vdots & & 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix} 0 \\ 0 \\ \vdots \\ \vdots \end{matrix}}\right\} r \\ \vphantom{\begin{matrix} 0 \\ \vdots \\ 0 \end{matrix}} \end{array} \in \mathscr{M}_n(\mathbb{K})$$

c'est-à-dire que K_r est la matrice J_r dont les coefficients sont décalés d'une rangée vers la droite. Puisque K_r est de rang $r \leqslant n-1$, elle est équivalente à J_r donc à toute matrice de rang r par transitivité de l'équivalence. Dans certains exercices, quand on a une matrice non inversible, donc de rang $r \leqslant n-1$, il peut être intéressant de dire que cette matrice est équivalente à K_r (cf. par exemple exercices 52, 53, 54).

Contrairement à la notation J_r, la notation K_r est personnelle (et donc non universelle).

VII Matrices de passage et changement de base, cas particulier des endomorphismes

On suppose dans cette partie que $\dim(E) = n$.

VII.1 Endomorphismes vs applications linéaires d'un espace dans lui-même

Le titre de ce paragraphe peut prêter à sourire car un endomorphisme n'est rien d'autre qu'une AL d'un espace dans lui-même, mais il faut bien comprendre que dans le cadre des matrices, on emploie dans les deux cas un vocabulaire différent :

- Quand on parle d'une application linéaire (quelconque, i.e. d'un espace dans lui-même, ou non), les bases de départ et d'arrivée sont quelconques, et pas forcément supposées égales (dans le cas où l'espace d'arrivée et l'espace de départ sont les mêmes). Par exemple, on a vu (cf. paragraphe III.1) que la matrice de $\lambda\,\mathrm{Id}_E$ dans les bases $B_E = (e_1, \dots, e_n)$ et $B_{E'} = (e_1, 2e_2, \dots, ne_n)$ était donnée par :

$$\mathrm{Mat}_{B_{E'},B_E}(\lambda\,\mathrm{Id}_E) = \begin{array}{c} \begin{matrix} u(e_1) & u(e_2) & \cdots & u(e_n) \end{matrix} \\ \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda/2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda/n \end{pmatrix} \end{array} \begin{matrix} e_1 \\ 2e_2 \\ \vdots \\ ne_n \end{matrix} \neq \lambda \mathrm{I}_n$$

  Pour éviter toute confusion, quand on parle d'une AL d'un espace dans lui-même, on écrit les bases de départ et d'arrivée c'est-à-dire qu'on parlera par exemple de $u : (E, B_1) \to (E, B_2)$.

- Quand on parle d'un endomorphisme, il est sous-entendu que la base de départ et d'arrivée est la même, c'est-à-dire qu'on s'intéresse uniquement aux applications linéaires $u : (E, B_E) \to (E, B_E)$, et c'est donc le cadre de cette partie. On a par exemple vu (cf. paragraphe III.1) que la matrice de $\lambda\,\mathrm{Id}_E$ (vu comme endomorphisme) est la même dans toutes les bases, c'est-à-dire λI_n (ce n'est pas le cas si on considère $\lambda\,\mathrm{Id}_E$ comme une AL de E dans lui-même, voir ci-dessus). Rappelons d'ailleurs que quand les bases de départ et d'arrivée sont les mêmes, on note les matrices $\mathrm{Mat}_B(u)$ au lieu de $\mathrm{Mat}_{B,B}(u)$. Cela n'empêchera pas de changer de base, mais à chaque fois celle de départ et d'arrivée seront les mêmes.

VII.2 Formule de changement de base

Les endomorphismes étant des cas particuliers d'applications linéaires, les résultats de la partie VI sont encore valables pour des endomorphismes :

Théorème. Soit $u \in \mathscr{L}(E)$. Soient (B, B') deux bases de E. Alors :

$$\mathrm{Mat}_{B'}(u) = P_{B',B} \times \mathrm{Mat}_B(u) \times P_{B,B'}$$

En particulier :

$$\mathrm{Mat}_{B'}(u) = \left(P_{B,B'}\right)^{-1} \times \mathrm{Mat}_B(u) \times P_{B,B'}$$

Moyen mnémotechnique : chaque base est collée à elle-même.

Remarques :

- On peut adapter le diagramme du paragraphe VI.3

$$\begin{tikzcd}
E, B \arrow[d, "u"] & E, B' \arrow[l, "\mathrm{Id}_E"'] \arrow[d, "u"] \\
E, B \arrow[r, "\mathrm{Id}_E"'] & E, B'
\end{tikzcd}$$

et son équivalent matriciel :

$$\begin{tikzcd}
E, B \arrow[d, "\mathrm{Mat}_B(u)"'] & E, B' \arrow[l, "P_{B,B'}"'] \arrow[d, "\mathrm{Mat}_{B'}(u)"] \\
E, B \arrow[r, "P_{B',B}"'] & E, B'
\end{tikzcd}$$

- Là aussi, il y a **deux** matrices de changement de base (et non pas une comme pour les vecteurs), même s'il y a la même base au départ et à l'arrivée. La raison en est simple : on va de B' à B et on revient de B à B', il faut donc deux matrices de passage !
- Encore une fois, c'est ce que je vous demandais de faire dans le chapitre 21.

VII.3 Matrices semblables

Là aussi, cette formule justifie la définition de matrices semblables.

VII.3.a Définition

Définition. Soit $(M, N) \in \mathscr{M}_n(\mathbb{K})^2$. On dit que M est semblable à N s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M = P^{-1}MP$.

Proposition. La relation « être semblable à » (on dit aussi : la similitude) est une relation d'équivalence.

Démonstration. Analogue au cas des matrices équivalentes.

Remarque : Cette relation étant symétrique, on dira donc (le cas échéant) que deux matrices M et N sont semblables.

Remarque : On a vite fait de confondre matrices équivalentes et matrices semblables (d'autant plus que la similitude est une relation d'équivalence). Il est évident que deux matrices semblables sont équivalentes, mais nous verrons que la réciproque est (très) fausse (et c'est bien dommage). En attendant des (contre-)exemples, on peut déjà remarquer qu'il y a des différences :

Nous en verrons d'autres dans la suite.

- Contrairement aux matrices équivalentes, il n'y a qu'une matrice P.

- ⚠ Toutes les matrices sont de même taille (carrées d'ordre n). En particulier, cela n'a aucun sens de parler de matrices semblables pour des matrices qui ne sont pas carrées (alors que cela a du sens pour des matrices équivalentes).
- ⚠ Autant être clair : IL N'Y A PAS DE CARACTÉRISATION SIMPLE POUR SAVOIR SI DEUX MATRICES SONT SEMBLABLES !!! Tout ce qu'on a c'est de (nombreuses) conditions nécessaires. Par exemple :

Savoir si deux matrices sont semblables ou non est en général un problème difficile : cf. paragraphe VII.3.d.

Proposition. Deux matrices semblables ont même rang.

DÉMONSTRATION. Si M et N sont semblables alors il existe P inversible telle que $M = P^{-1}NP$ et on ne change pas le rang en multipliant par une matrice inversible.

Remarque : ⚠ ⚠ ⚠ ⚠ ⚠ ⚠ ⚠ LA RÉCIPROQUE EST FAUSSE !!!!!!!!!

Exemple : Montrons qu'une matrice de la forme λI_n n'est semblable qu'à elle-même. Soit $\lambda \in \mathbb{K}$ et soit M semblable à λI_n. Alors il existe $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M = P^{-1} \times \lambda I_n \times P$ donc $M = \lambda P^{-1}P = \lambda I_n$.

Il faudrait aussi prouver que λI_n est bien semblable à elle-même... ce qui découle de la réflexivité de la similitude (qui est une relation d'équivalence).

Cependant, si $\lambda \neq 0$, alors λI_n est équivalente à $J_n = I_n$ (on travaille avec des matrices carrées), et c'est le cas de toute matrice de rang n donc de toute matrice inversible. En particulier, par transitivité : toute matrice de rang n est équivalente à λI_n. On voit donc :

- que deux matrices équivalentes n'ont aucune raison d'être semblables.
- que deux matrices de même rang n'ont aucune raison d'être semblables. « Avoir le même rang » est une condition nécessaire pour être semblables, mais pas suffisante. Étudier le rang est donc surtout nécessaire pour conclure par la négative. En effet, par contraposée :

Corollaire. Si deux matrices n'ont pas le même rang, alors elles ne sont pas semblables.

Une conséquence immédiate :

Proposition. Si A et B sont semblables alors, pour tout k, A^k et B^k sont semblables.

DÉMONSTRATION. Si A et B sont semblables, alors il existe P inversible telle que $A = P^{-1}BP$. Par une récurrence immédiate, pour tout $k \in \mathbb{N}$, $A^k = P^{-1}B^kP$ ce qui est le résultat voulu.

Remarque : ⚠ Attention, on peut avoir A^2 et B^2 semblables sans avoir A et B semblables ! Par exemple, I_n et $-I_n$ ne sont pas semblables (car une matrice de la forme λI_n n'est semblable qu'à elle-même) alors que leurs carrés sont égaux donc semblables. Là aussi, ce résultat sert surtout par sa contraposée : si deux matrices ont des puissances qui ne sont pas semblables, alors elles ne sont pas semblables. Par exemple, si deux matrices nilpotentes N_1 et N_2 n'ont pas le même indice de nilpotence, alors elles ne sont pas semblables : en effet, si on note p_1 et p_2 les indices de nilpotence de N_1 et N_2, alors on peut supposer sans perte de généralité que $p_1 < p_2$ donc ${N_1}^{p_1} = 0 \neq {N_2}^{p_1}$ donc ${N_1}^{p_1}$ et ${N_2}^{p_2}$ ne sont pas semblables, donc N_1 et N_2 ne sont pas semblables.

⚠ Réciproque fausse ! Voir ci-dessous.

VII.3.b Interprétation

Là aussi, il découle de la formule de changement de base (pour les endomorphismes, donc avec les mêmes bases à l'arrivée et au départ) que deux matrices d'un endomorphismes dans des bases différentes sont semblables, et on prouve de même que dans le paragraphe VI.4 que la réciproque est vraie, c'est-à-dire que si on écrit $M = P^{-1}NP$, alors il existe B, B' deux bases de E et $u \in \mathscr{L}(E)$ tels que $M = \mathrm{Mat}_{B'}(u), P = P_{B,B'}$ et $N = \mathrm{Mat}_B(u)$. En conclusion :

Théorème. Deux matrices sont semblables lorsqu'elles représentent le même **endomorphisme** dans des bases différentes.

Puisque la matrice de $\lambda \operatorname{Id}_E$ est la même dans toutes les bases (c'est-à-dire λI_n), on retrouve le fait que λI_n n'est semblable qu'à elle-même.

Remarque : Il est sous-entendu (puisqu'on parle d'endomorphismes) que les bases sont les mêmes au départ et à l'arrivée. En d'autres termes, M et N sont semblables si et seulement s'il existe un endomorphisme $u \in \mathscr{L}(E)$ et deux bases B et B' de E tels que $M = \operatorname{Mat}_{B'}(u)$ et $N = \operatorname{Mat}_B(u)$. Encore une différence entre matrices équivalentes et matrices semblables : des matrices sont équivalentes lorsqu'elles représentent la même AL dans des couples de bases différents (pas forcément les mêmes bases au départ et à l'arrivée) et deux matrices sont semblables lorsqu'elles représentent le même endomorphisme dans des bases différentes (mais les mêmes au départ et à l'arrivée). On voit que la deuxième condition est beaucoup plus restrictive, on a beaucoup moins de liberté.

VII.3.c Trace d'une matrice et d'un endomorphisme

On a vu que deux matrices semblables ont le même rang. Donnons une autre condition nécessaire de similitude.

Là aussi, cette condition ne sera pas suffisante : puisqu'on vous dit qu'il n'y a pas de caractérisation simple des matrices semblables !

Définition. Soit $A \in \mathscr{M}_n(\mathbb{K})$. On appelle trace de A, notée $\operatorname{tr}(A)$, la somme des coefficients diagonaux de A, c'est-à-dire :

$$\operatorname{tr}(A) = \sum_{i=1}^{n} A_{i,i}$$

Remarque : On ne peut parler de trace que pour une matrice carrée, cela n'a aucun sens de parler de trace lorsque la matrice n'est pas carrée.

Exemple : Si $\lambda \in \mathbb{K}, \operatorname{tr}(\lambda I_n) = n\lambda$.

Proposition. La trace est linéaire, c'est-à-dire que $\operatorname{tr} : \mathscr{M}_n(\mathbb{K}) \to \mathbb{K}$ est une forme linéaire (non nulle).

DÉMONSTRATION. $\rightsquigarrow$ EXERCICE.

Activité : La trace étant une forme linéaire (non nulle puisque $\operatorname{tr}(I_n) = n$), son noyau est un hyperplan. Donnons une base de $\ker(\operatorname{tr})$. Puisque c'est un hyperplan de $\mathscr{M}_n(\mathbb{K})$, il est de dimension $n^2 - 1$: il suffit donc de trouver une famille libre de cardinal $n^2 - 1$. Cherchons déjà parmi les matrices élémentaires (i.e. la base canonique de $\mathscr{M}_n(\mathbb{K})$) : une matrice élémentaire E_i, est de trace nulle si et seulement si l'unique 1 de la matrice n'est pas sur la diagonale, c'est-à-dire si et seulement si $i \neq j$. Par conséquent, les $E_{i,j}$ pour $i \neq j$ forment une famille libre (car sous-famille d'une famille libre, la base canonique) à $n^2 - n$ éléments de $\ker(\operatorname{tr})$.

Toujours préciser qu'une forme linéaire est non nulle avant d'affirmer que son noyau est un hyperplan. C'est parfois plus facile à dire qu'à faire (cf. exercice 37).

Il reste $n - 1$ matrices de trace nulle à trouver : nous n'avons donné que des matrices avec une diagonale nulle. Or, une matrice peut être de trace nulle sans avoir une diagonale nulle (par exemple si elle a un terme égal à 1 et un autre égal à -1 et tous les autres nuls). Il faut donc donner des matrices (les plus simples possibles) de trace nulle avec une diagonale non nulle. Puisqu'une matrice avec un seul terme non nul sur la diagonale n'est pas de trace nulle, cherchons des matrices un peu moins simples, c'est-à-dire avec deux coefficients non nuls sur la diagonale : on pense donc à

$$E_{1,1} - E_{2,2} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, E_{1,1} - E_{3,3} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & -1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \dots, E_{1,1} - E_{n,n} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -1 \end{pmatrix}$$

Si on note B la concaténation des familles $(E_{i,j})_{i\neq j}$ et $(E_{1,1} - E_{i,i})_{i\in[\![2\,;n]\!]}$, montrons que B est une famille libre. Soient $(\lambda_{i,j})_{i\neq j}$ et $(\mu_2, \dots, \mu_n)$ des scalaires et supposons que

$$\sum_{i\neq j} \lambda_{i,j} E_{i,j} + \sum_{i=2}^{n} \mu_i (E_{1,1} - E_{i,i}) = 0$$

En regroupant les $E_{1,1}$, il vient :

$$\sum_{i\neq j} \lambda_{i,j} E_{i,j} - \sum_{i=2}^{n} \mu_i E_{i,i} + \left(\sum_{i=2}^{n} \mu_i\right) E_{1,1} = 0$$

Or, les matrices élémentaires forment la base canonique de $\mathscr{M}_n(\mathbb{K})$ donc une famille libre : on en déduit que tous les coefficients ci-dessus sont nuls. En particulier, tous les $\lambda_{i,j}$ sont nuls, ainsi que les μ_i : la famille est libre. De plus, elle contient $n^2 - n + n - 1 = n^2 - 1$ éléments et $\ker(\mathrm{tr})$ est de dimension $n^2 - 1$ donc c'est une base de $\ker(\mathrm{tr})$.

Remarque : Comme on vient de le voir, une matrice peut avoir une trace nulle sans être nulle (à part quand $n = 1$, mais c'est d'un intérêt limité...). Cependant, le cas de figure suivant se produit régulièrement : montrons que si $A \in \mathscr{M}_n(\mathbb{R})$ vérifie $\mathrm{tr}\left(A \times A^\top\right) = 0$, alors $A = 0$.

En effet :

$$\begin{aligned}
\mathrm{tr}\left(A \times A^\top\right) &= \sum_{i=1}^{n} \left(A \times A^\top\right)_{i,i} \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} A_{i,k} A^\top_{k,i} \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} A_{i,k} A_{i,k} \\
&= \sum_{i=1}^{n} \sum_{k=1}^{n} {A_{i,k}}^2
\end{aligned}$$

De plus, on a une somme de termes positifs donc, si $\mathrm{tr}\left(A \times A^\top\right) = 0$, alors, tous les termes sont nuls donc, pour tous i et k, $A_{i,k} = 0$ c'est-à-dire que A est la matrice nulle. Ce résultat sera particulièrement utile dans le chapitre 34, mais aussi, par exemple, dans l'exercice 37 : si A est non nulle et si on cherche une matrice M telle que $\mathrm{tr}\,(AM) \neq 0$, alors il faut penser à $A^\top$!

Proposition. Soit $(A, B) \in \mathscr{M}_{n,p}(\mathbb{K}) \times \mathscr{M}_{p,n}(\mathbb{K})$. Alors $\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA)$.

Remarque : On a vu qu'on ne peut parler de trace que pour une matrice carrée. On ne peut donc pas parler de $\mathrm{tr}\,(A)$ ni de $\mathrm{tr}\,(B)$, mais AB et BA sont bien carrées et ont la même trace. Cependant, ces deux matrices n'ont pas forcément la même taille : si A est de taille 3×4 et B de taille 4×3, alors AB est carrée de taille 3 et BA carrée de taille 4 (ce qui ne les empêche pas d'avoir la même trace).

DÉMONSTRATION.

$$\begin{aligned}
\operatorname{tr}(AB) &= \sum_{i=1}^{n}(AB)_{i,i} \\
&= \sum_{i=1}^{n}\sum_{k=1}^{p} A_{i,k}B_{k,i} \\
&= \sum_{k=1}^{p}\sum_{i=1}^{n} B_{k,i}A_{i,k} \\
&= \sum_{k=1}^{p}(BA)_{k,k} \\
&= \operatorname{tr}(BA)
\end{aligned}$$

□

Remarque : Attention, cela ne veut pas dire que la trace soit « commutative » ce qui n'aurait d'ailleurs aucun sens. Par exemple, si on a trois matrices A, B, C (disons carrées de taille n pour simplifier les choses, même si on a vu qu'on pouvait parler de trace d'un produit de matrices non carrées), que peut-on dire de $\operatorname{tr}(ABC)$? En posant $M = A$ et $N = BC$, alors il découle du résultat précédent que :

$$\begin{aligned}
\operatorname{tr}(ABC) &= \operatorname{tr}(MN) \\
&= \operatorname{tr}(NM) \\
&= \operatorname{tr}(BCA)
\end{aligned}$$

On montre de même que $\operatorname{tr}(ABC) = \operatorname{tr}(CAB)$. On généralise aisément à un plus grand nombre de matrices : pour faire simple, on peut « permuter circulairement » les matrices, c'est-à-dire passer la matrice de gauche à droite, éventuellement plusieurs fois, mais cela ne veut pas dire qu'on peut écrire les matrices dans l'ordre qu'on veut ! Par exemple, $\operatorname{tr}(ABC) = \operatorname{tr}(BCA) = \operatorname{tr}(CAB)$ mais on n'a pas $\operatorname{tr}(ABC) = \operatorname{tr}(ACB)$ en général !

Exemple :

$$\operatorname{tr}\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\0&0\end{pmatrix}\right) = 1 \qquad \text{et} \qquad \operatorname{tr}\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}1&1\\0&0\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}\right) = 2$$

Remarque : Il ne faut quand même pas rêver : en général, $\operatorname{tr}(A \times B) \neq \operatorname{tr}(A) \times \operatorname{tr}(B)$! Par exemple, $\operatorname{tr}(I_n \times I_n) = n \neq n^2 = \operatorname{tr}(I_n) \times \operatorname{tr}(I_n)$.

Bon, on a supposé sans le dire que $n \geqslant 2$ car la trace a très peu d'intérêt pour des matrices de taille 1...

L'intérêt principal de la trace est qu'elle donne un invariant de similitude plus simple à calculer et qui vérifie plus de propriétés (comme la linéarité) que le rang :

Proposition. Deux matrices semblables ont la même trace.

DÉMONSTRATION. Soient M et N deux matrices semblables. Il existe donc $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $M = P^{-1}NP$. Dès lors, d'après ce qui précède :

$$\begin{aligned}
\operatorname{tr}(M) &= \operatorname{tr}\left(P^{-1}NP\right) \\
&= \operatorname{tr}\left(NPP^{-1}\right) \quad (\operatorname{tr}(AB) = \operatorname{tr}(BA)) \\
&= \operatorname{tr}(N)
\end{aligned}$$

□

Remarque : LA RÉCIPROQUE EST FAUSSE ! ! ! ! ! ! ! ! !

Puisqu'on vous dit qu'il n'y a pas de caractérisation simple de la similitude ! Cependant, on dit que la trace (comme le rang d'ailleurs) est un invariant de similitude.

Exemple : Par exemple, la matrice

$$E_{1,1}-E_{2,2}=\begin{pmatrix}1&0&0&\dots&0\\0&-1&0&\dots&0\\0&0&0&\dots&0\\\vdots&&&\ddots&\vdots\\0&0&0&\dots&0\end{pmatrix}$$

est de trace nulle mais n'est pas semblable à la matrice nulle car est de rang 2 et la matrice nulle est de rang 0 (et on a déjà vu que deux matrices semblables ont même rang, réciproque fausse). On pouvait aussi dire (cf. paragraphe VII.3.a) que la matrice nulle n'est semblable qu'à elle-même. On a deux matrices ayant même trace mais non semblables : là aussi, ce résultat est surtout utile par sa contraposée :

Corollaire. Si deux matrices n'ont pas la même trace, alors elles ne sont pas semblables.

Remarque : Le fait que deux matrices semblables aient même trace permet de pousser encore plus loin le principe du « Matrices/AL : même combat ». En effet, deux matrices étant semblables si et seulement si elles représentent le même endomorphisme dans des bases différentes, on en déduit que toutes les matrices représentant le même endomorphisme ont la même trace. Par conséquent, on peut définir la trace d'un endomorphisme de la façon suivante :

Définition. Soit $u \in \mathscr{L}(E)$. On appelle trace de u, notée $\mathrm{tr}\,(u)$, la trace de toute matrice associée à u.

Exemples :

- $\mathrm{tr}\,(Id_E)=\mathrm{tr}\,(I_n)=n$ (on rappelle que E est supposé de dimension n).
- Si on note D l'endomorphisme dérivation dans la base canonique, on a vu que sa matrice canoniquement associée est

$$A=\begin{array}{c}\begin{array}{cccccc}D(1)&D(X)&D(X^2)&\cdots&\cdots&D(X^n)\end{array}\\ \begin{pmatrix}0&1&0&\cdots&\cdots&0\\0&0&2&&&0\\\vdots&\vdots&0&\ddots&&\vdots\\\vdots&\vdots&\vdots&\ddots&\ddots&\vdots\\\vdots&\vdots&\vdots&&\ddots&n\\0&0&0&\cdots&&0\end{pmatrix}\end{array}\begin{array}{c}\\1\\X\\X^2\\\vdots\\X^{n-1}\\X^n\end{array}$$

Par conséquent, $\mathrm{tr}\,(D)=0$.

Des résultats concernant la trace des matrices, on déduit immédiatement :

Proposition.

- La trace est linéaire, c'est-à-dire que $\mathrm{tr}:\mathscr{L}(E)\to\mathbb{K}$ est une forme linéaire non nulle.
- Pour tous u et v appartenant à $\mathscr{L}(E)$, $\mathrm{tr}\,(v\circ u)=\mathrm{tr}\,(u\circ v)$.

Rappelons que la composition d'endomorphisme est l'analogue du produit matriciel.

Proposition. La trace d'un projecteur est égale à son rang.

DÉMONSTRATION. Découle de la matrice d'un projecteur dans une base bien choisie vue dans le paragraphe III.2.

Remarque : Ce résultat est assez intuitif : un projecteur envoie un espace sur 0 et laisse un autre vecteur invariant. La diagonale (dans une base adaptée) est donc composée de 1 et de 0, et le rang et la trace sont égaux au nombres de 1 donc à la dimension de E_1 (avec les mêmes notations que précédemment).

On a un résultat analogue pour les symétries : toujours d'après le paragraphe III.2, la trace de la symétrie par rapport à E_1 parallèlement à E_2 est $\dim(E_1) - \dim(E_2)$. Puisque $\dim(E_1) + \dim(E_2) = n$ (car ils sont supplémentaires), connaître $\operatorname{tr} s$ permet de donner les dimensions de E_1 et E_2.

VII.3.d Exemples

Il est en général plus facile de prouver que des matrices ne sont pas semblables puisqu'on dispose de conditions NÉCESSAIRES pour que deux matrices sont semblables : elles doivent avoir le même rang et la même trace. Ainsi, si elles n'ont pas le même rang ou n'ont pas la même trace, alors elles ne sont pas semblables. Donnons pour débuter des matrices dont on peut affirmer directement qu'elles ne sont PAS semblables.

On rajoutera le déterminant à la liste des conditions nécessaires dans le chapitre 33.

Exemples :

- On a vu qu'une matrice nilpotente de taille 3 est soit semblable à la matrice nulle (lorsque son indice de nilpotence vaut 1), soit à la matrice $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ (lorsque son indice de nilpotence vaut 2), soit à la matrice $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ (lorsque son indice de nilpotence vaut 3), et ces deux matrices sont deux à deux non semblables car n'ont pas le même rang. Par conséquent, il y a trois classes de similitudes pour les matrices nilpotentes de taille 3.
- Les matrices
$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
ne sont pas semblables alors qu'elles ont même rang puisqu'elles n'ont pas la même trace.

Cependant, ce n'est pas toujours aussi simple : parfois, certaines matrices ont même rang et même trace et ne sont tout de même pas semblables. Il faut alors improviser, chercher « quelque-chose qui est vrai pour l'une et qui n'est pas vrai pour l'autre » (par exemple une équation vérifiée par la matrice, une puissance nulle etc.).

Exemples :

- Les matrices
$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
ne sont pas semblables alors qu'elles ont même rang et même trace. En effet, $A^2 \neq 0$ et $B^2 = 0$. Si elles sont semblables, il existe $P \in \mathrm{GL}_4(\mathbb{K})$ tel que $A = P^{-1}BP$ donc $A^2 = P^{-1}B^2P = 0$ ce qui est absurde.

- Attention, si deux matrices nilpotentes n'ont pas le même indice de nilpotence, elles ne sont pas semblables (voir cii-dessus), mais, contrairement à ce que les deux exemples ci-dessus peuvent laisser croire, la réciproque est fausse : ce n'est pas parce que deux matrices sont nilpotentes et ont même indice de nilpotence qu'elles sont semblables. Par exemple, montrons que

$$N_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad N_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

ne sont pas semblables alors qu'elles ont même rang (4), même trace (0) et même indice de nilpotence (3). Un calcul simple (il est plus rapide de passer par les applications linéaires associées que faire le calcul à la main, exo) donne :

$${N_1}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad {N_2}^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

qui ne sont pas semblables car elles n'ont pas le même rang, donc N_1 et N_2 ne sont pas semblables.

L'exemple précédent est plutôt difficile et on vous aidera dans ce cas de figure. Par conséquent, quand on se demande si deux matrices sont semblables, on regarde le rang, la trace (plus tard : le déterminant), on essaye de trouver une propriété vraie pour l'une et pas pour l'autre etc.

Si on ne trouve rien : c'est peut-être qu'elles sont semblables. Là, la situation est « plus simple » (enfin, si on veut) : le seul moyen de prouver qu'elles sont en effet semblables est de prouver qu'elles vérifient la définition : soit exhiber une matrice P qui convient, soit trouver deux bases dans lesquelles elles représentent le même endomorphisme u. C'est plus simple car il n'y a pas à se demander quelle méthode utiliser... mais c'est parfois plus compliqué car trouver la matrice P ou les deux bases et l'endomorphisme u est parfois difficile.

On s'en sort en général en traduisant les propriétés vérifiées par u et la base initiale, en traduisant les propriétés que doivent vérifier u et la base finale, et en cherchant des vecteurs vérifiant ces dernières propriétés. Parfois (dans les exercices les plus simples), il suffit de permuter les vecteurs de la base, mais il faut aussi parfois prendre des vecteurs dans le noyau (quand on cherche à avoir des colonnes nulles) ou des vecteurs qui sont des images des autres.

Exemples :

- Les matrices

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad \tilde{A} = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}$$

sont semblables. En effet, si on note $B = (e_1, e_2)$ la base canonique de $\mathbb{K}^2$ et u l'endomorphisme de $\mathbb{K}^2$ canoniquement associé à A alors :

$$A = \begin{array}{c} \begin{array}{cc} u(e_1) & u(e_2) \end{array} \\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \end{array}$$

c'est-à-dire que $u(e_1) = 0$ et $u(e_2) = e_1$, et on cherche une base $B' = (\varepsilon_1, \varepsilon_2)$ telle que

$$\tilde{A} = \begin{array}{c} \begin{array}{cc} u(\varepsilon_1) & u(\varepsilon_2) \end{array} \\ \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \varepsilon_1 \\ \varepsilon_2 \end{array}$$

c'est-à-dire telle que $u(\varepsilon_1) = 2\varepsilon_2$ et $u(\varepsilon_2) = 0$. On pense donc à prendre $\varepsilon_2 = e_1$ (on veut un vecteur dans le noyau) et on cherche donc ε_1 tel que $u(\varepsilon_1) = 2e_1$: on pense donc à prendre $\varepsilon_1 = 2e_2$. On prouve aisément que $B' = (2e_2, e_1)$ est une base, et la matrice de u dans cette base est $\tilde{A}$, donc les matrices sont bien semblables.

- Les matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{et} \qquad A' = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

sont semblables. En effet, si on note $B = (e_1, e_2, e_3, e_4)$ la base canonique de $\mathbb{K}^4$ et u l'endomorphisme canoniquement associé à A, alors on a :

$$A = \begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_2) & u(e_3) & u(e_4) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \\ e_4 \end{array}$$

c'est-à-dire que $u(e_1) = 0, u(e_2) = e_1, u(e_3) = 0$ et $u(e_4) = e_3$, et on cherche une base $B' = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ telle que

$$\begin{array}{c} \begin{array}{cccc} u(\varepsilon_1) & u(\varepsilon_2) & u(\varepsilon_3) & u(\varepsilon_4) \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{array} = A'$$

c'est-à-dire que $u(\varepsilon_1) = u(\varepsilon_2) = 0$, $u(\varepsilon_3) = \varepsilon_1$ et $u(\varepsilon_4) = \varepsilon_2$. On commence par chercher deux éléments (libres) du noyau : e_1 et e_3 conviennent donc on pose $\varepsilon_1 = e_1, \varepsilon_2 = e_3$. On cherche ensuite un vecteur $\varepsilon 3$ d'image $\varepsilon_1 = e_1$ donc $\varepsilon_3 = e_2$ convient. De même, on cherche un vecteur ε_4 d'image $\varepsilon_2 = e_3$ donc $\varepsilon_4 = e_4$ convient. Par conséquent, en intervertissant simplement les vecteurs de la base, la matrice de u dans la base $B' = (e_1, e_3, e_2, e_4)$ est :

Méthode classique à retenir ! Mais attention : quand on change l'ordre des vecteurs, il faut le faire en haut et à droite !

$$\begin{array}{c} \begin{array}{cccc} u(e_1) & u(e_3) & u(e_2) & u(e_4) \end{array} \\ \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_3 \\ e_2 \\ e_4 \end{array} = A'$$

- Montrer que les matrices

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \qquad \text{et} \qquad \tilde{A} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

sont semblables. Si on note $B = (e_1, e_2, e_3)$ la base canonique de $\mathbb{K}^3$ et u l'endomorphisme canoniquement associé à A, alors on a :

$$A = \begin{array}{c} \begin{array}{ccc} u(e_1) & u(e_2) & u(e_3) \end{array} \\ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \end{array}$$

On cherche une base $B' = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$ telle que

$$\tilde{A} = \begin{array}{c} \begin{array}{ccc} u(\varepsilon_1) & u(\varepsilon_2) & u(\varepsilon_3) \end{array} \\ \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{array}$$

donc une base vérifiant $u(\varepsilon_1) = 6\varepsilon_1, u(\varepsilon_2) = u(\varepsilon_3) = 0$. En particulier, ε_2 et ε_3 doivent être deux éléments (libres) du noyau. On remarque immédiatement (mais si on ne le voit pas, on cherche une base du noyau) que $u(e_2) = 2u(e_1)$ et $u(e_3) = 3u(e_1)$ donc $e_2 - 2e_1$ et $e_3 - 3e_1$ sont deux éléments du noyau, et on montre aisément qu'ils sont libres. On cherche enfin un élément ε_1 (non nul) tel que $u(\varepsilon_1) = 6\varepsilon_1$. On peut prendre un vecteur $\varepsilon_1 = (x, y, z)$ et donner une solution (non nulle) du problème, mais on peut aussi remarquer que $\varepsilon_1 = (1, 1, 1)$ convient (solution évidente ?). Par conséquent, les vecteurs $\varepsilon_1 = (1, 1, 1), \varepsilon_2 = (-2, 1, 0)$ et $\varepsilon_3 = (-3, 0, 1)$ sont de bons candidats... Il suffit de prouver qu'ils forment une famille libre à 3 éléments, donc une base, pour conclure : la matrice de u dans cette base est la matrice $\tilde{A}$, donc A et $\tilde{A}$ représentent le même endomorphisme dans des bases différentes, d'où le résultat.

- Soit $N \in \mathscr{M}_3(\mathbb{K})$ telle que $N^3 = 0 \neq N^2$. Montrer que N est semblable à

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

Raisonnement tout à fait analogue à celui du paragraphe III.2 : si u est un endomorphisme représenté par u, u est nilpotent d'indice 3 donc il existe x_0 tel que $u^2(x_0) \neq 0$ et on prouve comme précédemment (famille libre à 3 éléments donc base) que la matrice de u dans la base $(u^2(x_0), u(x_0), x_0)$ est la matrice A. A et N représentent donc le même endomorphisme dans des bases différentes donc A et N sont semblables.

- On a vu que toute matrice de projecteur (respectivement toute matrice de symétrie) est semblable à une matrice diagonale avec des 1 et des 0 (respectivement des 1 et des -1) sur la diagonale.

- Montrer que les matrices (carrées de taille n)

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & & & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & n-1 \\ 0 & 0 & 0 & \cdots & & 0 \end{pmatrix} \qquad \text{et} \qquad \widetilde{A} = \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & & & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & 0 & \cdots & & 0 \end{pmatrix}$$

sont semblables. Si on note $B = (e_1, \dots, e_n)$ la base canonique de $\mathbb{K}^n$, alors

$$A = \begin{array}{c} \begin{array}{cccccc} u(e_1) & u(e_2) & u(e_3) & \cdots & \cdots & u(e_n) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 2 & & & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & & \ddots & n-1 \\ 0 & 0 & 0 & \cdots & & 0 \end{pmatrix} \end{array} \begin{array}{c} \\ e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_{n-1} \\ e_n \end{array}$$

On cherche donc une base $\tilde{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n)$ dans laquelle la matrice de u serait $\tilde{A}$. On cherche donc une base $\tilde{B} = (\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_n)$ pour laquelle on aurait :

- ⋆ $f(\varepsilon_1) = 0$: $\varepsilon_1 = e_1$ convient.
- ⋆ $f(\varepsilon_2) = \varepsilon_1$. Puisque $f(e_2) = e_1$, alors $\varepsilon_2 = e_2$ convient.
- ⋆ $f(\varepsilon_3) = \varepsilon_2$. Puisque $f(e_3) = 2e_2$, alors, par linéarité de f, $f(e_3/2) = e_2 = \varepsilon_2$ donc $\varepsilon_3 = e_3/2$ convient.
- ⋆ $f(\varepsilon_4) = \varepsilon_3 = e_3/2$. Puisque $f(e_4) = 3e_3$, alors, par linéarité de f, $f(e_4/6) = e_3/2 = \varepsilon_3$ donc $\varepsilon_4 = e_4/6$ convient.

Ainsi, on pense à prendre, pour tout $i \in [\![1 ; n]\!]$, $\varepsilon_i = e_i/(i-1)!$. Par définition, et par linéarité de f,

$$\begin{aligned} f(\varepsilon_i) &= \frac{1}{(i-1)!} f(e_i) \\ &= \frac{1}{i-1} \times \frac{1}{(i-2)!} \times (i-1)e_{i-1} \\ &= \frac{1}{(i-2)!} e_{i-1} \\ &= \varepsilon_{i-1} \end{aligned}$$

On montre aisément que B' est une base de E, A et $\tilde{A}$ représentent le même endomorphisme dans des bases différentes donc sont semblables. On peut le vérifier en remarquant que :

$$P_{B,B'} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1/2 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1/(n-1)! \end{pmatrix} \quad \text{et} \quad P_{B',B} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & (n-1)! \end{pmatrix}$$

Un calcul simple donne alors $P_{B',B} A P_{B,B'} = \tilde{A}$.

- Un avant-goût du programme de deuxième année : soit

$$A = \begin{pmatrix} -4 & -4 & 1 \\ 3 & 1 & 1 \\ 0 & -6 & 5 \end{pmatrix}$$

Alors on prouve comme au paragraphe V.2 que $\text{rg}(A - \lambda I_3) = 3$ sauf pour $\lambda = -1, 1$ et 2. En faisant comme au chapitre 21, on trouve trois vecteurs X_1, X_2, X_3 tels que $AX_1 = -X_1, AX_2 = X_2$ et $AX_3 = X_3$ et on montre aisément que $B = (X_1, X_2, X_3)$ est une base de $\mathbb{K}^3$ (il suffit de prouver que c'est une famille libre car c'est une famille à trois éléments en dimension 3). Un calcul simple donne alors :

$$P_{B'B} A P_{B,P'} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

si bien que A et la matrice diagonale ci-dessus sont semblables : on dit que A est diagonalisable.

Puisqu'il y a le bon nombre de vecteurs, il suffit de prouver (au choix) que la famille est libre ou génératrice. Pour prouver qu'elle est libre, on se donne une combinaison linéaire $\lambda_1 \varepsilon_1 + \dots + \lambda_n \varepsilon_n = 0$ donc $\lambda_1 e_1 + \dots + \lambda_n e_n/(n-1)! = 0$ et la famille $(e_1, \dots, e_n)$ est libre donc tous les $\lambda_i/(i-1)!$ sont nuls donc tous les λ_i sont nuls donc la famille est libre. Pour prouver qu'elle est génératrice, il suffit de voir que, pour tout x, il existe $(\lambda_1, \dots, \lambda_n)$ tel que $x = \lambda_1 e_1 + \dots + \lambda_n e_n$ (car la famille $(e_1, \dots, e_n)$ est génératrice) donc $x = \lambda_1 \varepsilon_1 + \dots + (n-1)! \lambda_n \varepsilon_n$ ce qui permet de conclure.