CALCUL MATRICIEL

I. Opérations matricielles

a. L'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$.

Soient n et p deux entiers naturels non nuls. Une **matrice** à n lignes et p colonnes (ou "matrice de format (n,p)") à coefficients dans $\mathbb{K}$ est un tableau de np scalaires habituellement repérés par une double indexation: $A = (a_{i,j})$ avec $1 \leq i \leq n$, $1 \leq j \leq p$. Le nombre i est l'indice de ligne, j est l'indice de colonne.

De façon naturelle, on peut ajouter deux matrices de même format, et multiplier une matrice par un scalaire, ainsi: si $A = (a_{i,j})$ et $B = (b_{i,j})$ et $\lambda \in \mathbb{K}$, alors

$$A + B = (a_{i,j} + b_{i,j}) \qquad ; \qquad \lambda A = (\lambda a_{i,j}) \ .$$

Il est immédiat de vérifier alors les axiomes de la structure d'espace vectoriel. Plus précisément, l'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ est muni d'une structure de $\mathbb{K}$-espace vectoriel de dimension np, une base étant constituée des **matrices élémentaires** $E_{k,l}$ $(1 \leq k \leq n,\ 1 \leq l \leq p)$: la matrice $E_{k,l}$ a tous ses coefficients nuls sauf celui d'indices (k,l) qui vaut 1. On peut écrire aussi, en utilisant le **symbole de Kronecker**, que $E_{k,l}$ est la matrice de format (n,p) dont le coefficient d'indices (i,j) vaut $\delta_{i,k}\delta_{j,l}$. Si $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$, on peut écrire $A = \sum_{i,j} a_{i,j} E_{i,j}$, la somme étant prise sur tous les couples $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$. On peut aussi écrire $A = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j} E_{i,j}$ ou $A = \sum_{j=1}^{p}\sum_{i=1}^{n} a_{i,j} E_{i,j}$.

b. Produit matriciel

Soient $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$. On définit une matrice $C = (c_{i,k})$, notée $C = AB$, appelée **produit** de A et B, dans $\mathcal{M}_{n,q}(\mathbb{K})$, par

$$\forall (i,k) \in [\![1,n]\!] \times [\![1,q]\!] \qquad c_{i,k} = \sum_{j=1}^{p} a_{i,j} b_{j,k} \ .$$

Disposition du calcul:

Remarque. On ne peut définir le produit AB que si le nombre de colonnes de A est égal au nombre de lignes de B. On mémorisera alors la relation, que l'on peut appeler **règle des dominos** :

$$(n,p) \times (p,q) = (n,q) \ .$$

De cette formule de calcul, on déduit les propriétés suivantes :

▷ Si $X \in \mathcal{M}_{p,1}(\mathbb{K}) \simeq \mathbb{K}^p$ est une matrice-colonne et $A \in \mathcal{M}_{n,p}(\mathbb{K})$, alors $AX \in \mathcal{M}_{n,1}(\mathbb{K}) \simeq \mathbb{K}^n$ est une combinaison linéaire des colonnes de A. Précisément, si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, alors $AX = \sum_{j=1}^{p} x_j C_j(A)$ où, pour tout $j \in [\![1,p]\!]$, on a noté $C_j(A) = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{n,j} \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K}) \simeq \mathbb{K}^n$.

Exemple. $\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} a + 2b + 3c \\ d + 2e + 3f \end{pmatrix} = C_1(A) + 2\,C_2(A) + 3\,C_3(A).$

▷ La k-ème colonne de AB est le produit de A par la k-ème colonne de B, soit

$$C_k(AB) = A \cdot C_k(B) \qquad (1 \leq k \leq q) \ .$$

▷ La i-ème ligne de AB est le produit de la i-ème ligne de A par B, soit

$$L_i(AB) = L_i(A) \cdot B \qquad (1 \leq i \leq n) \ ,$$

tout cela en supposant toujours A de format (n, p) et B de format (p, q).

Propriété. La multiplication matricielle est bilinéaire, c'est-à-dire distributive (à gauche et à droite) par rapport aux combinaisons linéaires. Autrement dit, si $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $A' \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$, $B' \in \mathcal{M}_{p,q}(\mathbb{K})$, $\lambda \in \mathbb{K}$, $\mu \in \mathbb{K}$, alors

$$(\lambda A + \mu A')\, B = \lambda\, AB + \mu\, A'B \qquad \text{et} \qquad A(\lambda B + \mu B') = \lambda\, AB + \mu\, AB' \ .$$

Propriété. La multiplication matricielle est associative. Autrement dit, si $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{p,q}(\mathbb{K})$, et $C \in \mathcal{M}_{q,r}(\mathbb{K})$, alors $A(BC) = (AB)C$, que l'on peut écrire sans ambiguïté ABC.

Notons aussi l'associativité mixte $(\lambda A)B = A(\lambda B) = \lambda(AB)$, qui permet d'écrire λAB sans ambiguïté.

Attention! Le produit de deux matrices peut être nul sans qu'aucune des deux matrices ne soit nulle! Par exemple, avec des matrices élémentaires de formats compatibles, $E_{1,1}E_{2,2} = 0$.

c. L'ensemble $\mathcal{M}_n(\mathbb{K})$.

On note $\mathcal{M}_n(\mathbb{K})$ pour $\mathcal{M}_{n,n}(\mathbb{K})$, l'espace vectoriel (de dimension n^2) des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$. La multiplication matricielle est alors une loi interne associative dans $\mathcal{M}_n(\mathbb{K})$. Elle possède un élément neutre, la matrice identité (ou matrice unité) I_n: on a $AI_n = I_nA = A$ pour tout $A \in \mathcal{M}_n(\mathbb{K})$.

Il est utile de connaître les règles de multiplication des matrices élémentaires $E_{i,j}$, que l'on peut aussi appeler **règle des dominos** :

$$\forall (i, j, k, l) \in [\![1, n]\!]^4 \qquad E_{i,j}E_{k,l} = \delta_{j,k}\, E_{i,l} \ .$$

On peut retrouver cela facilement en introduisant les matrices-colonnes E_k:

$$E_k = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}^\top \in \mathcal{M}_{n,1}(\mathbb{K}) \simeq \mathbb{K}^n$$

(le coefficient 1 étant en k-ème position). On peut dire que E_k est le k-ème vecteur de la base canonique de $\mathbb{K}^n$. On a alors facilement $E_{i,j} = E_iE_j^\top$, tandis que $E_i^\top E_j$ est un scalaire que l'on peut interpréter (si $\mathbb{K} = \mathbb{R}$) comme le produit scalaire (canonique) des vecteurs E_i et E_j, donc $E_i^\top E_j = \delta_{i,j}$. Du coup, par associativité du produit matriciel,

$$E_{i,j}E_{k,l} = (E_iE_j^\top)\,(E_kE_l^\top) = E_i\,(E_j^\top E_k)\,E_l^\top = E_i\,(\delta_{j,k})\,E_l^\top = \delta_{j,k}\,E_iE_l^\top = \delta_{j,k}\,E_{i,l} \ .$$

Attention! Le produit matriciel dans $\mathcal{M}_n(\mathbb{K})$ n'est pas commutatif! Par exemple, $E_{1,2}E_{2,1} = E_{1,1}$ alors que $E_{2,1}E_{1,2} = E_{2,2}$.

Définition. Une matrice de la forme $A = \lambda I_n$ avec $\lambda \in \mathbb{K}$ est appelée **matrice scalaire** car, du point de vue du produit matriciel, elle se comporte comme un scalaire: si $M \in \mathcal{M}_n(\mathbb{K})$, on a $AM = (\lambda I_n)M = \lambda M = M(\lambda I_n) = MA$. En particulier, ces matrices commutent avec toutes les autres matrices de $\mathcal{M}_n(\mathbb{K})$.

Puissances d'une matrice carrée. Si $A \in \mathcal{M}_n(\mathbb{K})$, on définit A^p pour tout p entier naturel de façon récursive, par

$$A^0 = I_n \qquad \text{et} \qquad \forall p \in \mathbb{N} \quad A^{p+1} = A\,A^p\ .$$

De l'associativité du produit matriciel, on déduit les propriétés usuelles $A^{p+q} = A^p\,A^q$ et $(A^p)^q = A^{pq}$.

Définition. Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est dite **nilpotente** s'il existe un entier naturel k tel que $A^k = 0_n$. De telles matrices existent: par exemple, $E_{1,2}^2 = 0_n$.

Si les matrices A et B **commutent** ($AB = BA$), alors $(AB)^p = A^p\,B^p$.

Si A et B commutent toujours, alors on a la **formule du binôme de Newton** :

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

Propriété. Le sous-ensemble de $\mathcal{M}_n(\mathbb{K})$ constitué des **matrices diagonales** est stable par le produit matriciel. De plus, la multiplication de ces matrices est commutative. Ce sous-ensemble $\mathcal{D}_n(\mathbb{K})$ est aussi un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$, de dimension n.

Propriété. Le sous-ensemble de $\mathcal{M}_n(\mathbb{K})$ constitué des **matrices triangulaires supérieures** est stable par le produit matriciel. Ce sous-ensemble $\mathcal{T}_n^+(\mathbb{K})$ est aussi un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension $\dfrac{n(n+1)}{2}$. Même chose pour les matrices triangulaires inférieures.

d. Matrices carrées inversibles.

Définition. Une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est dite **inversible** s'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = BA = I_n$.

On montre que, si tel est le cas, la matrice B introduite ci-dessus est unique, on pose alors $B = A^{-1}$ et on dit que B est l'**inverse** de A. Si A est inversible, alors A^{-1} l'est aussi et $(A^{-1})^{-1} = A$.

Propriété. Si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$ inversibles, alors la matrice AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$. Plus généralement, si $A_1, \cdots, A_k$ sont des matrices carrées d'ordre n inversibles, alors la matrice-produit $A_1 \cdots A_k$ est inversible et

$$(A_1 \cdots A_k)^{-1} = A_k^{-1} \cdots A_1^{-1}\ .$$

Définition. L'ensemble des matrices carrées d'ordre n inversibles est appelé **groupe linéaire**, et noté $\mathrm{GL}_n(\mathbb{K})$.

Propriété. Soient A et B deux matrices carrées d'ordre n. Si on a $AB = I_n$, alors on a aussi $BA = I_n$, donc A est inversible et $B = A^{-1}$. Autrement dit, pour qu'une matrice A soit inversible, il suffit qu'elle soit “inversible à droite” (ou “inversible à gauche”).

En effet, l'application $\varphi : M \mapsto BM$ est un endomorphisme de l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$, injectif (si $BM = 0$, alors $ABM = 0$ soit $M = 0$) donc surjectif (dimension finie) et il existe alors $A' \in \mathcal{M}_n(\mathbb{K})$ telle que $\varphi(A') = I_n$, soit $BA' = I_n$. Reste à prouver $A = A'$. Mais on a, par associativité,

$$A = AI_n = A(BA') = (AB)A' = I_nA' = A' \ .$$

Exemple des matrices diagonales ou triangulaires.

La matrice diagonale $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$ est inversible *si et seulement si* les coefficients diagonaux λ_i sont tous non nuls, et on a a alors $D^{-1} = \mathrm{diag}\left(\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}\right)$.

Une matrice triangulaire supérieure $T = (t_{i,j})$ est inversible *si et seulement si* ses coefficients diagonaux $t_{i,i}$ sont tous non nuls, et dans ce cas la matrice inverse T^{-1} est aussi triangulaire supérieure, de coefficients diagonaux $\frac{1}{t_{i,i}}$, $1 \leq i \leq n$.

Exercice. Le démontrer!

e. Transposition.

Définition. Si $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ est une matrice, on peut définir une matrice $B = (b_{i,j}) \in \mathcal{M}_{p,n}(\mathbb{K})$ par la relation $b_{i,j} = a_{j,i}$ pour tout $(i,j) \in [\![1,p]\!] \times [\![1,n]\!]$. On dit que B est la **transposée** de A, et on note $B = A^\top$.

Linéarité de la transposition. Si A et B sont dans $\mathcal{M}_{n,p}(\mathbb{K})$, si α et β sont deux scalaires, alors

$$(\alpha A + \beta B)^\top = \alpha\, A^\top + \beta\, B^\top \ .$$

Proposition. Si $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$, alors

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

Proposition. Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice inversible, alors sa transposée $A^\top$ est aussi inversible, et

$$(A^\top)^{-1} = (A^{-1})^\top \ .$$

Définition. Une matrice carrée $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ est dite **symétrique** lorsqu'elle vérifie $A^\top = A$, autrement dit si $a_{j,i} = a_{i,j}$ pour tout $(i,j) \in [\![1,n]\!]^2$. Elle est dite **antisymétrique** lorsque $A^\top = -A$, autrement dit si $a_{j,i} = -a_{i,j}$ pour tout couple (i,j).

Proposition. L'ensemble $\mathcal{S}_n(\mathbb{K})$ des matrices carrées d'ordre n symétriques est un s.e.v. de $\mathcal{M}_n(\mathbb{K})$ de dimension $\frac{n(n+1)}{2}$. L'ensemble $\mathcal{A}_n(\mathbb{K})$ des matrices carrées d'ordre n antisymétriques est un s.e.v. de $\mathcal{M}_n(\mathbb{K})$ de dimension $\frac{n(n-1)}{2}$. Ces deux s.e.v. sont supplémentaires dans $\mathcal{M}_n(\mathbb{K})$:

$$\mathcal{M}_n(\mathbb{K}) = \mathcal{S}_n(\mathbb{K}) \oplus \mathcal{A}_n(\mathbb{K}) \ .$$

Plus précisément, si $M \in \mathcal{M}_n(\mathbb{K})$, la seule façon de décomposer M en $M = S + A$, avec S symétrique et A antisymétrique, est de poser

$$S = \frac{1}{2}(M + M^\top) \qquad \text{et} \qquad A = \frac{1}{2}(M - M^\top) \ .$$

Les matrices S et A sont respectivement appelées “partie symétrique” et “partie anti-symétrique” de la matrice A.

II. Matrices et applications linéaires

a. Matrice d’un vecteur, d’une famille finie de vecteurs dans une base.

• Si $\mathcal{B} = (e_1, \cdots, e_n)$ est une base d’un espace vectoriel E de dimension n, si x est un vecteur de E, ce vecteur se décompose de façon unique dans la base $\mathcal{B}$ sous la forme $x = \sum_{i=1}^{n} x_i e_i$, les scalaires x_i étant les coordonnées du vecteur x dans la base $\mathcal{B}$. La matrice-colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$ est appelée **matrice du vecteur x dans la base $\mathcal{B}$**, et elle est notée $X = \mathrm{Mat}_{\mathcal{B}}(x)$.

• Si $\mathcal{X} = (x_1, \cdots, x_p)$ est une famille de p vecteurs de E (*attention au changement de notations!*), chaque vecteur x_j $(1 \leq j \leq p)$ se décompose en $x_j = \sum_{i=1}^{n} a_{i,j} e_i$ dans la base $\mathcal{B}$. La matrice $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ est appelée **matrice de la famille de vecteurs $\mathcal{X}$ dans la base $\mathcal{B}$**, et elle est notée $A = \mathrm{Mat}_{\mathcal{B}}(\mathcal{X})$. Pour tout $j \in [\![1, p]\!]$, sa j-ème colonne est $X_j = \mathrm{Mat}_{\mathcal{B}}(x_j)$.

b. Correspondance entre matrice et application linéaire.

Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions p et n respectivement, soit $\mathcal{B} = (e_1, \cdots, e_p)$ une base de E, soit $\mathcal{C} = (f_1, \cdots, f_n)$ une base de F. Si u est une application linéaire de E vers F, on peut construire la matrice de la famille $u(\mathcal{B}) = \big(u(e_1), \cdots, u(e_p)\big)$ de vecteurs de F dans la base $\mathcal{C}$ de cet espace vectoriel. Cette matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est aussi appelée **matrice de l’application linéaire u relativement aux bases $\mathcal{B}$ et $\mathcal{C}$**, et on note $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) = \mathrm{Mat}_{\mathcal{C}}\big(u(\mathcal{B})\big)$. On observe ainsi sur les colonnes de A les images par u des vecteurs de la base $\mathcal{B}$, décomposés dans la base $\mathcal{C}$:

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

Proposition. L’application $\Phi_{\mathcal{B},\mathcal{C}} : u \mapsto \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ est un isomorphisme entre les espaces vectoriels $\mathcal{L}(E, F)$ et $\mathcal{M}_{n,p}(\mathbb{K})$, autrement dit il y a une correspondance bijective entre les applications linéaires de E vers F, et les matrices de format (n, p), ceci bien sûr une fois que l’on a fait le choix d’une base $\mathcal{B}$ de E et d’une base $\mathcal{C}$ de F. De plus, on a

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(\alpha u + \beta v) = \alpha\, \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) + \beta\, \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(v) \ .$$

Conséquence. On a $\dim\big(\mathcal{L}(E, F)\big) = \dim\big(\mathcal{M}_{n,p}(\mathbb{K})\big) = np = (\dim E) \times (\dim F)$.

Proposition. Avec les mêmes notations, soit $x = \sum_{j=1}^{p} x_j e_j$ un vecteur de E, il est alors représenté dans la base $\mathcal{B}$ de E par la matrice-colonne $X = \mathrm{Mat}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$.

Son image $y = u(x)$ par u est alors le vecteur de F représenté dans la base $\mathcal{C}$ par la matrice-colonne $Y = AX \in \mathcal{M}_{n,1}(\mathbb{K})$. On a donc la relation

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

En effet,

$$\mathrm{Mat}_{\mathcal{C}}\big(u(x)\big) = \mathrm{Mat}_{\mathcal{C}}\Big(\sum_{j=1}^{p} x_j \, u(e_j)\Big) = \sum_{j=1}^{p} x_j \, \mathrm{Mat}_{\mathcal{C}}\big(u(e_j)\big) = \sum_{j=1}^{p} x_j \, C_j(A) = AX \ .$$

La multiplication d'une matrice par une matrice-colonne correspond donc au calcul des coordonnées de l'image d'un vecteur par une application linéaire.

Remarque idiote (quoique). Si deux matrices A et B, de même format (n,p), vérifient $\forall X \in \mathcal{M}_{p,1}(\mathbb{K}) \quad AX = BX$, alors $A = B$. *On peut le justifier en disant qu'elles représentent alors la même application linéaire de $\mathbb{K}^p$ vers $\mathbb{K}^n$. On peut aussi dire que, si l'on note $(E_1, \cdots, E_p)$ la base canonique de $\mathcal{M}_{p,1}(\mathbb{K}) \simeq \mathbb{K}^p$, on a alors $AE_j = BE_j$ pour tout $j \in [\![1,p]\!]$, soit $C_j(A) = C_j(B)$, on déduit l'égalité des matrices A et B colonne par colonne.*

Proposition. Soient E, F, G trois $\mathbb{K}$-espaces vectoriels de dimension finie, munis respectivement des bases $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$. Soient $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$, on a alors $v \circ u \in \mathcal{L}(E,G)$ et

$$\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(v \circ u) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(v) \cdot \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) \ .$$

En résumé, le produit des matrices correspond à la composition des applications linéaires.

Diagramme.

Proposition. Soient E et F deux e.v. de même dimension n, soit $u \in \mathcal{L}(E,F)$, soit $\mathcal{B}$ une base de E, soit $\mathcal{C}$ une base de F. Alors u est un isomorphisme *si et seulement si* la matrice $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) \in \mathcal{M}_n(\mathbb{K})$ est inversible. Dans ce cas, on a alors $A^{-1} = \mathrm{Mat}_{\mathcal{C},\mathcal{B}}(u^{-1})$.

Cas particulier. Si E est un e.v. de dimension n, si $u \in \mathcal{L}(E)$ est un endomorphisme de E, si $\mathcal{B}$ est une base de E, la matrice $A = \mathrm{Mat}_{\mathcal{B},\mathcal{B}}(u) \in \mathcal{M}_n(\mathbb{K})$, plus simplement notée $\mathrm{Mat}_{\mathcal{B}}(u)$, est appelée **matrice de l'endomorphisme u relativement à la base $\mathcal{B}$**. Cette matrice A est inversible *si et seulement si* u est un automorphisme de E, i.e. $u \in \mathrm{GL}(E)$. Dans ce cas, $A^{-1} = \mathrm{Mat}_{\mathcal{B}}(u^{-1})$.

Exemple. Si $\dim(E) = n$, si $\lambda \in \mathbb{K}$, alors, dans toute base $\mathcal{B}$ de E, l'homothétie de rapport λ est représentée par la matrice scalaire λI_n, soit $\mathrm{Mat}_{\mathcal{B}}(\lambda \ \mathrm{id}_E) = \lambda \, I_n$.

c. Changements de bases

Soient $\mathcal{B} = (e_1, \cdots, e_n)$ et $\mathcal{B}' = (e'_1, \cdots, e'_n)$ deux bases d'un $\mathbb{K}$-espace vectoriel E de dimension n. On appelle **matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$** la matrice de la famille de vecteurs $\mathcal{B}'$ dans la base $\mathcal{B}$, on la note alors $P_{\mathcal{B},\mathcal{B}'}$.

Si on note $P = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ cette matrice, on lit sur la j-ème colonne de P $(1 \leq j \leq n)$ les coordonnées du vecteur e'_j dans la base $\mathcal{B}$:

$$\forall j \in [\![1, n]\!] \qquad e'_j = \sum_{i=1}^{n} a_{i,j} e_i \ .$$

On peut noter que l'on a aussi $P = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)$ puisque la j-ème colonne de P contient les coordonnées dans la base $\mathcal{B}$ du vecteur $e'_j = \mathrm{id}_E(e'_j)$. Ainsi,

$$\boxed{P_{\mathcal{B},\mathcal{B}'} = \mathrm{Mat}_{\mathcal{B}}(\mathcal{B}') = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E)}$$

Conséquence. La matrice P est inversible: $P \in \mathrm{GL}_n(\mathbb{K})$, et son inverse P^{-1} est la matrice de passage de la base $\mathcal{B}'$ à la base $\mathcal{B}$. Autrement dit, $\big(P_{\mathcal{B},\mathcal{B}'}\big)^{-1} = P_{\mathcal{B}',\mathcal{B}}$.

Remarque. Toute matrice carrée inversible peut être considérée comme une matrice de passage. En effet, si $P \in \mathrm{GL}_n(\mathbb{K})$, alors $P = P_{\mathcal{B}_0,\mathcal{C}}$, où $\mathcal{B}_0$ est la base canonique de $\mathbb{K}^n$, et $\mathcal{C} = \big(C_1(P), \cdots, C_n(P)\big)$ est la base de $\mathbb{K}^n$ constituée des vecteurs-colonnes de P.

Coordonnées d'un vecteur. Si $x \in E$ est un vecteur, de coordonnées $(x_1, \cdots, x_n)$ dans la base $\mathcal{B}$ et de coordonnées $(x'_1, \cdots, x'_n)$ dans la base $\mathcal{B}'$, alors en introduisant les matrices-colonnes $X = \mathrm{Mat}_{\mathcal{B}}(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $X' = \mathrm{Mat}_{\mathcal{B}'}(x) = \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix}$, on a la relation $X = PX'$, que l'on peut retenir sous la forme : "*on obtient les coordonnées dans l'ancienne base en multipliant par la matrice de passage les coordonnées dans la nouvelle base*". *En effet,*

$$X = \mathrm{Mat}_{\mathcal{B}}(x) = \mathrm{Mat}_{\mathcal{B}}\big(\mathrm{id}_E(x)\big) = \mathrm{Mat}_{\mathcal{B}',\mathcal{B}}(\mathrm{id}_E) \cdot \mathrm{Mat}_{\mathcal{B}'}(x) = P\,X' \ .$$

Matrice d'une application linéaire. Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions finies p et n respectivement, soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, soient $\mathcal{C}$ et $\mathcal{C}'$ deux bases de F, soit u une application linéaire de E vers F, soient les matrices $M = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u)$ et $M' = \mathrm{Mat}_{\mathcal{B}',\mathcal{C}'}(u)$ de format (n, p), soient les matrices de passage $P = P_{\mathcal{B},\mathcal{B}'} \in \mathrm{GL}_p(\mathbb{K})$ et $Q = P_{\mathcal{C},\mathcal{C}'} \in \mathrm{GL}_n(\mathbb{K})$. On a alors la relation $M' = Q^{-1}MP$.

On peut retrouver cette relation par le diagramme suivant:

On peut aussi la retrouver par le calcul suivant: notons X et X' les matrices-colonnes représentant un même vecteur x de E dans les bases $\mathcal{B}$ et $\mathcal{B}'$ respectivement, soit

$y = u(x) \in F$ lui aussi représenté par les matrices-colonnes Y et Y' dans $\mathcal{C}$ et $\mathcal{C}'$ respectivement. On a alors les relations $Y = MX$ et $Y' = M'X'$, mais aussi $X = PX'$ et $Y = QY'$. On en tire $Y' = M'X' = Q^{-1}MPX'$ et, ceci étant vrai pour toute matrice-colonne X', on en déduit que $M' = Q^{-1}MP$.

Matrice d'un endomorphisme. En particulier, si u est un endomorphisme de E, représenté dans la base $\mathcal{B}$ par une matrice M et dans la base $\mathcal{B}'$ par une matrice M', on a alors la relation $M' = P^{-1}MP$ avec $P = P_{\mathcal{B},\mathcal{B}'}$.

On peut retrouver cette relation par le diagramme suivant:

Définition. Soient A et B deux matrices carrées d'ordre n à coefficients dans $\mathbb{K}$. On dit que B est **semblable** à A s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Propriétés. Cette relation, dite "de **similitude**", dans $\mathcal{M}_n(\mathbb{K})$ est
- réflexive: toute matrice de $\mathcal{M}_n(\mathbb{K})$ est semblable à elle-même ;
- symétrique: si B est semblable à A, alors A est semblable à B ;
- transitive: si A est semblable à B et B est semblable à C, alors A est semblable à C.

d. Correspondance canonique

Si on se donne une matrice $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$, on peut définir l'application linéaire u_A **canoniquement associée** à cette matrice, c'est l'application linéaire de $\mathbb{K}^p$ vers $\mathbb{K}^n$: $u_A \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$, telle que $\mathrm{Mat}_{\mathcal{B}_0,\mathcal{C}_0}(u_A) = A$, en notant $\mathcal{B}_0$ et $\mathcal{C}_0$ les bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$ respectivement. En adoptant l'identification "canonique" de $\mathbb{K}^p$ avec $\mathcal{M}_{p,1}(\mathbb{K})$ qui consiste à identifier le vecteur $x = (x_1, \cdots, x_p)$ de $\mathbb{K}^p$ avec la matrice-colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ de $\mathcal{M}_{p,1}(\mathbb{K})$, on peut aussi définir u_A comme suit:

$$\begin{cases} u_A \ : \ \mathbb{K}^p \simeq \mathcal{M}_{p,1}(\mathbb{K}) & \to & \mathbb{K}^n \simeq \mathcal{M}_{n,1}(\mathbb{K}) \\ \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} & \mapsto & Y = AX = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \end{cases}$$

avec, pour tout $i \in [\![1, n]\!]$, $y_i = \displaystyle\sum_{j=1}^{p} a_{i,j}x_j$.

On définit alors le **noyau** et l'**image** de A comme étant le noyau et l'image de l'application linéaire canoniquement associée u_A, soit $\mathrm{Ker}\, A = \mathrm{Ker}(u_A)$ (c'est un s.e.v. de $\mathbb{K}^p$) et $\mathrm{Im}\, A = \mathrm{Im}(u_A)$ (c'est un s.e.v. de $\mathbb{K}^n$).

Il est intéressant de noter que $\mathrm{Ker}\, A$ est l'ensemble des solutions du système linéaire homogène d'écriture matricielle $AX = 0$ (système de n équations à p inconnues).

Par ailleurs, $\operatorname{Im} A$ est le s.e.v. de $\mathbb{K}^n$ engendré par les vecteurs-colonnes de la matrice A :

$$\operatorname{Im} A = \operatorname{Vect}\big(C_1(A), \cdots, C_p(A)\big) \subset \mathbb{K}^n \ .$$

Exemple. Déterminer l'image et le noyau de la matrice $A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & -1 & -1 & 2 \\ 1 & -1 & 0 & 3 \end{pmatrix}$.

Le **rang** de la matrice A est alors défini comme étant le rang de la famille de ses vecteurs-colonnes, c'est aussi la dimension de son image :

$$\operatorname{rg}(A) = \operatorname{rg}\big(C_1(A), \cdots, C_p(A)\big) = \dim\big(\operatorname{Im}(A)\big) \ .$$

On a aussi, bien sûr, $\operatorname{rg}(A) = \operatorname{rg}(u_A)$.

On a alors $\operatorname{rg}(A) \leq \min\{n, p\}$ pour tout $A \in \mathcal{M}_{n,p}(\mathbb{K})$.

Si A est une matrice de format (n, p), le **théorème du rang** peut s'énoncer comme suit :

$$\operatorname{rg}(A) + \dim(\operatorname{Ker} A) = p \ .$$

Dans le cas des matrices carrées, si $A \in \mathcal{M}_n(\mathbb{K})$, elle représente canoniquement un endomorphisme u_A de l'espace vectoriel $\mathbb{K}^n$. Cet endomorphisme u_A est bijectif (i.e. est un automorphisme de $\mathbb{K}^n$) *si et seulement si* la matrice A est inversible.

Proposition. Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice carrée, on a les équivalences :

$$A \in \mathrm{GL}_n(\mathbb{K}) \iff \operatorname{Ker} A = \{0\} \iff \operatorname{Vect}\big(C_1(A), \cdots, C_n(A)\big) = \mathbb{K}^n \iff \operatorname{rg}(A) = n \ .$$

Proposition. Il y a conservation du rang par multiplication (à gauche ou à droite) par des matrices inversibles. Autrement dit, si on a $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $Q \in \mathrm{GL}_n(\mathbb{K})$, $P \in \mathrm{GL}_p(\mathbb{K})$, alors

$$\operatorname{rg}(A) = \operatorname{rg}(QA) = \operatorname{rg}(AP) = \operatorname{rg}(QAP) \ .$$

On peut en déduire que le rang d'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est le rang de toute application linéaire qu'elle représente (d'un e.v. E de dimension p vers un e.v. F de dimension n, munis de bases). On peut en déduire aussi que les opérations élémentaires sur les lignes et colonnes conservent le rang, nous y reviendrons.

Proposition. Une matrice et sa transposée ont le même rang : $\operatorname{rg}(A^\top) = \operatorname{rg}(A)$.

Une preuve est donnée en annexe.

Le rang d'une matrice est donc aussi le rang de ses vecteurs-lignes : si $A \in \mathcal{M}_{n,p}(\mathbb{K})$, alors

$$\operatorname{rg}(A) = \operatorname{rg}\big(L_1(A), \cdots, L_n(A)\big) \ .$$

III. Systèmes linéaires et algorithme du pivot

a. Opérations élémentaires.

Définition. Les **opérations élémentaires** sur les lignes d'une matrice sont de trois types :

- échange de deux lignes L_i et L_j avec $i \neq j$, le codage est $L_i \longleftrightarrow L_j$;
- ajout à une ligne L_i d'un multiple d'une autre, codage $L_i \leftarrow L_i + \lambda L_j$ avec $i \neq j$;
- multiplication d'une ligne L_i par un scalaire λ non nul, codage $L_i \leftarrow \lambda L_i$.

On définit de même des opérations élémentaires sur les colonnes d'une matrice.

Ces opérations élémentaires peuvent être interprétées en termes de produit matriciel de la façon suivante : partons de la matrice identité I_n et observons les matrices (alors carrées d'ordre n) obtenues en lui faisant subir respectivement les opérations $L_i \longleftrightarrow L_j$, $L_i \leftarrow L_i + \lambda L_j$ et $L_i \leftarrow \lambda L_i$, les voici:

La première d'entre elles est $A = I_n - (E_{i,i} + E_{j,j}) + (E_{i,j} + E_{j,i})$, ou encore $A = E_{i,j} + E_{j,i} + \displaystyle\sum_{k \notin \{i,j\}} E_{k,k}$, on l'appelle **matrice de transposition**.

La deuxième est $B = I_n + \lambda E_{i,j}$, on l'appelle **matrice de transvection**.

La troisième est $C = I_n + (\lambda - 1)E_{i,i} = \lambda E_{i,i} + \displaystyle\sum_{j \neq i} E_{j,j} = \mathrm{diag}(1, \cdots, 1, \lambda, 1, \cdots, 1)$, on l'appelle **matrice de dilatation**.

On peut noter que toutes ces matrices sont des matrices carrées inversibles, et que leurs inverses, ainsi que leurs transposées, sont encore des matrices de la même forme.

Proposition. Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. La matrice A' (de même format) obtenue à partir de A en lui faisant subir une quelconque opération élémentaire sur les lignes est le produit (dans cet ordre) de la matrice obtenue à partir de la matrice identité I_n par la même opération élémentaire et de la matrice A. Autrement dit, si l'on note Φ une opération élémentaire sur les lignes (pour des matrices à n lignes), on a

$$\Phi(A) = \Phi(I_n) \cdot A \ .$$

De même, la matrice A'' (de même format) obtenue à partir de A en lui faisant subir une quelconque opération élémentaire sur les colonnes est le produit (dans cet ordre) de la matrice A et de la matrice obtenue à partir de la matrice identité I_p par la même opération élémentaire. Autrement dit, si l'on note Ψ une opération élémentaire sur les colonnes (pour des matrices à p colonnes), on a

$$\Psi(A) = A \cdot \Psi(I_p) \ .$$

Ce qui est écrit ci-dessus reste valable si le symbole Φ (respectivement Ψ) représente une succession d'opérations élémentaires sur les lignes (respectivement sur les colonnes). Lorsque le format des matrices est fixé, ces transformations Φ et Ψ sont des automorphismes de l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$.

Proposition. Toute matrice carrée d'ordre n inversible peut être transformée, par des opérations élémentaires sur les lignes, en la matrice identité I_n.

*La preuve est constructive, il s'agit de l'***algorithme de Gauss-Jordan***. Voici la description de cet algorithme:*

Remarque. L'algorithme de Gauss-Jordan fournit aussi un moyen de calculer l'inverse d'une matrice carrée inversible A^{-1}, en effectuant des opérations élémentaires sur les lignes de la "matrice augmentée" $(A|I_n) \in \mathcal{M}_{n,2n}(\mathbb{K})$ pour transformer A en I_n : les mêmes opérations sur les lignes transforment alors I_n en A^{-1}, donc $(A|I_n)$ en $(I_n|A^{-1})$. *En effet, si une succession d'opérations élémentaires sur les lignes transforme la matrice A en la matrice I_n, soit $\Phi(A) = I_n$, on a alors $\Phi(I_n) \cdot A = I_n$, soit $A^{-1} = \Phi(I_n)$.*

Proposition. Les opérations élémentaires sur les colonnes conservent l'image d'une matrice. Les opérations élémentaires sur les lignes conservent le noyau. Toutes les opérations élémentaires conservent le rang.

b. Systèmes linéaires

On appelle **système linéaire de n équations à p inconnues** tout système **(S)** pouvant se mettre sous la forme $AX = B$, où les matrices $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K})$ sont données, et où l'inconnue est la matrice-colonne $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ de $\mathcal{M}_{p,1}(\mathbb{K})$ que l'on identifiera au vecteur $x = (x_1, \cdots, x_p)$ de $\mathbb{K}^p$. Ce système peut aussi s'écrire $u_A(x) = b$, où $u_A \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ est l'application linéaire canoniquement associée à la matrice A, et $b = (b_1, \cdots, b_n) \in \mathbb{K}^n$. L'ensemble des solutions du système **(S)**, que nous noterons $\mathcal{S}$, peut s'écrire alors $\mathcal{S} = u_A^{-1}\big(\{b\}\big)$. *Toutes ces notations seront réutilisées par la suite.*

Le rang r de la matrice A (ou de l'application linéaire u_A) est aussi appelé **rang du système (S)**. Comme c'est aussi le rang de la matrice transposée $A^\top$, on peut voir ce rang r comme étant le "nombre d'équations indépendantes" dans **(S)**.

Le système **(S)** est dit **compatible** s'il admet au moins une solution, i.e. si $\mathcal{S} \neq \emptyset$, il est dit **incompatible** sinon. La **condition de compatibilité** du système peut s'écrire: $b \in \mathrm{Im}(u_A)$, ou encore $B \in \mathrm{Im}(A)$, ce qui signifie que B est combinaison linéaire des colonnes de la matrice A.

Le **système homogène** (ou "sans second membre") associé à **(S)** est le système $(\mathbf{S}_0)$ d'écriture matricielle $AX = 0$, ou d'écriture fonctionnelle $u_A(x) = 0$. Son ensemble de solutions est $\mathcal{S}_0 = \mathrm{Ker}\, u_A = \mathrm{Ker}\, A$, c'est un sous-espace vectoriel de $\mathbb{K}^p$.

Propriété fondamentale. Un système linéaire homogène est toujours compatible puisqu'il a toujours le vecteur nul de $\mathbb{K}^p$ comme solution. L'ensemble $\mathcal{S}_0$ de ses solutions est un sous-espace vectoriel de $\mathbb{K}^p$ de dimension $p - r$, où r est le rang du système. On a donc, pour les systèmes linéaires homogènes :

$$\begin{aligned} \dim(\mathcal{S}_0) &= \text{(nombre d'inconnues)} - \text{(rang)} \\ &= \text{(nombre d'inconnues)} - \text{(nombre d'équations indépendantes)}\ . \end{aligned}$$

Proposition. Si le système **(S)** est compatible, on obtient sa solution générale en ajoutant une de ses solutions "particulières" à la solution générale du système homogène $(\mathbf{S}_0)$ associé. Si $\mathcal{S} \neq \emptyset$ et $x^* \in \mathcal{S}$, alors

$$\mathcal{S} = x^* + \mathcal{S}_0 = \big\{x^* + x\ ;\ x \in \mathcal{S}_0\big\}\ .$$

Définition. Si la matrice A est carrée et inversible, i.e. si $A \in \mathrm{GL}_n(\mathbb{K})$, le système linéaire $AX = B$ est qualifié de **système de Cramer**, il admet alors une unique solution qui est $X = A^{-1}B$.

Exercice. Soit le système linéaire $\left\{\begin{array}{rcrcrcl} 2x & - & 4y & - & 6z & = & a \\ x & + & 2y & - & 3z & = & b \\ -2x & + & 8y & + & 6z & = & c \end{array}\right.$. Discuter de la compatibilité de ce système et, lorsqu'il est compatible, exprimer ses solutions.

ANNEXE: Une matrice et sa transposée ont le même rang (une preuve)

Soit $A = (a_{i,j}) \in \mathcal{M}_{p,q}(\mathbb{K})$, l'objectif est de montrer que $\mathrm{rg}(A^\top) = \mathrm{rg}(A)$.

On rappelle que le rang d'une matrice est le rang de la famille de ses vecteurs-colonnes.

On va commencer par montrer que **le rang d'une matrice A est la taille maximale des matrices carrées inversibles extraites de A**.

Soit $r = \mathrm{rg}(A)$, soit d la taille maximale des matrices carrées inversibles extraites de A.

Soit $M \in \mathcal{M}_d(\mathbb{K})$ une matrice carrée de taille d extraite de A, et inversible. Par commodité, on va supposer qu'elle est constituée des d premières lignes et d premières colonnes de A, cela ne change rien au raisonnement. Ses d colonnes $C_1(M)$, $\cdots$, $C_d(M)$ sont alors linéairement indépendantes dans $\mathbb{K}^d$, on en déduit que les d premières colonnes $C_1(A)$, $\cdots$, $C_d(A)$ de A, dont elles sont extraites, sont linéairement indépendantes dans $\mathbb{K}^p$. En effet, toute relation de dépendance linéaire entre les $C_j(A)$ dans $\mathbb{K}^p$ impliquerait, par projection, une relation de dépendance linéaire entre les $C_j(M)$ dans $\mathbb{K}^d$. Il y a donc dans la matrice A au moins d colonnes indépendantes, on en déduit l'inégalité $r \geq d$.

Par ailleurs, puisque $\mathrm{rg}(A) = r$, il existe r colonnes de A qui sont linéairement indépendantes, on peut supposer que ce sont les r premières. Notons $N \in \mathcal{M}_{p,r}(\mathbb{K})$ la matrice constituée de ces r colonnes, elle est de rang r puisque ses r colonnes sont indépendantes. Ses p lignes $L_1(N)$, $\cdots$, $L_p(N)$ forment une famille génératrice de $\mathbb{K}^r$: en effet, si ce n'était pas le cas, elles seraient incluses dans un hyperplan H de $\mathbb{K}^r$, il existerait donc des scalaires λ_1, $\cdots$, λ_r non tous nuls tels que, pour tout $i \in [\![1, p]\!]$, on ait $\displaystyle\sum_{j=1}^{r} \lambda_j a_{i,j} = 0$ (ceci traduit l'appartenance de $L_i(N)$ à un hyperplan H de $\mathbb{K}^r$ d'équation $\displaystyle\sum_{j=1}^{r} \lambda_j x_j = 0$). On aurait alors, dans $\mathbb{K}^p$, la relation $\displaystyle\sum_{j=1}^{r} \lambda_j\, C_j(N) = 0$, ce qui est contradictoire, les colonnes de N étant indépendantes. Donc la famille $\big(L_1(N), \cdots, L_p(N)\big)$ est de rang r. On peut donc, de N, extraire r lignes indépendantes et construire une matrice carrée M d'ordre r extraite de A, dont les lignes sont indépendantes. Ceci prouve que les colonnes de $M^\top$ sont indépendantes, donc $M^\top$ est inversible, donc M est inversible. On obtient ainsi une matrice carrée d'ordre r inversible extraite de la matrice A. On a prouvé $d \geq r$.

Finalement, $r = d$. Il est alors immédiat que $\mathrm{rg}(A^\top) = \mathrm{rg}(A)$ puisque la taille maximale d'une matrice carrée inversible extraite est clairement la même pour A et pour $A^\top$.