COURS DE MATHÉMATIQUES : ALGÈBRE LINÉAIRE III.

Table des matières

1. Opérations sur les matrices.

1.1. Somme.

Exemple 1.1 (somme de deux matrices)**.**

$$(1) : \begin{pmatrix}1 & -1 & 0 & 2\end{pmatrix} + \begin{pmatrix}3 & 4 & -2 & -5\end{pmatrix} = \begin{pmatrix}1+3 & -1+4 & 0+(-2) & 2+(-5)\end{pmatrix} = \begin{pmatrix}4 & 3 & -2 & -3\end{pmatrix}$$

$$(2) : \begin{pmatrix}2\\0\\1\\-1\end{pmatrix} + \begin{pmatrix}1\\4\\-2\\0\end{pmatrix} = \begin{pmatrix}2+1\\0+4\\1+(-2)\\-1+0\end{pmatrix} = \begin{pmatrix}3\\4\\-1\\-1\end{pmatrix}$$

$$(3) : \begin{pmatrix}-2 & 0 & 1\\3 & 2 & -4\end{pmatrix} + \begin{pmatrix}4 & 5 & 6\\-7 & -8 & -9\end{pmatrix} = \begin{pmatrix}-2+4 & 0+5 & 1+6\\3+(-7) & 2+(-8) & -4+(-9)\end{pmatrix} = \begin{pmatrix}2 & 5 & 7\\-4 & -6 & -13\end{pmatrix}$$

Définition 1.2 (somme de deux matrices)**.** Soient A et B deux matrices à p lignes et n colonnes. Soient a_{ij} et b_{ij} les coefficients de A et B situés à l'intersection de la i-ème ligne et de la j-ème colonne. On définit alors la matrice $A+B$ (appelée *somme de A et de B*) comme la matrice dont le coefficient ij est $a_{ij} + b_{ij}$.

Remarque 1.3 (identifications des matrices $p \times n$ avec des pn-vecteurs)**.** Une matrice $p \times n$ c'est un tableau de pn nombres réels, donc cela contient la même information qu'un pn-vecteur, mais pas rangé sous forme de suite. Alors l'opération $A + B$ définie ci-dessus correspond simplement à l'addition des pn-vecteurs , elle se fait composantes par composantes.

Date: 9 avril 2009.

Attention cependant : en général, il n'y a pas de manière naturelle de transformer une matrice $p \times n$ en pn-vecteur. Par exemple une matrice 2×2 peut correspondre à plusieurs vecteurs de $\mathbb{R}^4$:

$$\begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix} = A \longleftrightarrow (1,2,-1,0)?\ A \longleftrightarrow (1,-1,2,0)?\ A \longleftrightarrow (1,0,-1,2)?$$

Il y a pourtant un cas où on a une identification plus naturelle que les autres : dans le cas des matrices colonnes.

Définition 1.4 (colonnes et vecteurs colonnes)**.** Soit A une matrice $p \times n$. Rappelons que la *première colonne de* A est la matrice $p \times 1$ de coefficients (a_{i1}), la *deuxième colonne de* A est la matrice $p \times 1$ de coefficients (a_{i2}), etc...

Le *premier vecteur colonne de* A est le p-vecteur $(a_{11}, a_{21}, \ldots, a_{p1})$, le *deuxième vecteur colonne de* A est le p-vecteur $(a_{12}, a_{22}, \ldots, a_{p2})$, etc...

Ne pas confondre la j-ème colonne avec le j-ème vecteur colonne, même s'ils se correspondent, et contiennent exactement la même information - simplement cette information n'est pas sous la même forme.

Exemple 1.5. Soit $A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 3 & 1 \end{pmatrix}$. Alors la première colonne de A est $C_1 = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$, et le premier vecteur colonne de A est $(1, -1, 3)$. Et la matrice A contient la même information que la suite $((1,-1,3),(2,0,1))$ de ses deux vecteurs colonnes.

Définition 1.6 (image d'une matrice)**.** Soit A une matrice $p \times n$. L'*image de* A est par définition le sous-espace vectoriel de $\mathbb{R}^p$ engendré par les vecteurs colonnes. Nous le noterons $\mathrm{Im}(A)$.

Définition 1.7 (rang d'une matrice)**.** Soit A une matrice $p \times n$. Le *rang de* A est par définition le rang de la suite de ses vecteurs colonnes. Nous le noterons $\mathrm{rg}(A)$.

Ainsi $\mathrm{rg}(A) = \dim(\mathrm{Im}(A))$.

Définition 1.8 (noyau d'une matrice)**.** Soit A une matrice $p \times n$. Le *noyau de* A est par définition le sous-espace vectoriel de $\mathbb{R}^n$ formé par les solutions du système homogène associé à la matrice A. Nous le noterons $\mathrm{Ker}(A)$.

1.2. **Produit par un réel.**

Exemple 1.9 (produit d'une matrice par un réel)**.**

$$(1) : 3 \times \begin{pmatrix} 1 & -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 \times 1 & 3 \times (-1) & 3 \times 0 & 3 \times 2 \end{pmatrix} = \begin{pmatrix} 3 & -3 & 0 & 6 \end{pmatrix}$$

$$(2) : (-2) \times \begin{pmatrix} 2 \\ 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} (-2) \times 2 \\ (-2) \times 0 \\ (-2) \times 1 \\ (-2) \times (-1) \end{pmatrix} = \begin{pmatrix} -4 \\ 0 \\ -2 \\ 2 \end{pmatrix}$$

$$(3) : 4 \times \begin{pmatrix} -2 & 0 & 1 \\ 3 & 2 & -4 \end{pmatrix} = \begin{pmatrix} 4 \times (-2) & 4 \times 0 & 4 \times 1 \\ 4 \times 3 & 4 \times 2 & 4 \times (-4) \end{pmatrix} = \begin{pmatrix} -8 & 0 & 4 \\ 12 & 8 & -16 \end{pmatrix}$$

Définition 1.10 (produit d'une matrice par un réel)**.** Soit A une matrices à p lignes et n colonnes, soit λ un réel. Soient a_{ij} le coefficient de A situé à l'intersection de la i-ème ligne et de la j-ème colonne. On définit alors la matrice λA (appelée *produit de* A *par* λ) comme la matrice dont le coefficient ij est λa_{ij}.

Remarque 1.11. La multiplication par un réel des $p \times n$-matrices correspond à la multiplication par un réel des pn-vecteurs correspondants : la multiplication se fait composantes par composantes dans les deux cas.

Par exemple les matrices colonne à p lignes correspondent à l'espace vectoriel $\mathbb{R}^p$ avec ses deux opérations. On peut faire des combinaisons linéaires de matrices colonnes, et plus généralement de matrices $A_1, \dots, A_k$ (du moment qu'elles sont toutes de même taille $p \times n$).

Par exemple :

$$2\begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix} + 3\begin{pmatrix} -2 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} -4 & 3 & 1 \\ 3 & -2 & -4 \end{pmatrix}$$

ou

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

1.3. **Produit de deux matrices.**

Exemple 1.12 (quelques produits)**.**

Une ligne fois une colonne :

$$\begin{pmatrix} 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 3 \times 1 + (-1) \times 2 + 2 \times (-3) \end{pmatrix} = \begin{pmatrix} -5 \end{pmatrix}$$

Deux lignes fois une colonne :

$$\begin{pmatrix} 3 & -1 & 2 \\ 1 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 3 \times 1 + (-1) \times 2 + 2 \times (-3) \\ 1 \times 1 + 4 \times 2 + 1 \times (-3) \end{pmatrix} = \begin{pmatrix} -5 \\ 6 \end{pmatrix}$$

Deux lignes fois trois colonnes :

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

$$\begin{pmatrix} 3 \times 1 + (-1) \times 2 + 2 \times (-3) & 3 \times (-1) + (-1) \times 3 + 2 \times 1 & 3 \times 0 + (-1) \times 1 + 2 \times 2 \\ 1 \times 1 + 4 \times 2 + 1 \times (-3) & 1 \times (-1) + 4 \times 3 + 1 \times 1 & 1 \times 0 + 4 \times 1 + 1 \times 2 \end{pmatrix}$$

$$= \begin{pmatrix} -5 & -4 & 3 \\ 6 & 12 & 6 \end{pmatrix}$$

Définition 1.13. Soient A et B deux matrices. On suppose que A a autant de colonnes que B a de lignes. Mettons que A a p lignes et n colonnes et B a n lignes et m colonnes. Soient a_{ij} et b_{ij} les coefficients de A et B situés à l'intersection de la i-ème ligne et de la j-ème colonne. On définit alors la matrice $A.B$ (appelée *produit de A par B*) comme la matrice C à p lignes et m colonnes dont le coefficient c_{ij} est :

$$c_{ij} = \sum_{k=1}^{k=n} a_{ik}.b_{kj}.$$

Concrètement : **le coefficient ij du produit s'obtient en faisant le produit de la i-ème ligne de A par la j-ème colonne de B.**

En particulier le produit AX, où A est une matrice de taille $p \times n$ et X est un vecteur colonne à n lignes est encore un vecteur colonne, cette fois à p lignes.

Attention si une matrice A n'est pas *carrée* (autant de lignes que de colonnes) il est *impossible* de la multiplier par elle-même !

Proposition 1.14 (premières propriétés du produit de matrices)**.**

(1) *La multiplication par la matrice nulle donne la matrice nulle.*

(2) *Soit A une matrice ayant p lignes et n colonnes, et soit E_j la matrice colonne à p lignes ayant des 0 partout sauf à la j-ème ligne où il y a un 1. Alors $A.E_j$ est la j-ème colonne de A.*

(3) *On note I_n la matrice $n \times n$ qui a des 1 sur la diagonale et des 0 ailleurs. Si A une matrice ayant p lignes et n colonnes, alors $AI_n = A$. (Et $I_p.A = A$.) Egalement : $A(\lambda I_n) = \lambda A$ (noter que λI_n est la matrice $n \times n$ qui a des λ sur la diagonale et des 0 ailleurs).*

(4) *(associativité) $A(BC) = (AB)C$.*

(5) *(distributivité) Pour A une matrice $p \times n$ et B, B' deux matrices $n \times m$ on a $A(B + B') = AB + AB'$. Pour A, A' des matrices $p \times n$ et B une matrice $n \times m$ on a $(A + A').B = AB + A'B$.*

(6) *$A(\lambda B) = \lambda AB$.*

(7) *$A(\lambda B + \lambda' B') = \lambda AB + \lambda' AB'$.*

Attention ! ! ! : on peut avoir $AB = AC$ avec $A \neq 0$ et $B \neq C$.
Exemple : $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, C = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.
Attention ! ! ! : en général $AB \neq BA$.
Exemple : $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$.

Démonstration. (1) La matrice nulle, c'est la matrice dont tous les coefficients sont nuls. Donc la propriété est évidente.

(2) D'abord $A.E_j$ est bien une matrice colonne ayant p ligne. D'après la formule du produit son i-ème coefficient est $a_{i,1} \times 0 + \cdots + a_{i,j-1} \times 0 + a_{i,j} \times 1 + a_{i,j+1} \times 0 + \cdots + a_{i,n} \times 0$, autrement dit c'est a_{ij}. Donc $A.E_j$ a les mêmes coefficients que la j-ème colonne de A - à laquelle elle est donc égale.

(3) La j-ème colonne de AI_n est le produit de A avec la j-ème colonne de I_n - c'est à dire E_j. Donc la j-ème colonne de AI_n est la j-ème colonne de A. Il en résulte que $AI_n = A$.

Nous laissons la preuve de $I_p.A = A$ et $A(\lambda I_n) = \lambda A$ au lecteur.

(4) Supposons $A = (a_{ij})_{1 \leq i \leq p, 1 \leq j \leq m}, B = (b_{jk})_{1 \leq j \leq m, 1 \leq k \leq n}, C = (c_{k\ell})_{1 \leq k \leq n, 1 \leq \ell \leq l}$, de sorte que les produits $AB, BC, (AB)C, A(BC)$ sont bien définis.

Le coefficient i, k de AB et le coefficient j, ℓ de BC sont donnés par les formules :

$$[AB]_{i,k} = \sum_{j=1}^{j=m} a_{i,j} b_{j,k}, [BC]_{j,\ell} = \sum_{k=1}^{k=n} b_{j,k} c_{k,\ell}$$

Donc les coefficient i, ℓ de $(AB)C$ et de $A(BC)$ sont donnés par les formules :

$$[(AB)C]_{i,\ell} = \sum_{k=1}^{k=n(} \sum_{j=1}^{j=m} a_{i,j} b_{j,k}) c_{k,\ell}, [A(BC)]_{i,\ell} = \sum_{j=1}^{j=m} a_{i,j} (\sum_{k=1}^{k=n} b_{j,k} c_{k,\ell})$$

Ainsi les coefficients i, ℓ des deux matrices $(AB)C$ et $A(BC)$ sont tous deux égaux à la même somme double :

$$\sum_{1 \leq j \leq m, 1 \leq k \leq n} a_{i,j} b_{j,k} c_{k,\ell}$$

(5) Le coefficient i, k de $A(B+B')$ est

$$\sum_j a_{i,j}(b_{j,k}+b'_{j,k})$$

Et le coefficient i, k de $AB+AB'$ est

$$(\sum_j a_{i,j}b_{j,k})+(\sum_j a_{i,j}b'_{j,k})$$

qui est la première formule développée. On a bien $A(B+B')=AB+AB'$.

On montre de la même manière que $(A+A').B=AB+A'B$.

(6) Le coefficient i, k de $A(\lambda B)$ est

$$\sum_j a_{i,j}(\lambda b_{j,k})$$

Et le coefficient i, k de λAB est

$$\lambda(\sum_j a_{i,j}b_{j,k})$$

On a donc bien $A(\lambda B)=\lambda AB$.

(7) La relation $A(\lambda B+\lambda' B')=\lambda AB+\lambda' AB'$ est la combinaison des deux propriétés précédentes.

□

2. Interprêtation matricielle des systèmes linéaires.

Proposition 2.1. *Soit A une matrice $p\times n$ et soit B une colonne à p lignes. Considérons le système linéaire associé à la matrice A et de second membre B :*

$$\begin{cases} a_{11}x_1+\cdots+a_{1n}x_n &= b_1 \\ a_{21}x_1+\cdots+a_{2n}x_n &= b_2 \\ \qquad\dots & \\ a_{n1}x_1+\cdots+a_{nn}x_n &= b_n \end{cases}$$

Alors $(x_1,\dots,x_n)$ est une solution si et seulement si la colonne $X=\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}$ (qui correspond à $(x_1,\dots,x_n)$) est solution de l'équation matricielle $AX=B$.

Par exemple les solutions d'un système linéaire homogène (de matrice A) correspondent au noyau de $\mathrm{Ker}(A)$.

3. Applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$.

3.1. Définition et exemples.

Définition 3.1 (applications linéaires)**.** Soit $f:\mathbb{R}^n\to\mathbb{R}^p$ une application. On dit que f est *linéaire* si f conserve les combinaisons linéaires, autrement dit si pour tout choix de deux vecteurs u, v de $\mathbb{R}^n$ et tout choix de deux réels λ, μ on a $f(\lambda u+\mu v)=\lambda f(u)+\mu f(v)$.

Exemple 3.2.

1) Considérons l'application $f:\mathbb{R}^2\to\mathbb{R}$ définie par $f(x,y)=x$ (la première composante). Alors f est linéaire. En effet pour $u=(x,y), u'=(x',y')$ et $\lambda,\lambda'\in\mathbb{R}$ on a $\lambda u+\lambda' u'=(\lambda x+\lambda' x',\lambda y+\lambda' y')$, donc $f(\lambda u+\lambda' u')=\lambda x+\lambda' x'$. Or $\lambda x+\lambda' x'=\lambda f(u)+\lambda' f(u')$.

2) $f(x,y,z) = 2x - y + 3z$. Pour $u = (x,y,z), u' = (x',y',z')$ et $\lambda, \lambda' \in \mathbb{R}$ on a $\lambda u + \lambda' u' = (\lambda x + \lambda' x', \lambda y + \lambda' y', \lambda z + \lambda' z')$, donc

$$f(\lambda u + \lambda' u') = 2(\lambda x + \lambda' x') - (\lambda y + \lambda' y') + 3(\lambda z + \lambda' z') = \lambda(2x - y + 3z) + \lambda'(2x' - y' + 3z').$$

Et

$$\lambda(2x - y + 3z) + \lambda'(2x' - y' + 3z') = \lambda f(u) + \lambda' f(u').$$

3) Plus généralement si on se fixe des réels $a_1, \ldots, a_n$, alors l'application $(x_1, \ldots, x_n) \mapsto a_1 x_1 + \cdots + a_n x_n$ est linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$.

4) Les coordonnées sont linéaires.

Si $\mathcal{B}B$ est une base de $\mathbb{R}^n$ et si on note $(y_1(u), \ldots, y_n(u))$ les coordonnées d'un vecteur u dans cette base, alors chacune des applications $u \mapsto y_1(u), \ldots, u \mapsto y_n(u)$ est linéaire.

5) Projections et symétries.

Supposons que $\mathbb{R}^n = E \oplus E'$. Rappelons qu'alors tout vecteur u de $\mathbb{R}^n$ s'écrit d'une unique façon $u = v + v'$ avec $v \in E$ et $v' \in E'$. On pose $p(u) = v$, qu'on appelle le *projeté de u sur E parallèlement à E'*. Alors $u \mapsto p(u)$ est linéaire (appelée la *projection sur E parallèlement à (ou dans la direction de) E'*).

Ainsi $\mathbb{R}^2 = D \oplus D'$ avec $D = \mathrm{Vect}(1,1)$ et $D' = \mathrm{Vect}(-2,1)$.

De même l'application $s : u \mapsto 2p(u) - u$ est linéaire, on l'appelle la *symétrie par rapport à E dans la direction de E'*. Si $v \in E$ on a $s(v) = 2v - v = v$, et si $v' \in E'$ on a $s(v') = 0 - v' = -v'$. En général si $u = v + v'$ avec $v \in E$ et $v' \in E'$ on aura $s(u) = v - v'$.

3.2. **Application linéaire associée à une matrice.**

Exemple 3.3 (exemple d'application associée à une matrice)**.** Soit $A = \begin{pmatrix} 2 & 3 \\ 1 & -1 \end{pmatrix}$, de sorte que pour toute colonne $X = \begin{pmatrix} x \\ y \end{pmatrix}$ on a $AX = \begin{pmatrix} 2x + 3y \\ x - y \end{pmatrix}$. Alors l'application associée à A envoie (x,y) sur $(2x + 3y, x - y)$

Plus généralement, quand on a une matrice A à p lignes et n colonnes, on peut considérer la fonction F qui envoie une colonne X à n lignes sur la matrice colonne produit AX (qui, elle, a p lignes). En utilisant la correspondance naturelle entre les matrices colonnes et les vecteurs, on peut interpréter $F : X \mapsto AX$ comme une fonction $f : \mathbb{R}^n \to \mathbb{R}^p$, qui se trouve être linéaire :

Proposition 3.4 (l'application associée à une matrice)**.** *Soit A une matrice $p \times n$. L'application $f : \mathbb{R}^n \to \mathbb{R}^p$ définie par $f(x_1, \ldots, x_n) =$ le vecteur correspondant à la colonne produit $A \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}$ est linéaire, on l'appelle l'*application linéaire associée à A*.*

Démonstration. Dans les propriétés du produit de matrice, on a vu que $A(\lambda X + \lambda' X') = \lambda AX + \lambda' AX'$, autrement dit $F(\lambda X + \lambda' X') = \lambda F(X) + \lambda' F(X')$. Si on traduit cette relation en termes de f, on trouve exactement $F(\lambda(x_1, \ldots, x_n) + \lambda'(x'_1, \ldots, x'_n)) = \lambda F(x_1, \ldots, x_n) + \lambda' F(x'_1, \ldots, x'_n)$. □

Cela signifie que toute application $f : \mathbb{R}^n \to \mathbb{R}^p$ dont la i-ème composante est de la forme

$$f_i(x_1, \ldots, x_n) = a_{i1} x_1 + \cdots + a_{in} x_n$$

est toujours linéaire (puisqu'une telle application est l'application linéaire associée à $A = (a_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$). Il n'est donc pas nécessaire de le revérifier à chaque fois.

Par exemple $(x, y, z, t) \mapsto (3x - y + 5z - 4t, x + 2y + z + 3t, -x - y - z - t)$ est clairement linéaire (puisque associée à $A = \begin{pmatrix} 3 & -1 & 5 & -4 \\ 1 & 2 & 1 & 3 \\ -1 & -1 & -1 & -1 \end{pmatrix}$). Nous verrons bientôt que toute application linéaire $f : \mathbb{R}^n \to \mathbb{R}^p$ est de cette forme, donc est associée à une (unique) matrice A à p lignes et n colonnes.

Notons d'ores et déjà que par exemple $f(x, y) = x + y + 1$ **n'est pas linéaire**. En effet $f(0, 0) = 1$, or si f était linéaire elle devrait vérifier $f(0, 0) = 0$, d'après le résultat suivant :

Lemme 3.5.
Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Alors $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^p}$.

Démonstration. En effet $0_{\mathbb{R}^n} = 0_{\mathbb{R}^n} + 0_{\mathbb{R}^n}$, donc $f(0_{\mathbb{R}^n}) = f(0_{\mathbb{R}^n} + 0_{\mathbb{R}^n})$ et comme f est linéaire cela donne $f(0_{\mathbb{R}^n}) = f(0_{\mathbb{R}^n}) + f(0_{\mathbb{R}^n}) = 2f(0_{\mathbb{R}^n})$. Mais le seul vecteur $v \in \mathbb{R}^p$ qui vérifie $v = 2v$ c'est $v = 0_{\mathbb{R}^p}$. □

3.3. **Matrices d'une application linéaire.**

Définition 3.6 (matrice d'une application linéaire dans des bases données)**.**
Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Soit $\mathcal{B} = (u_1, \ldots, u_n)$ une base de $\mathbb{R}^n$ et $\mathcal{B}' = (u'_1, \ldots, u'_p)$ une base de $\mathbb{R}^p$.

Alors la matrice de f de la base $\mathcal{B}$ dans la base $\mathcal{B}'$ est une matrice $p \times n$, dont la j-ème colonne donne les coordonnées dans $\mathcal{B}'$ de $f(u_j)$. On la note $\mathrm{Mat}(f, \mathcal{B} \to \mathcal{B}')$.

Ainsi la matrice de f de $\mathcal{B}$ dans $\mathcal{B}'$ est l'unique matrice $A = (a_{ij})_{1 \leq i \leq p, 1 \leq j \leq n}$ telle que pour tout $j = 1, \ldots, n$ on a :

$$f(u_j) = a_{1j}u'_1 + a_{2j}u'_2 + \cdots + a_{nj}u'_n$$

Notation. Lorsque $p = n$ et la base de l'espace d'arrivée $\mathcal{B}'$ coïncide avec la base $\mathcal{B}$ au départ on note la matrice $\mathrm{Mat}(f; \mathcal{B})$ au lieu de $\mathrm{Mat}(f; \mathcal{B} \to \mathcal{B})$.

Exemple 3.7.
1) Soit $f(x, y, z) = (y - 2z, x + y + z, 3x - y - 2z)$, qui est facilement linéaire. Soit $\mathcal{B}$ la base canonique, déterminons $\mathrm{Mat}(f; \mathcal{B})$. On a $f(1, 0, 0) = (0, 1, 3)$, puis $f(010, 0) = (1, 1, -1)$ et enfin $f(0, 0, 1) = (-2, 1, -2)$. Donc

$$\mathrm{Mat}(f; \mathcal{B}) = \begin{pmatrix} 0 & 1 & -2 \\ 1 & 1 & 1 \\ 3 & -1 & -2 \end{pmatrix}$$

Pour rire (et comprendre qu'il faut faire très attention aux bases choisies) : remarquons que la suite des vecteurs colonnes $\mathcal{B}' = ((0, 1, 3), (1, 1, -1), (-2, 1, -2))$ est une base ! Calculons alors $\mathrm{Mat}(f; \mathcal{B} \to \mathcal{B}')$: on trouve ...

$$\mathrm{Mat}(f; \mathcal{B} \to \mathcal{B}') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I_3$$

2) Nouvel exemple : matrice d'une projection. Soit $D = \mathrm{Vect}(u_1 = (1, 0))$ et soit $D' = \mathrm{Vect}(u_2 = (-1, 1))$. Comme $\mathcal{B} = ((1, 0), (-1, 1))$ est une base de $\mathbb{R}^2$ (elle est évidemment libre...) nous savons que $D \oplus D' = \mathbb{R}^2$. Soit alors $p : \mathbb{R}^2 \to \mathbb{R}^2$ qui envoie tout vecteur $u = v + v'$ sur $p(u) = v$.

Il n'est pas immédiat de trouver $p(x, y)$. Mais il est immédiat de déterminer $\mathrm{Mat}(p; \mathcal{B})$. En effet trouver cette matrice, c'est écrire chacun des vecteurs $p(u_1), p(u_2)$ sur la base (u_1, u_2). Or $u_1 \in D$, donc $u_1 = u_1 + 0$ et $p(u_1) = u_1$. Alors que $u_2 \in D'$ donc $u_2 = 0 + u_2$ et $p(u_2) = 0$. D'où :

$$\mathrm{Mat}(p; \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

En fait comme (u_1, u_2) est une base on peut écrire les vecteurs de base canonique sur cette base : $e_1 = u_1$ et $e_2 = u_2 + u_1$, et se servir de cette décomposition pour trouver $\mathrm{Mat}(p; \mathcal{B}^2 \to \mathcal{B}^2)$. Alors $p(e_1) = p(u_1) = u_1 = e_1$ et $p(e_2) = p(u_2 + u_1) = p(u_2) + p(u_1) = 0 + u_1 = e_1$. Ainsi :

$$\mathrm{Mat}(p; \mathcal{B}^2 \to \mathcal{B}^2) = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Ensuite $p(x, y) = p(xe_1 + ye_2) = xp(e_1) + yp(e_2) = (x + y)e_1 = (x + y, 0)$.

Remarque 3.8 (la matrice dépend des bases considérées !)**.** Comme on vient de le voir sur les exemples **la matrice dépend très fortement des bases choisies**.

Proposition 3.9.

Considérons une matrice A à p lignes et n colonnes, et l'application $F : X \mapsto AX$ envoyant une colonne à n lignes sur une colonne à p lignes. Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ l'application linéaire correspondante.

Alors la matrice de f de la base canonique de $\mathbb{R}^n$ dans la base canonique de $\mathbb{R}^p$ est la matrice A de départ.

Démonstration. En effet, par définition de f, l'image $f(e_j)$ du j-ème vecteur de la base canonique correspond à la colonne AE_j. Et nous savons que AE_j est la j-ème colonne de A. □

3.4. **Equations matricielles d'une application linéaire.**

On remarque d'abord qu'une application linéaire préserve les combinaisons linéaires à un nombre arbitraire de termes (et non pas seulement à deux termes, comme dans la définition).

Lemme 3.10 (d'autres caractérisations de la linéarité)**.**

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application. Alors les propriétés suivantes sont équivalentes :

(1) *pour deux vecteurs quelconques u, v dans $\mathbb{R}^n$ et deux réels quelconques λ, λ' on a $f(\lambda u + \lambda' u') = \lambda f(u) + \lambda' f(u')$ (autrement dit f est linéaire) ;*

(2) *(1bis) pour deux vecteurs quelconques u, v dans $\mathbb{R}^n$ on a $f(u + v) = f(u) + f(v)$ (f est* additive*) et pour tout vecteur $u \in \mathbb{R}^n$ et tout réel λ on a $f(\lambda u) = \lambda f(u)$ (f est* homogène (de degré 1)*)*

(3) *pour tout entier $k \geq 2$, pour toute suite de k vecteurs $u_1, u_2, \dots, u_k$ dans $\mathbb{R}^n$ et toute suite de k réels $\lambda_1, \lambda_2, \dots, \lambda_k$ on a $f(\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k) = \lambda_1 f(u_1) + \lambda_2 f(u_2) + \cdots + \lambda_k f(u_k)$*

Démonstration. (1) ⇒ (1bis) en faisant d'abord $\lambda = \lambda' = 1$, puis $u' = 0_{\mathbb{R}^n}$. Et si (1bis) est vérifiée alors (1) est vérifié : pour montrer que $f(\lambda u + \lambda' u') = \lambda f(u) + \lambda' f(u')$ on applique d'abord l'additivité, puis deux fois l'homogénéité.

La propriété (2) entraîne (1), en prenant $k = 2$. Montrons par récurrence que si (1) est vraie alors pour tout $k \geq 2$, l'application f transforme une combinaison linéaire à k termes $\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k$ en la combinaison linéaire= $\lambda_1 f(u_1) + \lambda_2 f(u_2) + \cdots + \lambda_k f(u_k)$.

C'est vrai pour $k = 2$ par définition de la linéarité. Ensuite supposons que ce soit vrai au rang k et considérons une combinaison linéaire à $k = 1$ termes : $w = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k + \lambda_{k+1} u_{k+1}$. Posons $u = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k$, $\lambda = 1$, $u' = u_{k+1}$, $\lambda' = \lambda_{k+1}$, de sorte que $w = \lambda u + \lambda' u'$. D'après (1) on a donc $f(w) = \lambda f(u) + \lambda' f(u')$, soit $f(w) = f(\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k) + \lambda_{k+1} f(u_{k+1})$. Par l'hypothèse de récurrence on sait que $f(\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k) = \lambda_1 f(u_1) + \lambda_2 f(u_2) + \cdots + \lambda_k f(u_k)$. D'où la propriété au rang $k + 1$. □

Le résultat suivant montre comment calculer l'image d'un vecteur en utilisant la matrice de l'application linéaire.

Théorème 3.11.

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Soit $\mathcal{B}$ une base de $\mathbb{R}^n$ et soit $\mathcal{B}'$ une base de $\mathbb{R}^p$. Posons $A = Mat(f, \mathcal{B} \to \mathcal{B}')$.

Pour tout vecteur $u \in \mathbb{R}^n$ soit X la colonne des coordonnées de u dans $\mathcal{B}$. Notons $f(u) = v$, et soit Y la colonne des coordonnées de v dans $\mathcal{B}$.

Alors $Y = AX$.

Démonstration. Notons $\mathcal{B} = (u_1, \dots, u_n)$ et $\mathcal{B}' = (u'_1, \dots, u'_p)$. Ecrivons alors $u = x_1u_1 + \cdots + x_nu_n$ et $v = y_1u'_1 + \cdots + y_pu'_p$, de sorte que $X = \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \cdots \\ y_p \end{pmatrix}$.

Comme $v = f(u)$ et f est linéaire nous savons, d'après le lemme 3.10, que $v = x_1f(u_1)+\cdots+x_nf(u_n)$.

D'après la définition de la matrice A, la colonne des coordonnées de $f(u_1)$ dans la base $\mathcal{B}'$ est justement la première colonne de A. Ainsi $f(u_1) = a_{11}u'_1 + a_{21}u'_2 + \cdots + a_{p1}u'_p$. Et on a de même

$$f(u_2) = a_{12}u'_1 + a_{22}u'_2 + \cdots + a_{p2}u'_p, \dots, f(u_n) = a_{1n}u'_1 + a_{2n}u'_2 + \cdots + a_{pn}u'_p$$

Dans la relation $v = x_1f(u_1) + \cdots + x_nf(u_n)$ remplaçons maintenant chaque $f(u_j)$ par son expression sur la base $\mathcal{B}'$, puis mettons chaque vecteur u'_i en facteur. Nous obtenons alors :

$$\begin{aligned} v &= x_1(a_{11}u'_1 + a_{21}u'_2 + \cdots + a_{p1}u'_p) + \cdots + x_n(a_{1n}u'_1 + a_{2n}u'_2 + \cdots + a_{pn}u'_p) \\ &= (x_1a_{11} + \cdots + x_na_{1n})u'_1 + \cdots + (x_1a_{p1} + \cdots + x_na_{pn})u'_p \end{aligned}$$

Comme $\mathcal{B}'$ est une base nous avons identifié les coordonnées de v :

$$\begin{cases} y_1 &= x_1a_{11} + \cdots + x_na_{1n} = a_{11}x_1 + \cdots + a_{1n}x_n \\ & \cdots \\ y_p &= x_1a_{p1} + \cdots + x_na_{pn} = a_{p1}x_1 + \cdots + a_{pn}x_n \end{cases}$$

Ce système de p équations traduit exactement l'égalité matricielle $Y = AX$. □

Théorème 3.12 (une application linéaire est complètement déterminée par sa matrice)**.**

Soit A une matrice à p lignes et n colonnes. Soit $\mathcal{B}$ une base de $\mathbb{R}^n$ et soit $\mathcal{B}'$ une base de $\mathbb{R}^p$.

Alors il existe une unique application linéaire $f : \mathbb{R}^n \to \mathbb{R}^p$ telle que $Mat(f; \mathcal{B} \to \mathcal{B}') = A$.

Formule pour f :

Notons c_1 le vecteur de $\mathbb{R}^p$ dont les coordonnées dans $\mathcal{B}'$ correspondent au premier vecteur colonne. Introduisons de même le vecteur c_2 de $\mathbb{R}^p$ dont les coordonnées dans $\mathcal{B}'$ correspondent au deuxième vecteur colonne, ainsi que $c_3, \dots, c_n$ [c comme **c***olonne]. Alors f est l'application qui envoie le vecteur $u \in \mathbb{R}^n$ de coordonnées $(y_1, \dots, y_n)$ dans la base $\mathcal{B}$ sur le vecteur $y_1c_1 + \cdots + y_nc_n$. Autrement dit :*

$$f(y_1u_1 + \cdots + y_nu_n) = y_1c_1 + \cdots + y_nc_n$$

Démonstration. Posons $\mathcal{B} = (u_1, u_2, \dots, u_n)$.

• Unicité (et formule).

Supposons que $f : \mathbb{R}^n \to \mathbb{R}^p$ est une application linéaire et que $\mathrm{Mat}(f; \mathcal{B} \to \mathcal{B}') = A$. Alors $f(u_1)$ est le vecteur dont les coordonnées dans $\mathcal{B}'$ correspondent au premier vecteur colonne - c'est le vecteur que l'on a noté c_1. Ainsi $f(u_1) = c_1$. De même $f(u_2) = c_2, \dots f(u_n) = c_n$. Maintenant, comme f est linéaire, pour toute combinaison linéaire $u = y_1u_1 + \cdots + y_nu_n$ on aura :

$$f(u) = f(y_1u_1 + \cdots + y_nu_n) = y_1f(u_1) + \cdots + y_nf(u_n) = y_1c_1 + \cdots + y_nc_n.$$

On aurait aussi pu déduire l'unicité (et même la formule) de l'équation matricielle des applications linéaires (théorème 3.11 ci-dessus). En effet d'après cette équation nous savons que $f(u)$ est le vecteur dont les coordonnées dans $\mathcal{B}'$ sont données par la colonne $A\begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix}$. Or nous pouvons écrire

$$\begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix} = y_1 \begin{pmatrix} 1 \\ \cdots \\ 0 \end{pmatrix} + \cdots + y_n \begin{pmatrix} 0 \\ \cdots \\ 1 \end{pmatrix} = y_1 E_1 + \cdots + y_n E_n.$$

Par la linéarité du produit matriciel nous avons alors

$$A\begin{pmatrix} y_1 \\ \cdots \\ y_n \end{pmatrix} = y_1 AE_1 + \cdots + y_n AE_n$$

La colonne AE_j est simplement la j-ème colonne de A, et par définition l'unique vecteur de $\mathbb{R}^p$ dont les coordonnées dans $\mathcal{B}'$ sont AE_j est c_j. Un calcul facile montre que $y_1 AE_1 + \cdots + y_n AE_n$ est la colonne des coordonnées dans $\mathcal{B}'$ de la combinaison linéaire $y_1 c_1 + \cdots + y_n c_n$. Ainsi nous retrouvons la formule $f(u) = y_1 c_1 + \cdots + y_n c_n$.

• Existence.

Considérons l'application $f : \mathbb{R}^n \to \mathbb{R}^p$ définie comme suit. Pour tout vecteur u de $\mathbb{R}^n$ soit $(y_1, \ldots, y_n)$ les coordonnées de u dans $\mathcal{B}$: on pose $f(u) = y_1 c_1 + \cdots + y_n c_n$. Vérifions que f répond au problème.

D'abord f est linéaire. En effet si $u = y_1 u_1 + \cdots + y_n u_n$, si $u' = y'_1 u_1 + \cdots + y'_n u_n$ et si $\lambda, \lambda' \in \mathbb{R}$ nous devons déterminer $f(\lambda u + \lambda' u')$. Or il est facile de trouver les coordonnées $(z_1, \ldots, z_n)$ de $w = \lambda u + \lambda' u'$, car

$$w = \lambda(y_1 u_1 + \cdots + y_n u_n) + \lambda'(y'_1 u_1 + \cdots + y'_n u_n) = (\lambda y_1 + \lambda' y'_1)u_1 + \cdots + (\lambda y_n + \lambda' y'_n)u_n$$

, donc $z_1 = \lambda y_1 + \lambda' y'_1, \ldots, z_n = \lambda y_n + \lambda' y'_n$.

Par définition de f cela donne $f(w) = (\lambda y_1 + \lambda' y'_1)c_1 + \cdots + (\lambda y_n + \lambda' y'_n)c_n$. En distribuant puis en regroupant, enfin en factorisant à nouveau on obtient :

$$f(w) = (\lambda y_1 c_1 + \lambda' y'_1 c_1) + \cdots + (\lambda y_n c_n + \lambda' y'_n c_n) = \lambda(y_1 c_1 + \cdots + y_n c_n) + \lambda'(y'_1 c_1 + \cdots + y'_n c_n)$$

ce qui signifie que $f(\lambda u + \lambda' u') = \lambda f(u) + \lambda' f(u')$, autrement dit f est linéaire.

Reste à vérifier que $\mathrm{Mat}(f; \mathcal{B} \to \mathcal{B}') = A$. Or (pour tout $j = 1, \ldots, n$) on a $f(u_j) = c_j$ par définition même de f. Et par la définition de c_j cela veut exactement dire que la colonne des coordonnées de $f(u_j)$ dans $\mathcal{B}'$ est la j-ème colonne de A.

□

3.5. **Equation matricielle dans les bases canoniques.** On reprend les résultats précédents dans le cas où les bases sont les bases canoniques.

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire, soit A sa matrice dans les bases canoniques. D'après la formule du théorème 3.12 on obtient

$$f(x_1, \ldots, x_n) = x_1 c_1 + \cdots + x_n c_n$$

où c_i est le vecteur dont les coordonnées dans la base canonique correspondent à la i-éme colonne de A, bref c_i **est le i-éme vecteur-colonne de la matrice** A.

La i-ème composante de $f(x_1, \ldots, x_n)$ est donc $a_{i1}x_1 + \cdots + a_{in}x_n$, ce qui correspond au produit de la i-ème ligne de A par le vecteur colonne X correspondant à $(x_1, \ldots, x_n)$. Autrement dit la colonne correspondant à $f(x_1, \ldots, x_n)$ est AX. Donc f est la version vectorielle de l'application entre colonnes $X \mapsto AX$: ainsi f coïncide avec l'application linéaire déjà associée à une matrice.

Ceci prouve que toute application linéaire $f : \mathbb{R}^n \to \mathbb{R}^p$ est associée à une matrice convenablement choisie.

Une autre façon de faire cette identification :

$$A\begin{pmatrix}x_1\\x_2\\\dots\\x_n\end{pmatrix}=A(x_1\begin{pmatrix}1\\0\\\dots\\0\end{pmatrix}+x_2\begin{pmatrix}0\\1\\\dots\\0\end{pmatrix}+\dots+x_n\begin{pmatrix}0\\0\\\dots\\1\end{pmatrix})=x_1A\begin{pmatrix}1\\0\\\dots\\0\end{pmatrix}+x_2A\begin{pmatrix}0\\1\\\dots\\0\end{pmatrix}+\dots+x_nA\begin{pmatrix}0\\0\\\dots\\1\end{pmatrix}$$

Or on sait que $AE_i = C_i$, la i-ème colonne de A. Donc finalement :

$$A\begin{pmatrix}x_1\\x_2\\\dots\\x_n\end{pmatrix}=x_1C1+x_2C_2+\dots+x_nC_n$$

C'est la version en colonnes de l'équation

$$f(x_1,x_2,\dots,x_n)=x_1c_1+x_2c_2+\dots+x_nc_n$$

4. Etude de l'équation $f(u) = v$. Noyau et image d'une application linéaire.

Voici un problème général concernant les applications (linéaires) $f : \mathbb{R}^n \to \mathbb{R}^p$:

un vecteur v dans l'espace d'arrivée $\mathbb{R}^p$ étant fixé, l'équation $f(u) = v$ a t-elle une solution en u ? si oui, combien ? quel est l'ensemble de toutes les solutions ?

Définition 4.1 (noyau, image)**.**

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Alors le *noyau de* f est l'ensemble $\mathrm{Ker}(f) = \{u \in \mathbb{R}^n,\ f(u) = 0_{\mathbb{R}^p}\}$. Et l'*image de* f est l'ensemble $\mathrm{Im}(f) = \{f(u),\ u \in \mathbb{R}^n\} \subset \mathbb{R}^p$.

Ainsi $\mathrm{Im}(f)$ correspond à l'ensemble des vecteurs v pour lesquels l'équation $f(u) = v$ admet au moins une solution en u. Et $\mathrm{Ker}(f)$ est l'ensemble des solutions de l'équation $f(u) = 0_{\mathbb{R}^p}$.

Remarque 4.2. Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire, soit $v \in \mathbb{R}^p$. Alors l'équation $f(u) = v$ admet une solution si et seulement si v a un antécédent, c'est à dire $v \in \mathrm{Im}(f)$.

Donc ou bien v n'a aucun antécédent ($v \notin \mathrm{Im}(f)$), et dans ce cas l'équation $f(u) = v$ n'a aucune solution (en u). Ou bien v admet (au moins) un antécédent u_0 ($f(u_0) = v$ donc $v \in \mathrm{Im}(f)$). Dans ce cas l'ensemble de tous les antécédents de v est l'ensemble de tous les vecteurs de la forme $u = u_0 + w$, avec $w \in \mathrm{Ker}(f)$. En effet si $u = u_0+w$ avec $w \in \mathrm{Ker}(f)$ alors par linéarité $f(u) = f(u_0)+f(w) = v+0_{\mathbb{R}^p} = v$, donc u est un antécédent de v. Réciproquement si $f(u) = v$, alors $f(u) - f(u_0) = v - v = 0$. Or par linéarité $f(u) - f(u_0) = f(u - u_0)$, ce qui donne $f(u - u_0) = 0$ et donc $u - u_0 \in \mathrm{Ker}(f)$. Ainsi $u = (u - u_0) + u_0$ et nous venons de voir que $u - u_0 \in \mathrm{Ker}(f)$.

En utilisant l'interprêtation matricielle des systèmes, l'ensemble des solutions de l'équation $f(u) = v$ est aussi l'ensemble des solutions du système linéaire correspondant à l'équation matricielle $AU = V$. Nous en déduisons que l'ensemble des solutions d'un système linéaire est une solution particulière U_0 + l'ensemble des solutions du système linéaire homogène associé (souvenir, souvenir).

Proposition 4.3 (surjectivité et image, injectivité et noyau)**.**

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Alors :

1) f est surjective $\iff$ $Im(f) = \mathbb{R}^p$.

2) f est injective $\iff$ $Ker(f) = \{0\}$.

Rappelons qu'une application est surjective si et injective si .

Démonstration. 1) Par définition f est surjective $\iff$ tout vecteur $v \in \mathbb{R}^p$ a au moins un antécédent par f. Or justement $\mathrm{Im}(f)$ est l'ensemble des vecteurs $v \in \mathbb{R}^p$ qui ont au moins un antécédent.

2) Si f est injective, alors $0_{\mathbb{R}^p}$ a au plus un antécédent. Mais $0_{\mathbb{R}^n}$ est toujours un antécédent (d'après le lemme). Donc $\mathrm{Ker}(f) = \{0_{\mathbb{R}^n}$.

Réciproquement supposons $\mathrm{Ker}(f) = \{0_{\mathbb{R}^n}\}$. Soit u, u' deux vecteurs de $\mathbb{R}^n$ ayant même image v alors $0_{\mathbb{R}^p} = v - v = f(u) - f(u')$. Or f est linéaire, donc $f(u) - f(u') = f(u - u')$. On obtient $f(u - u') = 0_{\mathbb{R}^p}$, donc $u - u' \in \mathrm{Ker}(f)$. Par hypothèse $\mathrm{Ker}(f) = \{0_{\mathbb{R}^n}\}$, donc $u - u' = 0_{\mathbb{R}^n}$, autrement dit $u = u'$. On a donc montré que quand un vecteur v a deux antécédents u et u' alors nécessairement $u = u'$: c'est l'injectivité de f.

□

Lemme 4.4 (le noyau et l'image sont des sous-espaces vectoriels)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Alors Im(f) est un sous-espace vectoriel de $\mathbb{R}^p$ et Ker(f) est un sous-espace vectoriel de $\mathbb{R}^n$.*

Démonstration. Soient $v, v' \in \mathrm{Im}(f)$ et soient $\mu, \mu' \in \mathbb{R}$. Par définition de l'image il existe $u \in \mathbb{R}^n$ tel que $f(u) = v$. Et de même il existe $u' \in \mathbb{R}^n$ tel que $f(u') = v'$. Alors $f(\mu u + \mu' u') = \mu f(u) + \mu' f(u') = \mu v + \mu' v'$, de sorte que la combinaison linéaire $muv + \mu' v'$ est encore dans l'image. D'autre part $f(0_{\mathbb{R}^n} = 0_{\mathbb{R}^p}$, donc $0_{\mathbb{R}^p} \in \mathrm{Im}(f)$. Nous avons vérifié les deux axiomes de sous-espace vectoriel pour $\mathrm{Im}(f)$.

Soient maintenant $u, u' \in \mathrm{Ker}(f)$ et soient $\lambda, \lambda' \in \mathbb{R}$. Par définition du noyau on a $f(u) = f(u') = 0_{\mathbb{R}^p}$. Donc en utilisant la linéarité de f on obtient $f(\lambda u + \lambda' u') = \lambda f(u) + \lambda' f(u') = 0_{\mathbb{R}^p}$. Ainsi bien la combinaison linéaire $\lambda u + \lambda' u'$ appartient bien au noyau. D'autre part $f(0_{\mathbb{R}^n} = 0_{\mathbb{R}^p}$, donc $0_{\mathbb{R}^n} \in \mathrm{Ker}(f)$. Nous avons vérifié les deux axiomes de sous-espace vectoriel pour $\mathrm{Ker}(f)$.

□

Lemme 4.5. *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Soit A sa matrice dans les bases canoniques. Alors le noyau et l'image de f sont le noyau et l'image de A.*

Démonstration. On a vu que f correspond à $X \mapsto AX$.

Donc $u = (x_1, \ldots, x_n)$ est dans le noyau de f ssi $AX = 0$, ce qui définit le noyau de A !

$\mathrm{Im}(f)$ est $\{f(u), u \in \mathbb{R}^n\}$. Or on a vu que $f(u) = f(x_1, \ldots, x_n) = x_1 c_1 + \cdots + x_n c_n$ où $c_1, \ldots, c_n$ sont les vecteurs colonnes de A.

Donc tout vecteur de $\mathrm{Im}(f)$ est combinaison linéaire des vecteurs colonnes de A. Réciproquement une combinaison linéaire des vecteurs colonnes s'écrit $v = x_1 c_1 + \cdots + x_n c_n$ pour un certain choix de réels $x_1, \ldots, x_n$. Alors le vecteur u définit par $u = (x_1, \ldots, x_n)$ vérifie $f(u) = x_1 c_1 + \cdots + x_n c_n = v$, donc $v \in \mathrm{Im}(f)$.

Nous avons montré que $\mathrm{Im}(f)$ est l'ensemble des combinaisons linéaires des vecteurs colonnes - par définition ceci est $\mathrm{Im}(A)$. D'où $\mathrm{Im}(f) = \mathrm{Im}(A)$.

□

Corollaire 4.6. *Ker(f) est l'ensemble des solutions du système homogène de matrice A (d'où un système d'équations cartésiennes du noyau, $AX = 0$). Et Im(f) est le sous-espace vectoriel engendré par les vecteurs colonnes de A (d'où une partie génératrice de l'image).*

Définition 4.7. Le *rang de f* est $\mathrm{rg}(f) = \dim(\mathrm{Im}(f))$. D'après le lemme 4.5 c'est aussi le rang de la matrice de f dans les bases canoniques.

Pour trouver une base de Im(f) - et du même coup le rang de f - il suffit donc d'appliquer l'algorithme du rang à la suite des vecteurs colonnes.

Proposition 4.8 (image d'une suite génératrice, image d'une suite liée)**.**

Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Soit $(u_1, \ldots, u_k)$ une suite de vecteur et pour chaque $i = 1, \ldots, k$ notons $v_i = f(u_i)$.

1) Soit E le sous-espace vectoriel engendré par $(u_1, \ldots, u_k)$. Alors $f(E) = \{f(u), u \in E\}$ est un sous-espace vectoriel, et il est engendré par la suite des images $(v_1, \ldots, v_k)$.

2) Si $(u_1, \ldots, u_k)$ est liée alors $(v_1, \ldots, v_k)$ est liée.

2bis) Si $(v_1, \ldots, v_k)$ est libre alors $(u_1, \ldots, u_k)$ est libre.

Démonstration. 1) (On procède comme lorsque $E = \mathbb{R}^n$...)

Montrons que $f(E) = \mathrm{Vect}(v_1, \dots, v_k)$ - par double inclusion.

Si $v \in f(E)$ alors $v = f(u)$ pour un certain $u \in E$. Comme $(u_1, \dots, u_k)$ engendre E, on peut trouver des coefficients $\lambda_1, \dots, \lambda_k$ tels que $u = \lambda_1 u_1 + \cdots + \lambda_k u_k$. Comme f est linéaire on obtient $f(u) = f(\lambda_1 u_1 + \cdots + \lambda_k u_k) = \lambda_1 f(u_1) + \cdots + \lambda_k f(u_k)$, donc $v = \lambda_1 v_1 + \cdots + \lambda_k v_k$. Autrement dit $v \in \mathrm{Vect}(v_1, \dots, v_k)$.

Réciproquement si $v \in \mathrm{Vect}(v_1, \dots, v_k)$ alors v peut s'écrire $v = \lambda_1 v_1 + \cdots + \lambda_k v_k$ pour certains coefficients réels λ_i, et en posant $u = \lambda_1 u_1 + \cdots + \lambda_k u_k$ on voit que le calcul ci-dessus donne $f(u) = v$, donc $v \in f(E)$.

2) Si $(u_1, \dots, u_k)$ est liée alors il y a des coefficients réels non tous nuls $\lambda_1, \dots, \lambda_k$ tels que $\lambda_1 u_1 + \cdots + \lambda_k u_k = 0$. En appliquant f on obtient $f(\lambda_1 u_1 + \cdots + \lambda_k u_k) = f(0)$, donc par linéarité de f $\lambda_1 f(u_1) + \cdots + \lambda_k f(u_k) = 0$. Autrement dit $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$ ce qui prouve que $(v_1, \dots, v_k)$ est liée, puisque l'un au moins des coefficients $\lambda_1, \dots, \lambda_k$ n'est pas nul.

2bis) C'est une conséquence directe de 2) : si $(u_1, \dots, u_k)$ était liée alors $(v_1, \dots, v_k)$ serait liée, or elle est supposée libre.

□

Une application linéaire ne peut pas délier ce qui était lié, une application linéaire conserve les relations de dépendance linéaire.

Lemme 4.9. *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire et soit r son rang. Soit $(v_1, \dots, v_r)$ une base de Im(f). Pour chaque v_i soit u_i un vecteur de $\mathbb{R}^n$ tel que $f(u_i) = v_i$.*

Alors $(u_1, \dots, u_r)$ est libre. De plus si on pose $E =$ Vect$(u_1, \dots, u_r)$ alors $E \oplus$ Ker$(f) = \mathbb{R}^n$.

Démonstration. $(u_1, \dots, u_r)$ est libre car son image par f est $(v_1, \dots, v_r)$, une base de $\mathrm{Im}(f)$, donc en particulier une suite libre.

Montrons d'abord $E \cap \mathrm{Ker}(f) = \{0\}$ (la somme est directe).

Pour cela soit $u \in E \cap \mathrm{Ker}(f)$. D'une part $u \in E$ donc u peut s'écrire $u = y_1 u_1 + \cdots + y_r u_r$. D'autre part $u \in \mathrm{Ker}(f)$ donc $f(u) = 0$. Or par linéarité $f(u) = y_1 f(u_1) + \cdots + y_r f(u_r) = y_1 v_1 + \cdots + y_r v_r$. On obtient $0 = y_1 v_1 + \cdots + y_r v_r$ et comme $(v_1, \dots, v_r)$ est libre, cela donne $y_1 = \cdots = y_r = 0$, donc finalement $u = 0$. réciproquement bien sûr $0 \in E \cap \mathrm{Ker}(f)$ car une intersection de sous-espaces vectoriels est un sous-espace vectoriel.

Montrons enfin $E + \mathrm{Ker}(f) = \mathbb{R}^n$ (la somme est tout).

Pour $w \in \mathbb{R}^n$ considérons l'image $v = f(w)$. Alors $v \in \mathrm{Im}(f)$. Donc on peut décomposer v sur la base $(v_1, \dots, v_r)$: $v = y_1 v_1 + \cdots + y_r v_r$. Posons $u = y_1 u_1 + \cdots + y_r u_r$. Par linéarité on obtient $f(u) = y_1 f(u_1) + \cdots + y_r f(u_r) = y_1 v_1 + \cdots + y_r v_r = v$, donc $f(u) = f(w)$. Alors $f(w-u) = f(w) - f(u) = v - v = 0$. Donc $w - u \in \mathrm{Ker}(f)$, et ainsi $w = (w-u) + u$ appartient à $\mathrm{Ker}(f) + E$.

Nous avons montré que tout $w \in \mathbb{R}^n$ est dans $E + \mathrm{Ker}(f)$, soit l'inclusion $\mathbb{R}^n \subset E + \mathrm{Ker}(f)$. Mais l'inclusion inverse est toujours vraie. □

Le sous-espace E du lemme est supplémentaire du noyau. Donc d'après la formule de la dimension $\dim(\mathrm{Ker}(f)) + \dim(E) = n$. Mais justement E a même dimension que $\mathrm{Im}(f)$, le rang r. Nous obtenons donc :

Théorème 4.10 (théorème noyau-image)**.** *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire.*

Alors dim$(Ker(f))$ + dim$(Im(f)) = n$, autrement dit dim$(Ker(f))$ + rg$(f) = n$.

Exemple 4.11.

1) Pour $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$ les deux vecteurs colonnes sont $(1, 2)$ et $(3, 6)$, donc proportionnels. Une base de l'image est donc $(1, 2)$. Alors nécessairement le noyau est une droite vectorielle. La relation $(3, 6) = 3(1, 2)$ signifie que $f(e_2) = 3f(e_1)$ donc $f(3e_1 - e_2) = 0$, d'où le noyau contient le vecteur non

nul $(3,-1)$. Mais comme le noyau et de dimension 1, ce vecteur non nul en est une base. Dans ce cas $((1,2),(3,-1))$ est libre. Donc la droite image et la droite noyau sont supplémentaires dans $\mathbb{R}^2$. Mais ce n'est pas toujours le cas.

2) Ainsi pour $A = \begin{pmatrix} 3 & 9 \\ -1 & -3 \end{pmatrix}$ l'image est : $\mathrm{Vect}(3,-1)$; et le noyau est : $\mathrm{Vect}(3,-1)$, la même droite vectorielle. Donc même si la somme des deux dimension vaut bien 2, les deux sous-espaces ne sont pas supplémentaires.

Corollaire 4.12 (théorème noyau-image pour les matrices)**.**
Soit A une matrice $p \times n$. Alors $dim\,(Ker(A)) + rg(A) = n$

Démonstration. On applique le théorème noyau-image à l'application linéaire f de matrice A dans les bases canoniques. Puis on conclut avec les identifications $\mathrm{Im}(A) = \mathrm{Im}(f)$, $\mathrm{Ker}(A) = \mathrm{Ker}(f)$.
□

On rappelle qu'une application $f : \mathbb{R}^n \to \mathbb{R}^p$ est *bijective* si tout élément v admet un et un seul antécédent (noté $f^{-1}(v)$).

Corollaire 4.13. *Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Si f est injective alors $n \leq p$. Si est surjective alors $n \geq p$. Et si f est bijective alors $n = p$.*

Démonstration. Nous avons $\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = n$.

Si f est injective alors $\mathrm{Ker}(f) = \{0_{\mathbb{R}^n}\}$ donc $\dim(\mathrm{Ker}(f)) = 0$ et $\dim(\mathrm{Im}(f)) = n$. Comme $\mathrm{Im}(f) \subset \mathbb{R}^p$ il en résulte que $n \leq p$.

Si f est surjective alors $\mathrm{Im}(f) = \mathbb{R}^p$, donc $\dim(\mathrm{Im}(f)) = p$ et il en résulta que $\dim(\mathrm{Ker}(f)) = n - p$. Comme une dimension est toujours ≥ 0 il en résulte que $n \geq p$.

Le cas bijectif est la combinaison des deux cas injectif et surjectif.
□

En particulier il n'existe pas d'application linéaire injective de $\mathbb{R}^n$ dans $\mathbb{R}^p$ si $n > p$, ni d'application linéaire surjective si $n < p$.

On peut dire précisément ce qui se passe lorsqu'une application linéaire lie ce qui était libre :

Lemme 4.14 (relation de dépendance linéaire entre vecteurs colonnes $\Rightarrow$ vecteur du noyau)**.**
Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Soit $(u_1, \dots, u_k)$ une suite libre de $\mathbb{R}^n$ telle que la suite $(v_1 = f(u_1), \dots, v_k = f(u_k))$ est liée. Alors toute relation de dépendance linéaire $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$ non triviale (α_i non tous nuls) entre les v_i fabrique un vecteur $u = \alpha_1 u_1 + \cdots + \alpha_k u_k = 0$ qui est dans le noyau et est non nul.

Corollaire 4.15.
Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. Si f est injective et $(u_1, \dots, u_k)$ est une suite libre alors la suite $(v_1 = f(u_1), \dots, v_k = f(u_k))$ est libre.

Proposition 4.16 (base du noyau)**.**
Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire de rang r. Soit A sa matrice dans les bases canoniques. Soit $(c_1, \dots, c_n)$ la suite des vecteurs colonnes de A. On applique l'algorithme du rang à $(c_1, \dots, c_n)$, il reste donc r vecteurs. Autrement dit l'algorithme rejette $n - r$ vecteurs-colonnes, car ils sont combinaisons linéaires de vecteurs-colonnes précédents.

Pour chaque vecteur-colonne rejeté c_i, on obtient une relation de dépendance linéaire entre certains vecteurs colonnes de la forme $c_i = \alpha_1 c_1 + \cdots + \alpha_{i-1} c_{i-1}$, ce qui fournit un vecteur du noyau $u_i = e_i - (\alpha_1 e_1 + \cdots + \alpha_{i-1} e_{i-1})$. Alors la suite des vecteurs du noyau ainsi obtenue est libre (car échelonnée) et donc forme en fait une base du noyau.

Exemple 4.17.

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

Dans ce cas $f(x, y, z, t) = x - y + 2z - 3t, 2x + 3z - 2t)$. La suite des vecteurs colonnes est $(c_1 = (1, 1, 2), c_2 = (1, -1, 0), c_3 = (1, 2, 3), c_4 = (1, -3, -2))$.

On applique l'algorithme du rang : alors (c_1, c_2) est une base de $\mathrm{Im}(f)$, et on trouve $2c_3 = 3c_1 - c_2$, et $c_4 = -c_1 + 2c_2$. Ensuite comme f de rang 2 on sait d'après le TNI que $\dim(\mathrm{Ker}(f)) = 2$. La relation $2c_3 = 3c_1 - c_2$ s'écrit aussi $3c_1 - c_2 - 2c_3 = 0$, soit $(3e_1 - e_2 - 2e_3) = 0$, donc $v_1 = (3, -1, -2, 0) \in \mathrm{Ker}(f)$. En utilisant la deuxième relation obtenue (pour rejeter c_4) on trouve $f(e_1 - 2e_2 + e_4) = 0$ donc $v_2 = (1, -2, 0, 1) \in \mathrm{Ker}(f)$. La suite (v_1, v_2) est libre donc c'est une base de $\mathrm{Ker}(f)$.

5. Etude de l'équation matricielle $AU = V$.

Donnons nous une matrice A à p lignes et n colonnes, ainsi qu'une matrice colonne à p lignes V. On va traduire certaines propriétés de l'équation $AU = V$ (où U est une matrice colonne à n lignes recherchée) en des propriétés de la matrice A. On sait que cela est équivalent à rechercher les solutions du système linéaire associé à A, de second membre V. Ou encore : résoudre $AU = V$ revient à résoudre $f(u) = v$ pour f l'application linéaire dont la matrice dans les bases canoniques est A, et u, v les vecteurs de $\mathbb{R}^n, \mathbb{R}^p$ correspondant aux colonnes U, V.

(1) En termes des vecteurs v, l'ensemble des V tels que $AU = V$ possède au moins une solution correspond exactement à $\mathrm{Im}(A)$.

(2) De même, en termes des vecteurs u, l'ensemble des U tels que $AU = 0$ correspond exactement à $\mathrm{Ker}(A)$.

(3) L'équation a au moins une solution pour tout second membre V ssi $\mathrm{Im}(A) = \mathbb{R}^p$ (donc en particulier $n \geq p$). Par abus de langage on dit parfois que A est *surjective* (mais en fait c'est l'application linéaire associée à A qui l'est).

(4) L'équation a au plus une solution pour tout second membre V ssi $\mathrm{Ker}(A) = \{0_{\mathbb{R}^n}\}$ (donc en particulier $n \leq p$). Par abus de langage on dit parfois que A est *injective* (mais en fait c'est l'application linéaire associée à A qui l'est).

(5) L'équation a toujours une et une seule solution ssi $\mathrm{Im}(A) = \mathbb{R}^p$ et $\mathrm{Ker}(A) = \{0_{\mathbb{R}^n}\}$ (dans ce cas $p = n$ la matrice est dite *carrée de taille* n).

Théorème 5.1 (caractérisation des matrices inversibles)**.** *Soit A une matrice carrée de taille n. Les propriétés suivantes sont équivalentes :*

(1) *$Im(A) = \mathbb{R}^n$ et $Ker(A) = \{0_{\mathbb{R}^n}\}$*
(2) *$Im(A) = \mathbb{R}^n$*
(2bis) *la suite $(c_1, \ldots, c_n)$ des vecteurs colonne de A est une base $\mathbb{R}^n$*
(3) *$Ker(A) = \{0_{\mathbb{R}^n}\}$*
(4) *Il existe une matrice B carrée de taille n telle que $AB = I_n$*
(5) *Il existe une matrice B carrée de taille n telle que $BA = I_n$*
(6) *Il existe une matrice B carrée de taille n telle que $AB = I_n$ et $BA = I_n$.*

Lorsque ces conditions sont vérifiées on dit que A est inversible *et il existe alors exactement une matrice B vérifiant l'une des trois dernières conditions - qu'on appelle l'inverse de A, et qu'on note A^{-1}.*

Remarque : la matrice A^{-1} vérifie $A.A^{-1} = A^{-1}.A = I_n$, elle est donc également inversible et son inverse est la matrice A de départ (prendre l'inverse est une opération symétrique).

Démonstration. Les implications (1) $\Rightarrow$ (2) et (1) $\Rightarrow$ (3) sont évidentes.

Le théorème noyau-image pour les matrices donne ici $\dim(\mathrm{Im}(A)) + \dim(\mathrm{Ker}(A)) = n$. Si par exemple $\mathrm{Im}(A) = \mathbb{R}^n$ alors on en déduit $\dim(\mathrm{Ker}(A)) = 0$, d'où $\mathrm{Ker}(A) = \{0_{\mathbb{R}^n}\}$. Ainsi (2) $\Rightarrow$ (1).

On prouve de même (3) $\Rightarrow$ (1).

L'équivalence (2) $\iff$ (2*bis*) est évidente.

Montrons que (2) $\iff$ (4). Pour cela montrons d'abord (2) $\Rightarrow$ (4). Puisque $\mathrm{Im}(A) = \mathbb{R}^n$ pour chaque colonne $E_1, E_2, \ldots, E_n$ il existe des colonnes $B_1, B_2, \ldots, B_n$ telles que $AB_1 = E_1, AB_2 = E_, \ldots, AB_n = E_n$. Soit B la matrice carrée dont les colonnes sont $B_1, \ldots, B_n$. Alors la k-ème colonne de AB est AB_k, soit E_k - la k-ème colonne de I_n. Ceci montre que $AB = I_n$

Montrons (4) $\Rightarrow$ (2). Soit donc B une matrice carrée telle que $AB = I_n$. Alors pour résoudre $AU = V$, il suffit de poser $U = BV$: on a bien $AU = A(BV) = (AB)V = I_nV = V$.

Montrons maintenant (4) $\Rightarrow$ (5). Plus précisément si une matrice B vérifie la condition $AB = I_n$ alors elle vérifie la condition $BA = I_n$. En effet on a $(AB)A = I_nA = A$, donc $A(BA - I_n) = 0$ (matrice carrée nulle). Ainsi chaque vecteur colonne de $BA - I_n$ est dans $\mathrm{Ker}(A)$. Mais nous avons déjà vu que (4) $\iff$ (1) $\iff$ (3) : donc $\mathrm{Ker}(A) = \{0_{\mathbb{R}^n}\}$. Il en résulte que chaque vecteur colonne de $BA - I_n$ est nul, donc $BA - I_n = 0$, autrement dit $BA = I_n$.

Puisque (6) est la conjonction de (4) et (5) nous avons aussi du même coup (4) $\Rightarrow$ (6). Et comme (6) $\Rightarrow$ (4) est évidente cela nous donne (4) $\iff$ (6).

Pour boucler la boucle il suffit de montrer, par exemple, que (5) $\Rightarrow$ (3). Supposons donc qu'une matrice carrée B vérifie $BA = I_n$. Alors toute colonne U correspondant à un vecteur de $\mathrm{Ker}(A)$ vérifie $AU = 0$, donc $B(AU) = 0$. Or $B(AU) = (BA)U = I_nB = B$. Ainsi $U = 0$ et donc en fait $\mathrm{Ker}(A) = \{0_{\mathbb{R}^n}\}$.

Unicité de l'inverse. Si A inversible on sait qu'il existe B telle que (4) $AB = I_n$, et on a vu que toute telle matrice B vérifie aussi (5) $BA = I_n$. En fait si B' vérifie (5) alors $B'(AB) = B'I_n = B'$ et $(B'A)B = I_nB = B$, et par associativité du produit de matrice $B = B'$. Donc les matrices qui vérifient (4) sont égales aux matrices qui vérifient (5). Et donc elles sont toutes égales entre elles.

□

Exemple 5.2. $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, A = \begin{pmatrix} 2 & 4 & 3 \\ 1 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}$

Remarque 5.3.

1) Montrer qu'une matrice carrée est inversible est beaucoup plus rapide que de calculer l'inverse.

2) Comme on l'a vu dans la preuve ci-dessus, la matrice inverse A^{-1} permet de résoudre les systèmes $AU = V$: la solution U est donnée par $U = A^{-1}V$.

Différentes méthodes de calcul de l'inverse :

On sait que l'inverse de A est une (la...) matrice B telle que $AB = I_n$. Autrement dit la matrice dont les colonnes $B_1, \ldots, B_n$ sont solutions de $AB_1 = E_1, \ldots, AB_n = E_n$. Pour trouver $B = A^{-1}$, on peut :

1) Ou bien directement résoudre le système "à paramètres" $AU = V$ par la méthode du pivot.

A la fin on obtient U en fonction de V, en fait on s'aperçoit que U est écrit sous la forme $U = BV$ pour B une certaine matrice - justement l'inverse.

2) Ou bien trouver à la main (lorsque n est petit) chaque colonne B_j, enutilisant la remarque suivante.

Chaque B_j correspond au coordonnées de e_j dans la base $(c_1, \ldots, c_n)$ formée par les vecteurs colonnes. En général on sait que A est inversible parce qu'on a vérifié que $(c_1, \ldots, c_n)$ est une base, en montant que tout vecteur e_j peut s'écrire comme combinaison linéaire de $(c_1, \ldots, c_n)$. Les coefficients de cette combinaison donnent la colonne B_j.

6. Matrices de passage.

Définition 6.1. Soient $\mathcal{B}, \mathcal{B}'$ deux bases de $\mathbb{R}^n$. La matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$ est la matrice carrée de taille n dont la i-ème colonne correspond aux coordonnées du i-ème vecteur de $\mathcal{B}'$ dans $\mathcal{B}$. On la notera $P_{\mathcal{B}\to\mathcal{B}'}$.

Trouver la matrice de passage c'est trouver les coordonnées de chaque vecteur de $\mathcal{B}'$ (dite la nouvelle base) dans $\mathcal{B}$ (dite l'ancienne base). Donc en général c'est très facile. Ce qui est plus dur c'est de trouver les coordonnées de chaque vecteur de l'ancienne base $\mathcal{B}$ dans la nouvelle base $\mathcal{B}'$, autrement dit de déterminer $P_{\mathcal{B}'\to\mathcal{B}}$.

Remarque : par définition on a aussi l'interprêtation $P_{\mathcal{B}\to\mathcal{B}'} = \mathrm{Mat}(\mathrm{id}, \mathcal{B}' \to \mathcal{B})$.

Lemme 6.2 (changement de coordonnées)**.**
Soient $\mathcal{B}, \mathcal{B}'$ deux bases de $\mathbb{R}^n$, et soit $P = P_{\mathcal{B}\to\mathcal{B}'}$ la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$. Pour tout vecteur $u \in \mathbb{R}^n$ notons X le vecteur colonne des coordonnées de u dans l'ancienne base $\mathcal{B}$, et notons X' le vecteur colonne des coordonnées de u dans la nouvelle base $\mathcal{B}'$.

Alors $X = PX'$.

Démonstration. Décomposons u sur les deux bases, l'ancienne et la nouvelle $u = x_1u_1 + \cdots + x_ku_k$ et $u = x_1'u_1' + \cdots + x_k'u_k'$. Pour relier les x_j' aux x_i, nous allons utiliser le fait que ce sont les coordonnées d'un même vecteur u dans deux bases différentes.

Les colonnes de la matrice P donnent les coordonnées des u_i' dans $\mathcal{B}$, autrement dit :

$$u_1' = p_{11}u_1 + \cdots + p_{n1}u_n, u_2' = p_{12}u_1 + \cdots + p_{n2}u_n, \ldots, u_n' = p_{1n}u_1 + \cdots + p_{nn}u_n.$$

Reportons alors l'expression de chacun des u_j' dans la décomposition $u' = x_1'u_1' + \cdots + x_k'u_k'$, nous obtenons :

$$u = x_1'u_1' + \cdots + x_k'u_k' = x_1'(p_{11}u_1 + \cdots + p_{n1}u_n) + \cdots + x_n'(p_{1n}u_1 + \cdots + p_{nn}u_n)$$

On regroupe les multiples de u_1, puis de u_2 etc... :

$$u = (x_1'p_{11} + \cdots + x_n'p_{1n})u_1 + \cdots + (x_1'p_{n1} + \cdots + x_n'p_{nn})u_n$$

Or on a déjà $u = x_1u_1 + \cdots + x_ku_k$ donc par identification des coordonnées dans $\mathcal{B}$:

$$\begin{cases} x_1 = x_1'p_{11} + \cdots + x_n'p_{1n} = p_{11}x_1' + \cdots + p_{1n}x_n' \\ \cdots \\ x_n = x_1'p_{n1} + \cdots + x_n'p_{nn} = p_{n1}x_1' + \cdots + p_{nn}x_n' \end{cases}$$

Et ceci signifie exactement $X = PX'$.

□

Corollaire 6.3 (passage $\Rightarrow$ inversible)**.** *Toute matrice de passage est inversible, et l'inverse de $P_{\mathcal{B}\to\mathcal{B}'}$ est $P_{\mathcal{B}'\to\mathcal{B}}$.*

Démonstration. Notons P' la matrice $P_{\mathcal{B}'\to\mathcal{B}}$, puis posons $A = PP'$. Pour montrer le corollaire nous allons prouver que $A = I_n$. La formule de changement de coordonnées donne à la fois $X = PX'$ et $X' = P'X$ (pour tout vecteur $u \in \mathbb{R}^n$!). Alors $X = PX' = P(P'X) = (PP')X$. Si nous appliquons n fois cette formule avec $u = u_1, u = u_2, \ldots, u = u_n$ nous obtenons $AE_1 = E_1, AE_2 = E_2, \ldots, AE_n = E_n$. Donc les colonnes de A sont celles de I_n, donc $A = I_n$, donc P est inversible d'inverse P'. □

Application : pour exprimer les coordonnées X' en fonction de X, il suffit d'appliquer la formule $X' = P'X$ (avec $P' = P_{\mathcal{B}'\to\mathcal{B}}$), soit encore $P^{-1}X = X'$.

Dans la pratique on peut calculer $(P_{\mathcal{B}\to\mathcal{B}'})^{-1}$ soit par les méthodes générales de calcul de l'inverse, soit en utilisant les formules trouvées pour montrer que $\mathcal{B}'$ est bien une base, i.e. est génératrice, i.e. permet de fabriquer chaque vecteur de la base de départ $\mathcal{B}$.

Réciproquement :

Lemme 6.4 (inversible $\Rightarrow$ passage)**.** *Toute matrice inversible A est la matrice de passage de la base canonique à la base formée des vecteurs-colonnes de A.*

Théorème 6.5. *Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une application linéaire. Soient $\mathcal{B}, \mathcal{B}'$ deux matrices de $\mathbb{R}^n$. Posons $A = Mat(f, \mathcal{B})$ et $A' = Mat(f, \mathcal{B}')$.*

Alors $A' = P^{-1}AP$.

Démonstration. Pour simplifier nous notons P' l'inverse de P, et nous montrons $A' = P'AP$.

Soit u un vecteur quelconque de $\mathbb{R}^n$, X la colonne de ses coordonnées dans $\mathcal{B}$, X' la colonne de ses coordonnées dans $\mathcal{B}'$. De même soit Y la colonne des coordonnées de $f(v)$ dans $\mathcal{B}$, et soit Y' la colonne des coordonnées de $f(v)$ dans $\mathcal{B}'$.

Alors d'après le théorème 3.11 nous savons que $Y = AX$ et $Y' = A'X'$. D'autre part $P'APX' = P'AX$ d'après le Lemme 6.2.

Donc $P'APX' = P'Y$, et en appliquant à nouveau le Lemme 6.2 on obtient $P'APX' = Y'$, soit $P'APX' = A'X'$. Cette relation étant valable pour tout vecteur colonne X', on en déduit que $P'AP = A'$ (prendre par exemple $X' = E_1, E_2, \dots, E_n$).

□