Chapitre 17

Matrices et applications linéaires

Sommaire

17.1 Matrices et applications linéaires

17.1.1 Matrice d'une famille de vecteurs dans une base

Définition 17.1.1 (matrice d'une famille de vecteurs dans une base)
Soit E un $\mathbb{K}$-espace vectoriel, de dim $n \geqslant 1$, muni d'une base $\varepsilon = (\varepsilon_i)_{1 \leqslant i \leqslant n}$.
Soit $v = (v_j)_{1 \leqslant j \leqslant p}$ une famille de p vecteurs de E. Pour tout j de $\{1, \ldots, p\}$, posons $v_j = \sum_{i=1}^{n} a_{ij}\, \varepsilon_i$.
La matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de terme général a_{ij} est appelée *matrice de la famille v dans la base* (ε).
On pourra noter cette matrice $A = \mathrm{Mat}_\varepsilon(v)$.

La j-ième colonne de A est donc formée des composantes de v_j dans la base ε.

Par exemple, si $n = 3$, $j = 2$, et si $\begin{cases} v_1 = 3\varepsilon_1 + 5\varepsilon_2 + \varepsilon_3 \\ v_2 = 2\varepsilon_1 + 4\varepsilon_2 + 7\varepsilon_3 \end{cases}$ on a $A = \mathrm{Mat}_\varepsilon(v) = \begin{pmatrix} 3 & 2 \\ 5 & 4 \\ 1 & 7 \end{pmatrix}$

On notera $[v]_\varepsilon$ la matrice-colonne des coordonnées d'un vecteur v de E dans la base (ε).

Si $v = (v_j)_{1 \leqslant j \leqslant p}$, on peut donc écrire : $A = M_\varepsilon(v) = \left(\begin{array}{c|c|c|c} [v_1]_\varepsilon & [v_2]_\varepsilon & \cdots & [v_p]_\varepsilon \end{array}\right)$

17.1.2 Matrice d'une app$^{\mathrm{n}}$ linéaire dans un couple de bases

Définition 17.1.2 (matrice d'une application linéaire)
Soit E et F deux espaces vectoriels sur $\mathbb{K}$.
On suppose que $\dim(E) = p \geqslant 1$, et que E est muni d'une base $(e_j)_{1 \leqslant j \leqslant p}$.
On suppose que $\dim(F) = n \geqslant 1$, et que F est muni d'une base $(\varepsilon_i)_{1 \leqslant i \leqslant n}$.
Soit f une application linéaire de E dans F.
On appelle matrice de f dans les bases e et ε la matrice A de la famille $(f(e_j))_{1 \leqslant j \leqslant p}$ dans la base ε.
Cette matrice, élément de $\mathcal{M}_{n,p}(\mathbb{K})$, est notée $\mathrm{Mat}_{e,\varepsilon}(f)$.

Pour $1 \leqslant j \leqslant p$, la j-ième colonne de $\mathrm{Mat}_{e,\varepsilon}(f)$ est donc formée des composantes de $f(e_j)$ dans ε.
Supposons par exemple qu'une base de E soit $e = (e_1, e_2, e_3)$, et qu'une base de F soit $\varepsilon = (\varepsilon_1, \varepsilon_2)$.

Si $f : E \to F$ est l'application linéaire définie par $\begin{cases} f(e_1) = \varepsilon_1 + 2\varepsilon_2 \\ f(e_2) = 7\varepsilon_1 + 5\varepsilon_2 \\ f(e_3) = 3\varepsilon_1 \end{cases}$ alors $\mathrm{Mat}_{e,\varepsilon}(f) = \begin{pmatrix} 1 & 7 & 3 \\ 2 & 5 & 0 \end{pmatrix}$

Remarques

– Une application linéaire est complètement déterminée par sa matrice dans un couple de bases donné. Si $\dim(E) = p \geqslant 1$ et $\dim(F) = n \geqslant 1$, l'application de $\mathcal{L}(E, F)$ dans $\mathcal{M}_{n,p}(\mathbb{K})$ qui à f associe sa matrice dans un couple de bases donné est donc une bijection.

– Si $f : E \to F$ est linéaire et si change la base de E ou de F, la matrice de f est, en général, modifiée. On analysera plus loin cette dépendance en fonction du couple de bases.
En revanche, la matrice de l'application nulle est la matrice nulle, quel que soit le couple de bases.

– On retiendra que si A est la matrice d'une application linéaire, le nombre de colonnes de A est la dimension de l'espace de départ, et le nombre de lignes de A est la dimension de l'espace d'arrivée.

Cas particulier : matrice d'un endomorphisme dans *une* base

Soit f un endomorphisme de E, où $\dim(E) = n \geqslant 1$.

On munit souvent E de la même base $e = (e_i)_{1\leqslant i\leqslant n}$ *au départ* et *à l'arrivée*.

Plutôt que de noter $\mathrm{Mat}_{e,e}(f)$, on note $\mathrm{Mat}_e(f)$ et on parle de *matrice de f dans la base e*.

Cette matrice est bien sûr carrée d'ordre n.

Par exemple, soit f l'endomorphisme de $\mathbb{R}_3[X]$ défini par $f(P(\mathrm{X})) = (\mathrm{X}+1)P'(\mathrm{X}) + P(\mathrm{X})$.

On munit $\mathbb{R}_3[\mathrm{X}]$ de la base canonique $e = 1, \mathrm{X}, \mathrm{X}^2, \mathrm{X}^3$ (dans cet ordre !).

On constate que $\begin{cases} f(1) = 1 \\ f(\mathrm{X}) = 1 + 2\mathrm{X} \\ f(\mathrm{X}^2) = 2\mathrm{X} + 3\mathrm{X}^2 \\ f(\mathrm{X}^3) = 3\mathrm{X}^2 + 4\mathrm{X}^3 \end{cases}$. On en déduit que $\mathrm{Mat}_e(f) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}$.

17.1.3 Coordonnées de l'image d'un vecteur par une app$^{\mathrm{n}}$ linéaire

Proposition 17.1.1 (interprétation matricielle de l'égalité $v = f(u)$)
Soit E et F deux espaces vectoriels sur $\mathbb{K}$.
On suppose que $\dim(E) = p \geqslant 1$, et que E est muni d'une base $e = (e_j)_{1\leqslant j\leqslant p}$.
On suppose que $\dim(F) = n \geqslant 1$, et que F est muni d'une base $\varepsilon = (\varepsilon_i)_{1\leqslant i\leqslant n}$.
Soit f une application linéaire de E dans F, de matrice A dans les bases e et ε.
Pour tout u de E, l'égalité vectorielle $v = f(u)$ équivaut à l'égalité matricielle $[f(u)]_\varepsilon = A[u]_e$.

Réciproquement, si pour une application $f\colon E \to F$ il existe une matrice A telle que $[f(u)]_\varepsilon = A[u]_e$ pour tout vecteur u de E, alors f est linéaire et $A = \mathrm{Mat}_{e,\varepsilon}(f)$.

Exemples

Supposons qu'une base de E soit $e = (e_1, e_2, e_3)$, et qu'une base de F soit $\varepsilon = (\varepsilon_1, \varepsilon_2)$.

Si $f : E \to F$ est l'application linéaire définie par $\begin{cases} f(e_1) = \varepsilon_1 + 2\varepsilon_2 \\ f(e_2) = 7\varepsilon_1 + 5\varepsilon_2 \\ f(e_3) = 3\varepsilon_1 \end{cases}$ alors $\mathrm{Mat}_{e,\varepsilon}(f) = \begin{pmatrix} 1 & 7 & 3 \\ 2 & 5 & 0 \end{pmatrix}$

Si u a pour coordonnées (x, y, z) dans e, et si v a pour coordonnées (x', y') dans ε, alors :

$$v = f(u) \;\Leftrightarrow\; [v]_\varepsilon = A[u]_e \;\Leftrightarrow\; \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 7 & 3 \\ 2 & 5 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\Leftrightarrow\; \begin{cases} x' = x + 7y + 3z \\ y' = 2x + 5y \end{cases}$$

Réciproquement, soit $g : E \to F$ envoyant $u = xe_1 + ye_2 + ze_3$ sur $v = x'\varepsilon_1 + y'\varepsilon_2$ avec $\begin{cases} x' = x + 2y + 3z \\ y' = 9x + 8y + 7z \end{cases}$

Ce système s'écrit $\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 9 & 8 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, et on a par exemple $g(e_1) = \varepsilon_1 + 9\varepsilon_2$.

Ainsi g est linéaire et sa matrice dans les bases e et ε est $\mathrm{Mat}_{e,\varepsilon}(g) = \begin{pmatrix} 1 & 2 & 3 \\ 9 & 8 & 7 \end{pmatrix}$

Application identité et matrice identité

La matrice de Id_E dans une base e de E est la matrice identité I_n, quelle que soit la base e.

Mais attention ! la matrice de Id_E n'est pas I_n si dans E on utilise une certaine base e *au départ* et une autre base ε *à l'arrivée* (voir plus loin la question des « matrices de passage »).

Application linéaire canoniquement associée à une matrice

On sait qu'à toute application linéaire $f : E \to F$ (avec $\dim(E) = p \geqslant 1$ et $\dim(F) = n \geqslant 1$) correspond une matrice unique de $\mathcal{M}_{n,p}(\mathbb{K})$ dans un couple de bases donné.

Inversement, une matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ peut « représenter » une infinité d'applications linéaires :
- on a en effet d'abord le choix des espaces E (de dimension p) et F (de dimension n).
- on a ensuite le choix d'une base de E et d'une base de F.

Si rien n'est imposé, on prend souvent $E = \mathbb{K}^p$ et $F = \mathbb{K}^n$, munis de leurs bases canoniques. L'application linéaire $f\colon \mathbb{K}^p \to \mathbb{K}^q$ ainsi obtenue est dite « canoniquement associée » à la matrice A.

Par exemple, soit $A = \begin{pmatrix} 1 & 0 & -1 & 3 & 2 \\ 3 & 2 & -1 & 0 & 4 \\ 2 & -1 & 5 & -2 & 0 \end{pmatrix}$ dans $\mathcal{M}_{3,5}(\mathbb{R})$.

Elle est canoniquement associée à une application linéaire $f\colon \mathbb{K}^5 \to \mathbb{K}^3$.

Celle-ci est définie par $\begin{cases} x' = x - z + 3t + 2u \\ y' = 3x + 2y - z + 4u \\ z' = 2x - y + 5z - 2t \end{cases}$ c'est-à-dire par $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 & 3 & 2 \\ 3 & 2 & -1 & 0 & 4 \\ 2 & -1 & 5 & -2 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \\ u \end{pmatrix}$.

Forme linéaire canoniquement associée à une matrice-ligne

Une matrice ligne $A = \begin{pmatrix} a_1 & a_2 & \cdots & a_p \end{pmatrix}$ est canoniquement associée à une forme linéaire f sur $\mathbb{K}^p$.

Plus précisément, f est définie par $f(x_1, x_2, \ldots, x_p) = \begin{pmatrix} a_1 & a_2 & \cdots & a_p \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \displaystyle\sum_{k=1}^{p} a_k x_k$.

17.1.4 Propriétés opératoires

▷ Matrice d'une combinaison linéaire

Proposition 17.1.2 (matrice de $\lambda f + \mu g$)
Soit E et F deux espaces vectoriels sur $\mathbb{K}$, de dimensions respectives $p \geqslant 1$ et $n \geqslant 1$.
On suppose que E est muni d'une base $e = (e_j)_{1 \leqslant j \leqslant p}$, et que F est muni d'une base $\varepsilon = (\varepsilon_i)_{1 \leqslant i \leqslant n}$.
L'application qui à f associe $\mathrm{Mat}_{e,\varepsilon}(f)$ est un isomorphisme de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$.

Ainsi, $\begin{cases} \text{pour tous } f, g \text{ dans } \mathcal{L}(E, F) \\ \text{pour tous } \lambda, \mu \text{ de } \mathbb{K} \end{cases} : \quad \mathrm{Mat}_{e,\varepsilon}(\lambda f + \mu g) = \lambda \,\mathrm{Mat}_{e,\varepsilon}(f) + \mu \,\mathrm{Mat}_{e,\varepsilon}(g)$.

Cet isomorphisme implique $\dim \mathcal{L}(E, F) = \dim \mathcal{M}_{n,p}(\mathbb{K}) = np = \dim(E) \dim(F)$.

Cet isomorphisme n'est pas « canonique », en ce sens qu'il dépend des bases e, ε choisies dans E, F.

Complément : une base de $\mathcal{L}(E,F)$

On considère toujours l'isomorphisme de $\mathcal{L}(E,F)$ sur $\mathcal{M}_{n,p}(\mathbb{K})$ qui à f associe $\text{Mat}_{e,\varepsilon}(f)$.

On obtient une base de $\mathcal{L}(E,F)$ en prenant l'image réciproque de la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$.

Une base de $\mathcal{L}(E,F)$ est donc formée des $(f_{i,j})_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant p}}$, avec $\begin{cases}\forall\, k \in \{1,\ldots,p\}\setminus\{j\},\ f_{i,j}(e_k) = \overrightarrow{0}\\ f_{i,j}(e_j) = \varepsilon_i\end{cases}$

En d'autres termes, $f_{i,j}$ est définie par $f_{i,j}\left(\displaystyle\sum_{k=1}^{p} x_k e_k\right) = x_j \varepsilon_i$.

Cette base de $\mathcal{L}(E,F)$ n'est pas « canonique » car elle dépend du choix des bases e et ε.

▷ Matrice d'une composée

Proposition 17.1.3 (matrice de la composée $g \circ f$)
Soit E, F, G trois espaces vectoriels sur $\mathbb{K}$, tous les trois de dimension finie non nulle.
On suppose que E, F, G sont munis de bases notées respectivement α, β et γ.
Soit f une application linéaire de E dans F, et soit g une application linéaire de F dans G.
On note $A = \text{Mat}_{\alpha,\beta}(f)$ et $B = \text{Mat}_{\beta,\gamma}(g)$.
Alors la matrice de l'application linéaire $g \circ f$, dans les bases α de E et γ de G, est $C = BA$.
Autrement dit : $\text{Mat}_{\alpha,\gamma}(g \circ f) = \text{Mat}_{\beta,\gamma}(g)\,\text{Mat}_{\alpha,\beta}(f)$.

▷ Matrice d'une application nilpotente

Soit E un $\mathbb{K}$-espace vectoriel de dimension $n \geqslant 1$, muni d'une base e.
Soit f un endomorphisme de E, de matrice A dans la base e.
Alors l'application f est nilpotente si et seulement si la matrice A est nilpotente.
Si f est nilpotente, il existe une base (ε) de E telle que $\text{Mat}_\varepsilon(f)$ est strictement triangulaire supérieure.

▷ Matrice d'un isomorphisme

Proposition 17.1.4 (matrice de f^{-1})
Soit E et F deux $\mathbb{K}$-espaces vectoriels de même dimension $n \geqslant 1$.
On suppose que E est muni d'une base e, et que F est muni d'une base ε.
Soit f une application linéaire de E dans F, de matrice A dans les bases e et ε.
Alors f est un isomorphisme si et seulement si la matrice A est inversible.
Dans ce cas, la matrice de f^{-1} dans les bases ε et e est A^{-1}, ou encore : $\text{Mat}_{\varepsilon,e}(f^{-1}) = \Big(\text{Mat}_{e,\varepsilon}(f)\Big)^{-1}$

Cas particulier : matrice d'un automorphisme dans *une* base

Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$, muni d'une base e.
Soit f un endomorphisme de E, de matrice A dans la base e.
Alors f est un automorphisme de E si et seulement si A est inversible dans $\mathcal{M}_n(\mathbb{K})$.
En d'autres termes, on a l'équivalence : $f \in \text{GL}(E) \iff A \in \text{GL}_n(\mathbb{K})$.
En cas d'inversibilité, la matrice de f^{-1} dans la base e est A^{-1}, ou encore : $\text{Mat}_e(f^{-1}) = \Big(\text{Mat}_e(f)\Big)^{-1}$.

▷ Matrice des puissances d'une application linéaire

Proposition 17.1.5 (matrice de f^n)
Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$, muni d'une base e.
Soit f un endomorphisme de E, de matrice A dans la base e.
Pour tout entier naturel k, la matrice de f^k dans la base e est A^k.
Cette propriété s'étend aux k négatifs si f est un automorphisme de E, c'est-à-dire si A est inversible.

17.2 Changements de bases

17.2.1 Matrice de passage d'une base à une autre

Proposition 17.2.1 (inversibilité de la matrice d'une famille de vecteurs)
Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$, muni d'une base e.
Soit $v = (v_j)_{1 \leqslant j \leqslant n}$ une famille de n vecteurs de E (autant donc que la dimension de E).
Soit A la matrice de la famille v dans la base e. C'est un élément de $\mathcal{M}_n(\mathbb{K})$.
Alors la famille v est une base de E si et seulement si la matrice A est inversible.

Définition 17.2.1 (matrice de passage d'une base à une autre)
Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$, muni de deux bases $e = (e_i)_{1 \leqslant i \leqslant n}$ et $\varepsilon = (\varepsilon_j)_{1 \leqslant j \leqslant n}$.
La matrice de la famille ε dans la base e est appelée *matrice de passage* de la base e à la base ε.
Cette matrice est notée P_e^ε. D'après ce qui précède, elle est inversible.

Deux interprétations d'une matrice de passage

Avec les notations précédentes, la matrice de passage de e à ε est à la fois :
– la matrice de l'identité, de E muni de la base ε, vers E muni de la base e : $P_e^\varepsilon = \mathrm{Mat}_{\varepsilon,e}(\mathrm{Id}_E)$.
– la matrice dans la base e de l'automorphisme f de E défini par : $\forall j \in \{1, \ldots, n\},\ f(e_j) = \varepsilon_j$.

Opérations entre matrices de passage

– l'inverse de la matrice de passage P_e^ε de e à ε est la matrice de passage P_ε^e de ε à e.
– si α, β et γ sont trois bases de E, alors on a la relation : $P_\alpha^\gamma = P_\alpha^\beta\, P_\beta^\gamma$.

17.2.2 Effet d'un changement de base(s)

Proposition 17.2.2 (changement de base pour les coordonnées d'un vecteur)
Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$, muni de deux bases e et ε.
Soit P la matrice de passage de e à ε. Pour tout u de E : $\boxed{[u]_e = P\,[u]_\varepsilon}$

On retiendra le paradoxe : la matrice de passage de l'*ancienne* base à la *nouvelle* donne les *anciennes* coordonnées en fonction des *nouvelles*.

Proposition 17.2.3 (changement de base pour la matrice d'une application linéaire)
Soit E et F deux espaces vectoriels sur $\mathbb{K}$, de dimensions respectives $p \geqslant 1$ et $n \geqslant 1$.
On suppose que E est muni d'une ancienne *base e et d'une* nouvelle *base e'.*
*De même soit ε l'*ancienne *base de F, et soit ε' la* nouvelle *base de F.*
Soit P la matrice de passage de e à e'. Soit Q la matrice de passage de ε à ε'.
Soit f une application linéaire de E dans F.
*Soit A la matrice de f dans les bases e et ε (*ancienne *matrice de f).*
*Soit B la matrice de f dans les bases e' et ε' (*nouvelle *matrice de f).*
Alors on a l'égalité : $\boxed{B = Q^{-1}\,A\,P}$.

Cas particulier important

Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$.
Soit e l'*ancienne* base de E et e' la *nouvelle* base de E.
Soit f un endomorphisme de E, de matrice A dans la base e, et de matrice B dans la base e'.
Soit P la matrice de passage de e à e'. Alors on a l'égalité : $\boxed{B = P^{-1}\,A\,P}$.

17.2.3 Matrices équivalentes et matrices semblables

Définition 17.2.2 (matrices équivalentes)
Soit A et B deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$.
On dit que A et B sont équivalentes s'il existe G dans $\mathrm{GL}_p(\mathbb{K})$ et D dans $\mathrm{GL}_n(\mathbb{K})$, telles que $B = GAD$.

Remarques

On pourra noter $A \equiv B$ pour dire « A est équivalente à B ».

On définit ainsi une relation ... d'équivalence dans $\mathcal{M}_{np}(\mathbb{K})$:

– réflexivité : toute matrice A est équivalente à elle-même (écrire $A = I_nAI_p$).
– symétrie : si $A \equiv B$, alors $B \equiv A$. En effet $B = GAD \Rightarrow A = G^{-1}BD^{-1}$.
– transitivité : si $\begin{cases} A \equiv B \\ B \equiv C \end{cases}$ alors $A \equiv C$. En effet $\begin{cases} B = GAD \\ C = G'BD' \end{cases} \Rightarrow C = (G'G)A(DD')$.

Caractérisation de l'équivalence des matrices

– Soit E et F deux espaces vectoriels sur $\mathbb{K}$, de dimensions respectives $p \geqslant 1$ et $n \geqslant 1$.
  Soit f une application linéaire de E dans F.
  Soit A la matrice de f dans un certain couple de bases (e dans E et ε dans F).
  Soit B la matrice de f dans un autre couple de bases (e' dans E et ε' dans F).
  On sait que $B = Q^{-1}AP$ où P et Q sont les deux matrices de passages.
  Les matrices A et B sont donc équivalentes dans $\mathcal{M}_{n,p}(\mathbb{K})$.

– Réciproquement, soit A et B deux matrices équivalentes dans $\mathcal{M}_{n,p}(\mathbb{K})$.
Soit E et F deux espaces vectoriels sur $\mathbb{K}$, de dimensions respectives $p \geqslant 1$ et $n \geqslant 1$.
On se donne un couple de bases, e dans E et ε dans F.
On sait que dans ce couple de bases, A représente une unique application linéaire $f : E \to F$.
Alors il existe un couple de bases, e' dans E et ε' dans F, tel que $B = \mathrm{Mat}_{e',\varepsilon'}(f)$.

– On peut résumer de la façon suivante :

Proposition 17.2.4 (une caractérisation de l'équivalence des matrices)
Deux matrices A et B sont équivalentes dans $\mathcal{M}_{n,p}(\mathbb{K})$ si et seulement si elles sont susceptibles de représenter une même application linéaire f (chacune dans un certain couple de bases).

17.2.4 Réduction de la matrice de f à une forme J_r

Définition 17.2.3 (matrices J_r)
On se place dans l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$, avec $n \geqslant 1$ et $p \geqslant 1$.
Soit r un entier vérifiant $0 \leqslant r \leqslant \min(n,p)$.
On note J_r la matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, de coefficients a_{ij} définis par $\begin{cases} a_{i,i} = 1 \text{ pour } 1 \leqslant i \leqslant r \\ a_{i,j} = 0 \text{ dans tous les autres cas} \end{cases}$

Par exemple, dans $\mathcal{M}_{4,5}(\mathbb{K})$: $J_2 = \begin{pmatrix} 1&0&0&0&0\\0&1&0&0&0\\0&0&0&0&0\\0&0&0&0&0 \end{pmatrix}$ et $J_3 = \begin{pmatrix} 1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&0 \end{pmatrix}$

On devrait noter $J_r(n,p)$, mais en général les valeurs de n et p sont connues sans ambiguïté.
La matrice J_0 est la matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$.
Si $n = p$, la matrice J_n est la matrice identité I_n.
Voici une fonction `Python` pour former les matrices J_r. Les arguments sont les entiers r, n et p (ces deux derniers étant facultatifs avec la valeur 5 par défaut) :

```
>>> def J(r, n=5, p=5):
        return np.array([[i==j<r for j in range(p)] for i in range(n)]).astype(int)
```

Voici deux exemples d'utilisation de la fonction `J` :

```
>>> J(3)  # ici r=3, et n=p=5 par défaut
array([[1, 0, 0, 0, 0],
       [0, 1, 0, 0, 0],
       [0, 0, 1, 0, 0],
       [0, 0, 0, 0, 0],
       [0, 0, 0, 0, 0]])
```

```
>>> J(2,4,6)  # ici r=2, n=4, p=6
array([[1, 0, 0, 0, 0, 0],
       [0, 1, 0, 0, 0, 0],
       [0, 0, 0, 0, 0, 0],
       [0, 0, 0, 0, 0, 0]])
>>>
```

Proposition 17.2.5 (réduction de la matrice de f à une forme J_r)
Soient E et F deux espaces vectoriels sur $\mathbb{K}$, de dimension finie non nulle.
Soit f une application linéaire de E dans F, de rang r.
Alors il existe une base e de E et une base ε de F telles que $\mathrm{Mat}_{e,\varepsilon}(f) = J_r$.

Attention !

Soit E un espace vectoriel de dimension finie non nulle. Soit f un endomorphisme de E, de rang r.
La proposition précédente **ne dit pas** qu'il existe *une* base e de E dans laquelle $\mathrm{Mat}_e(f) = J_r$.
Donc même si f est un endomorphisme d'un espace vectoriel E, la proposition précédente énonce l'existence **de deux bases**, une base e « au départ » et une base ε « à l'arrivée ».

17.2.5 Matrices semblables

Définition 17.2.4 (matrices semblables)
Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$.
On dit que A et B sont semblables s'il existe P dans $\mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Remarques

Cette notion ne concerne que les matrices carrées.
Deux matrices semblables sont équivalentes, mais la réciproque est fausse.
La seule matrice semblable à la matrice nulle de $\mathcal{M}_n(\mathbb{K})$ est la matrice nulle.
Pour tout scalaire λ, la seule matrice semblable à λI_n est la matrice λI_n.

Caractérisation de la similitude des matrices

– Soit E un espace vectoriel sur $\mathbb{K}$, de dimension $n \geqslant 1$. Soit f un endomorphisme de E.
  Soit A la matrice de f dans une base e de E. Soit B la matrice de f dans une autre base e' de E.
  On sait que $B = P^{-1}AP$ où P est la matrice de passage de e à e'.
  Les matrices A et B sont donc semblables dans $\mathcal{M}_n(\mathbb{K})$.
– Réciproquement, soit A et B deux matrices semblables dans $\mathcal{M}_n(\mathbb{K})$.
  Soit E un espace vectoriel sur $\mathbb{K}$, de dimension n, muni d'une base e.
  On sait que, dans cette base e, la matrice A représente un unique endomorphisme f de E.
  Dans ces conditions, si e' est la base de E définie par $P = P_{e,e'}$, on a $B = \mathrm{Mat}_{e'}(f)$.
– On peut résumer de la façon suivante :

  Proposition 17.2.6 (une caractérisation de la similitude des matrices)
  Deux matrices A, B de $\mathcal{M}_n(\mathbb{K})$ sont semblables si et seulement si elles sont susceptibles de représenter un même endomorphisme f d'un espace vectoriel E de dimension n (chacune dans une certaine base).

Puissances de matrices semblables

Si $B = P^{-1}AP$, alors pour tout entier naturel n on a : $B^n = P^{-1}A^nP$.
Cette relation s'étend aux exposants négatifs si A (et donc B) est inversible.
On peut donc calculer B^n si A^n est plus facile à obtenir, notamment si A est diagonale.

17.3 Trace d'une matrice, d'un endomorphisme

17.3.1 Trace d'une matrice carrée

Définition 17.3.1 (trace d'une matrice carrée)
Soit A une matrice carrée d'ordre n, à coefficients dans $\mathbb{K}$, de terme général a_{ij}.
On appelle *trace* de A, et on note $\mathrm{tr}(A)$ ou $\mathrm{Tr}(A)$, la somme $\sum_{i=1}^{n} a_{ii}$ des coefficients diagonaux de A.

Proposition 17.3.1 (linéarité de la trace)
L'application « trace » est une forme linéaire *sur l'espace vectoriel $\mathcal{M}_n(\mathbb{K})$.*
Ainsi pour toutes matrices A et B de $\mathcal{M}_n(\mathbb{K})$, et tous α, β de $\mathbb{K}$, on a : $\mathrm{tr}(\alpha A + \beta B) = \alpha\,\mathrm{tr}(A) + \beta\,\mathrm{tr}(B)$.

Proposition 17.3.2 (trace d'un produit de deux matrices)
Soient A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ et B une matrice de $\mathcal{M}_{p,n}(\mathbb{K})$.
La matrice AB est donc carrée d'ordre n, tandis que BA est carrée d'ordre p.
Les matrices carrées AB et BA ont la même trace : $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

L'égalité $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ est vraie en particulier pour toutes matrices A, B de $\mathcal{M}_n(\mathbb{K})$.

On ne doit pas généraliser abusivement à des produits de plus de deux matrices.

On peut par exemple écrire : $\mathrm{tr}(ABC) = \mathrm{tr}\Big(A(BC)\Big) = \mathrm{tr}\Big((BC)A\Big) = \mathrm{tr}(BCA)$.

On peut également écrire $\mathrm{tr}(ABC) = \mathrm{tr}(CAB)$ (en échangeant AB et C).

Mais en général on a : $\mathrm{tr}(ABC) \neq \mathrm{tr}(ACB)$.

Proposition 17.3.3 (invariance de la trace par similitude)
Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$.
Si A et B sont semblables, alors elles ont la même trace.

Pour démontrer ce résultat, on écrit $\mathrm{tr}(P^{-1}AP) = \mathrm{tr}((P^{-1}A)P) = \mathrm{tr}(P(P^{-1}A)) = \mathrm{tr}(A)$.

Mais on évitera d'écrire : $\mathrm{tr}(P^{-1}AP) = \mathrm{tr}(P^{-1}PA) = \mathrm{tr}(A)$!

La réciproque de la proposition précédents est évidemment fausse.

Attention à ne pas confondre...

On ne confondra pas la *trace* $\mathrm{tr}(A)$ d'une matrice carrée A avec sa *transposée* $A^\top$.

– la trace n'est définie que pour les matrices carrées, et la trace de A est un *scalaire*.

– la transposée d'une matrice A de type (n, p) est une *matrice* de type (p, n).

– on a la relation $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, mais on a la relation $({}^\top AB) = B^\top A^\top$.

17.3.2 Trace d'un endomorphisme

Définition 17.3.2 (trace d'un endomorphisme en dimension finie)
Soit E un espace vectoriel sur $\mathbb{K}$, de dimension finie $n \geqslant 1$. Soit f un endomorphisme de E.
La *trace* de f, notée $\mathrm{tr}(f)$ ou $\mathrm{Tr}(f)$, est la trace de la matrice de f dans une base *quelconque* de E.

Cette définition est légitime : bien que la matrice de f dépende en général de la base choisie dans l'espace vectoriel E, la trace de cette matrice n'en dépend pas (en effet toutes les matrices susceptibles de représenter l'application f dans une base e de E sont semblables entre elles).

Les deux propriétés suivantes sont des conséquences immédiates de la définition :

– si f et g sont deux endomorphismes de E, on a : $\mathrm{tr}(\lambda f + \mu g) = \lambda\,\mathrm{tr}(f) + \mu\,\mathrm{tr}(g)$.

– si f et g sont dans $\mathcal{L}(E)$, on a : $\mathrm{tr}(gf) = \mathrm{tr}(fg)$ (rappelons qu'on note gf pour $g \circ f$).

Trace d'une projection vectorielle (d'un projecteur)

Si p est la projection vectorielle de E sur un sous-espace F de dimension r, alors $\mathrm{tr}(p) = \mathrm{rg}(p) = r$.

Pour s'en persuader, il suffit de choisir un supplémentaire G de F dans E et de se placer dans une base adaptée à la somme directe $E = F \oplus G$: la matrice de p dans cette base est diagonale, les r premiers coefficients diagonaux valant 1 et les $n - r$ derniers valant 0.

Trace d'une rotation vectorielle en dimension 3

On se place dans l'espace vectoriel euclidien orienté $\mathbb{R}^3$.

Soit r la rotation vectorielle d'angle θ autour d'un vecteur unitaire w. Alors $\mathrm{tr}(r) = 1 + 2\cos\theta$.

Il suffit en effet de se placer dans une base orthonormée directe u, v, w.

La matrice de r dans cette base est $A = \begin{pmatrix} \cos\theta & -\sin\theta & 1 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $\mathrm{tr}(A) = 1 + 2\cos\theta$.

Si on connait la matrice B de la rotation r dans une base quelconque (même non orthonormée), on sait qu'on a l'égalité $\mathrm{tr}(B) = 1 + 2\cos\theta$. On en déduit immédiatement le cosinus de l'angle de la rotation r. Pour calculer son sinus il faut *orienter* l'axe de la rotation.

17.4 Noyau, image et rang d'une matrice

17.4.1 Application linéaire canoniquement associée

On revient ici sur des remarques déjà formulées dans la section 16.2.4.

Considérons l'application qui à $u = (x_1, x_2, \ldots, x_p)$ associe la matrice colonne $U = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$.

C'est un isomorphisme « canonique » de $\mathbb{K}^p$ sur $\mathcal{M}_{p,1}(\mathbb{K})$.

Il permet d'identifier un p-uplet u avec la matrice colonne U correspondante, de hauteur p.

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. L'application $\varphi\colon U \mapsto V = AU$ est linéaire de $\mathcal{M}_{p,1}(\mathbb{K})$ dans $\mathcal{M}_{n,1}(\mathbb{K})$.

À isomorphisme près, il s'agit donc d'une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$.

Si par exemple $A = \begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & 4 & 1 & 5 \\ 1 & 6 & 2 & 0 \end{pmatrix}$, et $U = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, alors $V = AU = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, où $\begin{cases} y_1 = 2x_1 + x_2 + 3x_4 \\ y_2 = 4x_2 + x_3 + 5x_4 \\ y_3 = x_1 + 6x_2 + 2x_3 \end{cases}$

Dans cet exemple, on peut donc identifier :

– l'application linéaire $\varphi\colon U \mapsto V = AU$ de $\mathcal{M}_{4,1}(\mathbb{K})$ dans $\mathcal{M}_{3,1}(\mathbb{K})$.

– l'application linéaire $f\colon u = (x_1, x_2, x_3, x_4) \mapsto v = (y_1, y_2, y_3)$ de $\mathbb{K}^4$ dans $\mathbb{K}^3$.

Définition 17.4.1 (application linéaire canoniquement associée à une matrice)
Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$.

– l'application $\varphi\colon U \mapsto V = AU$ est linéaire de $\mathcal{M}_{p,1}(\mathbb{K})$ dans $\mathcal{M}_{n,1}(\mathbb{K})$.

– si on munit $\mathbb{K}^p$ et $\mathbb{K}^n$ de leurs bases canoniques, A est la matrice d'un unique f de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$.

L'application linéaire φ (ou l'application linéaire f) est dite canoniquement associée à la matrice A.

Compte-tenu de ce qui a été dit plus haut, on peut donc quasiment considérer que f et φ sont une seule et même application linéaire.
Dans la suite, on utilisera indifféremment l'une ou l'autre des deux terminologies.

17.4.2 Noyau, image et rang d'une matrice

Définition 17.4.2 (image et noyau d'une matrice)
Soit A une matrice, élément de $\mathcal{M}_{n,p}(\mathbb{K})$.
Soit $f : \mathbb{K}^p \to \mathbb{K}^n$ l'application linéaire canoniquement associée à la matrice A.
On appelle *image* de A, et on note $\text{Im }(A)$, l'image $\text{Im }(f)$ de l'application linéaire f.
On appelle *noyau* de A, et on note $\text{Ker}(A)$, l'image $\text{Ker}(f)$ de l'application linéaire f.
On appelle *rang* de A, et on note $\text{rg}\,(A)$, le rang $\text{rg}\,(f)$ de l'application linéaire f.
Le *théorème de la dimension* nous assure que $\text{rg}\,(A) + \dim(\text{Ker}(A)) = p$.

▷ Premières remarques

– Si A est dans $\mathcal{M}_{n,p}(\mathbb{K})$, alors ($\text{rg}\,(A) \leqslant n$ et $\text{rg}\,(A) \leqslant p$).

– Le rang d'une matrice A est nul si et seulement si A est elle-même la matrice nulle.

– Le rang d'une matrice A est égal à 1 si et seulement A possède une colonne non nulle et si les autres colonnes de A lui sont proportionnelles (voir ci-après).

– Le rang de A est celui de *toute* application linéaire susceptible d'être représentée par A.

▷ Image d'une matrice, et vecteur colonnes

Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. Notons $\text{C}_1, \text{C}_2, \ldots, \text{C}_p$ les p colonnes de A.

Pour toute colonne $U = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ de $\mathcal{M}_{p,1}(\mathbb{K})$, on peut écrire $AU = \left(\begin{array}{c|c|c|c} \text{C}_1 & \text{C}_2 & \cdots & \text{C}_p \end{array}\right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \sum_{j=1}^{p} x_j \text{C}_j$.

Ainsi l'image de A, c'est-à-dire l'ensemble des AU, est le sous-espace engendré par $\text{C}_1, \text{C}_2, \ldots, \text{C}_p$
On retiendra donc : *les colonnes d'une matrice A forment une famille génératrice de* $\text{Im }(A)$.

▷ Noyau d'une matrice, et vecteurs-ligne

Soit $A = (a_{i,j})$ un élément de $\mathcal{M}_{n,p}(\mathbb{K})$. Notons $L_1, L_2, \ldots, L_n$ les n lignes de A.

On peut donc écrire $A = \begin{pmatrix} L_1 \\ \hline L_2 \\ \hline \vdots \\ \hline L_n \end{pmatrix}$. On se donne ensuite $U = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ quelconque dans $\mathcal{M}_{p,1}(\mathbb{K})$.

Avec ces notations, on peut caractériser les vecteurs U de $\text{Ker}(A)$:

$$U \in \text{Ker}(A) \Leftrightarrow \begin{pmatrix} L_1 \\ \hline L_2 \\ \hline \vdots \\ \hline L_n \end{pmatrix} U = \overrightarrow{0} \Leftrightarrow \begin{pmatrix} L_1\,U \\ L_2\,U \\ \vdots \\ L_n\,U \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Leftrightarrow \Big(\forall i \in \{1, \ldots, n\},\ L_i U = 0\Big)$$

Mais $L_i = \begin{pmatrix} a_{i,1} & a_{i,2} & \ldots & a_{i,p} \end{pmatrix}$ donc l'égalité $L_i U = 0$ s'écrit $\displaystyle\sum_{j=1}^{p} a_{ij} x_j = 0$.

Ainsi l'égalité $AU = \overrightarrow{0}$ se traduit par un système de n équations linéaires impliquant les lignes de A.

On retiendra donc : *les lignes d'une matrice A permettent de former un système d'équations de* $\text{Ker}(A)$.

▷ Illustration par un exemple

On demande de déterminer le noyau, l'image, et le rang de la matrice $A = \begin{pmatrix} 2 & -1 & 1 & 5 \\ -1 & 2 & 3 & -4 \\ 3 & 0 & 5 & 6 \end{pmatrix}$

On obtient $\text{Ker}(A)$ en résolvant $AU = 0$, où U est dans $\mathcal{M}_{4,1}(\mathbb{K})$.

$$AU = 0 \Leftrightarrow A \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} 2x - y + z + 5t = 0 & (E_1) \\ -x + 2y + 3z - 4t = 0 & (E_2) \\ 3x + 5z + 6t = 0 & (E_3) \end{cases} \Leftrightarrow \begin{cases} 2x - y + z + 5t = 0 \\ -x + 2y + 3z - 4t = 0 \end{cases}$$

On voit bien comment les trois lignes de A ont permis d'écrire un système d'équations du noyau.
On s'est ramené à un système de deux équations car $(E_3) = 2(E_1) + (E_2)$.

Le dernier système obtenu équivaut à $\begin{cases} 2x - y = -z - 5t \\ x - 2y = 3z - 4t \end{cases}$ c'est-à-dire $x = -\dfrac{5}{3}z - 2t$ et $y = -\dfrac{7}{3}z + t$.

Cela équivaut à $(x, y, z, t) = \dfrac{z}{3}(-5, -7, 1, 0) + t(-2, 1, 0, 1)$, avec $(z, t) \in \mathbb{R}^2$.

Ainsi : $U = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \text{Ker}(A) \Leftrightarrow X = \dfrac{z}{3} \begin{pmatrix} -5 \\ -7 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$ où z et t sont deux scalaires quelconques.

Finalement, on peut dire que $\text{Ker}(A)$ est le plan engendré par $V = \begin{pmatrix} -5 \\ -7 \\ 1 \\ 0 \end{pmatrix}$ et $W = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}$.

On peut aussi invoquer l'application linéaire $f : \mathbb{K}^4 \to \mathbb{K}^3$ canoniquement associée à A. La différence (très mince) consiste à considérer $u = (x, y, z, t)$ plutôt que la matrice colonne U, et à dire que le noyau est le plan engendré par $v = (-5, -7, 1, 0)$ et $w = (-2, 1, 0, 1)$ dans $\mathbb{K}^4$.

Le théorème de la dimension donne $\dim \mathbb{K}^4 = \dim \text{Im}\ (f) + \dim \text{Ker}(f)$.

Or $\dim \text{Ker}(f) = 2$. On en déduit $\dim \text{Im}\ (f) = 2$, c'est-à-dire $\dim \text{Im}\ (A) = 2$.

Ainsi l'application linéaire f, et donc la matrice A, sont de rang 2.

Les vecteurs $\begin{cases} f(e_1) = (2, -1, 3) \\ f(e_2) = (-1, 2, 0) \end{cases}$ sont dans le plan Im (f) et sont libres. Ils en forment donc une base.

En termes quasiment identiques : les colonnes $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$ et $\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$ forment une base de Im (A).

Le sous-espace Im (f) est un plan de $\mathbb{R}^3$, donc le noyau d'une forme linéaire non nulle sur $\mathbb{R}^3$.

Il existe donc $(a, b, c) \neq \overrightarrow{0}$ dans $\mathbb{R}^3$ tel que : $(x, y, z) \in \text{Im}\ (f) \Leftrightarrow ax + by + cz = 0$.

En utilisant $\begin{cases} f(e_1) = (2, -1, 3) \\ f(e_2) = (-1, 2, 0) \end{cases}$ on trouve $\begin{cases} 2a - b + 3c = 0 \\ -a + 2b = 0 \end{cases}$ donc $\begin{cases} a = 2b \\ c = -b \end{cases}$

Les coefficients de l'équation $ax + by + cz = 0$ sont uniques à un facteur multiplicatif non nul près.

En choisissant $b = 1$, on voit que l'équation de Im (f) dans la base canonique est : $2x + y - z = 0$.

Mais à y regarder de plus près, on a déjà rencontré cette relation !

En effet, on a vu que les lignes $(E_1), (E_2), (E_3)$ du système $AU = 0$ étaient liées par : $(E_3) = 2(E_1)+(E_2)$.

Cette relation signifie que toutes les colonnes $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de Im (A) vérifient $z = 2x + y$.

Le même exemple, avec Python

Attention : la fonction `matrix_rank` renvoie le rang d'une matrice (au sens que nous avons donné à ce terme). Il ne faut pas confondre avec la fonction `rank` qui donne le nombre d'indices utilisés pour décrire le tableau A (ici `rank(A)` donne 2 car il faut deux indices pour décrire A).

```
>>> print(a)
[[ 2 -1  1  5]
 [-1  2  3 -4]
 [ 3  0  5  6]]
```

Voici la **bonne** fonction !

```
>>> np.linalg.matrix_rank(a)
2
```

Voici la **mauvaise** fonction !

```
>>> np.rank(a)
2
```

▷ Inversibilité et noyau, pour une matrice carrée

Proposition 17.4.1
Soit A une matrice carrée. Alors A est inversible si et seulement si son noyau est réduit à $\overrightarrow{0}$.

Pour prouver que A est inversible dans $\mathcal{M}_n(\mathbb{K})$, il suffit donc de prouver que $AX = \overrightarrow{0} \Rightarrow X = \overrightarrow{0}$.
Évidemment, cela ne donne pas A^{-1} (il faudrait pour cela résoudre $AX = B$ en $X = A^{-1}B$).

▷ Effet du produit par une matrice inversible

Proposition 17.4.2 (invariance du noyau par multiplication à gauche)
Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, et soit Q dans $\text{GL}_n(\mathbb{K})$ (donc Q est carrée inversible d'ordre n).
Alors les matrices A et QA ont le même noyau. En particulier elles ont le même rang.

Proposition 17.4.3 (invariance de l'image par multiplication à droite)
Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$, et soit P dans $\text{GL}_p(\mathbb{K})$ (donc P est carrée inversible d'ordre p).
Alors les matrices A et AP ont la même image. En particulier elles ont le même rang.

17.4.3 Matrices équivalentes et rang

On se place dans l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$, avec $n \geqslant 1$ et $p \geqslant 1$.

On rappelle qu'on note J_r la matrice de coefficients a_{ij} définie par $\begin{cases} a_{i,i} = 1 \text{ pour } 1 \leqslant i \leqslant r \\ a_{i,j} = 0 \text{ dans tous les autres cas} \end{cases}$

Par exemple, dans $\mathcal{M}_{4,5}(\mathbb{K})$, on a $J_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ et $J_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

Proposition 17.4.4 (caractérisation du rang avec les matrices J_r)
Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$.
Alors A est de rang r si et seulement si elle est équivalente à la matrice J_r.

Proposition 17.4.5 (rapport entre équivalence et rang des matrices)
Soit A et B deux matrices de $\mathcal{M}_{n,p}(\mathbb{K})$.
Alors A et B sont équivalentes si et seulement si elles ont le même rang.

On sait qu'en écrivant « A est équivalente à B », on définit une relation d'équivalence sur $\mathcal{M}_{n,p}(\mathbb{K})$.

La proposition précédente dit que dans $\mathcal{M}_{n,p}(\mathbb{K})$, il y a exactement $1 + \min(n, p)$ classes d'équivalence pour cette relation : chaque classe d'équivalence est en effet formée des matrices A dont le rang a une valeur donnée r, avec $0 \leqslant r \leqslant \min(n, p)$.

En particulier, dans $\mathcal{M}_n(\mathbb{K})$ on a l'inégalité $\mathrm{rg}\,(A) \leqslant n$ et l'équivalence ($\mathrm{rg}\,(A) = n \Leftrightarrow A$ est inversible).

Proposition 17.4.6 (invariance du rang par transposition)
Le rang d'une matrice A est égal au rang de sa matrice transposée $A^\top$.

Proposition 17.4.7 (interprétations du rang d'une matrice)
Soit A dans $\mathcal{M}_{n,p}(\mathbb{K})$. Le rang de A est égal au nombre maximum de colonnes libres dans A.
Le rang de A est aussi égal au nombre maximum de lignes libres dans A.

17.4.4 Rang et matrices extraites

Définition 17.4.3 (matrices extraites)
Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$.
Soit I une famille strictement croissante d'indices de ligne, c'est-à-dire d'éléments de $\{1, \ldots, n\}$.
Soit J une famille strictement croissante d'indices de colonne, c'est-à-dire d'éléments de $\{1, \ldots, p\}$.
Soit B la matrice formée par les intersections des lignes L_i et C_j de A, avec i dans I et j dans J.
On dit que B est une *matrice extraite* de A.

Par exemple, si $\begin{cases} I = (1, 3, 4) \\ J = (2, 3, 5, 6) \end{cases}$ et $A = \begin{pmatrix} 2 & \boxed{5} & \boxed{3} & 1 & \boxed{6} & \boxed{7} \\ 4 & 1 & 5 & 6 & 8 & 9 \\ 1 & \boxed{0} & \boxed{4} & 5 & \boxed{1} & \boxed{3} \\ 7 & \boxed{9} & \boxed{6} & 2 & \boxed{1} & \boxed{5} \\ 1 & 5 & 8 & 2 & 7 & 4 \end{pmatrix}$ alors $B = \begin{pmatrix} 5 & 3 & 6 & 7 \\ 0 & 4 & 1 & 3 \\ 9 & 6 & 1 & 5 \end{pmatrix}$

Voici une fonction `sousmatrice` permettant d'extraire d'une matrice A une sous-matrice B en précisant un ensemble I d'indices de ligne, puis un ensemble J d'indices de colonne.

```
>>> def sousmatrice(a, I, J):
        return np.array([[a[i-1, j-1] for j in J] for i in I])
```

Et voici une illustration de cette fonction, en reprenant l'exemple précédent :

```
>>> a
array([[2, 5, 3, 1, 6, 7],
       [4, 1, 5, 6, 8, 9],
       [1, 0, 4, 5, 1, 3],
       [7, 9, 6, 2, 1, 5],
       [1, 5, 8, 2, 7, 4]])
```

```
>>> sousmatrice(a,[1,3,4],[2,3,5,6])
array([[5, 3, 6, 7],
       [0, 4, 1, 3],
       [9, 6, 1, 5]])
>>>
```

Proposition 17.4.8 (rang d'une matrice extraite)
Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$, et soit B une matrice extraite de A. Alors $\text{rg}\,(B) \leqslant \text{rg}\,(A)$.

Conséquence importante : si de A on peut extraire une matrice de rang r, alors $\text{rg}\,(A) \geqslant r$.

Proposition 17.4.9 (caractérisation du rang par les matrices extraites inversibles)
Soit A un élément de $\mathcal{M}_{n,p}(\mathbb{K})$.
Alors $\text{rg}\,(A)$ est l'ordre maximum d'une matrice carrée inversible extraite de A.

17.5 Calcul effectif du rang

17.5.1 Matrices échelonnées

Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. On note $\mathrm{L}_1, \ldots, \mathrm{L}_n$ les lignes successives de A.

Pour chaque ligne L_i de A, soit $d(i)$ le plus petit indice j, s'il existe, tel que $a_{ij} \neq 0$.

On dit que A est *échelonnée* supérieurement s'il existe un entier r de $\{0, \ldots, n\}$ tel que :
- pour tout indice i inférieur ou égal à r, la ligne L_i est non nulle.
- pour tout indice i strictement supérieur à r, la ligne L_i est nulle.
- la suite $d(1), d(2), \ldots, d(r)$ est strictement croissante.

Proposition 17.5.1 (rang d'une matrice échelonnée)
Avec les notations précédentes, la matrice échelonnée A est de rang r.
Les r coefficients non nuls situés aux positions $(i, d(i))$ sont appelés les pivots *de A.*

Par exemple, $A = \begin{pmatrix} 0 & \boxed{1} & 3 & 4 & 0 & 1 & 5 \\ 0 & 0 & \boxed{3} & 5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \boxed{4} & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & \boxed{9} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ est échelonnée, avec quatre pivots, donc $\text{rg}\,(A) = 4$.

La matrice nulle de $\mathcal{M}_{n,p}(\mathbb{K})$ est un cas particulier de matrice échelonnée (avec zéro pivot !).

On définit comme précédemment les matrices « échelonnées inférieurement ». En fait une matrice A est échelonnée inférieurement si et seulement si sa transposée est échelonnée supérieurement.

17.5.2 Opérations élémentaires

Définition 17.5.1 (opérations élémentaires sur les lignes d'une matrice)
Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. Notons $\mathrm{L}_1, \mathrm{L}_2, \ldots, \mathrm{L}_n$ les lignes de A.
On appelle *opération élémentaire* sur les lignes de A l'une des opérations suivantes :
- multiplier une ligne L_i par un scalaire **non nul** α : on note $\mathrm{L}_i \leftarrow \alpha \mathrm{L}_i$.
- ajouter à l'une des lignes L_i un multiple d'une **autre** ligne L_j : on note $\mathrm{L}_i \leftarrow \mathrm{L}_i + \beta \mathrm{L}_j$.
- échanger deux lignes L_i et L_j : on note $\mathrm{L}_i \leftrightarrow \mathrm{L}_j$.

On définit de même les opérations élémentaires sur les colonnes de la matrice A.

Ces opérations sont notées : $\mathrm{C}_i \leftarrow \alpha \mathrm{C}_i$ (avec $\alpha \neq 0$), $\mathrm{C}_i \leftarrow \mathrm{C}_i + \beta \mathrm{C}_j$ (avec $j \neq i$), et $\mathrm{C}_i \leftrightarrow \mathrm{C}_j$.

▷ Nécessiter d'utiliser un pivot non nul

Dans les opérations $\mathrm{L}_i \leftarrow \alpha \mathrm{L}_i$ et $\mathrm{C}_i \leftarrow \alpha \mathrm{C}_i$, on dit que α joue le rôle de « pivot ».

Dans ces deux opérations il est absolument indispensable que α soit non nul.

On fera notamment attention au cas où α dépend d'un paramètre : pour les valeurs de ce paramètre qui annuleraient α, l'opération se traduirait par $\mathrm{L_i} \leftarrow 0$ ou $\mathrm{C}_i \leftarrow 0$ et serait « illégale ».

On note souvent $\mathrm{L}_i \leftarrow \alpha \mathrm{L}_i + \beta \mathrm{L}_j$ la composée de $\mathrm{L}_i \leftarrow \alpha \mathrm{L}_i$ (avec $\alpha \neq 0$) puis de $\mathrm{L}_i \leftarrow \mathrm{L}_i + \beta \mathrm{L}_j$.

▷ Interprétation matricielle des opérations élémentaires sur les lignes

Proposition 17.5.2 (interprétation matricielle d'une opération élémentaire sur les lignes)
On se place dans $\mathcal{M}_{n,p}(\mathbb{K})$ et on considère une opération élémentaire particulière sur les lignes.
Notons φ l'application qui à toute matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette opération.
Alors il existe une matrice inversible Q d'ordre n telle que : $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}),\ \varphi(A) = QA$.

- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, toute opération élémentaire sur les *lignes* équivaut donc à une multiplication à *gauche* par une certaine matrice inversible Q d'ordre n. Pour savoir quelle est cette matrice inversible Q, il suffit d'appliquer cette opération élémentaire à la matrice identité I_n.
- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, on considère une succession donnée d'opérations élémentaires sur les lignes.
  Soit φ l'application qui à A dans $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette succession d'opérations.
  Alors il existe une matrice inversible Q d'ordre n telle que : $\forall A \in \mathcal{M}_n(\mathbb{K}),\ \varphi(A) = QA$.
- En particulier, si on applique à A une opération élémentaire sur les lignes (ou une succession de telles opérations), on transforme A en une matrice B telle que $\mathrm{Ker}(B) = \mathrm{Ker}(A)$.
  D'une façon plus familière, les opérations élémentaires sur les lignes « conservent le noyau ».

▷ Interprétation matricielle des opérations élémentaires sur les colonnes

Proposition 17.5.3 (interprétation matricielle d'une opération élémentaire sur les colonnes)
On se place dans $\mathcal{M}_{n,p}(\mathbb{K})$ et on considère une opération élémentaire particulière sur les colonnes.
Notons φ l'application qui à toute matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette opération.
Alors il existe une matrice inversible P d'ordre p telle que : $\forall A \in \mathcal{M}_{n,p}(\mathbb{K}),\ \varphi(A) = AP$.

- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, toute opération élémentaire sur les *colonnes* équivaut donc à une multiplication à *droite* par une certaine matrice inversible P d'ordre p. Pour savoir quelle est cette matrice inversible P, il suffit d'appliquer cette opération élémentaire à la matrice identité I_p.
- Dans $\mathcal{M}_{n,p}(\mathbb{K})$, on considère une succession donnée d'opérations élémentaires sur les colonnes.
  Soit φ l'application qui à A dans $\mathcal{M}_{n,p}(\mathbb{K})$ associe sa transformée par cette succession d'opérations.
  Alors il existe une matrice inversible P d'ordre p telle que : $\forall\, A \in \mathcal{M}_n(\mathbb{K}),\ \varphi(A) = AP$.
- En particulier, si on applique à A une opération élémentaire sur les colonnes (ou une succession de telles opérations), on transforme A en une matrice B telle que $\mathrm{Im}\ (B) = \mathrm{Im}\ (A)$.
  D'une façon plus familière, les opérations élémentaires sur les colonnes « conservent l'image ».

Proposition 17.5.4 (conservation du rang par des opérations élémentaires)
Toute opération élémentaire (ou toute suite d'opérations élémentaires), sur les lignes et/ou sur les colonnes, transforme une matrice A en une matrice de même rang.

On peut donc dire que les opérations élémentaires (sur les colonnes, sur les lignes) « conservent le rang ».

17.5.3 Calcul du rang par la méthode du pivot

Rappelons les trois acceptions du mot « rang » :

- Soit E, F deux espaces vectoriels sur $\mathbb{K}$, avec $\dim(E) = p \geqslant 1$, et $\dim(F) = n \geqslant 1$.
  Soit $f : E \to F$ une application linéaire. Le rang de f est la dimension de $\mathrm{Im}\ (f)$.
- Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$.
  Alors $\mathrm{rg}\,(A) = \mathrm{rg}\,(f)$, où $f : \mathbb{K}^p \to \mathbb{K}^n$ est l'application linéaire canoniquement associée à A.
- Soit E un $\mathbb{K}$ espace vectoriel, avec $\dim(E) = n$. Soit $(v_j)_{1\leqslant j\leqslant p}$ une famille de p vecteurs de E.
  Soit $F = \mathrm{Vect}(v)$ le sous-espace de E engendré par les vecteurs v_j.
  Le rang de la famille $v = (v_j)_{1\leqslant j\leqslant p}$ est la dimension du sous-espace $F = \mathrm{Vect}(v)$.

En fait, chacune de ces trois définitions se rattache immédiatement aux deux autres :

- Le rang d'une matrice A est celui de la famille de ses vecteurs-colonne, et celui de la famille de ses vecteurs-ligne, et celui de toute application linéaire susceptible d'être représentée par A.
- Le rang d'une famille de vecteurs de E est celui de la matrice des coordonnées de ces vecteurs dans une base e quelconque de E.
- Le rang d'une application linéaire $f : E \to F$ est le rang de la matrice de f dans un couple de bases quelconques (e dans E et ε dans F). C'est aussi le rang de la famille des images des vecteurs d'une base quelconque e de E.

Finalement tout peut se ramener au calcul du rang d'une matrice.

Proposition 17.5.5 (passage à une forme échelonnée supérieure)
Soit A une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$. Par une succession bien choisie d'opérations élémentaires sur les lignes, on peut transformer A en une matrice échelonnée supérieurement.

Les opérations élémentaires ne modifiant pas le rang de la matrice initiale, on peut ainsi calculer le rang de A : c'est celui de la matrice échelonnée finale, c'est-à-dire le nombre de ses pivots non nuls.

Exemple de calcul de rang

On demande de calculer le rang de $A = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 \\ 3 & 7 & 11 & 15 & 19 \\ 2 & 6 & 1 & 0 & 11 \\ 1 & 3 & 14 & 21 & 16 \end{pmatrix}$

Au moyen des opérations $\begin{cases} \mathrm{L}_2 \leftarrow \mathrm{L}_2 - 3\mathrm{L}_1 \\ \mathrm{L}_3 \leftarrow \mathrm{L}_3 - 2\mathrm{L}_1 \\ \mathrm{L}_4 \leftarrow \mathrm{L}_4 - \mathrm{L}_1 \end{cases}$ on passe de A à $A_1 = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 \\ 0 & -8 & -16 & -24 & -32 \\ 0 & -4 & -17 & -26 & -23 \\ 0 & -2 & 5 & 8 & -1 \end{pmatrix}$

Au moyen des opérations $\begin{cases} \mathrm{L}_3 \leftarrow 2\mathrm{L}_3 - \mathrm{L}_2 \\ \mathrm{L}_4 \leftarrow 4\mathrm{L}_4 - \mathrm{L}_2 \end{cases}$ on passe de A_1 à $A_2 = \begin{pmatrix} 1 & 5 & 9 & 13 & 17 \\ 0 & -8 & -16 & -24 & -32 \\ 0 & 0 & 18 & 28 & 14 \\ 0 & 0 & -36 & -56 & -28 \end{pmatrix}$

Au moyen de l'opération $\mathrm{L}_4 \leftarrow \mathrm{L}_4 + 2\mathrm{L}_3$, on passe de A_2 à $A_3 = \begin{pmatrix} \boxed{1} & 5 & 9 & 13 & 17 \\ 0 & \boxed{-8} & -16 & -24 & -32 \\ 0 & 0 & \boxed{18} & 28 & 14 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

La matrice finale, donc la matrice initiale A, est de rang 3.

17.5.4 Calcul de l'inverse par la méthode du pivot

Proposition 17.5.6 (inversion d'une matrice par opérations élémentaires)
Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors A est inversible si et seulement si il est possible, par une succession bien choisie d'opérations élémentaires sur les lignes, de passer de la matrice A à la matrice I_n.
La même succession d'opérations, dans le même ordre, transforme alors I_n en la matrice A^{-1}.

▷ Principe de la méthode

Soit A la matrice de $\mathcal{M}_n(\mathbb{K})$ dont on veut calculer l'inverse.

On place A et I_n côte à côte dans un tableau $(A \mid I_n)$ à n lignes et $2n$ colonnes.

On procède ensuite à une succession d'opérations élémentaires sur les lignes de ce tableau.

Cette suite d'opérations correspond à la multiplication à gauche par une matrice inversible P.

On passe donc du tableau $(A \mid I_n)$ au tableau $P(A \mid I_n) = (PA \mid PI_n)$.

On choisit la succession d'opérations sur les lignes de A de manière à transformer A en $PA = I_n$.

Cela signifie que $P = A^{-1}$, et le tableau final est donc $(I_n \mid A^{-1})$.

▷ Dans la pratique

– On part du tableau $(A \mid \mathrm{I}_n) = \left(\begin{array}{cccc|ccccc} a_{11} & a_{12} & \dots & a_{1p} & 1 & 0 & \dots & 0 & 0 \\ a_{21} & a_{22} & \dots & a_{2p} & 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \vdots & 0 & \ddots & \ddots & 1 & 0 \\ a_{n1} & a_{n2} & \dots & a_{n,p} & 0 & 0 & \dots & 0 & 1 \end{array}\right)$

– Il y a au moins un coefficient non nul sur la première colonne sinon A ne serait pas inversible.
Au besoin après un échange de deux lignes, on peut supposer $a_{11} \neq 0$.

– On applique les opérations élémentaires : $\mathrm{L}_i \leftarrow a_{11}\mathrm{L}_i - a_{i1}\mathrm{L}_1$, pour $i = 2, \ldots, n$.
Dans cette succession d'opérations, le coefficient non nul a_{11} est appelé le *pivot*.

Le tableau prend alors la forme suivante : $\left(\begin{array}{cccc|cccccc} a_{11} & a_{12} & \ldots & a_{1p} & 1 & 0 & \ldots & & 0 & 0 \\ 0 & b_{22} & \ldots & b_{2p} & -a_{21} & a_{11} & \ddots & & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & a_{11} & 0 \\ 0 & b_{n2} & \ldots & b_{n,p} & -a_{n1} & 0 & \ldots & & 0 & a_{11} \end{array}\right)$

– La partie gauche du tableau est une matrice qui a le même rang que A, et qui est donc inversible.

Il y a donc au moins un coefficient non nul dans la deuxième colonne et à partir de la deuxième ligne (sinon les deux premières colonnes seraient proportionnelles et A ne serait pas inversible).

Au besoin après un échange de L_2 avec une ligne L_i d'indice $i \geqslant 3$, on peut donc supposer $b_{22} \neq 0$.

Avec le pivot b_{22}, on annule tous les coefficients de la deuxième colonne (sans toucher à la ligne L_2).

– On continue ainsi jusqu'à obtenir à la place de A une matrice diagonale.

On termine en divisant chaque ligne par le coefficient diagonal obtenu.

À la place qu'occupait I_n se trouve maintenant A^{-1}.

– Il n'est pas utile (au contraire) de rendre égaux à 1 les coefficients diagonaux de la moitié gauche du tableau avant d'y avoir obtenu une matrice diagonale, pour éviter d'introduire des coefficients fractionnaires difficiles à manipuler : puisqu'on demande souvent d'inverser des matrices à coefficients entiers, autant garder les coefficients entiers le plus longtemps possible !

▷ Un exemple assez technique

Soit $A = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 & 1/5 \\ 1/3 & 1/4 & 1/5 & 1/6 \\ 1/4 & 1/5 & 1/6 & 1/7 \end{pmatrix}$. Pour calculer A^{-1}, on forme $\left(\begin{array}{cccc|cccc} 1 & 1/2 & 1/3 & 1/4 & 1 & 0 & 0 & 0 \\ 1/2 & 1/3 & 1/4 & 1/5 & 0 & 1 & 0 & 0 \\ 1/3 & 1/4 & 1/5 & 1/6 & 0 & 0 & 1 & 0 \\ 1/4 & 1/5 & 1/6 & 1/7 & 0 & 0 & 0 & 1 \end{array}\right)$

Pour limiter la technicité de l'exercice, il est préférable de travailler sur des coefficients entiers.

Les opérations $\begin{cases} \mathrm{L}_1 \leftarrow 12\mathrm{L}_1 \\ \mathrm{L}_2 \leftarrow 60\mathrm{L}_2 \\ \mathrm{L}_3 \leftarrow 60\mathrm{L}_3 \\ \mathrm{L}_4 \leftarrow 420\mathrm{L}_4 \end{cases}$ conduisent au tableau $\left(\begin{array}{cccc|cccc} 12 & 6 & 4 & 3 & 12 & 0 & 0 & 0 \\ 30 & 20 & 15 & 12 & 0 & 60 & 0 & 0 \\ 20 & 15 & 12 & 10 & 0 & 0 & 60 & 0 \\ 105 & 84 & 70 & 60 & 0 & 0 & 0 & 420 \end{array}\right)$

On va utiliser le pivot 12 en position $(1, 1)$.

Les opérations $\begin{cases} \mathrm{L}_2 \leftarrow 2\mathrm{L}_2 - 5\mathrm{L}_1 \\ \mathrm{L}_3 \leftarrow 3\mathrm{L}_3 - 5\mathrm{L}_1 \\ \mathrm{L}_4 \leftarrow 4\mathrm{L}_4 - 35\mathrm{L}_1 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 12 & 6 & 4 & 3 & 12 & 0 & 0 & 0 \\ 0 & 10 & 10 & 9 & -60 & 120 & 0 & 0 \\ 0 & 15 & 16 & 15 & -60 & 0 & 180 & 0 \\ 0 & 126 & 140 & 135 & -420 & 0 & 0 & 1680 \end{array}\right)$

On va utiliser le pivot 10 en position $(2, 2)$.

Les opérations $\begin{cases} \mathrm{L}_1 \leftarrow 5\mathrm{L}_1 - 3\mathrm{L}_2 \\ \mathrm{L}_3 \leftarrow 2\mathrm{L}_3 - 3\mathrm{L}_2 \\ \mathrm{L}_4 \leftarrow 5\mathrm{L}_4 - 63\mathrm{L}_2 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 60 & 0 & -10 & -12 & 240 & -360 & 0 & 0 \\ 0 & 10 & 10 & 9 & -60 & 120 & 0 & 0 \\ 0 & 0 & 2 & 3 & 60 & -360 & 360 & 0 \\ 0 & 0 & 70 & 108 & 1680 & -7560 & 0 & 8400 \end{array}\right)$

On va utiliser le pivot 2 en position $(3, 3)$.

Les opérations $\begin{cases} L_1 \leftarrow L_1 + 5L_3 \\ L_2 \leftarrow L_2 - 5L_3 \\ L_4 \leftarrow L_4 - 35L_3 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 60 & 0 & 0 & 3 & 540 & -2160 & 1800 & 0 \\ 0 & 10 & 0 & -6 & -360 & 1920 & -1800 & 0 \\ 0 & 0 & 2 & 3 & 60 & -360 & 360 & 0 \\ 0 & 0 & 0 & 3 & -420 & 5040 & -12600 & 8400 \end{array}\right)$

On va utiliser le pivot 3 en position $(4, 4)$.

Les opérations $\begin{cases} L_1 \leftarrow L_1 - L_4 \\ L_2 \leftarrow L_2 + 2L_4 \\ L_3 \leftarrow L_3 - L_4 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 60 & 0 & 0 & 0 & 960 & -7200 & 14400 & -8400 \\ 0 & 10 & 0 & 0 & -1200 & 12000 & -27000 & 16800 \\ 0 & 0 & 2 & 0 & 480 & -5400 & 12960 & -8400 \\ 0 & 0 & 0 & 3 & -420 & 5040 & -12600 & 8400 \end{array}\right)$

On a transformé la partie gauche du tableau en une matrice diagonale dont les coefficients diagonaux sont non nuls (cette matrice est donc inversible, ce qui prouve que la matrice initiale A est inversible).

On peut maintenant transformer la partie gauche du tableau en la matrice identité.

Les opérations $\begin{cases} L_1 \leftarrow L_1/60 \\ L_2 \leftarrow L_2/10 \\ L_3 \leftarrow L_3/2 \\ L_4 \leftarrow L_4/3 \end{cases}$ conduisent à $\left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 16 & -120 & 240 & -140 \\ 0 & 1 & 0 & 0 & -120 & 1200 & -2700 & 1680 \\ 0 & 0 & 1 & 0 & 240 & -2700 & 6480 & -4200 \\ 0 & 0 & 0 & 1 & -140 & 1680 & -4200 & 2800 \end{array}\right)$

On a ainsi montré que A est inversible et que $A^{-1} = \begin{pmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{pmatrix}$

▷ **Le même exemple en `Python`**

On importe le module `numpy` (renommé comme d'habitude en `np`), et on fabrique la matrice A :

```
>>> import numpy as np
>>> A = np.array([[1/(i+j+1) for j in range(4)] for i in range(4)]); A
[[ 1.          0.5         0.33333333  0.25      ]
 [ 0.5         0.33333333  0.25        0.2       ]
 [ 0.33333333  0.25        0.2         0.16666667]
 [ 0.25        0.2         0.16666667  0.14285714]]
```

Par sécurité, on fait une sauvegarde de A dans la variable $A\_$.
On borde ensuite A par la matrice identité d'ordre 4 :

```
>>> A_ = A.copy(); A = np.hstack((A,np.eye(4))); print(A)
[[ 1.          0.5         0.33333333  0.25        1.   0.   0.   0. ]
 [ 0.5         0.33333333  0.25        0.2         0.   1.   0.   0. ]
 [ 0.33333333  0.25        0.2         0.16666667  0.   0.   1.   0. ]
 [ 0.25        0.2         0.16666667  0.14285714  0.   0.   0.   1. ]]
```

On multiplie chaque ligne par un coefficient bien choisi, ce qui permet de passer en coefficients entiers :

```
>>> A[0]=12*A[0]; A[1]=60*A[1]; A[2]=60*A[2]; A[3]=420*A[3]
>>> A = np.asarray(A,dtype='int'); print(A)
[[ 12   6   4   3  12   0   0   0]
 [ 30  20  15  12   0  60   0   0]
 [ 20  15  12  10   0   0  60   0]
 [105  84  70  60   0   0   0 420]]
```

On effectue les opérations élémentaires avec le pivot qui est en position $(1, 1)$:

```
>>> A[1]=2*A[1]-5*A[0]; A[2]=3*A[2]-5*A[0]; A[3]=4*A[3]-35*A[0]; print(A)
[[  12    6    4    3   12    0    0    0]
 [   0   10   10    9  -60  120    0    0]
 [   0   15   16   15  -60    0  180    0]
 [   0  126  140  135 -420    0    0 1680]]
```

On effectue les opérations élémentaires avec le pivot qui est en position $(2, 2)$:

```
>>> A[0]=5*A[0]-3*A[1]; A[2]=2*A[2]-3*A[1]; A[3]=5*A[3]-63*A[1]; print(A)
[[   60     0   -10   -12   240  -360     0     0]
 [    0    10    10     9   -60   120     0     0]
 [    0     0     2     3    60  -360   360     0]
 [    0     0    70   108  1680 -7560     0  8400]]
```

On effectue les opérations élémentaires avec le pivot qui est en position $(3, 3)$:

```
>>> A[0]=A[0]+5*A[2]; A[1]=A[1]-5*A[2]; A[3]=A[3]-35*A[2]; print(A)
[[    60      0      0      3    540  -2160   1800      0]
 [     0     10      0     -6   -360   1920  -1800      0]
 [     0      0      2      3     60   -360    360      0]
 [     0      0      0      3   -420   5040 -12600   8400]]
```

On effectue les opérations élémentaires avec le pivot qui est en position $(4, 4)$:

```
>>> A[0]=A[0]-A[3]; A[1]=A[1]+2*A[3]; A[2]=A[2]-A[3]; print(A)
[[    60      0      0      0    960  -7200  14400  -8400]
 [     0     10      0      0  -1200  12000 -27000  16800]
 [     0      0      2      0    480  -5400  12960  -8400]
 [     0      0      0      3   -420   5040 -12600   8400]]
```

On divise chaque ligne par un coefficient bien choisi, pour obtenir l'identité à gauche :

```
>>> A[0] = A[0]/60; A[1] = A[1]/10; A[2] = A[2]/2; A[3] = A[3]/3; print(A)
[[    1     0     0     0    16  -120   240  -140]
 [    0     1     0     0  -120  1200 -2700  1680]
 [    0     0     1     0   240 -2700  6480 -4200]
 [    0     0     0     1  -140  1680 -4200  2800]]
```

On extrait le bloc de droite : c'est la matrice $B = A^{-1}$.

On vérifie le résultat en effectuant le produit AB avec la copie de sauvegarde de A. On obtient bien la matrice identité (on a arrondi ici à douze décimales, donc c'est très précis) :

```
>>> B = A[:,4:]; print(B)
[[   16  -120   240  -140]
 [ -120  1200 -2700  1680]
 [  240 -2700  6480 -4200]
 [ -140  1680 -4200  2800]]
```

```
>>> (A_.dot(B)).round(12)
array([[ 1.,  0.,  0.,  0.],
       [ 0.,  1.,  0.,  0.],
       [ 0.,  0.,  1., -0.],
       [ 0.,  0.,  0.,  1.]])
```

17.6 Systèmes linéaires

17.6.1 Généralités et définitions

On appelle *système linéaire* de n équations, à p inconnues et à coefficients dans $\mathbb{K}$ tout système d'équations de la forme :

$$(S)\quad \begin{cases} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1j}\,x_j + \cdots + a_{1p}\,x_p = b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2j}\,x_j + \cdots + a_{2p}\,x_p = b_2 \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{i1}\,x_1 + a_{i2}\,x_2 + \cdots + a_{ij}\,x_j + \cdots + a_{ip}\,x_p = b_i \\ \quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \cdots + a_{nj}\,x_j + \cdots + a_{np}\,x_p = b_n \end{cases}$$

▷ Matrice du système, et seconds membres

Les a_{ij} sont appelés les *coefficients du système.*

La matrice $A = (a_{i,j})$, élément de $\mathcal{M}_{n,p}(\mathbb{K})$, est appelée *matrice du système.*

Le rang de la matrice A est appelé le *rang du système.*

Les coefficients $b_1, \ldots, b_n$ sont appelés *seconds membres* du système.

Le système (S) est dit *triangulaire* (ou “en escaliers”, ou “en cascades”) si la matrice A est triangulaire.

Il est dit *carré* si $n = p$. La matrice A est alors un élément de $\mathcal{M}_n(\mathbb{K})$.

▷ Inconnues et solutions

On dit que l'élément $(x_1, x_2, \ldots, x_p)$ de $\mathbb{K}^p$ est le p-uplet des *inconnues* du système.

On dit que $(x_1, x_2, \ldots, x_p)$ est une *solution* du système si les valeurs $x_1, x_2, \ldots, x_p$ satisfont à chacune des égalités figurant dans (S).

L'ensemble des solutions d'un système linéaire peut être vide ou non vide.

S'il est non vide, on verra qu'il est soit réduit à un seul élément de $\mathbb{K}^p$, soit infini.

Deux systèmes linéaires sont dits *équivalents* s'ils ont le même ensemble de solutions.

▷ Système homogène associé

Un système linéaire est dit *homogène* si ses seconds membres b_i sont nuls.

À un système (S) on *associe* un système homogène (H) en annulant les seconds membres.

Un système homogène possède au moins la solution $(0, 0, \ldots, 0)$ dite *solution triviale.*

17.6.2 Interprétations d'un système linéaire

On reprend le système (S) défini au paragraphe précédent.

▷ Interprétation matricielle

Soit $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{K})$ la matrice de (S), $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ (seconds membres) et $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ (inconnues).

Le système (S) s'écrit $\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1p} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix}$, c'est-à-dire $AX = B$.

Si A est carrée inversible (système dit « de Cramer »), (S) possède une solution unique $X = A^{-1}B$.

Le système homogène (H) associé à (S) s'écrit : $AX = 0$.

▷ Interprétation en termes d'applications linéaires

Soit f l'application linéaire de $\mathbb{K}^p$ vers $\mathbb{K}^n$, de matrice A dans les bases canoniques.
Soit $x = (x_1, x_2, \dots, x_p)$ dans $\mathbb{K}^p$ et $b = (b_1, b_2, \dots, b_n)$ dans $\mathbb{K}^n$.

Le système (S) équivaut alors à l'égalité vectorielle $f(x) = b$.
Résoudre (S), c'est donc chercher l'image réciproque du vecteur b de $\mathbb{K}^n$ par l'application linéaire f.
Le système (S) admet au moins une solution si et seulement si b est dans l'image de f.

Le système homogène associé (H) équivaut à $f(x) = \overrightarrow{0}$.
Résoudre (H), c'est donc trouver $\operatorname{Ker}(f)$.

▷ Interprétation en termes de combinaisons linéaires

Pour $1 \leqslant j \leqslant p$, soit $C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$ la j-ième colonne de A, et $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ la colonne des seconds membres.

Le système (S) s'écrit : $x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + \cdots + x_j \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix} + \cdots + x_p \begin{pmatrix} a_{1p} \\ a_{2p} \\ \vdots \\ a_{np} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$.

Autrement dit, (S) s'écrit : $x_1C_1 + \cdots + x_jC_j + \cdots + x_pC_p = B$.

Résoudre (S), c'est trouver toutes les façons d'écrire B comme combinaison linéaire de $C_1, C_2, \dots, C_p$.

Le système (S) a donc au moins une solution si et seulement si B est dans $\operatorname{Vect}(C_1, C_2, \dots, C_p)$.

Résoudre (H), c'est trouver tous les p-uplets $(x_1, \dots, x_p)$ tels que $\displaystyle\sum_{k=1}^{p} x_k C_k = \overrightarrow{0}$.

Le système (H) a d'autres solutions que la solution triviale si et seulement si $C_1, C_2, \dots, C_p$ sont liés.

▷ Interprétation en termes d'intersections d'hyperplans affines

Revenons à la définition initiale du système (S).

Considérons la i-ème équation (E_i) : $a_{i1}\, x_1 \;+ a_{i2}\, x_2 \;+ \cdots \;+ a_{ij}\, x_j + \cdots + a_{ip}\, x_p = b_i$.

On commence par éliminer le cas particulier où tous les coefficients $a_{i,j}$ de cette ligne sont nuls : si $b_i = 0$, alors l'équation (E_i) se réduit à $0 = 0$ et elle peut évidemment être supprimée, mais si $b_i \neq 0$ l'équation (E_i) (et donc le système (S)) n'a aucune solution.

On suppose donc que chacune des p-uplets $(a_{i1}, a_{i2}, \ldots, a_{ip})$, avec $1 \leqslant i \leqslant n$ est non identiquement nul.

Chaque application $\varphi_i \colon \mathbb{K}^p \to \mathbb{K}$ définie par $\varphi_i(x_1, x_2, \ldots, x_p) = a_{i1}\, x_1 + a_{i2}\, x_2 + \cdots + a_{ij}\, x_j + \cdots + a_{ip}\, x_p$ est alors une forme linéaire non nulle sur $\mathbb{K}^p$.

L'équation (E_i) s'écrit $\varphi_i(x) = b$, et elle définit un hyperplan affine $\mathcal{H}_i$ de $\mathbb{K}^p$.

L'équation correspondante du système homogène associé (H) s'écrit $\varphi_i(x) = 0$, et elle définit l' hyperplan vectoriel H_i de $\mathbb{K}^p$ qui est la direction de $\mathcal{H}_i$.

Avec cette interprétation, résoudre (S) c'est déterminer l'intersection éventuelle des hyperplans affines $\mathcal{H}_i$, et résoudre (H) c'est déterminer l'intersection des hyperplans vectoriels H_i.

17.6.3 Structure de l'ensemble des solutions

On reprend le système (S) défini dans les deux sous-sections précédentes.

▷ Structure de l'ensemble des solutions du système homogène

Utilisons l'interprétation de (H) en termes d'applications linéaires : $x = (x_1, x_2, \ldots, x_p)$ est solution du système homogène (H) si et seulement si $f(x) = \overrightarrow{0}$ c'est-à-dire si et seulement si x est dans $\mathrm{Ker}(f)$.

Or le théorème du rang nous dit que $\mathrm{Ker}(f)$ est un sous-espace vectoriel de $\mathbb{K}^p$ de dimension $p - \mathrm{rg}\,(f)$, où $\mathrm{rg}\,(f)$ est le rang de l'application linéaire f c'est-à-dire le rang de la matrice A.

On peut donc énoncer :

Proposition 17.6.1 (structure de l'ensemble des solutions du système linéaire homogène (H))
Soit (H) un système linéaire homogène de n équations, à p inconnues et à coefficients dans $\mathbb{K}$.
Soit $AX = 0$ l'écriture matricielle de ce système, avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$ et X dans $\mathcal{M}_{p,1}(\mathbb{K})$.
Soit r le rang du système, c'est-à-dire le rang de la matrice A.
Alors l'ensemble des solutions X de (H) est un espace vectoriel de dimension $p - r$.

▷ Structure de l'ensemble des solutions du système (S)

Supposons que (S) possède au moins une solution x_0. Soit x un vecteur de $\mathbb{K}^p$.

On a les équivalences : (x est solution de (S)) $\Leftrightarrow\ f(x) = b\ \Leftrightarrow\ f(x) = f(x_0)\ \Leftrightarrow\ f(x - x_0) = \overrightarrow{0}$.

Ainsi les solution de (S) sont les $x = x_0 + h$ où h est une solution quelconque de (H).

L'ensemble des solutions $\mathcal{S}_S$ de (S) est alors un sous-espace affine dont la direction est l'ensemble des solutions $\mathcal{S}_H$ de (H) (et dont la dimension est donc $p - r$).

Attention : le calcul précédent est fait sous réserve que le système (S) admette au moins une solution !

Définition 17.6.1 (système linéaire compatible)
Soit (S) un système linéaire homogène de n équations, à p inconnues et à coefficients dans $\mathbb{K}$.
Soit $AX = B$ l'écriture matricielle de (S), avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$, B dans $\mathcal{M}_{n,1}(\mathbb{K})$, X dans $\mathcal{M}_{p,1}(\mathbb{K})$.
On dit que le système (S) est *compatible* si l'ensemble de ses solutions est non vide.
Cela équivaut, avec l'interprétation matricielle, à dire que B est dans l'image de A.

On peut donc énoncer :

Proposition 17.6.2 (structure de l'ensemble des solutions d'un système linéaire compatible)
Soit (S) un système linéaire homogène de n équations, à p inconnues et à coefficients dans $\mathbb{K}$.
Soit $AX = B$ l'écriture matricielle de (S), avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$, B dans $\mathcal{M}_{n,1}(\mathbb{K})$, X dans $\mathcal{M}_{p,1}(\mathbb{K})$.
Soit r le rang du système, c'est-à-dire le rang de la matrice A.
On suppose que le système (S) est compatible.
Alors l'ensemble des solutions X de (S) est un sous-espace affine de dimension $p - r$.
Il s'obtient en ajoutant à une solution particulière de (S) la solution générale de (H).

▷ **Remarques**

On garde bien sûr les notations précédentes, le système (S) s'écrivant indifféremment $f(x) = b$ (avec f dans $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$) ou $AX = B$ (avec A dans $\mathcal{M}_{n,p}(\mathbb{K})$).

Le système (S) est *compatible* si et seulement si b appartient à $\text{Im}\ (f)$.

Cela dépend de b, mais si f est surjective (c'est-à-dire $\text{rg}\,(A) = n$) alors (S) est *toujours* compatible.

Si f est injective (c'est-à-dire $\text{rg}\,(A) = p$) alors (S) possède *au plus* une solution, quelque soit le second membre b. Le système homogène (H), quant à lui, possède alors uniquement la solution triviale.

L'ensemble des solutions d'un système linéaire quelconque (S) est donc ou bien vide (système *incompatible*) ou bien réduit à un seul élément (cas où f est bijective), ou bien infini (en tant que sous-espace affine de dimension supérieure ou égale à 1).

17.6.4 Systèmes de Cramer

On se place ici dans le cas d'un système carré (n équations et n inconnues).
Matriciellement, le système (S) s'écrit donc $AX = B$, avec A dans $\mathcal{M}_n(\mathbb{K})$, et X, B dans $\mathcal{M}_{n,1}(\mathbb{K})$.

Proposition 17.6.3 (solution unique d'un système de Cramer)
Soit (S) un système linéaire de n équations, à n inconnues et à coefficients dans $\mathbb{K}$.
Soit $AX = B$ son écriture matricielle, avec A dans $\mathcal{M}_n(\mathbb{K})$, B dans $\mathcal{M}_{n,1}(\mathbb{K})$, X cherché dans $\mathcal{M}_{n,1}(\mathbb{K})$.
Soit r le rang du système, c'est-à-dire le rang de la matrice A.
Le système (S) possède une solution unique si et seulement $r = n$, c'est-à-dire si et seulement si la matrice A est inversible. Cette solution est $X = A^{-1}B$.
On dit alors que le système (S) est un « système de Cramer »

Avec l'interprétation en termes d'applications linéaires : $f(x) = b \Leftrightarrow x = f^{-1}(b)$.
Si on utilise l'interprétation de (S) en termes de combinaisons linéaires, l'unique solution $(x_1, x_2, \ldots, x_n)$ de (S) est le n-uplet des coordonnées du vecteur B (colonne des seconds membres) dans la base de $\mathbb{K}^n$ formée par les colonnes $C_1, C_2, \ldots, C_n$ de la matrice A.

Résolution d'un système de Cramer avec Python

On utilise la fonction `solve` dans `numpy.linalg` :

```
>>> import numpy as np                # toujours s'assurer que numpy est importé!
>>> a = np.array([[1,3,2],[4,1,5],[6,2,7]]); a  # matrice A carrée d'ordre 3
array([[1, 3, 2],
       [4, 1, 5],
       [6, 2, 7]])
>>> np.linalg.det(a)                  # déterminant non nul, matrice inversible
7.0000000000000071
>>> b = np.array([1,4,2]); b          # le vecteur B des seconds membres
array([1, 4, 2])
>>> x = np.linalg.solve(a, b); x      # la solution X du système
array([-6.42857143, -1.71428571,  6.28571429])
>>> np.dot(a,x)                       # on vérifie qu'on a bien AX=B
array([ 1.,  4.,  2.])
```

Voici ce qui se passe si la matrice carrée A n'est pas inversible.
Dans l'exemple ci-dessous la matrice A est de rang 2 (on a en effet $\mathrm{L}_3 = \mathrm{L}_2 - 2\mathrm{L}_1$).

```
>>> a = np.array([[1,3,2],[4,1,5],[2,-5,1]]); a
array([[ 1,  3,  2],
       [ 4,  1,  5],
       [ 2, -5,  1]])
>>> x = np.linalg.solve(a, b);     # avec les mêmes seconds membres
    [...]
    numpy.linalg.linalg.LinAlgError: Singular matrix
>>> np.linalg.det(a)               # effectivement le déterminant est nul
0.0
```

17.6.5 Résolution par la méthode du pivot de Gauss

Les aspects techniques de la résolution des systèmes linéaires par la méthode du pivot de Gauss ont été traités dans le chapitre 2 (« Calculs algébriques ») de ce cours.