Outils Mathématiques (SV2)
Année 2004-2005

Angela Pasquale

Département et Laboratoire de Mathématiques, Université de Metz
E-mail address: pasquale@math.univ-metz.fr

Table des matières

CHAPITRE 1

Calcul matriciel

Le calcul matriciel donne un outil théorique et pratique puissant pour étudier phénomènes comme par exemple l'évolution discrète des systèmes biologiques et écologiques ou la transmission des caractères génétiques d'une génération aux suivantes.

Exemple 1. Nous étudions un modèle simplifié pour l'évolution du paysage d'une région du midi de la France. Le paysage de la région peut évoluer dans l'un des quatre états suivants :

V Terrain cultivé : vignes, vergers, ...

Pe Pelouse

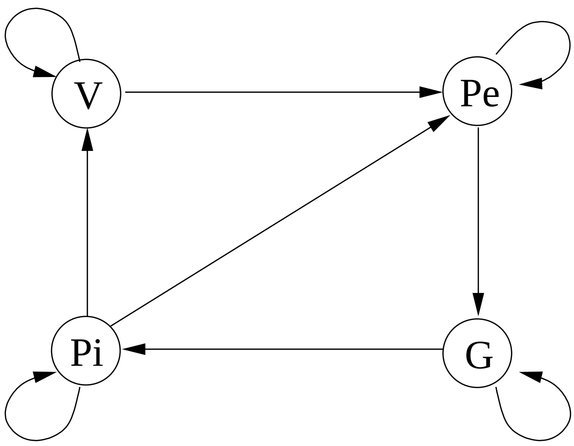
G Garrigue

Pi Pinède

Dans une période de dix ans le paysage peut changer selon le règles suivantes :

- étant V, il peut le rester ou peut se transformer par abandon en Pe ;
- étant Pe, il peut le rester ou peut se reconstruire en G ;
- étant G, il peut le rester ou peut se reconstruire en Pi ;
- étant Pi, il peut le rester ou peut être transformé en V ou peut se transformer en Pe par incendie ou dégradation.

On peut représenter le système par le diagramme suivant :



Chaque flèche représente une transition possible d'un état à un autre, et la direction de la flèche donne le sens dans lequel la transition a lieu.

Nous pouvons aussi représenter le diagramme sous forme d'un tableau dans lequel les lignes et les colonnes indiquent les quatre états.

	V	Pe	G	Pi
V	1	0	0	1
Pe	1	1	0	1
G	0	1	1	0
Pi	0	0	1	1

On écrit :

1 si une transition de l'état mentionné en haut de la colonne de cette case vers celui mentionné au début de la ligne, est possible.

0 si la transition n'est pas possible.

Un tableau de ce type s'appelle une matrice. Généralement, une matrice est notée par une lettre et les chiffres que la forment sont entourés par des parenthèses. Si on note la matrice ci-dessus par T, alors la notation habituelle est

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

T est la *matrice de transition* du modèle considéré.

Selon ce modèle, il n'est pas possible de passer directement de l'état G à l'état Pe, mais cette transition est possible en 2 étapes :

$$\text{G} \to \text{Pi} \to \text{Pe}$$

ou bien en 3 étapes, par exemple

$$\text{G} \to \text{Pi} \to \text{V} \to \text{Pe}$$

ou

$$\text{G} \to \text{Pi} \to \text{Pi} \to \text{Pe}$$

En général, dénombrer tous les transitions différentes possibles en n étapes d'un état à un autre en n étapes, est difficile. On verra que ce nombre peut être donné facilement par le calcul matriciel.

1. Matrices

DÉFINITION 1. Une *matrice* (réelle) $m \times n$ est un tableau à m lignes et n colonnes d'éléments de $\mathbb{R}$. On note

$$A = \left(a_{i,j}\right)_{1 \leq i \leq m; 1 \leq j \leq n}$$

la matrice dont $a_{i,j}$ est l'élément de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne. – On peut aussi écrire simplement $A = (a_{i,j})$ quand il n'y a pas ambiguïté sur la dimension de A –.

Les nombres $a_{i,j}$ s'appellent les *coefficients* de A. L'ensemble des matrices réelles $m \times n$ est noté $\mathcal{M}_{m,n}(\mathbb{R})$.

EXEMPLE 2. $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 6 \\ 3 & 1 & 1 & 3 \end{pmatrix}$ est une matrice 3×4 et $a_{2,4} = 6$.

1.1. Cas particuliers. Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$.

- Si tous les coefficients de A sont nuls, on dit que A est la *matrice nulle* ; on la note $0_{m,n}$ (ou plus simplement 0 quand il n'y a pas ambiguïté).
- Lorsque $m = n$ on dit que A est une *matrice carrée d'ordre n*. L'ensemble des matrices carrées d'ordre n se note $\mathcal{M}_n(\mathbb{R})$.

Définition 2. Un vecteur de $\mathbb{R}^m$ est une matrice $m \times 1$. Dans ce cas on note $v = (v_j)_{1\leq j\leq m}$.

Exemple 3. $v = \begin{pmatrix} 1 \\ 0 \\ \sqrt{2} \end{pmatrix}$ est un vecteur de $\mathbb{R}^3$.

Définition 3. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n avec $A = (a_{i,j})$. Les coefficients $a_{j,j}$ de la matrice s'appellent les *éléments diagonaux* de A. Lorsque $a_{i,j} = 0$ pour tous $i \neq j$, on dit que A est une *matrice diagonale*. La matrice diagonale d'ordre n avec $a_{j,j} = 1$ pour tout j s'appelle la *matrice unité* ; on la note I_n.

Exemple 4. La matrice $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ est diagonale et ses éléments diagonaux sont $a_{1,1} = 1$, $a_{2,2} = 3$ et $a_{3,3} = 0$.

Exemple 5. $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

2. Opérations sur les matrices

2.1. Addition de deux matrices $m \times n$. Soient $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B = (b_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$. La *somme de A et B*, notée $A + B$, est la matrice $m \times n$ définie par :

$$A + B = \Big(c_{i,j}\Big)_{1\leq i\leq m;1\leq j\leq n} \qquad \text{avec } c_{i,j} = a_{i,j} + b_{i,j} \text{ pour tout } i, j\,.$$

Exemple 6. Soient $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}$. Alors $A + B = \begin{pmatrix} 1 & 1 & 4 \\ 5 & 5 & 4 \end{pmatrix}$.

Remarque 1. (1) L'addition de matrices est *commutative* :

$$A + B = B + A \text{ pour tout } A, B \in \mathcal{M}_{m,n}(\mathbb{R}).$$

Elle est aussi *associative* :

$$(A + B) + C = A + (B + C) \text{ pour tout } A, B, C \in \mathcal{M}_{m,n}(\mathbb{R}).$$

(2) $O_{m,n}$ est l'*élément neutre* pour l'addition dans $\mathcal{M}_{m,n}(\mathbb{R})$, càd $A + O_{m,n} = A$ pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$.

(3) La matrice $-A = (-a_{i,j})$ s'appelle l'*opposée de A*. Elle vérifie $A + (-A) = 0_{m,n}$.

Exemple 7. Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix}$, alors $-A = \begin{pmatrix} -1 & -2 & -3 \\ 0 & 1 & 2 \end{pmatrix}$

2.2. Produit d'une matrice par un scalaire. Soient $A = (a_{i,j}) \in \mathcal{M}_{m,n}(\mathbb{R})$ et $\lambda \in \mathbb{R}$. La matrice produit de λ et A, notée λA est la matrice $m \times n$ définie par :

$$\lambda A = \Big(c_{i,j}\Big)_{1\leq i\leq m; 1\leq j\leq n} \qquad \text{avec } c_{i,j} = \lambda a_{i,j} \text{ pour tout } i, j\,.$$

Exemple 8. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & -2 \end{pmatrix}$. Alors $3A = \begin{pmatrix} 3 & 6 & 9 \\ 3 & 0 & -6 \end{pmatrix}$.

Remarque 2. Propriétés : Si $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ et $\lambda, \mu \in \mathbb{R}$, alors :

(1) $\lambda(A + B) = \lambda A + \lambda B$;

(2) $(\lambda + \mu)A = \lambda A + \mu A$;

(3) $1A = A$;

(4) $0A = 0_{m,n}$.

2.3. Produit d'une matrice $m\times n$ et d'une matrice $n\times p$. Soient $A = \Big(a_{i,j}\Big)_{1\leq i\leq m; 1\leq j\leq n} \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B = \Big(b_{j,l}\Big)_{1\leq j\leq n; 1\leq l\leq p} \in \mathcal{M}_{n,p}(\mathbb{R})$. Le *produit de A et de B*, noté AB, est la matrice $m \times p$ définie par

$$AB = \Big(c_{i,l}\Big)_{1\leq i\leq m; 1\leq l\leq p} \qquad \text{avec } c_{i,l} = \sum_{j=1}^{n} a_{i,j} b_{j,l} \text{ pour tout } i, j\,.$$

Le coefficient de AB d'indice i, l est donc obtenu en sommant les produits des éléments de la $i^{\text{ème}}$ ligne de A par les éléments correspondants de la $l^{\text{ème}}$ colonne de B.

Exemple 9. Soient $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix}$. Alors

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 1 & 1\cdot 2 + 2\cdot 0 & 1\cdot 1 + 2\cdot 0 \\ 3\cdot 1 + 4\cdot 1 & 3\cdot 2 + 4\cdot 0 & 3\cdot 1 + 4\cdot 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 7 & 6 & 3 \end{pmatrix}$$

$$2 \times \underbrace{2 \qquad 2}_{=} \times 3 \qquad\qquad\qquad 2 \times 3$$

Remarque 3. (1) Afin que le produit AB existe, il faut que le nombre de colonnes de A soit égal au nombre de lignes de B.

(2) Le produit AB existe toujours si A et B sont carrées du même ordre n.

(3) Si AB et BA existent, en général, $AB \neq BA$.

Exemple 10. Si $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, alors $AB = \begin{pmatrix} 1 & 3 \\ 3 & 7 \end{pmatrix}$ et $BA = \begin{pmatrix} 4 & 6 \\ 3 & 4 \end{pmatrix}$.

Remarque 4. Propriétés :

(a) Pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{n,p}(\mathbb{R})$ et $C \in \mathcal{M}_{p,q}(\mathbb{R})$: $A(BC) = (AB)C$.

(b) Pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $B, C \in \mathcal{M}_{n,p}(\mathbb{R})$: $A(B + C) = AB + AC$.

(c) Pour tout $A, B \in \mathcal{M}_{m,n}(\mathbb{R})$ et $C \in \mathcal{M}_{n,p}(\mathbb{R})$: $(A + B)C = AC + BC$.

(d) Pour tout $A \in \mathcal{M}_{m,n}(\mathbb{R})$, $B \in \mathcal{M}_{n,p}(\mathbb{R})$ et $\lambda \in \mathbb{R}$: $(\lambda A)B = \lambda(AB) = A(\lambda B)$.

(e) Pour tout $A \in \mathcal{M}_n(\mathbb{R})$: $AI_n = I_n A = A$.

2.4. Transposée. Soit $A = \left(a_{i,j}\right)_{1\le i\le m\,;1\le j\le n} \in \mathcal{M}_{m,n}(\mathbb{R})$. La *transposée* de A, notée A^T, est la matrice $A^T \in \mathcal{M}_{n,m}(\mathbb{R})$ dont le (i,j)-ème coefficient est $a_{j,i}$.

Exemple 11. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{pmatrix}$. Alors $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 0 \\ 3 & 1 \end{pmatrix}$.

Remarque 5. Propriétés : Soit $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $\lambda \in \mathbb{R}$. Alors :

(1) $(A^T)^T = A$;
(2) $(\lambda A)^T = \lambda A^T$;
(3) Si $B \in \mathcal{M}_{m,n}(\mathbb{R})$, alors $(A+B)^T = A^T + B^T$;
(4) Si $B \in \mathcal{M}_{n,p}(\mathbb{R})$, alors $(AB)^T = B^T A^T$.

2.5. Puissance $k^{\text{ème}}$ d'une matrice. Soit $A \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n. Le produit

$$\underbrace{A \cdot A \ldots A}_{k\text{-fois}}$$

s'appelle la *puissance $k^{\text{ème}}$ de A*, notée A^k. En particuliers : $A^1 = A$ et $A^{k+1} = A^k A = AA^k$.

On pose, par convention, $A^0 = I_n$.

Remarque 6. En général : $(AB)^k \neq A^k B^k$ pour tout $A, B \in \mathcal{M}_n(\mathbb{R})$.

Exemple 12. On considère la matrice de transition pour le système de l'exemple 1 :

$$T = \begin{array}{c c} \swarrow & \begin{array}{cccc} \text{V} & \text{Pe} & \text{G} & \text{Pi} \end{array} \\ \begin{array}{c} \text{V} \\ \text{Pe} \\ \text{G} \\ \text{Pi} \end{array} & \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

Alors

$$T^2 = TT = \begin{array}{c c} \swarrow & \\ \begin{array}{c} \\ \text{Pe} \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array} \begin{array}{c c} \swarrow & \begin{array}{cccc} & & \text{G} & \end{array} \\ & \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{array} = \begin{array}{c c} \swarrow & \begin{array}{cccc} & & \text{G} & \end{array} \\ \begin{array}{c} \\ \text{Pe} \\ \\ \\ \end{array} & \begin{pmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 3 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix} \end{array}$$

La $3^{\text{ème}}$ colonne de la deuxième matrice T dans le produit représente les transitions possibles avec début en G ; la $2^{\text{ème}}$ ligne de la première matrice T dans le produit représente les transitions possibles avec arrivée en Pe. Le coefficient d'indice $(2,3)$ de T^2 est donné par

$$1 \cdot 0 + 1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1 = 0 + 0 + 0 + 1 = 1$$

nombre chemins G $\to$ V $\to$ Pe
nombre chemins G $\to$ Pe $\to$ Pe
nombre chemins G $\to$ G $\to$ Pe
nombre chemins G $\to$ Pi $\to$ Pe

Donc le coefficient d'indice $(2,3)$ de T^2 représente le numéro de chemins de G (=le $3^{\text{ème}}$ état) à Pe (=le $2^{\text{ème}}$ état) en 2 étapes.

De même, si on considère T^3, on a :

$$T^3 = T^2 \cdot T = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 2 & 1 & 1 & 3 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 3 & 3 \\ 3 & 2 & 4 & 6 \\ 3 & 3 & 2 & 4 \\ 1 & 3 & 3 & 2 \end{pmatrix}$$

Le coefficient $(2,3)$ de T^3 est 4. Donc il y a 4 chemins possibles de G à Pe en 3 étapes.

En général : le coefficient (i,j) de T^k donne le nombre des chemins du $j^{\text{ème}}$ état au $i^{\text{ème}}$ état en k étapes.

3. Chaînes de Markov

Jusqu'au présent nous avons seulement considéré l'évolution d'un système discrète dans l'hypothèse que le système est dans l'un d'un nombre finis d'états (par ex. le paysage de la région du midi de la France de l'exemple 1 est un système qui peut être terrain cultivé, ou bien pelouse, ou bien garrigue, ou bien pinède). On se a donné les transitions possibles d'un état à l'autre pendant une période fixée de temps (égale à dix ans dans l'exemple considéré). Les transitions possibles ont été réunies dans la matrice de transition du système. A partir des puissances de la matrice de transition on peut donc déduire tous les transitions qui peuvent avoir lieu dans deux périodes, dans trois périodes, ou, en général, dans k périodes, où k est un nombre naturel arbitraire.

On étudie maintenant la situation dans laquelle le paysage de la région est reparti entre les états différents on considère combien de terrain passe (ou puisse passer) entre les états différents. Les outils nécessaires pour cet étude sont les chaînes de Markov.

Un *processus stochastique* est un modèle mathématique qui évolue dans le temps d'une façon probabilistique. Le but de l'étude d'un processus stochastique est d'obtenir, par l'analyse d'un modèle mathématique, une estimation probabilistique du comportement et de l'évolution temporelle d'un système.

Définition 4. Une *chaîne de Markov (en temps discret)* est un processus stochastique particuliers dans lequel :

(1) Un système est observé en temps discret $t = 0, 1, 2, \dots$.

(2) Le système est reparti entre les états d'une collection finie d'états possibles.

(3) Pour chaque temps t, les états dans lesquels le système se trouve au temps $t+1$ dépendent seulement des états dans lesquels le système se trouve au temps t.

On va expliquer cette définition au moyen d'un exemple.

Exemple 13. Le système considéré est le paysage de la région du midi de la France de l'exemple 1. Il peut se trouver dans un nombre fini d'états différents, qui sont : V, Pe, G, Pi. On étude l'évolution du système pour des valeurs discrètes du temps, chaque dix ans après un temps initial $t = 0$ (par exemple, si $t = 0$ est l'an 2000, alors $t = 1$ serait l'an 2010, $t = 2$ serait l'an 2020, etc.).

La condition (3) de la définition de chaînes de Markov est plus compliquée. On suppose d'abord que on connaît pour chaque instant t considéré et pour chaque pair $i, j \in \{1, 2, 3, 4\}$ le nombre

$$p_{i,j}(t) := \begin{array}{l} \text{probabilité que le système soit dans le } j^{\text{ème}} \text{ état au temps } t \\ \text{et dans le } i^{\text{ème}} \text{ état au temps } t+1. \end{array}$$

Dans l'exemple considéré, $p_{1,2}(0)$ est donc la probabilité que le système soit dans le $2^{\text{ème}}$ état au temps $t = 0$ (=en 2000) et dans le 1^{er} état au temps $t = 1$ (=en 2010).

Les nombres $p_{i,j}(t)$ sont définis comme suit : on pose

$$\begin{aligned} Q_j(t) &:= \text{portion du système qui peut être dans le } j^{\text{ème}} \text{ état au temps } t\,; \\ Q_{i,j}(t) &:= \begin{array}{l} \text{portion du système qui peut être dans le } j^{\text{ème}} \text{ état au temps } t \\ \text{et dans le } i^{\text{ème}} \text{ état au temps } t+1\,; \end{array} \end{aligned}$$

Alors

$$p_{i,j}(t) = \frac{Q_{i,j}(t)}{Q_j(t)}\,.$$

On remarque que $p_{i,j}(t) \in [0,1]$. La condition (3) de la définition de chaînes de Markov corresponde à supposer que ces probabilités ne dépendent pas de t. Donc $p_{i,j} \in \mathbb{R}$ et on obtient une matrice $P = (p_{i,j})$ (qui et une matrice carrée d'ordre 4 dans l'exemple).

Remarque 7. Le nombre $p_{i,j} = \frac{Q_{i,j}(t)}{Q_j(t)}$ est indépendant de t, mais $Q_{i,j}(t)$ et $Q_j(t)$ sont généralement dépendants de t.

Définition 5. La matrice P s'appelle la *matrice des probabilités de transition.* Dans un système avec N états possibles, elle est une matrice carrée d'ordre N.

Remarque 8. Soit $P = (p_{i,j})_{1\leq i,j\leq N}$ une matrice des probabilités de transition. Alors P a les propriétés suivantes :

(1) Pour tout $i,j \in \{1,2,\ldots,N\}$ on a $0 \leq p_{i,j} \leq 1$;

(2) Pour tout $j \in \{1,2,\ldots,N\}$ on a $\sum_{i=1}^{N} p_{i,j} = 1$.

La condition (2) signifie que la somme des coefficients de chaque colonne est égale à 1.

Preuve de (2) : La portion du système qui au temps t est dans le $j^{\text{ème}}$ état doit se partager entre les N états possibles au temps $t+1$, càd $\sum_{i=1}^{N} Q_{i,j}(t) = Q_j(t)$, d'où

$$\sum_{i=1}^{N} p_{i,j} = \frac{\sum_{i=1}^{N} Q_{i,j}(t)}{Q_j(t)} = 1\,.$$

□

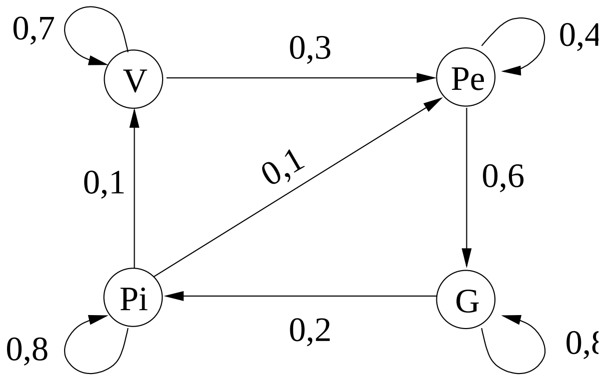
Définition 6. Une matrice carrée $P \in \mathcal{M}_n(\mathbb{R})$ qui satisfait aux conditions (1) et (2) ci-dessus s'appelle une *matrice stochastique.*

Exemple 14. On suppose que dans l'exemple 1 on a la matrice des probabilités de transition suivante :

$$P = \begin{array}{r} \swarrow \\ \text{V} \\ \text{Pe} \\ \text{G} \\ \text{Pi} \end{array} \begin{array}{c} \begin{array}{cccc} \text{V} & \text{Pe} & \text{G} & \text{Pi} \end{array} \\ \begin{pmatrix} 0,7 & 0 & 0 & 0,1 \\ 0,3 & 0,4 & 0 & 0,1 \\ 0 & 0,6 & 0,8 & 0 \\ 0 & 0 & 0,2 & 0,8 \end{pmatrix} \end{array}$$

On remarque que la somme des coefficients de chaque colonne est en effet égale à 1.

La matrice P correspond au diagramme des probabilités de transition suivant :



Ce diagramme est construit dans la même façon que le diagramme de transition, mais on ajoute à chaque flèche (=chaque transition possible) la probabilité de transition correspondante.

On suppose maintenant que au début (pour $t = 0$, càd en 2000) le paysage de la région a une distribution uniforme entre V et Pe. On décrit cette distribution par le vecteur

$$X_0 = \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix} \begin{array}{c} \mathrm{V} \\ \mathrm{Pe} \\ \mathrm{G} \\ \mathrm{Pi} \end{array} .$$

La distribution probable du paysage au temps $t = 1$ (càd en 2010) est donnée par

$$X_1 = PX_0 = \begin{pmatrix} 0,7 & 0 & 0 & 0,1 \\ 0,3 & 0,4 & 0 & 0,1 \\ 0 & 0,6 & 0,8 & 0 \\ 0 & 0 & 0,2 & 0,8 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0,35 \\ 0,35 \\ 0,30 \\ 0 \end{pmatrix} \begin{array}{c} \mathrm{V} \\ \mathrm{Pe} \\ \mathrm{G} \\ \mathrm{Pi} \end{array} .$$

Donc, en 2010, la distribution probable est :

- le 0,35% de la région est terrain cultivé V ;
- le 0,35% de la région est couvert de pelouse Pe ;
- le 0,30% de la région est couvert de garrigue G ;
- Il n'y a pas de pinède Pi dans la région.

Définition 7. Un vecteur dont les coordonnées sont non négatives avec somme égale à 1 s'appelle un *vecteur de probabilité*.

On remarque que X_0 et X_1 dans l'exemple ci-dessus sont deux vecteurs de probabilité. En effet, on a la propriété suivante.

Lemme 2. *Si X est un vecteur de probabilité $N \times 1$ et P est une matrice stochastique d'ordre N, alors PX est un vecteur de probabilité $N \times 1$.*

Soit X_t le vecteur de distribution au temps t d'un système a N états dont l'évolution est déterminée par une matrice de probabilités de transition P (qui est donc une matrice carrée d'ordre N). Alors

$$X_{t+1} = PX_t \,.$$

Comme P est une matrice associée au système et elle ne dépende pas du temps t, on trouve que l'état du système au temps t, donné par X_{t+1}, peut être complètement déterminé à partir de la connaissance de l'état du système au temps t, donné par X_t. Ceci explique la condition (3) dans la définition de chaînes de Markov.

Pour le calcul de X_t on a donc :

$t = 0$	X_0
$t = 1$	$X_1 = PX_0$
$t = 2$	$X_2 = PX_1 = P(PX_0) = P^2X_0$
$\vdots$	$\vdots$
$t = k$	$X_k = PX_{k-1} = \cdots = P^kX_0$
$\vdots$	$\vdots$

En conclusion :

Une chaîne de Markov est donnée par un système qui peut se trouver dans un nombre fini d'états $S_1, \ldots, S_N$, ensemble avec une matrice stochastique $P \in \mathcal{M}_n(\mathbb{R})$ (la matrice des probabilités de transition) qui décrit l'évolution du système en temps discret $t = 0, 1, 2, \ldots$. Plus précisément : si X_0 est le vecteur de probabilité qui donne la distribution du système entre $S_1, \ldots, S_N$ au temps $t = 0$, alors le vecteur de distribution qui donne la distribution au temps t est

$$X_t = PX_{t-1} = P^tX_0 \,.$$

Donner la matrice P est équivalent à donner son diagramme des probabilités de transition. En autre, si $k = 0, 1, 2, \ldots$, alors P^k est la matrice des probabilités de transition du système en k étapes, càd le coefficient i, j de P^k donne la probabilité que le système soit dans le $j^{\text{ème}}$ état au temps t et dans le $i^{\text{ème}}$ état au temps $t + k$ (pour chaque temps t).

Ici on a construit X_t à partir de X_0 et P. Nous pouvons même considérer le problème suivant. On suppose que on connaît la distribution du paysage de la région de l'exemple 1 au temps t, par exemple pour $t = 2$:

$$X_2 = \begin{pmatrix} 1/3 \\ 1/4 \\ 1/4 \\ 1/6 \end{pmatrix} .$$

Peut on déduire laquelle à été la distribution X_0 du paysage à l'instant initial $t = 0$? La relation entre X_0 et X_2 est $X_2 = (P^2)X_0$. Si on écrit

$$X_0 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

(vecteur inconnu), cette relation devient

$$\begin{pmatrix} 0,49 & 0 & 0,02 & 0,15 \\ 0,33 & 0,16 & 0,02 & 0,15 \\ 0,18 & 0,72 & 0,64 & 0,06 \\ 0 & 0,12 & 0,32 & 0,64 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/4 \\ 1/4 \\ 1/6 \end{pmatrix} .$$

Deux vecteurs sont égaux si et seulement si leurs composants correspondants sont égaux, donc les relations suivantes doivent être satisfaites au même temps :

$$\begin{cases} 0,49\, x_1 + 0,02\, x_2 + 0,15\, x_3 = 1/3 \\ 0,33\, x_1 + 0,16\, x_2 + 0,02\, x_3 + 0,15\, x_4 = 1/4 \\ 0,18\, x_1 + 0,72\, x_2 + 0,64\, x_3 + 0,06\, x_4 = 1/4 \\ 0,12\, x_2 + 0,32\, x_3 + 0,64\, x_4 = 1/6 \end{cases}$$

Ce qu'on a obtenu est un système de 4 équations linéaires dans les 4 inconnues x_1, x_2, x_3, x_4. La résolution des systèmes linéaires est le thème centrale du prochain chapitre.

CHAPITRE 3

Systèmes d'équations linéaires

Définition 1. On appelle *système linéaire de m équations à n inconnues* $x_1, x_2, \dots, x_n$ toute famille d'équations de la forme

$$\left\{\begin{array}{ccccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & +\cdots+ & a_{1n}\,x_n & = & b_1 \\ a_{21}\,x_1 & + & a_{22}\,x_2 & +\cdots+ & a_{2n}\,x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & \vdots & \vdots \\ a_{m1}\,x_1 & + & a_{m2}\,x_2 & +\cdots+ & a_{mn}\,x_n & = & b_m \end{array}\right. \tag{S}$$

où $a_{ij} \in \mathbb{R}$ sont appelés les *coefficients du système* et $b_j \in \mathbb{R}$ sont appelés les *coefficients du second membre*. Si les coefficients du second membre sont tous nuls ($b_j = 0$ pour tout j), on dit que le système est *homogène*. On appelle *solution* du système (S) tout n-uplet $(x_1^0, x_2^0, \dots, x_n^0)$ qui satisfait toutes les équations dans (S). *Résoudre* un système signifie trouver l'ensemble des solutions de ce système.

Exemple 1. Les solutions d'une équation linéaire $a_{11}x_1 + a_{12}x_2 = b_1$ à deux inconnues x_1, x_2 forment une droite dans le plan (x_1, x_2). Les solutions du système de 2 équations linéaires en 2 inconnues

$$\left\{\begin{array}{ccccc} a_{11}\,x_1 & + & a_{12}\,x_2 & = & b_1 \\ a_{21}\,x_1 & + & a_{22}\,x_2 & = & b_2 \end{array}\right. \tag{S'}$$

sont les points d'intersection de deux droites du plan (x_1, x_2). On sait que deux droites dans le plan peuvent avoir :

(1) exactement un point (x_1^0, x_2^0) d'intersection ;

(2) pas d'intersection ;

(3) une infinité de points d'intersection lorsque les droites sont coïncidentes.

Dans le cas (1), le système (S') a exactement une seule solution, le point (x_1^0, x_2^0) ; dans le cas (2), le système (S') n'a pas de solution ; enfin, dans le cas (3), toutes les points de la droite sont solutions du système (S').

Comme dans l'exemple 1, un système linéaire arbitraire peut admettre exactement une solution, ou bien il peut n'avoir pas de solution, ou bien il peut avoir une infinité de solutions.

On dit qu'un système linéaire est *compatible* lorsqu'il admet au moins une solution, et *incompatible* sinon.

Deux systèmes sont dit *équivalents* s'ils admettent le même ensemble de solutions (peut être l'ensemble vide s'il n'y a pas de solution).

La technique utilisée pour résoudre un système d'équations comme ce de l'exemple 1 est une technique de substitution. On fait d'abord des manipulations sur les équations pour éliminer x_1 ou x_2 de la seconde équation. Si on a éliminé pour exemple x_1, on va substituer la valeur obtenue pour x_2 dans la première équation et en fin on va résoudre une équation du type $ax_1 = b$. La méthode que on va utiliser dans le cas général est aussi une méthode de substitution après des manipulations.

On considère le système (S) donné au début du chapitre. Nous appliquons aux lignes du système une suite finie d'opérations qui transforment (S) dans un système équivalent qui est plus simple à résoudre. On notera par L_i la $i^{\text{ème}}$ ligne du système (S). Les opérations que nous pouvons appliquer sont données par la définition suivante :

Définition 2. On appelle une *opération élémentaire sur les lignes* de (S) l'une des opérations suivantes :

O_1 : Échanger deux lignes.
On écrira : $L_i \leftrightarrow L_j$ si on échange les $i^{\text{ème}}$ et la $j^{\text{ème}}$ lignes.

O_2 : Ajouter à une ligne un multiple d'un autre ligne.
On écrira : $L'_i = L_i + \gamma L_j$ si la nouvelle $i^{\text{ème}}$ ligne L'_i est obtenue en ajoutant γ fois la $j^{\text{ème}}$ ligne L_j à la $i^{\text{ème}}$ ligne L_i.

O_3 : Diviser (tous les coefficients d')une ligne par un même nombre non nul.
On écrira : $L'_i = \frac{1}{\gamma} L_i$ si la nouvelle $i^{\text{ème}}$ ligne L'_i est obtenue en divisant tous les coefficients de L_i par $\gamma \neq 0$.

Une suite opportune d'opérations élémentaires permet d'obtenir un système qui est plus simple à résoudre. Les systèmes qui sont les plus faciles à résoudre sont ceux où il y a beaucoup de zéros dans les coefficients. Les systèmes les plus faciles sont en effet ceux de la forme

$$\left\{\begin{array}{ccc} x_1 & = & b_1 \\ x_2 & = & b_2 \\ \vdots & \vdots & \vdots \\ x_n & = & b_n \end{array}\right.$$

qui sont déjà résolus. On peut écrire un tel système dans la forme

$$\left\{\begin{array}{ccccccc} 1 \cdot x_1 & + & 0 \cdot x_2 & + \cdots + & 0 \cdot x_n & = & b_1 \\ 0 \cdot x_1 & + & 1 \cdot x_2 & + \cdots + & 0 \cdot x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & \vdots & \vdots \\ 0 \cdot x_1 & + & 0 \cdot x_2 & + \cdots + & 1 \cdot x_n & = & b_n \end{array}\right.$$

Tous les coefficients sont nuls sauf ceux qui sont sur la diagonale, qui sont égaux à 1. Un système de cette façon s'appelle échelonné réduit. Le but de l'algorithme de solution qui on va donner est d'obtenir un système échelonné réduit qui est équivalent au système donné. L'opération O_1 permet de choisir l'ordre des lignes le plus convenaient ; O_2 permet de annuler des coefficients dans les colonnes ; O_3 permet d'obtenir les coefficients non nuls égaux à 1. La propriété fondamentale est la suivante :

Propriété 1. *Si on obtient un système (S') à partir de (S) par une suite finie d'opérations élémentaires sur les lignes, alors (S) et (S') sont équivalents.*

Avant de montrer l'algorithme de réduction à un système échelonné réduit, on remarque que les opérations élémentaires sont des manipulations des coefficients. Donc il n'y a pas besoin de transcrire à chaque pas toutes les inconnues, à condition que nous maintenons les coefficients dans une structure ordonnée. Plus précisément, on peut travailler avec la matrice "augmentée" associée au système donné.

La système (S) peut être écrit dans la forme matricielle

$$AX = B$$

où

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

est la matrice dont les termes sont les coefficients du système (la *matrice associée* au système),

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

est le *vecteur du second membre* du système, et

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

est le *vecteur des inconnues.*

DÉFINITION 3. La matrice augmentée associée au système (S) est la matrice obtenue de A en y ajoutant B comme $(n+1)$-ème colonne. Elle est notée $(A|B)$:

$$(A|B) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

1. Matrices échelonnées

DÉFINITION 4. On dit qu'une matrice $C = (c_{ij})_{1 \leq i \leq m\,;1 \leq j \leq p}$ est *échelonnée* lorsqu'il existe un entier $r \in \{0, 1, \ldots, m\}$ tel que

(1) Pour tout indice $i \leq r$ la $i^{\text{ème}}$ ligne de C est non nulle et pour tout indice $i > r$ la $i^{\text{ème}}$ ligne de C est nulle ;
(2) Si $d(i)$ est le plus petit indice j tel que le coefficient c_{ij} est non nul, alors la suite $d(1)$, $d(2)$, $\ldots$, $d(r)$ est strictement croissante : $d(1) < d(2) < \cdots < d(r)$

EXEMPLE 2. (1) La matrice $C = \begin{pmatrix} 0 & \boxed{2} & 0 & 1 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ est échelonnée : $r = 2$ et $d(1) = 2 < d(2) = 3$.

(2) Les matrices nulles et les matrices identité sont échelonnées.

(3) La matrice $C = \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & \boxed{2} & 1 \\ 0 & 0 & \boxed{1} & 0 \end{pmatrix}$ n'est pas échelonnée : $r = 3$ et $d(1) = 1$, $d(2) = 3$, $d(3) = 3$, mais $1 < 3 \not< 3$.

DÉFINITION 5. Les coefficients $c_{1,d(1)}, c_{2,d(2)}, \ldots, c_{r,d(r)}$ de la matrice échelonnée C sont appelés *pivots.*

Exemple 3. Soit C la matrice (non échelonnée) de l'exemple 3(3). On veut obtenir un zéro comme coefficient c_{33}, maintenant occupé par 1, en appliquant une opération élémentaire O_2. On échange d'abord L_2 et L_3 car le coefficient 1 est "plus simple" du coefficient 2.

$$C = \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \underset{L_2 \leftrightarrow L_3}{\rightsquigarrow} \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \underset{L_3' = L_3 - 2L_2}{\rightsquigarrow} \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

En général on a :

Propriété 2. *Toute matrice peut être transformée en une matrice échelonnée par une suite d'opérations O_1 et O_2.*

Définition 6. Si C' est une matrice échelonnée obtenue de C à partir d'opérations élémentaires, alors on appelle C' une *forme échelonnée* de C.

Définition 7. Le *rang* d'une matrice C, noté $\mathrm{rg}\, C$, est le nombre de pivots d'une forme échelonnée de C. Ce nombre ne dépend pas de la forme échelonnée choisie.

La propriété 2 est démontrée par l'algorithme suivant :

Algorithme de réduction d'une matrice à forme échelonnée :

On part de la matrice $m \times p$

$$C = \begin{pmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,p} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,p} \\ \vdots & \vdots & & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,p} \end{pmatrix}$$

Étape 1 :
On regarde la première colonne de C :

- Si tous les coefficients de la colonne sont nuls, alors on passe à la colonne successive.
- Sinon : on choisit i tel que $c_{i,1} \neq 0$. Par exemple, on pourrait choisir le premier $i \in \{1, \ldots, m\}$ tel que $c_{i,1} \neq 0$; un choix plus efficace pour la solution "manuelle" du système est le "plus simple" $c_{i,1}$ (on doit diviser par $c_{i,1}$).

  On échange $L_1 \leftrightarrow L_i$.

  Le coefficient $c_{1,1}$ de la nouvelle matrice est maintenant non nul. On pose pour i de 2 à m :

  $$L_i' = L_i - \frac{c_{i,1}}{c_{1,1}} L_1.$$

  De cette façon, le coefficient $(i, 1)$ de la nouvelle $i^{\text{ème}}$ ligne est :

  $$c_{i,1}' = c_{i,1} - \frac{c_{i,1}}{c_{1,1}} c_{1,1} = 0 \,,$$

Exemple 4. On considère la matrice 3×4

$$C = \begin{pmatrix} 3 & 2 & -1 & 15 \\ 2 & 3 & 1 & 10 \\ 1 & 1 & 1 & 6 \end{pmatrix}$$

Étape 1 :
On regarde la première colonne de C.

Il y a des coefficients non nuls dans la première colonne. Le coefficient $c_{i,1}$ le plus simple pour une division par $c_{i,1}$ est $c_{3,1} = 1$.

$$C \underset{L_1 \leftrightarrow L_3}{\rightsquigarrow} \begin{pmatrix} 1 & 1 & 1 & 6 \\ 2 & 3 & 1 & 10 \\ 3 & 2 & -1 & 15 \end{pmatrix}$$

Dans la nouvelle matrice on a $c_{1,1} = 1$. On pose :

$$L_2' = L_2 - 2L_1$$
$$L_3' = L_2 - 3L_1$$

d'où :

$$\underset{\substack{L_2' = L_2 - 2L_1 \\ L_3' = L_2 - 3L_1}}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & 1 & -1 & -2 \\ 0 & -1 & -4 & -3 \end{pmatrix}.$$

càd tous les coefficients de la première colonne, sauf le premier, sont nuls. On obtient donc une matrice de la forme

$$\left(\begin{array}{c|ccc} \boxed{c_{1,1}} & c_{1,2} & \cdots & c_{1,p} \\ \hline 0 & & & \\ \vdots & & C^{(1)} & \\ 0 & & & \end{array}\right)$$

Étape 2 :
On répète Étape 1 sur la matrice

$$\left(\begin{array}{c|c} 0 & \\ \vdots & C^{(1)} \\ 0 & \end{array}\right)$$

(à chaque pas, on continue à écrire la première ligne de la matrice obtenue à la fin de l'étape 1. Telle ligne restera inchangée).

Étapes successives :
On répète l'étape jusqu'on obtient une matrice échelonnée, qui est une forme échelonnée de la matrice C donnée.

On a donc obtenu une matrice de la forme

$$\left(\begin{array}{c|ccc} \boxed{1} & 1 & 1 & 6 \\ \hline 0 & 1 & -1 & -2 \\ 0 & -1 & -4 & -3 \end{array}\right).$$

Étape 2 :
On répète Étape 1 sur la matrice

$$\left(\begin{array}{c|ccc} 0 & 1 & -1 & -2 \\ 0 & -1 & -4 & -3 \end{array}\right). \qquad (*)$$

La première colonne de

$$C^{(1)} := \begin{pmatrix} 1 & -1 & -2 \\ -1 & -4 & -3 \end{pmatrix}$$

contient des coefficients non nuls. On choisit le coefficient 1 dans sa première colonne, qui est le coefficient le plus simple pour une division. Le coefficient est déjà dans la première ligne de la matrice (*). D'où une inversion de lignes n'est pas nécessaire. On applique maintenant une opération O_2 pour mettre à zéro le coefficient -1 de la première colonne de $C^{(1)}$. En écrivant toujours la première ligne (inchangée) de la matrice trouvée à la fin de l'étape 1, on a donc :

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

$$\underset{L_3'=L_3+L_2}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & -5 & -5 \end{pmatrix}$$

La dernière matrice est déjà en forme échelonnée.

La forme échelonnée de C trouvée est donc

$$\begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & \boxed{-5} & -5 \end{pmatrix}$$

2. Matrices échelonnée

DÉFINITION 8. Une matrice échelonnée et appelée *réduite* lorsque :

- ses pivots sont égaux à 1 ;
- tous les coefficients d'une colonne d'un pivot, sauf le pivot même, sont nuls.

Exemple 5. (1) La matrice

$$C = \begin{pmatrix} \boxed{1} & 2 & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \end{pmatrix}$$

est échelonnée réduite.

(2) La matrice

$$C = \begin{pmatrix} 0 & \boxed{2} & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

est échelonnée, mais pas réduite. Dans cet exemple, aucune des deux conditions pour être réduite est satisfaite, car le pivot $c_{1,2}$ est égal à 2 et $c_{1,4}$ n'est pas un pivot mais il est un coefficient non nul dans la colonne d'un pivot.

On remarque que on peut passer de C à une matrice échelonnée réduite C' au moyen d'opérations élémentaires sur les lignes :

$$C = \begin{pmatrix} 0 & \boxed{2} & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \underset{L_1'=L_1-L_3}{\rightsquigarrow} \begin{pmatrix} 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \underset{L_1'=L_1/2}{\rightsquigarrow} \begin{pmatrix} 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} =: C'$$

La propriété que chaque matrice échelonnée peut être transformée dans une matrice échelonnée réduite par des opérations élémentaires est vraie en général. Elle peut être montrée par l'algorithme suivant.

<u>Algorithme de réduction d'une matrice en forme échelonnée à une matrice en forme échelonnée réduite :</u>

L'idée de cet algorithme est d'utiliser l'opération élémentaire O_3 pour mettre les pivots égaux à 1, et d'utiliser les pivots pour annuler (au moyen de O_2) les autres coefficients des colonnes des pivots.

On part de la matrice échelonnée C avec pivots $c_{1,d(1)}$, $c_{2,d(2)}$, …, $c_{r,d(r)}$. Donc :

$$C = \left(\begin{array}{ccc|c|c} & & & c_{1,d(r)} & * \\ & ** & & \vdots & \vdots \\ & & & c_{r-1,d(r)} & \vdots \\ \hline 0 & \cdots & 0 & \boxed{c_{r,d(r)}} & * \\ \hline 0 \cdots & & & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 \cdots & & & \cdots & 0 \end{array}\right)$$

Étape 1 :
On regarde le dernier pivot $c_{r,d(r)}$.
Si $c_{r,d(r)} \neq 1$ on pose : $L_r' = \dfrac{1}{c_{r,d(r)}}\, L_r$.
On a maintenant $c_{r,d(r)} = 1$.

Exemple 6. On considère la matrice échelonnée obtenue dans l'exemple 4. Donc

$$\begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & \boxed{\text{-5}} & -5 \end{pmatrix}$$

Étape 1 :
On a $d(3) = 3$. On regarde le dernier pivot $c_{3,3} = -5$. On pose $L_3' = \dfrac{1}{-5}\, L_3$:

$$C \underset{L_3'=\frac{1}{-5}\,L_3}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

Le dernier pivot de la nouvelle matrice est maintenant $c_{3,3} = 1$.

On regarde les coefficients dans la colonne de $c_{r,d(r)}$. Pour i de 1 à $r-1$ on pose

$$L_i' = L_i - c_{i,d(r)} L_r \,.$$

De cette façon, le coefficient $(i, d(r))$ de la nouvelle $i^{\text{ème}}$ ligne est :

$$c'_{i,d(r)} = c_{i,d(r)} - c_{i,d(r)} \cdot 1 = 0 \,,$$

càd tous les coefficients de la colonne du pivot $c_{r,d(r)}$, sauf le pivot même, sont nuls. On obtient donc une matrice de la forme

$$C = \left(\begin{array}{cc|c|c|c|c}
 & \vdots & c_{1,d(r-1)} & \# & 0 & *' \\
**' & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & & c_{r-2,d(r-2)} & \vdots & \vdots & \vdots \\
\hline
0 \ \cdots & 0 & \boxed{c_{r-1,d(r-1)}} & \# & 0 & \vdots \\
\hline
0 & & \cdots & 0 & \boxed{1} & *' \\
\hline
0 \ \cdots & & & & \cdots & 0 \\
\vdots & & & & & \vdots \\
0 \ \cdots & & & & \cdots & 0
\end{array}\right)$$

Étape 2 :
On applique l'étape 1 au pivot $c_{r-1,d(r-1)}$.

Étapes successives :
On répète la procédure jusqu'on a obtenu une matrice échelonnée réduite.

On met à zéro les coefficients de la $3^{\text{ème}}$ colonne différents du pivot :

$$\begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

$$\underset{\substack{L_2'=L_2+L_3 \\ L_1'=L_1-L_3}}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 1 & 0 & 5 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

Étape 2 :
On applique l'étape 1 au pivot précédent, qui est $c_{2,2} = 1$. Comme ce pivot est déjà égal à 1, on peut passer à mettre à zéro le coefficient non nul $c_{1,2}$ de la $2^{\text{ème}}$ colonne :

$$\begin{pmatrix} \boxed{1} & 1 & 0 & 5 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

$$\underset{L_1'=L_1-L_2}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 0 & 0 & 6 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

La dernière matrice est une forme échelonnée réduite de la matrice C donnée.

De l'algorithme on déduit la propriété suivante :

PROPRIÉTÉ 3. *Toute matrice échelonnée peut être transformée en une matrice échelonnée réduite par une suite finie d'opérations élémentaires* O_2 *et* O_3.

Ainsi : Toute matrice peut être transformée en une matrice échelonnée réduite par une suite finie d'opérations élémentaires O_1, O_2 *et* O_3.

3. Méthode d'élimination de Gauss pour la résolution d'un système d'équations linéaires

La méthode d'élimination de Gauss consiste à transformer le système

$$\text{(S)}: \qquad AX = B$$

de matrice augmentée $(A|B)$ en un système équivalent (S') dont la matrice augmentée $(A|B)'$ est égale à une forme échelonnée réduite de $(A|B)$. Le solutions de (S) sont les solutions de (S'), qui peuvent être calculées facilement.

Exemple 7. (1) Le système

$$\left\{\begin{array}{rcrcrcr} 3x_1 & + & 2x_2 & - & x_3 & = & 15 \\ 2x_1 & + & 3x_2 & + & x_3 & = & 10 \\ x_1 & + & x_2 & + & x_3 & = & 6 \end{array}\right. \tag{S}$$

a matrice augmentée

$$(A|B) = \left(\begin{array}{ccc|c} 3 & 2 & -1 & 15 \\ 2 & 3 & 1 & 10 \\ 1 & 1 & 1 & 6 \end{array}\right).$$

$(A|B)$ est la matrice considérée dans les exemples 5 et 7, où on a montré que

$$C \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right) = (A|B)'.$$

Le système (S) est donc équivalent au système

$$\left\{\begin{array}{ccccc} x_1 & & & = & 6 \\ & x_2 & & = & -1 \\ & & x_3 & = & 1 \end{array}\right. . \tag{S'}$$

Donc (S) admet la solution unique $(x_1, x_2, x_3) = (6, -1, 1)$.

(2) Le système d'équations linéaires

$$\left\{\begin{array}{rcrcrcrcr} x & + & y & + & z & & & = & 6 \\ 2x & + & 3y & + & z & + & w & = & 10 \end{array}\right. \tag{S}$$

a matrice augmentée

$$(A|B) = \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 6 \\ 2 & 3 & 1 & 1 & 10 \end{array}\right).$$

On a :

$$(A|B) \underset{L_2'=L_2-2L_1}{\rightsquigarrow} \left(\begin{array}{cccc|c} \boxed{1} & 1 & 1 & 0 & 6 \\ 0 & \boxed{1} & -1 & 1 & -2 \end{array}\right) \underset{L_1'=L_1-L_2}{\rightsquigarrow} \left(\begin{array}{cccc|c} \boxed{1} & 0 & 2 & -1 & 8 \\ 0 & \boxed{1} & -1 & 1 & -2 \end{array}\right) =: (A|B)'$$

On pose

$$A' = \begin{pmatrix} \boxed{1} & 0 & 2 & -1 \\ 0 & \boxed{1} & -1 & 1 \end{pmatrix} \qquad \text{et} \qquad B' = \begin{pmatrix} 8 \\ -2 \end{pmatrix}$$

Ainsi $(A|B)' = (A'|B')$. Le système (S) est équivalent au système (S') dont la matrice augmentée est la forme échelonnée réduite $(A|B)'$ de $(A|B)$, càd

$$\left\{\begin{array}{ccccccccc} x & & & + & 2z & - & w & = & 8 \\ & & y & - & z & + & w & = & -2 \end{array}\right. \tag{S'}$$

Les inconnues z et w sont *libres*, càd associées à colonnes sans pivot. Pour résoudre (S'), on fixe arbitrairement la valeur des inconnues libres z et w et on calcule ensuite la valeur des autres inconnues en fonction de z et w. On pose donc $z = r$ et $w = s$ avec $r \in R$ et $s \in \mathbb{R}$. On obtient

$$\left\{\begin{array}{ccccccccccccc} x & = & 8 & - & 2z & - & w & = & 8 & - & 2r & - & s \\ y & = & -2 & + & z & - & w & = & -2 & + & r & - & s \\ z & = & r \\ w & = & s \end{array}\right. \qquad \text{avec } r, s \in \mathbb{R}$$

Le système admet donc une infinité de solutions dépendante de 2 paramètres réels r et s (écrit : ∞^2 solutions). Solutions particulières sont obtenues en attribuant des valeurs à r et s. Par exemple, en posant $r = 1$ et $s = 0$, on obtient la solution particulière :

$$(x, y, z, w) = (6, -1, 1, 0)\,.$$

On remarque que :

$$\begin{aligned} \text{nombre paramètres} &= \text{nombre inconnues libres} \\ &= \text{nombre inconnues} - \text{nombre pivots de } A' \end{aligned}$$

(3) Pour le système d'équations linéaires

$$\left\{\begin{array}{ccccccc} x & + & y & + & z & = & 4 \\ 2x & + & 3y & + & 2z & = & 10 \\ 3x & + & 4y & + & 3z & = & 12 \end{array}\right. \tag{S}$$

on a

$$(A|B) = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 2 & 3 & 2 & 10 \\ 3 & 4 & 3 & 12 \end{array}\right) \underset{\substack{L_2' = L_2 - 2L_1 \\ L_3' = L_3 - 3L_1}}{\rightsquigarrow} \left(\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{array}\right)$$

$$\underset{L_3' = L_3 - L_2}{\rightsquigarrow} \left(\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 4 \\ 0 & \boxed{1} & 0 & 2 \\ 0 & 0 & 0 & \boxed{-2} \end{array}\right) = (A'|B')$$

La dernière matrice est une forme échelonnée (non réduite) de $(A|B)$. Ici n'est pas nécessaire de chercher la forme échelonnée réduite : le système (S) est équivalent au système dont la matrice augmentée est $(A'|B')$. Si on écrit sa dernière ligne sous forme d'équation, on trouve :

$$0 \cdot x + 0 \cdot y + 0 \cdot y = -2\,,$$

ce qui est impossible. Par conséquent, le système n'admet pas de solutions.

Le problème d'incompatibilité du système dans l'exemple (3) vient du fait qu'il y a un pivot dans B'. La notion de rang permit de caractériser l'existence et la nature des solutions d'un système linéaire.

Soit $(A|B)$ la matrice augmentée du système $(S) : AX = B$, avec forme échelonnée $(A'|B')$. Alors A' est une forme échelonnée de A. On pose :

$$\begin{aligned} p &:= \text{rg } A &&= \text{nombre de pivots de } A \\ q &:= \text{rg } (A|B) &&= \text{nombre de pivots de } (A'|B') \end{aligned}$$

Si A est de type $m \times n$, alors A' est aussi de type $m \times n$, donc $p \leq \min\{m, n\}$. Comme B' est un vecteur colonne, on a $q \in \{p, p+1\}$.

La caractèrisation de la compatibilité/incompatibilité d'un système linéaires est donnée par le théorème suivant :

Théorème 1. *Soit $AX = B$ un système linéaire de m équations à n inconnues. Soient $p = \text{rg } A$ et $q = \text{rg } (A|B)$. Ce système*

- *n'admet pas de solutions (càd est incompatible) si $q > p$ (càd $q = p + 1$) ;*
- *admet de solutions (càd est compatible) si $q = p$, et plus précisement il admet*
  (a) *une solution unique si $q = p = n$;*
  (b) *∞^{n-p} solutions si $q = p < n$.*

4. Résolution simultanée de systèmes linéaires ayant la même matrice des coefficients

Soient

$$B_1 = \begin{pmatrix} b_{1,1} \\ \vdots \\ b_{m,1} \end{pmatrix}, \quad \ldots, \quad B_s = \begin{pmatrix} b_{1,s} \\ \vdots \\ b_{m,s} \end{pmatrix}$$

s vecteurs $m \times 1$. La résolution des s systèmes d'équations linéaires

$$AX_1 = B_1, \quad \ldots \quad, AX_s = B_s$$

(où $X_1, \ldots, X_s$ sont vecteurs d'inconnues) peut être faire simultanément en appliquant la méthode de Gauss à la matrice A augmentée de toutes les matrices B_i, càd à la matrice

$$(A|B) \qquad \text{où} \qquad B = \begin{pmatrix} b_{1,1} & \ldots & b_{1,s} \\ \vdots & & \vdots \\ b_{m,1} & \ldots & b_{m,s} \end{pmatrix}.$$

Plus précisément, la solution simultanée est obtenue en appliquent opérations élémentaires O_1, O_2 et O_3 sur les lignes de $(A|B)$ pour obtenir une matrice $(A'|B')$, où A' est une forme échelonnée réduite de A.

Exemple 8. On veut résoudre les systèmes linéaires

$$\left\{ \begin{array}{rcrcl} 2x & + & 3y & = & 1 \\ 4x & + & 3y & = & 0 \end{array} \right. \tag{S_1}$$

et

$$\left\{ \begin{array}{rcrcl} 2x & + & 3y & = & 0 \\ 4x & + & 3y & = & 1 \end{array} \right. \tag{S_2}$$

Ici

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix}, \qquad B_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

et

$$(A|B) = \left(\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 3 & 0 & 1 \end{array}\right) \underset{L_2'=L_2-2L_1}{\leadsto} \left(\begin{array}{cc|cc} \boxed{2} & 3 & 1 & 0 \\ 0 & \boxed{-3} & -2 & 1 \end{array}\right)$$

$$\underset{L_1'=L_1+L_2}{\leadsto} \left(\begin{array}{cc|cc} \boxed{2} & 0 & -1 & 1 \\ 0 & \boxed{-3} & -2 & 1 \end{array}\right) \underset{\substack{L_1'=L_1/2 \\ L_2'=-L_2/3}}{\leadsto} \left(\begin{array}{cc|cc} \boxed{1} & 0 & -1/2 & 1/2 \\ 0 & \boxed{1} & 2/3 & -1/3 \end{array}\right)$$

Les systèmes (S_1) et (S_2) sont respectivement équivalents aux systèmes de matrices augmentées

$$(A'|B_1') = \left(\begin{array}{cc|c} \boxed{1} & 0 & -1/2 \\ 0 & \boxed{1} & 2/3 \end{array}\right)$$

et

$$(A'|B_2') = \left(\begin{array}{cc|c} \boxed{1} & 0 & 1/2 \\ 0 & \boxed{1} & -1/3 \end{array}\right).$$

(S_1) a donc solution unique $(x, y) = (-1/2, 2/3)$ et (S_2) a solution unique $(x, y) = (1/2, -1/3)$.

5. Déterminants

Définition 9. Le *déterminant* d'une matrice carrée A d'ordre n est le nombre réel dét A associé à A comme suit :

- si A est une matrice *triangulaire*, càd de la forme

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

  alors

$$\text{dét}\, A := a_{1,1}a_{2,2}\cdots a_{n,n}$$

  est le produit des éléments diagonaux de A ;
- si A est arbitraire, alors

$$\text{dét}\, A = (-1)^{\ell}\, \text{dét}\, A'$$

  où :
  (a) A' est une forme échelonnée de A, obtenue de A par une suite finie d'opérations sur le lignes de type O_1 et O_2 (Rem : A' est une matrice triangulaire)
  (b) ℓ est le nombre de opérations O_1 (échanges de lignes) utilisées pour passer de A à A'.

Remarque 1. Il est important de remarquer que dét A est en effet indépendant de la forme échelonnée A' de A choisie. Cette propriété ne peut par être vue facilement de la définition donnée. Il y a des définitions de déterminant équivalentes à la définition ci-dessus qui permettent de vérifier la propriété facilement. Ces définitions équivalents ne seront pas étudiées dans ce cours.

Exemple 9. (1) La matrice $A = \begin{pmatrix} 2 & 3 \\ 1 & 3 \end{pmatrix}$ n'est pas triangulaire. On en cherche une forme échelonnée A' (qui serait une matrice triangulaire) :

$$A \underset{L_1 \leftrightarrow L_2}{\rightsquigarrow} \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} \underset{L_2' = L_2 - 2L_1}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 3 \\ 0 & \boxed{-3} \end{pmatrix} =: A'.$$

Une opération O_1 a été utilisée, d'où $\ell = 1$. Donc : dét $A = (-1)^1 \cdot 1 \cdot (-3) = 3$.

(2) On montre que le déterminant de la matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est donné par la formule

$$\text{dét} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Preuve.

(a) Si $a = 0$, alors

$$A = \begin{pmatrix} 0 & b \\ c & d \end{pmatrix} \underset{L_1 \leftrightarrow L_2}{\rightsquigarrow} \begin{pmatrix} \boxed{c} & d \\ 0 & \boxed{b} \end{pmatrix} =: A',$$

d'où dét $A = (-1)^1 bc = -bc$.

(b) Si $a \neq 0$, alors on peut diviser par a :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \underset{L_2' = L_2 - \frac{c}{a} L_1}{\rightsquigarrow} \begin{pmatrix} a & b \\ 0 & d - \dfrac{cb}{a} \end{pmatrix}$$

qui donne $\text{dét}\, A = (-1)^0 a\big(d - \dfrac{cb}{a}\big) = ad - bc.$

$\square$

(3) La formule de (2) pour la matrice $A = \begin{pmatrix} 2 & 3 \\ 1 & 3 \end{pmatrix}$ de (1) donne

$$\text{dét}\, A = 2 \cdot 3 - 3 \cdot 1 = 6 - 3 = 3\,.$$

(4) Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{pmatrix}$. On a :

$$A \underset{L_3' = L_3 - L_1}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{2} & 1 \\ 0 & -2 & -1 \end{pmatrix} \underset{L_3' = L_3 + L_2}{\rightsquigarrow} \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{2} & 1 \\ 0 & 0 & 0 \end{pmatrix} =: A'$$

d'où $\text{dét}\, A = 1 \cdot 2 \cdot 0 = 0\,.$

Remarque 2. La *règle de Sarrus* donne une méthode pratique pour le calcul du déterminant d'une matrice carrée d'ordre 3

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}.$$

La méthode consiste à ajouter les deux premières colonnes de A à la droite de celui-ci :

$$\begin{matrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \end{matrix}$$

$\text{dét}\, A$ est la somme des produits des éléments situés sur une même flèche, chaque produit étant affecté du signe indiqué.

Exemple 10. La règle de Sarrus appliquée à la matrice de l'exemple 9(4) donne

$$\text{dét}\, A = 1 \cdot 2 \cdot 2 + 2 \cdot 1 \cdot 1 + 3 \cdot 0 \cdot 0 - 3 \cdot 2 \cdot 1 - 1 \cdot 1 \cdot 0 - 2 \cdot 0 \cdot 2 = 4 + 2 - 6 = 0\,.$$

Propriété 4. *Soient $A, B \in \mathcal{M}_n(\mathbb{R})$. Alors :*

(a) *$\text{dét}(AB) = (\text{dét}\, A)(\text{dét}\, B)$;*

(b) *$\text{dét}(A^T) = \text{dét}\, A$;*

(c) *Si B résulte de l'échange de deux lignes (ou deux colonnes) de A, alors $\det B = -\det A$;*

(d) *Si tout les éléments d'une ligne (ou d'une colonne) de A sont zéros, alors $\text{dét}\, A = 0$.*

6. Matrices inversibles

DÉFINITION 10. Soit A une matrice carrée d'ordre n. S'il existe une matrice (carrée d'ordre n) D telle que

$$AD = DA = I_n\,,$$

alors A est dite *inversible*. D est la *matrice inverse* de A. En général, la matrice inverse de A est notée A^{-1}.

PROPRIÉTÉ 5. *Soit $(S) : AX = B$ un système linéaire de n équations à n inconnues –d'où A est une matrice carrée d'ordre N–. Si A est inversible, alors (S) admet une solution unique, donnée par $X = A^{-1}B$.*

Preuve. $AX = B \Rightarrow X = I_nX = (A^{-1}A)X = A^{-1}(AX) = A^{-1}B$. □

THÉORÈME 2. *Soit A une matrice carrée d'ordre n. Les propriétés suivantes sont équivalentes :*

(a) *A est inversible ;*

(b) *$\text{dét}\, A \neq 0$;*

(c) *La forme échelonnée réduite de A est I_n.*

Si A est inversible, alors sa matrice inverse A^{-1} peut être calculée par la méthode suivante : on cherche la forme échelonnée réduite de la matrice augmentée $(A|I_n)$ au moyen d'une suite finie d'opérations élémentaires sur les lignes O_1, O_2 et O_3. Telle forme échelonnée réduite est de la forme $(I_n|X)$. La matrice X à droite est A^{-1}.

Preuve. On justifie la méthode donnée pour la recherche de la matrice inverse de la matrice inversible A. Soit $B_j = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ le vecteur formé de zéros à exception de l'élément de la $j^{\text{ème}}$ ligne, qui est 1. On note par $X_1, \ldots, X_n$ les colonnes inconnues de la matrice inverse X de A. La condition $AX = I_n$ est équivalente aux n systèmes d'équations linéaires

$$AX_1 = B_1\,, \cdots, AX_n = B_n\,.$$

La matrice inconnue X est donc déterminée par la méthode de résolution simultanée de ces n systèmes ayant la même matrice des coefficients A. Voir §4. □

EXEMPLE 11. La matrice $A = \begin{pmatrix} 2 & 3 \\ 4 & 3 \end{pmatrix}$ de l'exemple 8 est inversible car $\text{dét}\, A = 2 \cdot 3 - 3 \cdot 4 = 6 - 12 = -6 \neq 0$. Dans l'exemple 8 on a montre que

$$(A|I_2) \rightsquigarrow \left(I_2 \;\middle|\; \begin{matrix} -1/2 & 1/2 \\ 2/3 & -1/3 \end{matrix} \right).$$

Donc

$$A^{-1} = \begin{pmatrix} -1/2 & 1/2 \\ 2/3 & -1/3 \end{pmatrix}.$$

Bibliographie