

Mathématiques (L2)
Année 2006-2007

Angela Pasquale

DÉPARTEMENT ET LABORATOIRE DE MATHÉMATIQUES, UNIVERSITÉ PAUL VERLAINE – METZ
E-mail address: `pasquale@math.univ-metz.fr`

Table des matières

Chapitre 1. Intégration	5
1. Propriétés de l'intégrale et méthodes de calcul	6
2. Interprétation géométrique de l'intégrale définie de f : calcul des aires	10
3. Calcul de volumes	11
4. Modèles mathématiques et équations différentielles	13
Chapitre 2. Nombres complexes	17
1. Nombres complexes et rotations	20
2. Équations de second degré à coefficients réels	20
Chapitre 3. Calcul matriciel	21
1. Matrices	22
2. Opérations sur les matrices	23
3. Chaînes de Markov	26
Chapitre 4. Systèmes d'équations linéaires	31
1. Matrices échelonnées	33
2. Matrices échelonnées réduites	35
3. Méthode d'élimination de Gauss pour la résolution de systèmes d'équations linéaires	38
Bibliographie	41

CHAPITRE 1

Intégration

Dans ce chapitre on introduit la notion d'intégrale d'une fonction. Nous ne donnera pas la définition générale et mathématiquement la plus rigoureuse, mais on donnera plutôt une définition plus pratique qui nous permettra de calculer les intégrales. Plus précisément, l'opération d'intégration d'une fonction sera introduite comme l'opération réciproque de la dérivation.

Lorsqu'une fonction décrit un système (par exemple la croissance d'une population ou la taille d'un individu dans une population), la dérivée première sert à étudier la variation du système. Mais on peut aussi réaliser l'opération contraire, c'est-à-dire retrouver la fonction décrivant le système à partir d'une mesure de la variation du système même. Cette procédure inverse, qui consiste à trouver une fonction connaissant sa dérivée, s'appelle intégration. La technique de retrouver une fonction primitive à partir de sa dérivée est fréquemment utilisée lors de l'élaboration des modèles mathématiques des systèmes.

On étudiera les propriétés des intégrales et leurs applications, notamment le calcul de l'aire d'une surface du plan, le calcul du volume de certains solides, et la résolution de certaines équations différentielles élémentaires de premier ordre. Dans la dernière section, on s'intéressera aux intégrales impropres.

DÉFINITION 1. Soit $f : I \rightarrow \mathbb{R}$ où I est un intervalle. On dit qu'une fonction $F : I \rightarrow \mathbb{R}$ est une *primitive* de f sur I lorsque

- (1) F est dérivable sur I ;
- (2) $F'(x) = f(x)$ pour tout $x \in I$.

REMARQUE 1. Si F est une primitive de f et $C \in \mathbb{R}$ est la fonction constante, alors $F + C$ est dérivable et $(F + C)' = F' + C' = F' = f$, c'est-à-dire aussi $F + C$ est primitive de f .

Réciproquement, toute primitive d'une fonction f est de la forme $F + C$ où F est une primitive fixée de f et $C \in \mathbb{R}$ est une constante arbitraire.

EXEMPLE 1. $F(x) = \frac{1}{3}x^3$ est une primitive de $f(x) = x^2$ sur \mathbb{R} car $F'(x) = (\frac{1}{3}x^3)' = x^2$. Aussi $\frac{1}{3}x^3 + 1$ est une primitive de f car $(\frac{1}{3}x^3 + 1)' = (\frac{1}{3}x^3)' + 1' = x^2$. Toute primitive de $f(x) = x^2$ est de la forme $\frac{1}{3}x^3 + C$ avec $C \in \mathbb{R}$. Plus généralement, soit $n \neq 1$: une primitive de $f(x) = x^n$ sur \mathbb{R} est $F(x) = \frac{1}{n+1}x^{n+1}$.

EXEMPLE 2. Comme $(\ln x)' = 1/x$ pour tout $x \in]0; +\infty[$, la fonction $\ln x$ est une primitive de $1/x$ sur $]0; +\infty[$. Toute primitive de $1/x$ sur $]0; +\infty[$ est de la forme $\ln x + C$ avec $C \in \mathbb{R}$. De même, on montre que toute primitive de $1/x$ sur $] -\infty, 0[\cup]0; +\infty[$ est de la forme $\ln |x| + C$ avec $C \in \mathbb{R}$.

DÉFINITION 2. On dit que f est *intégrable* si elle admet une primitive. Dans ce cas on note $\int f(x) dx$ l'une quelconque des primitives de f , définie à une constante près que l'on écrit toujours explicitement. $\int f(x) dx$ s'appelle l'*intégrale indéfinie* de f .

EXEMPLE 3. Pour tout $n \neq -1$ on a

$$\int x^n dx = \frac{1}{n+1} x^{n+1} + C, \quad C \in \mathbb{R},$$

$$\int \frac{1}{x} dx = \ln|x| + C, \quad C \in \mathbb{R}.$$

DÉFINITION 3. Soit F une primitive de f sur I et soit $[a; b]$ une partie de I . L'intégrale (définie) de f sur l'intervalle $[a; b]$, notée $\int_a^b f(x) dx$, est le nombre réel qui est égale à $F(b) - F(a)$. Donc

$$\int_a^b f(x) dx = \int_a^b F'(x) dx = F(b) - F(a).$$

La différence $F(b) - F(a)$ s'appelle la *variation de F entre a et b* ; elle est aussi notée $[F(x)]_a^b$.

EXEMPLE 4.

$$\int_0^1 x^3 dx = \left[\frac{1}{4} x^4 \right]_0^1 = \frac{1}{4} \cdot 1^4 - \frac{1}{4} \cdot 0^3 = \frac{1}{4}.$$

REMARQUE 2. L'intégrale $\int_a^b f(x) dx$ est indépendant du choix de la primitive de f . Par exemple, aussi $F(x) = \frac{1}{4}x^4 + 2$ est une primitive de $f(x) = x^3$, et $[F(x)]_0^1 = \left(\frac{1}{4} \cdot 1^4 + 2\right) - \left(\frac{1}{4} \cdot 0^3 + 2\right) = \frac{1}{4}$.

DÉFINITION 4. Si $a < b$, on définit

$$\int_b^a f(x) dx = - \int_a^b f(x) dx.$$

1. Propriétés de l'intégrale et méthodes de calcul

Les règles d'intégration se déduisent des règles de dérivation, dont elles sont les opérations inverses.

- (1) L'intégrale de la somme de deux fonctions est la somme de leurs intégrales, c'est-à-dire pour $f, g : [a; b] \rightarrow \mathbb{R}$ on a

$$\int_a^b [f(x) + g(x)] dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

REMARQUE 3. Si $F' = f$ et $G' = g$, alors $(F + G)' = F' + G'$. Ceci justifie la règle ci-dessus.

EXEMPLE 5. On veut calculer $\int_0^\pi (x + \sin x) dx$. Ici $[a; b] = [0; \pi]$ et $f(x) = x$, $g(x) = \sin x$. D'après la propriété ci-dessus on a

$$\begin{aligned} \int_0^\pi (x + \sin x) dx &= \int_0^\pi x dx + \int_0^\pi \sin x dx \\ &= \left[\frac{x^2}{2} \right]_0^\pi + [-\cos x]_0^\pi \\ &= \frac{\pi^2}{2} + (-\cos \pi) - (-\cos 0) \\ &= \frac{\pi^2}{2} + 2. \end{aligned}$$

- (2) L'intégrale du produit d'une fonction par une constante est le produit de la constante par l'intégrale de la fonction, c'est-à-dire pour $f : [a; b] \rightarrow \mathbb{R}$ et $C \in \mathbb{R}$ on a

$$\int_a^b [Cf(x)] dx = C \int_a^b f(x) dx.$$

REMARQUE 4. Cette règle est due à la propriété $(CF)' = C \cdot F'$ pour la primitive F de f .

EXEMPLE 6.

$$\int_0^1 (3x) dx = 3 \int_0^1 x dx = 3 \cdot \left[\frac{1}{2}x^2 \right]_0^1 = \frac{3}{2}.$$

- (3) Si $f(x) \geq 0$ pour tout $x \in [a; b]$, alors

$$\int_a^b f(x) dx \geq 0.$$

- (4) (Formule de Chasles) Si $f : [a; b] \rightarrow \mathbb{R}$ et $c \in [a; b]$, alors

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

REMARQUE 5. Si $F' = f$, alors $\int_a^b f(x) dx = F(b) - F(a)$ et $\int_a^c f(x) dx + \int_c^b f(x) dx = (F(c) - F(a)) + (F(b) - F(c)) = F(b) - F(a)$.

- (5) (Intégration par parties) Pour $f, g : [a; b] \rightarrow \mathbb{R}$ on a

$$\int_a^b [f(x)g'(x)] dx = [f(x)g(x)]_a^b - \int_a^b [f'(x)g(x)] dx.$$

REMARQUE 6. La règle d'intégration par parties se déduit de la règle de Leibniz $(fg)' = f'g + fg'$, qui donne avec (1) :

$$\begin{aligned} [f(x)g(x)]_a^b &= \int_a^b (f(x)g(x))' dx \\ &= \int_a^b (f'(x)g(x) + f(x)g'(x)) dx \\ &= \int_a^b f'(x)g(x) dx + \int_a^b f(x)g'(x) dx. \end{aligned}$$

EXEMPLE 7. Calculer $\int_1^e \ln x dx$.

Comme $\ln x = \ln x \cdot 1$ et $x' = 1$, on peut utiliser (5) avec

$$\begin{aligned} f(x) &= \ln x & f'(x) &= \frac{1}{x} \\ g'(x) &= 1 & g(x) &= x. \end{aligned}$$

On obtient :

$$\begin{aligned}
 \int_1^e \ln x \, dx &= \int_1^e 1 \cdot \ln x \, dx \\
 &= \int_1^e (x)' \cdot \ln x \, dx \\
 &= [x \cdot \ln x]_1^e - \int_1^e x \cdot (\ln x)' \, dx \\
 &= e \cdot \ln e - 1 \cdot \ln 1 - \int_1^e 1 \, dx \\
 &= e - 0 - [x]_1^e \\
 &= e - (e - 1) = 1
 \end{aligned}$$

EXEMPLE 8. (Intégration par parties itérée) Pour calculer $\int_0^\pi e^x \cos x \, dx$ en intégrant par parties, on pose

$$\begin{aligned}
 f(x) &= \cos x & f'(x) &= -\sin x \\
 g'(x) &= e^x & g(x) &= e^x
 \end{aligned}$$

(on pourra remarquer qu'ici ce n'est pas important laquelle des deux fonctions on va choisir comme f ou g). Donc

$$\begin{aligned}
 \int_0^\pi e^x \cos x \, dx &= [e^x \cos x]_0^\pi + \int_0^\pi e^x \sin x \, dx \\
 &= (e^\pi \cos \pi) - (e^0 \cos 0) + \int_0^\pi e^x \sin x \, dx \\
 &= -e^\pi - 1 + \int_0^\pi e^x \sin x \, dx.
 \end{aligned}$$

Pour évaluer $\int_0^\pi e^x \sin x \, dx$, on applique nouvellement l'intégration par parties avec

$$\begin{aligned}
 f(x) &= \sin x & f'(x) &= \cos x \\
 g'(x) &= e^x & g(x) &= e^x
 \end{aligned}$$

d'où

$$\begin{aligned}
 \int_0^\pi e^x \sin x \, dx &= [e^x \sin x]_0^\pi - \int_0^\pi e^x \cos x \, dx \\
 &= - \int_0^\pi e^x \cos x \, dx.
 \end{aligned}$$

On déduit donc que

$$\int_0^\pi e^x \cos x \, dx = -e^\pi - 1 - \int_0^\pi e^x \cos x \, dx,$$

d'où

$$\int_0^\pi e^x \cos x \, dx = -\frac{e^\pi + 1}{2}.$$

(6) (Changement de variable)

$$\int_a^b g(f(x))f'(x) \, dx = \int_{f(a)}^{f(b)} g(y) \, dy.$$

REMARQUE 7. Soit G une primitive de g . La dérivée de la fonction composée donne

$$[G(f(x))] = G'(f(x)) \cdot f(x) = g(f(x)) \cdot f(x),$$

d'où :

$$\begin{aligned} \int_a^b g(f(x)) \cdot f(x) dx &= \int_a^b G'(f(x)) \cdot f(x) dx \\ &= \int_a^b [G(f(x))] dx \\ &= [G(f(x))]_a^b \\ &= [G(y)]_{f(a)}^{f(b)} \\ &= \int_{f(a)}^{f(b)} g(y) dy. \end{aligned}$$

REMARQUE 8 (Utilisation de la notation différentielle). Soit $y = f(x)$ une fonction. La *différentielle* de f est $dy = \frac{dy}{dx} dx = \frac{df}{dx} dx = f'(x) dx$. La notation différentielle peut être utilisée pour se rappeler la règle de changement de variables :

$$\begin{aligned} \int_a^b \underbrace{g(f(x))}_y \underbrace{f'(x) dx}_{dy} &= \int_{f(a)}^{f(b)} g(y) dy \\ &\downarrow \\ &\begin{cases} x = a \mapsto y = f(a) \\ x = b \mapsto y = f(b) \end{cases} \end{aligned}$$

EXEMPLE 9. On détermine $\int_0^2 x e^{x^2} dx$ en utilisant (6). On remarque que $e^{x^2} = g(f(x))$ avec $g(y) = e^y$ et $f(x) = x^2$. On a $f'(x) = 2x$; en outre, $f(0) = 0$ et $f(2) = 4$. Donc :

$$\begin{aligned} \int_0^2 x e^{x^2} dx &= \frac{1}{2} \int_0^2 e^{x^2} \cdot (2x) dx \\ &= \frac{1}{2} \int_0^4 e^y dy \\ &= \frac{1}{2} [e^y]_0^4 \\ &= \frac{1}{2} (e^4 - 1). \end{aligned}$$

EXEMPLE 10. Calculer $\int_e^{e^2} \frac{1}{x \ln x} dx$. Comme $\frac{1}{x} = (\ln x)'$, on pose $y = \ln x$. Alors

$$\begin{aligned} \int_e^{e^2} \frac{1}{x \ln x} dx &= \int_1^2 \frac{1}{y} dy \\ &\downarrow \\ &\begin{cases} dy = (\ln x)' dx = \frac{1}{x} dx \\ x = e \mapsto y = \ln e = 1 \\ x = e^2 \mapsto y = \ln e^2 = 2 \end{cases} \\ &= [\ln y]_1^2 = \ln 2. \end{aligned}$$

2. Interprétation géométrique de l'intégrale définie de f : calcul des aires

Si $f(x) \geq 0$ pour tout $x \in [a, b]$, alors $\int_a^b f(x) dx$ est l'aire de la surface limitée par le graphe de f , l'axe des abscisses et les droites $x = a$ et $x = b$. Voir figure 1.

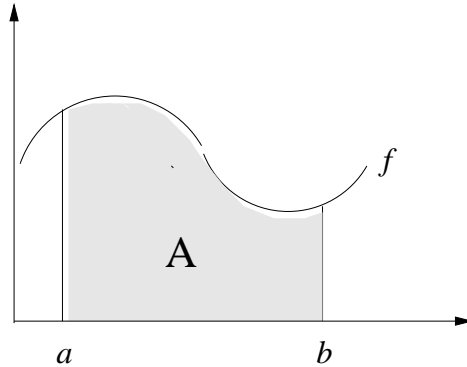


FIG. 1. $\int_a^b f(x) dx = \text{aire de la surface A}$

Si f prend des valeurs négatives, l'aire est affectée du signe moins sur les intervalles où $f < 0$.

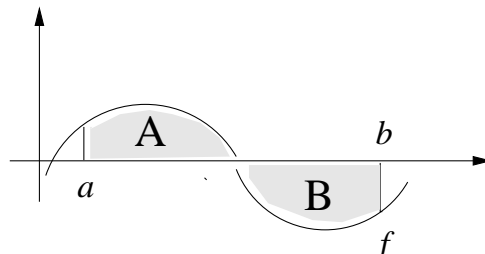


FIG. 2. $\int_a^b f(x) dx = (\text{aire de la surface A}) - (\text{aire de la surface B})$

2.1. Calcul des aires des surfaces du plan.

EXEMPLE 11. Calculer l'aire de la surface entre les paraboles $y = x^2$ et $y = 2x - x^2$.

Les points d'intersection des deux courbes sont donnés par les solutions du système $\begin{cases} y = x^2 \\ y = 2x - x^2 \end{cases}$, qui sont $O(0, 0)$ et $P(1, 1)$.

On a :

$$\begin{aligned}
 \text{aire}(A) &= \text{aire}(B) - \text{aire}(C) \\
 &= \int_0^1 (2x - x^2) dx - \int_0^1 x^2 dx \\
 &= \int_0^1 (2x - 2x^2) dx \\
 &= 2 \int_0^1 (x - x^2) dx \\
 &= 2 \left\{ \left[\frac{1}{2}x^2 \right]_0^1 - \left[\frac{1}{3}x^3 \right]_0^1 \right\} \\
 &= 2(1/2 - 1/3) = 1/3.
 \end{aligned}$$

3. Calcul de volumes

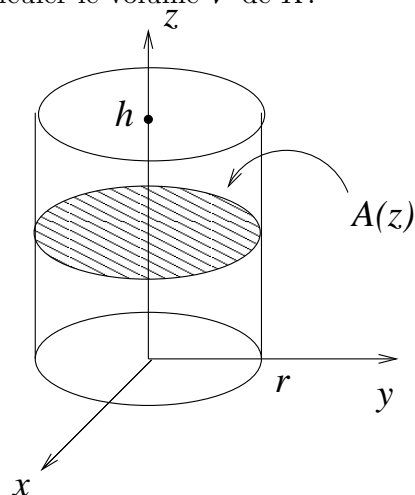
Pour des solides réguliers, comme par exemple le cube, la sphère ou le cylindre droit, on connaît des formules géométriques pour calculer leur volume. Pour calculer le volume de certains solides qui sont moins réguliers on peut utiliser des techniques d'intégration.

Soient $a < b$ et soit K un solide de \mathbb{R}^3 limité par les plans $z = a$ et $z = b$ tel que l'aire de la section de cote z est une fonction intégrable $A(z)$. Le volume V du solide K est alors

$$V = \int_a^b A(z) dz.$$

REMARQUE 9. En général, l'aire $A(z)$ est soit calculée au moyen de formules géométriques, soit elle aussi est le résultat du calcul d'une intégrale.

EXEMPLE 12. Soit K un cylindre droit de hauteur h et ayant pour base un disque de rayon r . Calculer le volume V de K .



$A(z) = \pi r^2$ est indépendant de z , d'où
 $V = \int_0^h \pi r^2 dz = \pi r^2 \int_0^h dz = \pi r^2 h$,
ce qui donne la formule usuelle.

EXEMPLE 13. Calculer le volume V d'une sphère de rayon r .

La section de cote z est un disque de rayon $\sqrt{r^2 - z^2}$ dont l'aire est $A(z) = \pi(r^2 - z^2)$. Donc le volume de la sphère est

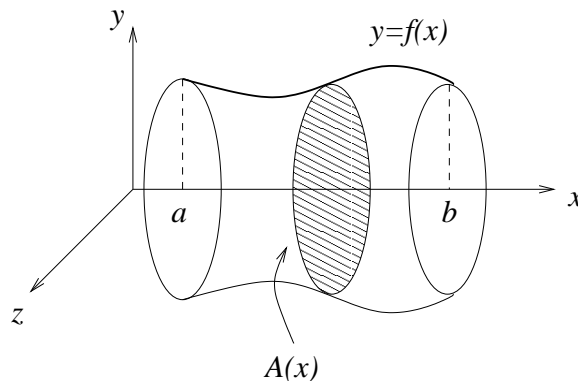
$$\begin{aligned}
V &= \int_{-r}^r A(z) dz = \pi \int_{-r}^r (r^2 - z^2) dz \\
&= \pi \int_{-r}^r r^2 dz - \pi \int_{-r}^r z^2 dz \\
&= \pi r^2 [z]_{-r}^r - \pi \left[\frac{1}{3} z^3 \right]_{-r}^r \\
&= 2\pi r^3 - \frac{1}{3} \pi r^3 \\
&= \frac{4}{3} \pi r^3
\end{aligned}$$

3.1. Le volume d'un solide de révolution. Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction telle que $f(x) \geq 0$ pour tout x . Le volume V du solide engendré par la rotation autour de l'axe des abscisses de la surface limitée par celui-ci, la courbe d'équation $y = f(x)$ et les droites d'équation $x = a$ et $x = b$, est donné par la formule

$$V = \pi \int_a^b [f(x)]^2 dx.$$

En effet, $A(x) = \pi[f(x)]^2$, d'où

$$V = \int_a^b A(x) dx = \pi \int_a^b [f(x)]^2 dx.$$



EXEMPLE 14. Calculer le volume V du solide engendré par la rotation autour de l'axe des abscisses de la surface limitée par celui-ci et la courbe d'équation $y = x^2$ pour $0 \leq x \leq 2$.

On a

$$V = \pi \int_0^2 x^4 dx = \pi \left[\frac{1}{5} x^5 \right]_0^2 = \frac{32}{5} \pi.$$

EXEMPLE 15. Calculer le volume V du solide de révolution engendré par la rotation autour de l'axe des ordonnées de la surface limitée par les courbes d'équation $y = \sqrt{x}$ et $y = x$.

Les points d'intersection entre les deux courbes sont $O(0, 0)$ et $P(1, 1)$. Le volume cherché V peut être calculé comme $V = V_1 - V_2$ où V_1 est le volume du cône engendré par la rotation autour de l'axe des ordonnées de la surface limitée par la courbe $y = x$ pour $0 \leq y \leq 1$, et V_2 est le volume

du solide engendré par la rotation autour de l'axe des ordonnées de la surface limitée par la courbe $y = \sqrt{x}$ pour $0 \leq y \leq 1$.

Pour le calcul de V_2 , on observe que l'aire de la section de cote y est πy^2 , le rayon étant $x = y^2$ (obtenu en résolvant $y = \sqrt{x}$ pour une valeur positive de x).

Donc :

$$V_1 = \pi \int_0^1 y^2 dy = \frac{\pi}{3},$$

$$V_2 = \pi \int_0^1 y^4 dy = \frac{\pi}{5},$$

et

$$V = \frac{\pi}{3} - \frac{\pi}{5} = \frac{2\pi}{15}.$$

4. Modèles mathématiques et équations différentielles

Un modèle est une description d'un phénomène (p.ex. biologique) qui permet d'expliquer le phénomène même et d'en prévoir certains aspects. La description du phénomène est précisée par des variables. Des exemples importants de modèles sont ceux qui donnent l'évolution d'une population dans le temps. Ici une des variables les plus importantes est le nombre d'individus $N(t)$ au temps t .

EXEMPLE 16 (La loi de croissance exponentielle). Dans ce modèle, l'hypothèse de base est : pendant un petit intervalle Δt , la variation ΔN d'individus est proportionnelle à N et à Δt , c'est-à-dire

$$\Delta N = kN\Delta t \quad \text{ou} \quad \frac{\Delta N}{\Delta t} = kN,$$

où k est une constante.

Si $k > 0$, alors la population est croissante ; si $k < 0$, alors elle est décroissante ; si $k = 0$, elle reste constante.

Si l'on suppose que Δt peut prendre des valeurs arbitrairement petites, on obtient

$$\frac{dN}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\Delta N}{\Delta t} = kN.$$

La variable N est donc la solution de l'équation

$$N'(t) = kN(t).$$

Une équation comme $N'(t) = kN(t)$, qui donne un lien entre N et sa dérivée N' s'appelle une *équation différentielle de 1er ordre*.

4.1. Résolution d'une équation différentielle de 1er ordre à variables séparables.

Une *équation différentielle de 1er ordre à variables séparables* est une équation différentielle de la forme $\frac{dx}{dt} = f(t)g(x)$ où $f(t)$ et $g(x)$ sont des fonctions données et la fonction inconnue est $x = x(t)$.

L'équation différentielle de l'exemple 16 est de ce type, avec $x = N$, $f(t) = k$ et $g(x) = x$.

En supposant $g(x) \neq 0$, on divise par $g(x)$ et on obtient

$$\frac{1}{g(x)} x'(t) = f(t).$$

Si $x = x(t)$ est une solution, alors

$$\frac{1}{g(x(t))} x'(t) = f(t),$$

d'où

$$\int \frac{1}{g(x(t))} x'(t) dt = \int f(t) dt.$$

On pose $y = x(t)$. Alors $dy = x'(t) dt$ et la formule de changement de variables donne

$$\int g(y) dy = \int f(t) dt.$$

Il faut donc résoudre ces intégrales indéfinies.

EXEMPLE 17. On veut résoudre l'équation différentielle de la loi de croissance exponentielle $N'(t) = kN(t)$ du nombre d'individus $N(t)$ d'une population au temps t . En supposant que $N(t) > 0$, on divise par $N(t)$ et on obtient

$$\frac{N'(t)}{N(t)} = k$$

d'où

$$\int \frac{N'(t)}{N(t)} dt = \int k dt.$$

On pose $y = N(t)$. On a $dy = N'(t) dt$ et

$$\int \frac{1}{y} dy = \int k dt.$$

Comme y représente le nombre d'individus, on peut supposer que $y > 0$. En intégrant, on déduit

$$\ln y + C_1 = kt + C_2$$

où C_1 et C_2 sont des constantes réelles arbitraires. Donc aussi $C_3 := C_2 - C_1$ est une constante réelle arbitraire. En prenant l'exponentiel de chaque côté de la dernière équation, on a

$$y = e^{kt+C_3},$$

c'est-à-dire

$$N(t) = Ce^{kt}$$

où $C := e^{C_3}$ est une constante positive arbitraire. L'équation $N'(t) = kN(t)$ correspond donc à la croissance exponentielle $N(t) = Ce^{kt}$.

EXEMPLE 18. On suppose que le volume $V(t)$ d'une cellule au temps t change selon la loi $V'(t) = \sin t$. Déterminer $V(t)$ si le volume au temps $t = 0$ est $V(0) = 3$ unités de volume.

L'équation différentielle donnée peut être résolue directement par $V(t) = -\cos t + C$ avec $C \in \mathbb{R}$. La condition initiale $V(0) = 3$ nous permet de déterminer la valeur de la constante C . En effet $3 = V(0) = -\cos 0 + C = -1 + C$ donne $C = 4$. Donc $V(t) = -\cos t + 4$.

EXEMPLE 19. On suppose maintenant que le volume $V(t)$ de la cellule au temps t change selon la loi $V'(t) = \sin t[V(t)]^2$, et on cherche $V(t)$ si le volume au temps $t = 0$ est $V(0) = 3$ unités de volume.

On remarque que $V(t) \equiv 0$ est une solution de l'équation différentielle donnée, mais elle ne satisfait pas la condition initiale donnée. En supposant $V(t) \neq 0$, on a

$$\frac{V'(t)}{[V(t)]^2} = \sin t$$

d'où

$$\int \frac{V'(t)}{[V(t)]^2} dt = \int \sin t dt.$$

On pose $y = V(t)$. La formule de changement de variables entraîne

$$\int \frac{1}{y^2} dy = \int \sin t dt.$$

c'est-à-dire

$$\frac{1}{y} = \cos t + C \quad C \in \mathbb{R},$$

d'où la solution générale

$$V(t) = y = \frac{1}{\cos t + C}, \quad C \in \mathbb{R}.$$

La valeur de la constante C est fixée par la condition initiale :

$$3 = V(0) = \frac{1}{\cos 0 + C} = \frac{1}{1 + C}$$

qui entraîne $C = -2/3$. En conclusion,

$$V(t) = y = \frac{1}{\cos t - 2/3}.$$

CHAPITRE 2

Nombres complexes

C'est bien connu que l'équation $x^2 + 1 = 0$ n'a pas de solutions dans l'ensemble des nombres réels, car une solution x devrait satisfaire $x^2 = -1$, c'est-à-dire son carré doit au même temps être positive (car $x^2 \geq 0$ pour tout $x \in \mathbb{R}$) et négative (égale à -1). Le problème est que l'ensemble des nombres réels est "trop petit". Ceci est un problème qu'on connaît déjà en cherchant des solutions entières : il n'y a pas de nombre entier x qui est solution de l'équation $x^2 - 3 = 0$, car 3 n'est pas le carré d'un nombre $x \in \mathbb{Z}$. L'équation $x^2 - 3 = 0$ admet les solutions $\pm\sqrt{3}$ si on "enlarge" l'ensemble des nombres entiers aux nombres réels. L'ensemble des nombres où l'équation $x^2 + 1 = 0$ (ainsi que chaque équation de la forme $ax^2 + bx + c = 0$) admet 2 solutions est l'ensemble des nombres complexes. Les nombres complexes constituent à la fois un outil mathématique puissant et une théorie mathématique importante. Une application remarquable est que ces nombres permettent de décrire une rotation du plan comme une simple multiplication.

DÉFINITION 5. Un *nombre complexe* est un nombre de la forme $z = a + bi$ où a et b sont des réels quelconques et i est un symbole tel que $i^2 = -1$. Le nombre a est appelé la *partie réelle* de z , notée $\operatorname{Re} z$; le nombre b est appelé la *partie imaginaire* de z , notée $\operatorname{Im} z$. La forme $a + ib$ s'appelle la *forme algébrique* du nombre complexe z . Si $b = 0$, alors $z = a$ est un nombre réel; si $a = 0$, alors $z = bi$ s'appelle un *nombre imaginaire pur*.

On note par \mathbb{C} l'ensemble des nombres complexes.

DÉFINITION 6 (Égalité de deux nombres complexes). Deux nombres complexes $z = a + bi$ et $w = c + di$ sont égaux lorsque $a = c$ et $b = d$.

Tout nombre complexe peut être représenté comme un point du *plan complexe*, c'est-à-dire un plan muni d'un repère orthonormé. Le point $z = a + bi$ correspond au point d'abscisse a et ordonnée b . Les points de l'axe des abscisses sont donc les nombres réels et ceux de l'axe des ordonnées les nombres imaginaires purs.

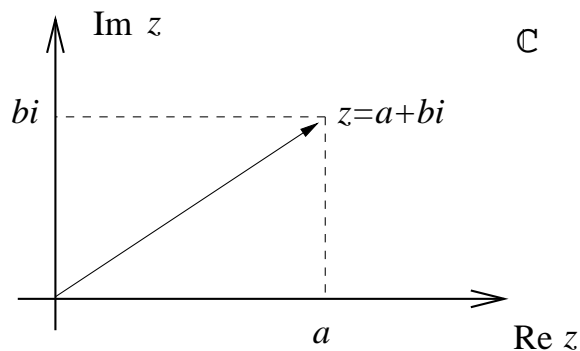


FIG. 1. Représentation du point $z = a + bi$ dans le plan complexe

Opérations entre nombres complexes

DÉFINITION 7. La *somme* de $z = a + bi \in \mathbb{C}$ et $w = c + di \in \mathbb{C}$ est le nombre complexe $z + w = (a + c) + (b + d)i$ obtenu en sommant les parties réelles et les parties imaginaires de z et w .

EXEMPLE 20. $(3 + i) + (\sqrt{2} - 2i) = (3 + \sqrt{2}) - i$.

REMARQUE 10. Si on représente z et w sur le plan complexe, alors $z + w$ est la somme des vecteurs z et w par la règle du parallélogramme.

DÉFINITION 8. Le *produit* de $z = a + bi \in \mathbb{C}$ et $w = c + di \in \mathbb{C}$ est le nombre complexe zw obtenu en appliquant les lois distributives, associatives et commutatives usuelles des nombres réels ainsi que la règle $i^2 = -1$: donc

$$zw = (a + bi)(c + id) = a(c + id) + bi(c + id) = ac + adi + bci + bd(i^2) = (ac - bd) + (ad + bc)i.$$

EXEMPLE 21. $(3 + i)(\sqrt{2} - 2i) = 3(\sqrt{2} - 2i) + i(\sqrt{2} - 2i) = (3\sqrt{2} + 2) + (\sqrt{2} - 6)i$.

DÉFINITION 9. Le *conjugué* de $z = a + bi \in \mathbb{C}$ est le nombre complexe $\bar{z} = a - bi$.

REMARQUE 11. \bar{z} est le symétrique de z par rapport à l'axe des abscisses.

DÉFINITION 10. Le *module* de $z = a + bi \in \mathbb{C}$ est le nombre réel non-négatif $|z| = \sqrt{a^2 + b^2}$.

REMARQUE 12. $|z|$ est la distance entre z et l'origine O du repère.

EXEMPLE 22. Le module de $z = 3 + 2i$ est $|3 + 2i| = \sqrt{9 + 4} = \sqrt{13}$.

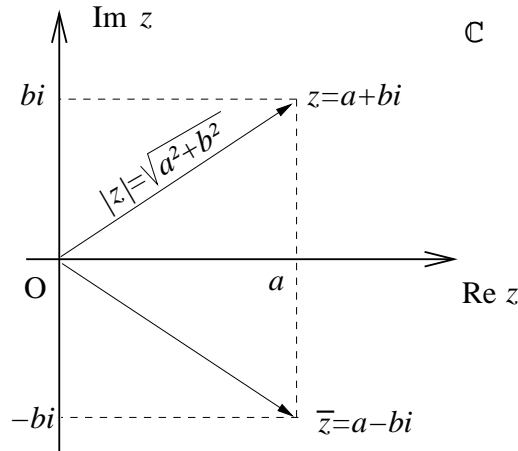


FIG. 2. Le conjugué et le module du point $z = a + bi$

Propriétés du conjugué : Pour tout $z = a + bi \in \mathbb{C}$ et $w \in \mathbb{C}$ on a :

- (a) $z + \bar{z} = 2 \operatorname{Re} z$ et $z - \bar{z} = 2i \operatorname{Im} z$
- (b) $\overline{z + w} = \bar{z} + \bar{w}$;
- (c) $\overline{z\bar{w}} = \bar{z}; \bar{\bar{w}}$;
- (d) $z\bar{z} = (a + ib)(a - ib) = a^2 + b^2 = |z|^2$.

REMARQUE 13. On peut diviser deux nombres complexes comme dans l'exemple suivant :

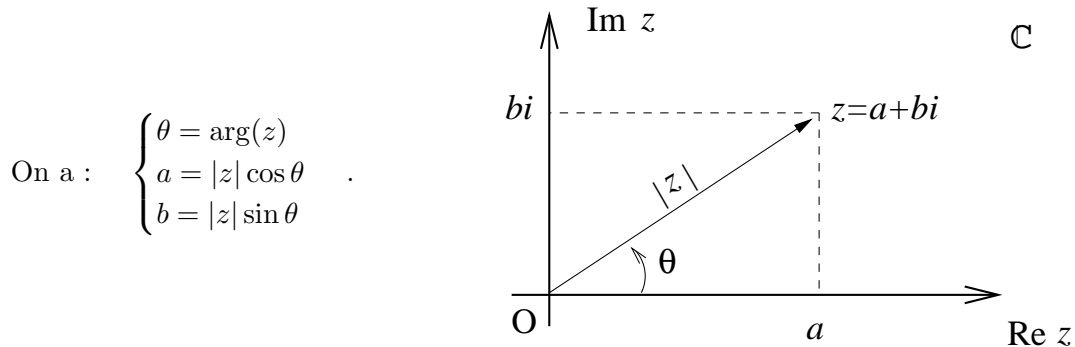
$$\frac{3+2i}{1+4i} = \frac{(3+2i)(1-4i)}{(1+4i)(1-4i)} = \frac{11-10i}{1+16} = \frac{11}{17} - \frac{10}{17}i.$$

En général, si $w \neq 0$, alors

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} = \frac{z\bar{w}}{|w|^2}.$$

En particulier, $\frac{1}{w} = \frac{\bar{w}}{|w|^2}$. On remarque aussi que $\overline{\frac{z}{w}} = \frac{\bar{z}}{\bar{w}}$.

DÉFINITION 11. L'argument $\arg(z)$ de $z \in \mathbb{C} \setminus \{0\}$ est toute mesure θ (en radians) de l'angle entre z et l'axe des abscisses (voir figure). Donc $\arg(z)$ est défini à $2k\pi$ près, avec $k \in \mathbb{Z}$.



On pose $e^{i\theta} = \cos \theta + i \sin \theta$. La forme trigonométrique du nombre $z = a + bi$ est

$$z := r(\cos \theta + i \sin \theta) = re^{i\theta}$$

où $r = |z|$ et $\theta = \arg(z)$.

REMARQUE 14. (a) On a $e^{i\theta}e^{i\sigma} = (\cos \theta + i \sin \theta)(\cos \sigma + i \sin \sigma) = (\cos \theta \cos \sigma - \sin \theta \sin \sigma) + (\sin \theta \cos \sigma + \cos \theta \sin \sigma)i = \cos(\theta + \sigma) + \sin(\theta + \sigma)i = e^{i(\theta + \sigma)}$.

(b) Si $z = re^{i\theta}$ et $w = se^{i\sigma}$, alors

$$zw = re^{i(\theta + \sigma)} \quad \text{et} \quad \frac{1}{z} = \frac{1}{r}e^{-i\theta}.$$

En outre, pour tout $n \in \mathbb{N}$ on a la formule de Moivre

$$z^n = r^n e^{in\theta}.$$

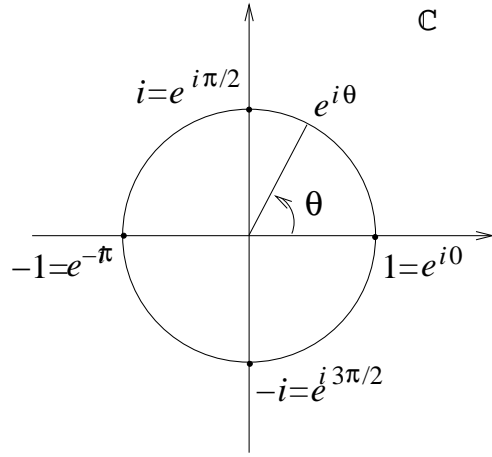
EXEMPLE 23. (1) La forme trigonométrique de $z = 1 + i$ est $z = \sqrt{2}e^{i\pi/4}$ car $\arg(z) = \frac{\pi}{4}$ et $|z| = \sqrt{2}$.

(2) La forme algébrique de $z = 2e^{i\pi/6}$ est $z = 2 \cos(\pi/6) + 2 \sin(\pi/6)i = \sqrt{3} + i$.

(3) Pour $z = 2e^{i\pi/6}$ et $w = 3e^{i\pi/4}$ on a $zw = 6e^{i(\pi/6 + \pi/4)} = 6e^{i\pi/2} = 6i$ et $\frac{z}{w} = \frac{2}{3}e^{i(\pi/3 - \pi/4)} = \frac{2}{3}e^{i\pi/12}$.

REMARQUE 15. .

Soit $z \in \mathbb{C}$. Alors $|z| = 1$ si et seulement si $z = e^{i\theta}$ avec $\theta \in \mathbb{R}$.



REMARQUE 16. On a $e^{i\theta} = \cos \theta + \sin \theta i$ et $e^{-i\theta} = \cos \theta - \sin \theta i = \overline{e^{i\theta}}$, d'où les formules d'Euler

$$\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \text{et} \quad \sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

1. Nombres complexes et rotations

Soient $z = a + bi = re^{i\theta} \in \mathbb{C}$ et $\alpha \in [0, 2\pi[$. Le produit $e^{i\alpha}z = re^{i(\theta+\alpha)}$ donne la rotation de centre 0 du point z d'un angle α en sens antihoraire.

EXEMPLE 24. La multiplication par $i = e^{i\pi/2}$ correspond à une rotation d'angle $\pi/2$

EXEMPLE 25. Soit r la rotation de centre 0 et angle $5\pi/6$ et soit P le point de coordonnées $(\sqrt{3}, 1)$. Déterminer $Q := r(P)$.

On identifie P avec $z := \sqrt{3} + i \in \mathbb{C}$. On a $|z| = 2$ et, si $\theta = \arg(z)$, alors $\cos \theta = \sqrt{3}/2$ et $\sin \theta = 1/2$, d'où $\theta = \pi/6$. Comme $e^{i5\pi/6} \cdot 2e^{i\pi/6} = -2$, on obtient $Q = (-2, 0)$.

2. Équations de second degré à coefficients réels

On veut résoudre dans \mathbb{C} une équation de la forme

$$ax^2 + bx + c = 0 \quad \text{avec} \quad a, b, c \in \mathbb{R}. \quad (*)$$

On remarque que $(\pm i)^2 = i^2 = -1$, d'où $\sqrt{-1} = \pm i$. Si $\alpha \geq 0$, alors $\sqrt{-\alpha} = \sqrt{-1}\sqrt{\alpha} = \pm i\sqrt{\alpha}$.

Les solutions de (*) sont donc :

- Si $\Delta > 0$, on a les 2 solutions réelles distinctes

$$z_1 = \frac{-b + \sqrt{\Delta}}{2a}, \quad z_2 = \frac{-b - \sqrt{\Delta}}{2a}.$$

- Si $\Delta = 0$, on a les 2 solutions réelles coïncidentes

$$z_1 = z_2 = \frac{-b}{2a}.$$

- Si $\Delta < 0$, on a les 2 solutions complexes conjugués

$$z_1 = \frac{-b + i\sqrt{\Delta}}{2a}, \quad z_2 = \frac{-b - i\sqrt{\Delta}}{2a}.$$

EXEMPLE 26. Les solutions dans \mathbb{C} de l'équation $z^2 + z + 1 = 0$ sont

$$\frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1 \pm \sqrt{-3}}{2} = \frac{-1 \pm i\sqrt{3}}{2}.$$

CHAPITRE 3

Calcul matriciel

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

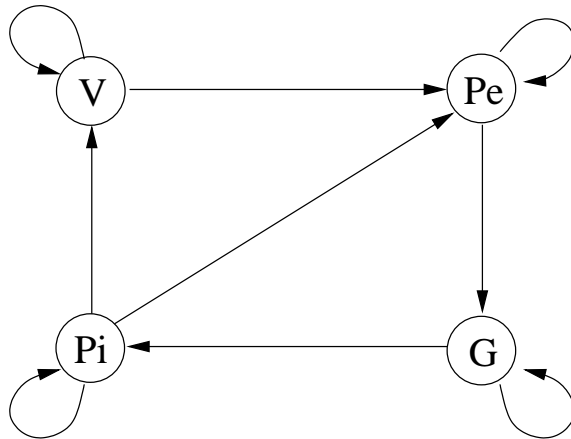
EXEMPLE 27. Nous étudions un modèle simplifié pour l'évolution du paysage d'une région du midi de la France. Le paysage de la région peut évoluer dans l'un des quatre états suivants :

- Ⓟ Terrain cultivé : vignes, vergers, ...
- Ⓟ Pelouse
- Ⓟ Garrigue
- Ⓟ Pinède

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

- étant Ⓟ, il peut le rester ou peut se transformer par abandon en Ⓟ ;
- étant Ⓟ, il peut le rester ou peut se reconstruire en Ⓟ ;
- étant Ⓟ, il peut le rester ou peut se reconstruire en Ⓟ ;
- étant Ⓟ, il peut le rester ou peut être transformé en Ⓟ ou peut se transformer en Ⓟ 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 majuscule et les chiffres qui 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 \textcircled{G} à l'état \textcircled{Pe} , mais cette transition est possible en 2 étapes :

$$\textcircled{G} \rightarrow \textcircled{Pi} \rightarrow \textcircled{Pe}$$

ou bien en 3 étapes, par exemple

$$\textcircled{G} \rightarrow \textcircled{Pi} \rightarrow \textcircled{V} \rightarrow \textcircled{Pe}$$

ou

$$\textcircled{G} \rightarrow \textcircled{Pi} \rightarrow \textcircled{Pi} \rightarrow \textcircled{Pe}$$

En général, dénombrer toutes 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 12. Une *matrice* (réelle) $m \times n$ est un tableau à m lignes et n colonnes d'éléments de \mathbb{R} . On note

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

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 28. $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 6 \\ 3 & 1 & 1 & 3 \end{pmatrix}$ 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 d'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 13. 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 29. $v = \begin{pmatrix} 1 \\ 0 \\ \sqrt{2} \end{pmatrix}$ est un vecteur de \mathbb{R}^3 .

DÉFINITION 14. 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 30. 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 31. $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

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 :

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

EXEMPLE 32. 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 17. (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) $0_{m,n}$ est l'*élément neutre* pour l'addition dans $\mathcal{M}_{m,n}(\mathbb{R})$, c'est-à-dire $A + 0_{m,n} = A$ 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 33. Si $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix}$, alors $-A = \begin{pmatrix} -1 & -2 & -3 \\ 0 & 1 & 2 \end{pmatrix}$

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 $\textcircled{\text{G}}$ à $\textcircled{\text{Pe}}$ 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'à présent nous avons seulement considéré l'évolution d'un système discret dans l'hypothèse où le système est dans l'un d'un nombre fini 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 s'est 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 toutes les transitions qui peuvent avoir lieu dans deux périodes, dans trois périodes, ou, en général, dans k périodes, où k est un entier naturel arbitraire.

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

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

DÉFINITION 15. Une *chaîne de Markov (en temps discret)* est un processus stochastique particulier dans lequel :

- (1) Le système est observé en temps discret $t = 0, 1, 2, \dots$
- (2) Le système est réparti 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 39. 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 : $\textcircled{\text{V}}$, $\textcircled{\text{Pe}}$, $\textcircled{\text{G}}$, $\textcircled{\text{Pi}}$. On étudie l'évolution du système pour des valeurs discrètes du temps, tous les 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 l'on connaît pour chaque instant t considéré et pour chaque paire $i, j \in \{1, 2, 3, 4\}$ 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^{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

$$Q_j(t) := \text{proportion du système qui peut être dans le } j^{\text{ème}} \text{ état au temps } t ;$$

$$Q_{i,j}(t) := \text{proportion du système qui peut être dans le } j^{\text{ème}} \text{ état au temps } t \text{ et dans le } i^{\text{ème}} \text{ état au temps } t + 1 ;$$

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 revient 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 est une matrice carrée d'ordre 4 dans l'exemple).

REMARQUE 23. 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 16. La matrice P s'appelle la *matrice des probabilités de transition*. Dans un système avec N états possibles, c'est une matrice carrée d'ordre N .

REMARQUE 24. 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 $i, j \in \{1, 2, \dots, N\}$ on a $0 \leq p_{i,j} \leq 1$;
- (2) Pour tout $j \in \{1, 2, \dots, N\}$ 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 $t + 1$, c'est-à-dire $\sum_{i=1}^N Q_{i,j}(t) = Q_j(t)$, d'où

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

□

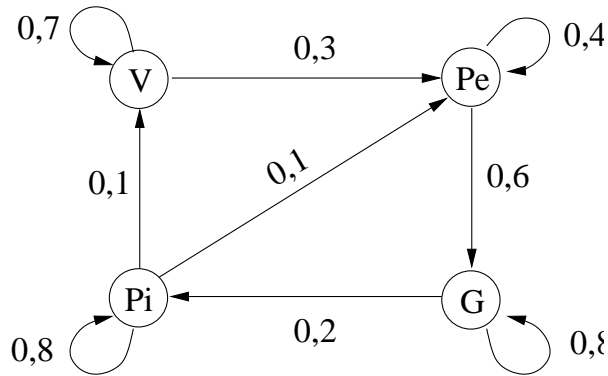
DÉFINITION 17. Une matrice carrée $P \in \mathcal{M}_n(\mathbb{R})$ qui satisfait aux conditions (1) et (2) ci-dessus s'appelle une *matrice stochastique*.

EXEMPLE 40. On suppose que dans l'exemple 1 on a la matrice des probabilités de transition suivante :

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

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 de 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 qu'au début (pour $t = 0$, c'est-à-dire en 2000) le paysage de la région a une répartition uniforme entre \textcircled{V} et \textcircled{Pe} . On décrit cette distribution par le vecteur

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

La distribution probable du paysage au temps $t = 1$ (c'est-à-dire en 2010) est donnée par

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

Donc, en 2010, la distribution probable est :

- 35% de la région est un terrain cultivé \textcircled{V} ;
- 35% de la région est couvert de pelouse \textcircled{Pe} ;
- 30% de la région est couvert de garrigue \textcircled{G} ;
- Il n'y a pas de pinède \textcircled{Pi} dans la région.

DÉFINITION 18. 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 4. 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 à 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 ne dépend 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.

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

Pour le calcul de X_t on a donc :

En conclusion :

Une chaîne de Markov est donnée par un système qui peut se trouver dans un nombre fini d'états S_1, \dots, S_N , muni d'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 $t = 0, 1, 2, \dots$. Plus précisément : si X_0 est le vecteur de probabilité qui donne la distribution du système entre S_1, \dots, 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 outre, si $k = 0, 1, 2, \dots$, alors P^k est la matrice des probabilités de transition du système en k étapes, c'est-à-dire 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 en déduire quelle a 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 coordonnées correspondantes sont égales, 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 à 4 inconnues x_1, x_2, x_3, x_4 . La résolution des systèmes linéaires est le thème central du prochain chapitre.

Systemes d'equations lineaires

DÉFINITION 1. On appelle *systeme lineaire de m equations à n inconnues* x_1, x_2, \dots, x_n toute famille d'equations de la forme

$$\begin{cases} a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1 \\ a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n = b_2 \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n = b_m \end{cases} \quad (\text{S})$$

où les nombres $a_{ij} \in \mathbb{R}$ sont appelés les *coefficients du systeme* 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 systeme est *homogene*. On appelle *solution* du systeme (S) tout n -uplet $(x_1^0, x_2^0, \dots, x_n^0)$ qui satisfait toutes les equations dans (S). *Resoudre* un systeme signifie trouver l'ensemble des solutions de ce systeme.

EXEMPLE 1. Les solutions d'une equation lineaire $a_{11}x_1 + a_{12}x_2 = b_1$ à deux inconnues x_1, x_2 forment une droite dans le plan (x_1, x_2) . Les solutions du systeme de 2 equations lineaires en 2 inconnues

$$\begin{cases} a_{11} x_1 + a_{12} x_2 = b_1 \\ a_{21} x_1 + a_{22} x_2 = b_2 \end{cases} \quad (\text{S}')$$

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 infinite de points d'intersection lorsque les droites sont confondues.

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

Comme dans l'exemple 1, un systeme lineaire arbitraire peut admettre exactement une solution, ou bien il peut n'avoir pas de solution, ou bien il peut avoir une infinite de solutions.

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

Deux systemes sont dit *equivalents* 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 systeme d'equations comme celui de l'exemple 1 est une technique de substitution. On fait d'abord des manipulations sur les equations pour éliminer x_1 ou x_2 de la seconde equation. Si on a éliminé par exemple x_1 , on va substituer la valeur obtenue pour x_2 dans la première equation et enfin on va résoudre une equation du type $ax_1 = b$. La méthode que l'on va utiliser dans le cas général est aussi une méthode de substitution après des manipulations.

On considère le systeme (S) donné au début du chapitre. Nous appliquons aux lignes du systeme une suite finie d'opérations qui transforment (S) en un systeme 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 dans 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 la $i^{\text{ème}}$ et la $j^{\text{ème}}$ ligne.

O_2 : Ajouter à une ligne un multiple d'une 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

$$\begin{cases} x_1 = b_1 \\ x_2 = b_2 \\ \vdots \\ x_n = b_n \end{cases}$$

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

$$\begin{cases} 1 \cdot x_1 + 0 \cdot x_2 + \cdots + 0 \cdot x_n = b_1 \\ 0 \cdot x_1 + 1 \cdot x_2 + \cdots + 0 \cdot x_n = b_2 \\ \vdots \\ 0 \cdot x_1 + 0 \cdot x_2 + \cdots + 1 \cdot x_n = b_n \end{cases}$$

Tous les coefficients sont nuls sauf ceux qui sont sur la diagonale, qui sont égaux à 1. Un tel système s'appelle échelonné réduit. Le but de l'algorithme 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 favorable ; O_2 permet d'annuler des coefficients dans les colonnes ; O_3 permet d'obtenir les coefficients non nuls égaux à 1. La propriété fondamentale est la suivante :

PROPOSITION 5. 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 maintenions 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 sous la forme matricielle

$$AX = B$$

où

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

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 $(A|B)$:

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

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 $r \in \{0, 1, \dots, m\}$ 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 $d(1), d(2), \dots, d(r)$ est strictement croissante : $d(1) < d(2) < \dots < d(r)$

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 $c_{1,d(1)}, c_{2,d(2)}, \dots, c_{r,d(r)}$ 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" que le coefficient 2.

$$C = \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & \textcircled{1} & 0 \end{pmatrix} \xrightarrow{L_2 \leftrightarrow L_3} \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & \textcircled{1} & 0 \\ 0 & 0 & 2 & 1 \end{pmatrix} \xrightarrow{L'_3 = L_3 - 2L_2} \begin{pmatrix} \boxed{3} & 1 & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

En général on a :

PROPOSITION 6. 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 $\text{rg } C$, 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 suivante.
- Sinon : on choisit i tel que $c_{i,1} \neq 0$. Par exemple, on pourrait choisir le premier $i \in \{1, \dots, m\}$ 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 $c_{3,1} = 1$.

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

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

$$L'_2 = L_2 - 2L_1$$

$$L'_3 = L_3 - 3L_1$$

d'où :

$$\begin{matrix} L'_2 = L_2 - 2L_1 \\ L'_3 = L_3 - 3L_1 \end{matrix} \xrightarrow{\quad} \begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & 1 & -1 & -2 \\ 0 & -1 & -4 & -3 \end{pmatrix}.$$

c'est-à-dire tous les coefficients de la première colonne, sauf le premier, sont nuls. On obtient donc une matrice de la forme

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

Étape 2 :

On répète l'Étape 1 sur la matrice

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

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

Étapes suivantes :

On répète l'étape jusqu'à obtenir 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 l'Étape 1 sur la matrice

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

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 :

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

$$\xrightarrow{L'_3=L_3+L_2} \left(\begin{array}{c|ccc} \boxed{1} & 1 & 1 & 6 \\ \hline 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & -5 & -5 \end{array} \right)$$

La dernière matrice a déjà une forme échelonnée.

La forme échelonnée de C trouvée est donc

$$\left(\begin{array}{c|ccc} \boxed{1} & 1 & 1 & 6 \\ \hline 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & \boxed{-5} & -5 \end{array} \right)$$

2. Matrices échelonnées réduites

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

- ses pivots sont égaux à 1 ;
- tous les coefficients d'une colonne d'un pivot, excepté le pivot, 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 n'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 l'on peut passer de C à une matrice échelonnée réduite C' au moyen d'opérations élémentaires sur les lignes :

$$C = \begin{pmatrix} 0 & \boxed{2} & 0 & 1 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{L'_1=L_1-L_3} \begin{pmatrix} 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{L'_1=L_1/2} \begin{pmatrix} 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} =: C'$$

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.

Algorithme de réduction d'une matrice en forme échelonnée à une matrice en forme échelonnée réduite :

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 $c_{1,d(1)}, c_{2,d(2)}, \dots, c_{r,d(r)}$. Donc :

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

Étape 1 :

On regarde le dernier pivot $c_{r,d(r)}$.

Si $c_{r,d(r)} \neq 1$ on pose : $L'_r = \frac{1}{c_{r,d(r)}} L_r$.

On a maintenant $c_{r,d(r)} = 1$.

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

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

Étape 1 :

On a $d(3) = 3$. On regarde le dernier pivot $c_{3,3} = -5$. On pose $L'_3 = \frac{1}{-5} L_3$:

$$C \xrightarrow{L'_3 = \frac{1}{-5} L_3} \begin{pmatrix} \boxed{1} & 1 & 1 & 6 \\ 0 & \boxed{1} & -1 & -2 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

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 :

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

c'est-à-dire tous les coefficients de la colonne du pivot $c_{r,d(r)}$, sauf le pivot, sont nuls. On obtient donc une matrice de la forme

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

Étape 2 :

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

Étapes suivantes :

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

De l'algorithme on déduit la propriété suivante :

PROPOSITION 7. *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 .

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

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

$$\begin{array}{l} \rightsquigarrow \\ L'_2 = L_2 + L_3 \\ L'_1 = L_1 - L_3 \end{array} \begin{pmatrix} \boxed{1} & 1 & 0 & 5 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

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 annuler le coefficient non nul $c_{1,2}$ de la 2^{ème} colonne :

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

$$\begin{array}{l} \rightsquigarrow \\ L'_1 = L_1 - L_2 \end{array} \begin{pmatrix} \boxed{1} & 0 & 0 & 6 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 1 \end{pmatrix}$$

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

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

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

$$(S) : \quad AX = B$$

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

EXEMPLE 7. (1) Pour le système

$$\begin{cases} 3x_1 + 2x_2 - x_3 = 15 \\ 2x_1 + 3x_2 + x_3 = 10 \\ x_1 + x_2 + x_3 = 6 \end{cases} \quad (S)$$

on obtient la 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 l'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

$$\begin{cases} x_1 & & & = & 6 \\ & x_2 & & = & -1 \\ & & x_3 & = & 1 \end{cases} \quad (S')$$

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

(2) Le système d'équations linéaires

$$\begin{cases} x + y + z & = & 6 \\ 2x + 3y + z + w & = & 10 \end{cases} \quad (S)$$

a la matrice augmentée

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

On a :

$$(A|B) \xrightarrow{L'_2=L_2-2L_1} \left(\begin{array}{cccc|c} \boxed{1} & 1 & 1 & 0 & 6 \\ 0 & \boxed{1} & -1 & 1 & -2 \end{array} \right) \xrightarrow{L'_1=L_1-L_2} \left(\begin{array}{cccc|c} \boxed{1} & 0 & 2 & -1 & 8 \\ 0 & \boxed{1} & -1 & 1 & -2 \end{array} \right) =: (A|B)'$$

On pose

$$A' = \left(\begin{array}{cccc} \boxed{1} & 0 & 2 & -1 \\ 0 & \boxed{1} & -1 & 1 \end{array} \right) \quad \text{et} \quad B' = \begin{pmatrix} 8 \\ -2 \end{pmatrix}$$

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 $(A|B)$, c'est-à-dire

$$\begin{cases} x & + & 2z & - & w & = & 8 \\ & y & - & z & + & w & = & -2 \end{cases} \quad (S')$$

Les inconnues z et w sont *libres*, c'est-à-dire associées à des 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 $r \in \mathbb{R}$ et $s \in \mathbb{R}$. On obtient

$$\begin{cases} x = 8 - 2z - w = 8 - 2r - s \\ y = -2 + z - w = -2 + r - s \\ z = r \\ w = s \end{cases} \quad \text{avec } r, s \in \mathbb{R}$$

Le système admet donc une infinité de solutions dépendantes de 2 paramètres réels r et s (on écrit : ∞^2 solutions). Les 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

$$\begin{cases} x + y + z = 4 \\ 2x + 3y + 2z = 10 \\ 3x + 4y + 3z = 12 \end{cases} \quad (\text{S})$$

on a

$$\begin{aligned} (A|B) &= \left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 2 & 3 & 2 & 10 \\ 3 & 4 & 3 & 12 \end{array} \right) \xrightarrow[\substack{L'_2=L_2-2L_1 \\ L'_3=L_3-3L_1}]{\sim} \left(\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 \end{array} \right) \\ &\xrightarrow[\substack{L'_3=L_3-L_2}]{\sim} \left(\begin{array}{ccc|c} \boxed{1} & 1 & 1 & 4 \\ 0 & \boxed{1} & 0 & 2 \\ 0 & 0 & 0 & \boxed{-2} \end{array} \right) = (A'|B') \end{aligned}$$

La dernière matrice est une forme échelonnée (non réduite) de $(A|B)$. Ici il 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 permet de caractériser l'existence et la nature des solutions d'un système linéaire.

Soit $(A|B)$ la matrice augmentée du système $(S) : AX = B$, de 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 caractérisation 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 (c'est-à-dire est incompatible) si $q > p$ (c'est-à-dire $q = p + 1$);*
- *admet de solutions (c'est-à-dire est compatible) si $q = p$, et plus précisément il admet*
 - (a) *une solution unique si $q = p = n$;*
 - (b) *∞^{n-p} solutions si $q = p < n$.*

Bibliographie

- [BB] J.-P. et F. Bertrandias, *Mathématiques pour les sciences de la vie, de la nature et de la santé*, Presses universitaires de Grenoble, 1997
- [Li] S. Lipschutz, *Algèbre linéaire - Cours et problèmes* (2nd /'ed.). Série Schaum, McGraw-Hill Inc. (eds.), 1994, 453 pages.
- [GRL] R. N. Greenwell, N. P. Ritchey et M. L. Lial, *Calculus for the Life Sciences*, Addison-Wesley, 2003, 880 pages.
- [Str] G. Strang, *Linear algebra and its applications*, Academic Press, 1988.