Algèbre et géométrie S1
Résolution de systèmes linéaires

Joël Riou

15 septembre 2024

Table des matières

Bibliographie

COLIN, Jean-Jacques et Jean-Marie MORVAN. *Géométrie affine et euclidienne.* (516.071 COL géo). Cépaduès, 2017.

LAY, David C. *Algèbre linéaire et applications.* (512.5 LAY alg). Pearson, 2012.

LIPSCHUTZ, Seymour et Marc LIPSON. *Algèbre linéaire.* (512.5 LIP alg). Dunod, 2003.

MONIER, Jean-Marie. *Géométrie MPSI.* (516.071 MON géo). Dunod, 2006.

I. Points du plan et de l'espace

Sommaire

R désigne l'ensemble des nombres réels.

Pour tout entier naturel $n \in \mathbf{N}$, on notera $\mathbf{R}^n$ l'ensemble des n-uplets $(x_1, \ldots, x_n)$ de nombres réels.

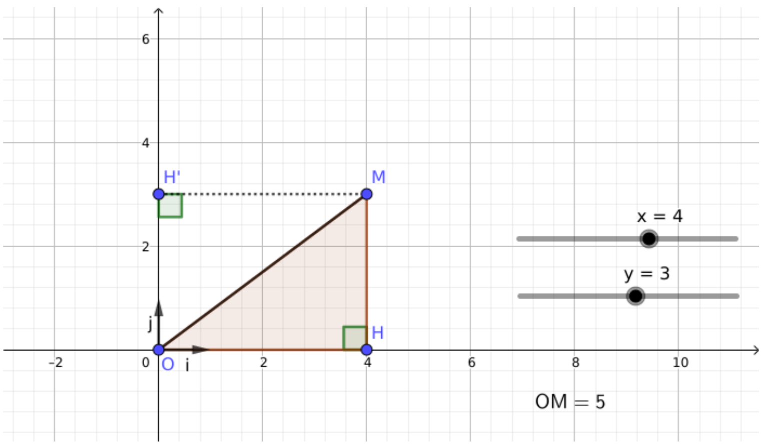
Nous allons ici nous concentrer sur les cas $n = 2$ et $n = 3$, c'est-à-dire $\mathbf{R}^2$ et $\mathbf{R}^3$.

I.1. Plan

I.1.1. Coordonnées cartésiennes

Par convention, on utilisera les lettres x et y pour désigner les éléments de $\mathbf{R}^2$ comme des couples (x, y).

On peut *représenter* les éléments de $\mathbf{R}^2$ dans le plan (du tableau, de l'écran, d'une feuille, etc...) en utilisant un repère $(O, \vec{\imath}, \vec{\jmath})$ orthonormal :



https://www.geogebra.org/m/hhgrp3uu

Le point O est l'origine du repère. Que le repère soit orthonormal signifie que $\vec{\imath}$ et $\vec{\jmath}$ sont deux vecteurs orthogonaux et de longueur 1.

On a une correspondance entre les points du plan et les éléments de $\mathbf{R}^2$. Plus précisément, le point M correspond au couple $(x, y) \in \mathbf{R}^2$ si et seulement si la relation vectorielle suivante est vérifiée :

$$\overrightarrow{OM} = x\,\vec{\imath} + y\,\vec{\jmath}\,.$$

Si cette identité est vérifiée, on dit que $(x, y) \in \mathbf{R}^2$ sont les coordonnées du point M.

La coordonnée x se lit sur la figure en *projetant* orthogonalement le point M sur l'axe (Ox) (la droite passant par O et dirigée par le vecteur $\vec{\imath}$), ce qui donne le point H. La distance OH est toujours positive (ou nulle) et égale à x ou $-x$ suivant le position de H sur l'axe (Ox) : plus précisément, OH est la valeur absolue de x, ce qui est noté $OH = |x|$.

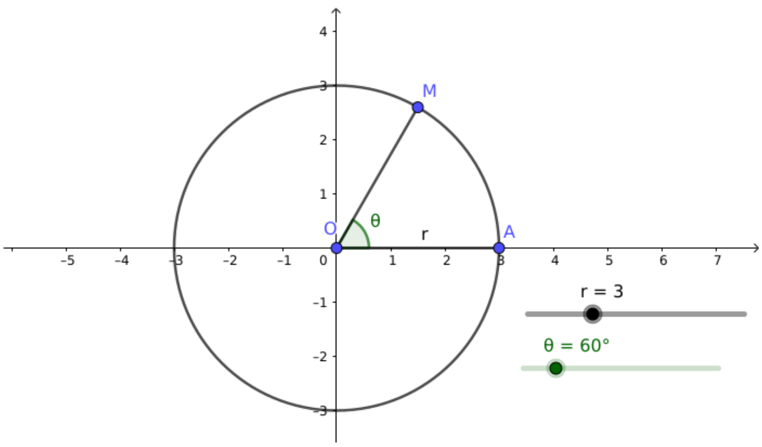
De même, en projetant M sur l'axe (Oy) (passant par O et dirigé par le vecteur $\vec{\jmath}$), on obtient le point H' et la distance OH' est donnée par la formule $OH' = |y|$. On dit que y est l'*ordonnée* du point M.

Il faut bien comprendre la distinction qu'il y a entre des points du plan pensés de façon géométrique et des couples (x, y) de réels qui relèvent d'une pensée plus algébrique. Néanmoins, par abus de langage, on identifiera parfois un point M avec le couple (x, y) qui lui correspond, et un des enjeux de ce cours est de faire des allers-retours entre des problèmes géométriques (intersections de droites, etc.) et des problèmes algébriques (résolution de systèmes d'équations, etc.).

(On peut remarquer que le théorème de Pythagore appliqué au triangle OHM se traduit par l'identité $OM^2 = OH^2 + HM^2$, c'est-à-dire que $OM^2 = x^2 + y^2$. Autrement dit, la distance OM est égale à $\sqrt{x^2 + y^2}$. Les questions métriques (ou de distances) ne seront pas développées dans ce cours.)

I.1.2. Coordonnées polaires

Un autre point de vue (qui ne sera pas développé dans ce cours) consiste à décrire les points du plan en coordonnées polaires. Un point M est décrit en termes de la distance $r := OM \in \mathbf{R}_+$ et de l'angle orienté θ qui permet de caractériser la position du point M sur le cercle de centre O et de rayon r. (Précisément, si A est le point de coordonnées cartésiennes $(r, 0)$, alors θ est l'angle orienté $(\overrightarrow{OA}, \overrightarrow{OM})$. On remarque que pour que θ soit bien défini, il faut que $M \neq O$.)



https://www.geogebra.org/m/qpnfpscu

(Sur la figure, les angles sont écrits en *degrés*, un « tour complet » correspondant à $360°$. En mathématiques, il est plus conventionnel de « mesurer » les angles en radians : $180° = \pi$.)

Le point M peut être décrit par ses coordonnées polaires (r, θ) ou par ses coordonnées cartésiennes $(x, y) \in \mathbf{R}^2$. Pour faire le lien entre les deux, le plus simple est d'exprimer x et y

en fonction de r et θ grâce aux fonctions trigonométriques :

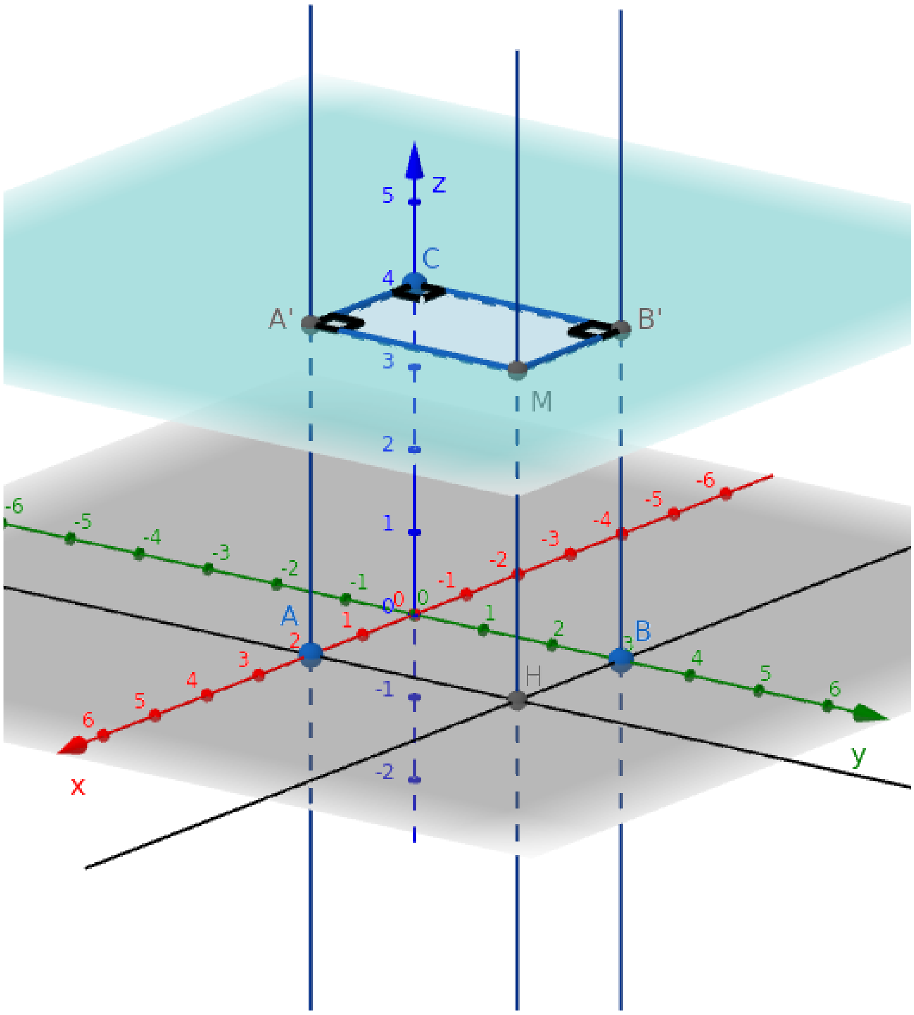
$$\left\{\begin{array}{lll} x & = & r\cos\theta \\ y & = & r\sin\theta \end{array}\right.$$

De façon similaire, un nombre complexe $z \in \mathbf{C}$ peut être écrit sous la forme $z = x + iy$ avec $(x, y) \in \mathbf{R}^2$ ou bien sous la forme $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$.

I.2. Espace

I.2.1. Coordonnées cartésiennes

L'essentiel de ce qui a été dit dans le paragraphe I.1.1 vaut dans l'espace, si ce n'est qu'au lieu d'avoir deux coordonnées x et y, on a trois coordonnées x, y et z. Autrement dit, on s'intéresse à des triplets $(x, y, z) \in \mathbf{R}^3$, que l'on peut représenter dans l'espace :



https://www.geogebra.org/m/c2hd5z7d

Sur cette figure, les axes (Ox), (Oy) et (Oz) sont respectivement en rouge, vert et bleu.

On a représenté un point M de coordonnées $(x, y, z) \in \mathbf{R}^3$, ce qui signifie que

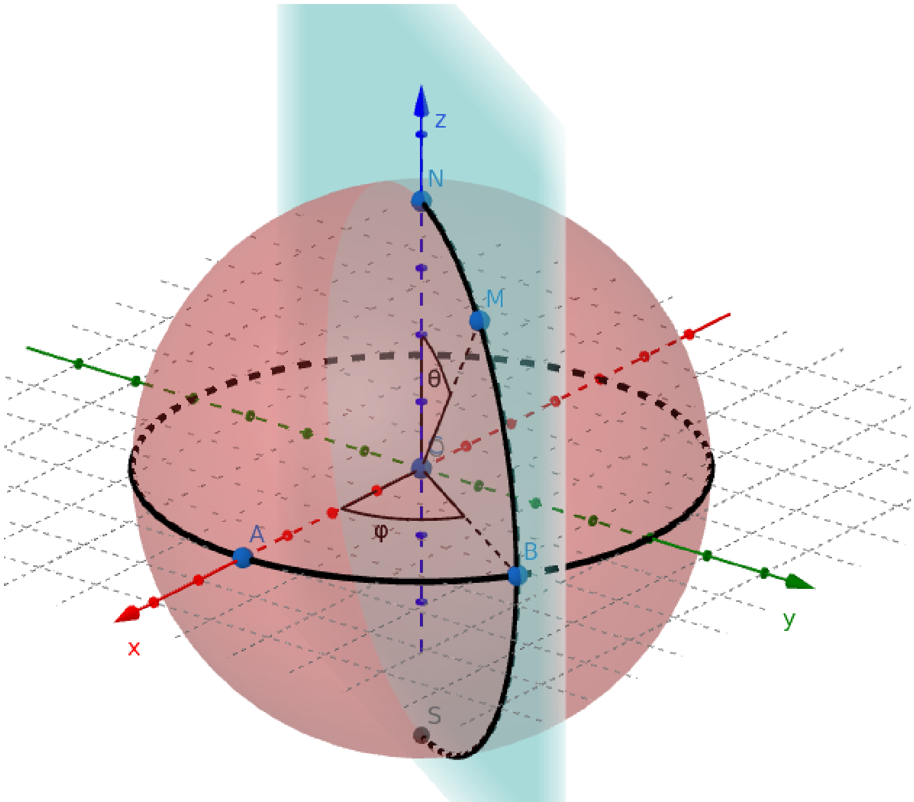
$$\overrightarrow{OM} = x\,\vec{\imath} + y\,\vec{\jmath} + z\vec{k}\,,$$

où O est l'origine, et $\vec{\imath}$, $\vec{\jmath}$ et $\vec{k}$ sont trois vecteurs (de longeur 1, et non représentés sur la figure ci-dessus) dirigeant les trois axes (Ox), (Oy) et (Oz).

Sur la figure, quelques autres points sont représentés, voici leurs coordonnées : $A(x, 0, 0)$, $B(0, y, 0)$, $H(x, y, 0)$. L'origine O et les points A, B et H sont situés dans un plan que l'on peut appeler le plan horizontal passant par l'origine. Un plan parallèle à ce plan a été représenté, et il contient les points $C(0, 0, z)$, $A'(x, 0, z)$, $B'(0, y, z)$ et $M(x, y, z)$. (On remarque que les quatre points C, A', B' et M ont la même troisième coordonnée z.)

I.2.2. Coordonnées sphériques

Le contenu de ce paragraphe est donné à titre purement culturel, et ne sera pas exigible pour les examens.



https://www.geogebra.org/m/er7hvxaw

Les points A, N, S, B et M sont tous sur une sphère $\mathscr{S}$ de centre l'origine O et de rayon $r \in \mathbf{R}_+$.

Les coordonnées cartésiennes de A sont $(r, 0, 0)$. Celles du « pôle Nord » sont $(0, 0, r)$. Celles du « pôle Sud » sont $(0, 0, -r)$.

L'intersection du plan « équatorial » (Oxy) et de la sphère $\mathscr{S}$ est un cercle $\mathscr{C}$ contenant le point A. Si B est un point de ce cercle $\mathscr{C}$, on peut noter φ l'angle (orienté) $(\overrightarrow{OA}, \overrightarrow{OB})$. Les coordonnées cartésiennes du point B sont $(r\cos\varphi, r\sin\varphi, 0)$. (En effet, avec la terminologie de §I.1.2, on pourrait dire, que *considéré dans le plan équatorial*, les coordonnées polaires de B sont (r, φ).)

Étant donné le point B, on peut considérer le (demi-)méridien $\mathscr{M}$, qui est le demi-cercle de diamètre $[NS]$ passant par le point B. Un point M placé sur ce demi-méridien est uniquement déterminé par l'angle géométrique $\theta = \widehat{NOM} \in [0, \pi]$.

L'angle φ est la *longitude* du point M, tandis que θ est sa *colatitude*. (La latitude de M est $\frac{\pi}{2} - \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.)

On peut montrer que les coordonnées cartésiennes (x, y, z) de M sont données par les formules suivantes :

$$\left\{\begin{array}{lcl} x & = & r \sin\theta \cos\varphi \\ y & = & r \sin\theta \sin\varphi \\ z & = & r \cos\theta \end{array}\right.$$

II. Ensembles de points dans le plan

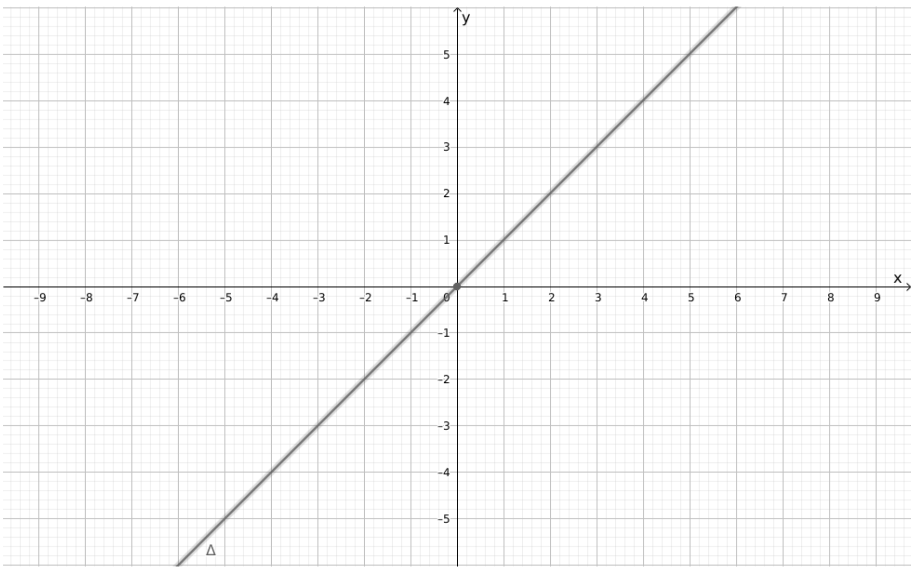
Sommaire

Dans le chapitre précédent, il a notamment été question d'éléments de $\mathbf{R}^2$, que l'on peut se représenter comme des points du plan. Pour l'instant, nous ne considérions que des points individuels. Nous allons maintenant étudier des *ensembles* de points.

II.1. Ensembles définis par une équation (cartésienne)

II.1.1.

Si $(x, y) \in \mathbf{R}^2$, c'est-à-dire que x et y sont deux réels, on peut se demander s'il est vrai que $y = x$: c'est soit vrai, soit faux. On peut noter Δ l'ensemble des éléments $(x, y) \in \mathbf{R}^2$ tels que $y = x$. Voici une représentation graphique de cet ensemble Δ :



Représentation de la droite Δ d'équation $y = x$.

Plus généralement, si $\mathscr{P}(x, y)$ est une condition portant sur deux nombres réels arbitraires x et y (ci-dessus, $\mathscr{P}(x, y)$ était la condition « $y = x$ »), on dit aussi que $\mathscr{P}(x, y)$ est un prédicat, alors on peut introduire un ensemble $\mathscr{C}_{\mathscr{P}}$ formé des couples $(x,y) \in \mathbf{R}^2$ tels que $\mathscr{P}(x,y)$ (soit vrai).

En termes plus formalistes, on pourrait écrire que $\mathscr{C}_{\mathscr{P}}$ est le sous-ensemble de $\mathbf{R}^2$ caractérisé par :

$$\forall (x,y) \in \mathbf{R}^2 \qquad (x,y) \in \mathscr{C}_{\mathscr{P}} \iff \mathscr{P}(x,y) \text{ est vrai.}$$

Ici, « $\forall$ » est le *quantificateur* « pour tout », qui signifie ici que ce qui suit est vrai pour tout couple $(x,y) \in \mathbf{R}^2$, ou de façon équivalente « quel que soit le couple (x,y) appartenant à $\mathbf{R}^2$ ». Ainsi, l'ensemble $\mathscr{C}_{\mathscr{P}}$ est précisément constitué des éléments (x,y) de $\mathbf{R}^2$ tels que la condition $\mathscr{P}(x,y)$ soit vérifiée.

II.1.2.

Dans ce cours, nous allons considérer plus particulièrement des sous-ensembles de $\mathbf{R}^2$ définis par une équation cartésienne. On considère $f\colon \mathbf{R}^2 \to \mathbf{R}$ une application. Avec les notations, on peut introduire la condition $\mathscr{P}(x,y)$ signifiant « $f(x,y) = 0$ ». On peut alors noter Z_f le sous-ensemble de $\mathbf{R}^2$ défini par cette condition, c'est-à-dire qu'un couple $(x,y) \in \mathbf{R}^2$ appartient à Z_f si et seulement si $f(x,y) = 0$. Ceci se note aussi sous la forme suivante :

$$Z_f = \{(x,y) \in \mathbf{R}^2, f(x,y) = 0\} \ ,$$

ce qui se lit « Z_f est l'ensemble des $(x,y) \in \mathbf{R}^2$ tels que $f(x,y) = 0$. ».

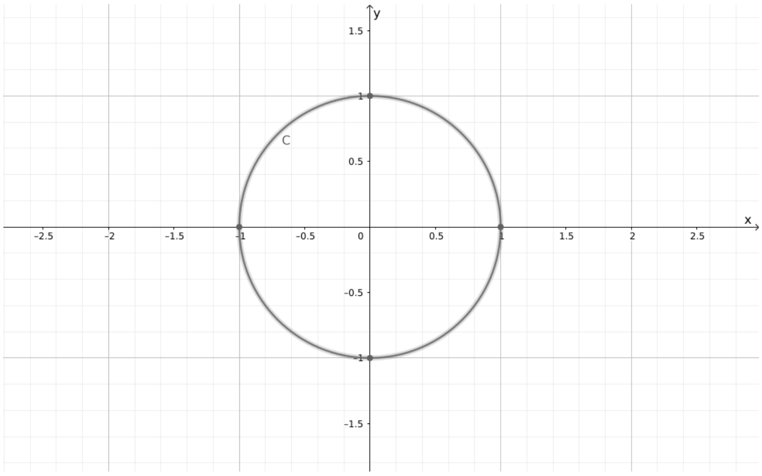
En considérant le cas particulier de l'application $f\colon \mathbf{R}^2 \to \mathbf{R}$ définie par $f(x,y) = x - y$, on retrouve l'ensemble Δ défini ci-desus.

Plus généralement, si $f_1\colon \mathbf{R}^2 \to \mathbf{R}$ et $f_2\colon \mathbf{R}^2 \to \mathbf{R}$ sont deux applications, on peut considérer l'ensemble des $(x,y) \in \mathbf{R}^2$ tels que $f_1(x,y) = f_2(x,y)$. Bien sûr, en posant $f(x,y) := f_1(x,y) - f_2(x,y)$, on remarque que $f_1(x,y) = f_2(x,y)$ si et seulement si $f(x,y) = 0$. Ainsi, les ensembles définis par une équation du type « $f_1(x,y) = f_2(x,y)$ » peuvent aussi être décrits comme des ensembles définis par une équation de la forme « $f(x,y) = 0$ ».

II.2. Exemples d'ensembles définis par une équation

II.2.1. Cercle

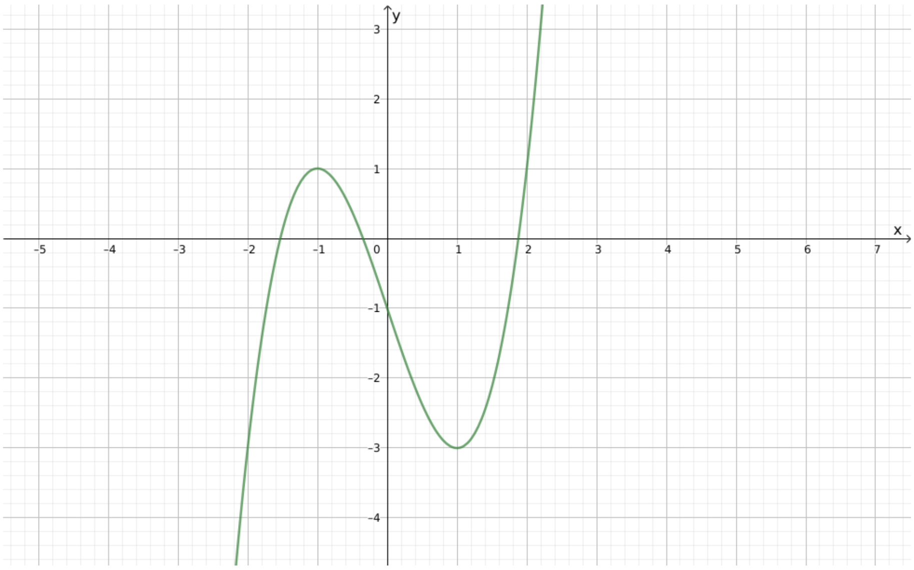
Dans le plan, si M est un point dont les coordonnées sont $(x,y) \in \mathbf{R}^2$, le carré de la distance OM entre M et l'origine O est $OM^2 = x^2 + y^2$. Ainsi, la distance OM est égale à 1 si et seulement si $x^2 + y^2 = 1$. Le cercle de centre O et de rayon 1 (appelé « cercle unité ») est défini par l'équation $x^2 + y^2 = 1$.



Représentation du cercle $\mathscr{C}$ d'équation $x^2 + y^2 = 1$.

II.2.2. Courbes représentatives de fonctions

Supposons que $f : \mathbf{R} \to \mathbf{R}$ soit une application. La « courbe » représentative de f est le sous-ensemble $\mathscr{C}_f$ du plan défini par l'équation « $y = f(x)$ ». (Le mot « courbe » est en principe réservé pour le cas où la fonction f vérifie des propriétés de régularité que vous étudierez dans le cours de *Calculus*, par exemple quand f est de classe $\mathscr{C}^1$.)



Courbe représentative $\mathscr{C}_f$ de la fonction f définie par $f(x) = x^3 - 3x + 1$,
autrement dit de la courbe d'équation $y = x^3 - 3x + 1$.

Théorème II.2.2.1. *Supposons que f et g soient deux applications $\mathbf{R} \to \mathbf{R}$. Les deux conditions suivantes sont équivalentes :*

(i) $f = g$, *c'est-à-dire que pour tout $x \in \mathbf{R}$, $f(x) = g(x)$;*

(ii) Les « courbes » représentatives $\mathscr{C}_f$ et $\mathscr{C}_g$ sont égales, c'est-à-dire qu'elles contiennent exactement les mêmes éléments de $\mathbf{R}^2$.

Il est évident que $(i) \Longrightarrow (ii)$. Montrons $(ii) \Longrightarrow (i)$. Supposons que $\mathscr{C}_f = \mathscr{C}_g$, et montrons que pour tout $x \in \mathbf{R}$, on a $f(x) = g(x)$. Soit $t \in \mathbf{R}$. Considérons le point $M := (t, f(t)) \in \mathbf{R}^2$.

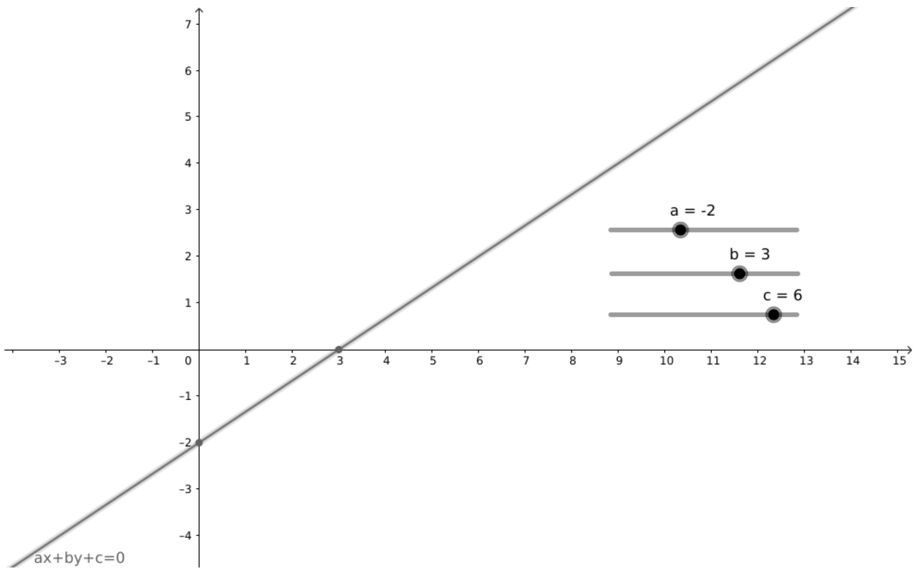
Comme $f(t) = f(t)$, ce point M appartient à l'ensemble $\mathscr{C}_f$ défini par l'équation « $y = f(x)$ ». Comme $\mathscr{C}_f = \mathscr{C}_g$, le point $M = (t, f(t))$ appartient aussi à l'ensemble défini par l'équation « $y = g(x)$ », donc en faisant la substitution $x := t$ et $y := f(t)$ dans cette équation, on obtient $f(t) = g(t)$. On a montré que $f(t) = g(t)$ pour tout $t \in \mathbf{R}$, donc $f = g$.

II.3. Droites définies par une équation

II.3.1. Définition

Dans ce cours, on appelera droite un sous-ensemble $\mathscr{D}$ de $\mathbf{R}^2$ qui soit défini par une équation de la forme $ax + by + c = 0$ où (a, b, c) sont des coefficients réels *fixés*, avec $a \neq 0$ ou $b \neq 0$. Autrement dit, pour que $\mathscr{D}$ soit une droite, il faut et il suffit qu'il existe une fonction $f\colon \mathbf{R}^2 \to \mathbf{R}$ de la forme $f(x, y) = ax + by + c$ (avec a, b, c réels tels que a et b ne soient pas tous les deux nuls) telle que $\mathscr{D}$ soit le sous-ensemble de $\mathbf{R}^2$ défini par l'équation $f(x, y) = 0$.

Par exemple, la droite Δ d'équation $y = x$ considérée plus haut est effectivement définie par l'équation $ax + by + c = 0$ avec $(a, b, c) = (1, -1, 0)$.



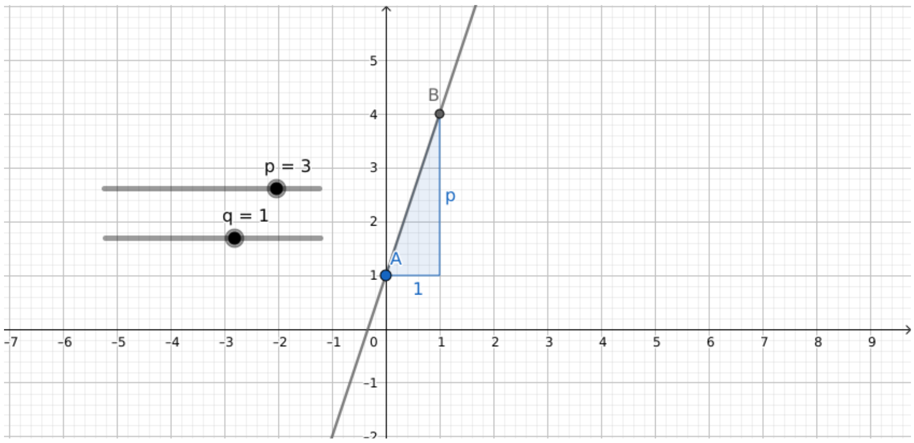
Droite définie par une équation $ax + by + c = 0$
(cas particulier de la droite d'équation $-2x + 3y + 6 = 0$)
`https://www.geogebra.org/m/huzxnnzj`

II.3.2. Équation réduite d'une droite

Définition II.3.2.1. *Soit $(a, b, c) \in \mathbf{R}^3$. On suppose que $(a, b) \neq (0, 0)$. On note $\mathscr{D}$ la droite d'équation $ax + by + c = 0$. L'équation réduite (ou standard, ou canonique) associée est :*

— *l'équation « $y = -\frac{a}{b}x - \frac{c}{b}$ » si $b \neq 0$;*
— *l'équation « $x = -\frac{c}{a}$ » sinon.*

Il est important de remarquer que l'ensemble des points du plan définis par l'équation réduite n'est autre que la droite $\mathscr{D}$ définie par l'équation originale $ax + by + c = 0$.



Droite d'équation réduite de la forme $y = px + q$ avec $(p, q) \in \mathbf{R}^2$.

Dans le cas d'une équation réduite de la forme $y = px + q$ avec $(p, q) \in \mathbf{R}^2$ comme ci-dessus, on lit le coefficient q comme étant « l'ordonnée à l'origine », c'est-à-dire que si on note A le point d'intersection de la droite avec l'axe (Oy), les coordonnées de A sont $(0, q)$. Le coefficient p se lit comme étant la *pente* de la droite, et est communément appelé « coefficient directeur ». Plus précisément, si (x, y) et (x', y') sont deux points de la droite considérée, on a $y = px + q$ et $y' = px' + q$, donc $y' - y = p(x' - x)$. Si on note alors $\Delta x := x' - x$ et $\Delta y := y' - y$ les « accroissements » des abscisses et ordonnées respectivement, on a la relation $\Delta y = p \cdot \Delta x$.

Théorème II.3.2.2. *Soit $(a, b, c) \in \mathbf{R}^3$. Soit $(a', b', c') \in \mathbf{R}^3$. On suppose que $(a, b) \neq (0, 0)$ et que $(a', b') \neq (0, 0)$. On note $\mathscr{D}$ la droite définie par l'équation $ax + by + c = 0$ et $\mathscr{D}'$ celle définie par l'équation $a'x + b'y + c' = 0$. Les conditions suivantes sont équivalentes :*

(i) $\mathscr{D} = \mathscr{D}'$;
(ii) les deux équations réduites associées sont identiques ;
(iii) il existe un $\lambda \in \mathbf{R}^\star$ (c'est-à-dire un nombre réel non nul), tel que $a' = \lambda a$, $b' = \lambda b$ et $c' = \lambda c$.

Commençons par remarquer que $(iii) \Longrightarrow (i)$. En effet, si on suppose (iii), alors pour tous $(x, y) \in \mathbf{R}$, $a'x + b'y + c' = \lambda(ax + by + c)$, et donc, puisque $\lambda \neq 0$, il est évident que $ax + by + c = 0$ si et seulement si $a'x + b'y + c = 0$, donc $\mathscr{D} = \mathscr{D}'$.

Le point-clef de la démonstration consiste à montrer $(i) \Longrightarrow (ii)$. Supposons donc (i). Plaçons-nous pour commencer dans le cas particulier où $b \neq 0$, de sorte que $\mathscr{D}$ admette pour équation réduite $y = f(x)$ où $f \colon \mathbf{R} \to \mathbf{R}$ est la fonction définie par $f(x) = -\frac{a}{b}x - \frac{c}{b}$. Autrement dit, $\mathscr{D}$ est la courbe représentative $\mathscr{C}_f$ de cette fonction f. Comme $\mathscr{D} = \mathscr{D}'$, j'affirme que l'équation réduite de $\mathscr{D}'$ n'est pas de la forme $x = x_0$ avec $x_0 \in \mathbf{R}$: en effet, en tant que courbe représentative d'une fonction $\mathscr{D}'$ contient des points d'abscisse arbitraire. L'équation réduite de $\mathscr{D}'$ est donc de la forme $y = \lambda x + \mu$ avec $(\lambda, \mu) \in \mathbf{R}$. On a donc $b' \neq 0$, et $\mathscr{D}'$ est la courbe représentative de la fonction $g \colon \mathbf{R} \to \mathbf{R}$ définie par $g(x) = -\frac{a'}{b'}x - \frac{c'}{b'}$. Comme $\mathscr{D} = \mathscr{D}'$, les courbes représentatives de f et g sont égales, donc $f = g$ d'après le théorème II.2.2.1. Les deux équations réduites de $\mathscr{D}$ et $\mathscr{D}'$ sont donc les mêmes.

Sous l'hypothèse (i), on a vu que $b \neq 0$ impliquait $b' \neq 0$. Par symétrie des rôles entre les droites $\mathscr{D}$ et $\mathscr{D}'$, on obtient réciproquement que si $b' \neq 0$ alors $b \neq 0$. Ainsi $b \neq 0$ si et seulement si $b' \neq 0$, et on a vu que si $b \neq 0$, on avait bien (ii). Supposons maintenant $b = 0$. D'après ce qui précède, on a aussi $b' = 0$. Les équations réduites de $\mathscr{D}$ et $\mathscr{D}'$ sont toutes les deux de la forme $x = x_0$ (pour $\mathscr{D}$) et $x = x'_0$ (pour $\mathscr{D}'$). En considérant par exemple le point

$(x_0, 0) \in \mathscr{D}$, l'hypothèse $\mathscr{D} = \mathscr{D}'$ permet de conclure que $x_0 = x'_0$, donc les équations réduites de $\mathscr{D}$ et $\mathscr{D}'$ sont les mêmes. Ceci achève la démonstration de l'implication $(i) \Longrightarrow (ii)$.

Il reste à montrer que $(ii) \Longrightarrow (iii)$. Supposons (ii). Supposons pour commencer que $b \neq 0$ (et donc $b' \neq 0$). On peut alors dire que $\mathscr{D}$ et $\mathscr{D}'$ sont les courbes représentatives de la même fonction $f\colon \mathbf{R} \to \mathbf{R}$ définie par $f(x) = -\frac{a}{b}x - \frac{c}{b} = -\frac{a'}{b'}x - \frac{c'}{b'}$. En calculant $f(0) = -\frac{c}{b} = -\frac{c'}{b'}$, on obtient que $\frac{c'}{b'} = \frac{c}{b}$. Par ailleurs, en considérant $f(1) - f(0) = -\frac{a}{b} = -\frac{a'}{b'}$, on obtient que $\frac{a'}{b'} = \frac{a}{b}$. On peut poser $\lambda := \frac{b'}{b}$, et on a alors de façon évidente $a' = \lambda a$, $b' = \lambda b$ et $c' = \lambda c$, donc la condition (iii) est vraie. Il reste à considérer le cas où $b = b' = 0$. On a alors $a \neq 0$ et $a' \neq 0$. La droite $\mathscr{D}$ admet pour équation réduite $x = -\frac{c}{a}$. L'équation réduite de $\mathscr{D}'$ est $x = -\frac{c'}{a'}$, qui est la même équation que pour $\mathscr{D}$, donc $\frac{c}{a} = \frac{c'}{a'}$. En posant $\lambda := \frac{a'}{a}$, on a bien $a' = \lambda a$, $b' = \lambda b$ et $c' = \lambda c$. Dans tous les cas (iii) est vrai, ce qui achève la démonstration du théorème.

(On peut remarquer que dans la condition (iii), le coefficient λ, s'il existe, est forcément unique.)

Remarque II.3.2.3. *En étudiant plus précisément la démonstration ci-dessus de l'implication* $(i) \Longrightarrow (ii)$*, on peut montrer que l'on pourrait ajouter au théorème ci-dessus une condition équivalente supplémentaire :*
(i') $\mathscr{D} \subset \mathscr{D}'$;

II.4. Ensembles définis par un paramétrage

II.4.1. Définition

Supposons que l'on ait défini deux applications $X\colon \mathbf{R} \to \mathbf{R}$ et $Y\colon \mathbf{R} \to \mathbf{R}$. La « courbe » paramétrée associée à (X, Y) est l'ensemble $\mathscr{C}$ formé des couples $(x, y) \in \mathbf{R}^2$ tels qu'il existe $t \in \mathbf{R}$ tel que $x = X(t)$ et $y = Y(t)$. En termes formels, $\mathscr{C}$ est le sous-ensemble de $\mathbf{R}^2$ caractérisé par :

$$\forall (x, y) \in \mathbf{R}^2 \qquad (x, y) \in \mathscr{C} \iff \exists t \in \mathbf{R}\ (x = X(t) \text{ et } y = Y(t))$$

(Dans cette caractérisation de $\mathscr{C}$, le point-clef ici est la présence du *quantificateur* « $\exists$ » qui signifie « il existe ».)

Autrement dit, les points de $\mathscr{C}$ sont ceux de la forme $(X(t), Y(t))$ pour $t \in \mathbf{R}$. Cela se note parfois :

$$\mathscr{C} := \{(X(t), Y(t)), t \in \mathbf{R}\}$$

Si pour tout $t \in \mathbf{R}$, on note $M(t) := (X(t), Y(t)) \in \mathbf{R}^2$, la courbe $\mathscr{C}$ est formée des points $M(t)$ pour t parcourant $\mathbf{R}$.

Par exemple, si on a posé $X(t) := t$ et $Y(t) := t$ pour tout $t \in \mathbf{R}$, alors $\mathscr{C}$ est la droite d'équation « $y = x$ ».

Si $X(t) := \cos t$ et $Y(t) := \sin t$ pour tout $t \in \mathbf{R}$, alors $\mathscr{C}$ est le cercle de centre 0 et de rayon 1. En effet, si $t \in \mathbf{R}$, alors on sait que $\cos^2 t + \sin^2 t = 1$, donc $(\cos t, \sin t)$ appartient bien au cercle d'équation « $x^2 + y^2 = 1$ », et réciproquement, on peut montrer que si $(x, y) \in \mathbf{R}^2$ est tel que $x^2 + y^2 = 1$, alors il existe $t \in \mathbf{R}$ tel que $x = \cos t$ et $y = \sin t$.

Pour décrire une courbe paramétrée, on peut rédiger comme suit :
« La courbe paramétrée ci-dessus est définie par :

$$\left\{ \begin{array}{lcl} X(t) & = & \cos t \\ Y(t) & = & \sin t \end{array} \right. \qquad \text{pour } t \in \mathbf{R}\,. »$$

Conventionnellement, on écrit même plutôt $x(t) = \ldots$ et $y(t) = \ldots$ plutôt que $X(t) = \ldots$ et $Y(t) = \ldots$, mais ceci peut entraîner des confusions.

Il est important de bien comprendre le rôle du paramètre $t \in \mathbf{R}$: pour chaque $t \in \mathbf{R}$, on dispose d'un point $M(t) := (X(t), Y(t)) \in \mathbf{R}^2$. Ainsi, on exprime les coordonnées x et y d'un point de la courbe $\mathscr{C}$ en fonction du paramètre t.

(Dans le paragraphe ci-dessus, on pourrait restreindre les valeurs du paramètre t, en demandant que t appartienne à I, un intervalle de $\mathbf{R}$ plutôt qu'à $\mathbf{R}$ tout entier. Il suffirait alors que $X(t)$ et $Y(t)$ soient définis sur l'intervalle I ; par exemple, le cercle unité pourrait aussi bien être défini par le paramétrage $x(t) = \cos(t)$ et $y(t) = \sin(t)$ pour $t \in [0, 2\pi[$. Nous n'aurons pas besoin de cette généralisation.)

II.4.2. Paramétrages de droites

Nous allons décrire des paramétrages des droites dans le plan. (Les droites ont été définies en termes d'équations cartésiennes au §II.3.1).

Proposition II.4.2.1. *Soit $(a, b) \in \mathbf{R}^2$. On note $\mathscr{D} \subset \mathbf{R}^2$ la droite d'équation $y = ax + b$. Alors $\mathscr{D}$ admet le paramétrage suivant :*

$$\left\{ \begin{array}{rcl} X(t) & = & t \\ Y(t) & = & at + b \end{array} \right. \qquad \textit{pour } t \in \mathbf{R}\,.$$

En effet, si on note $\mathscr{C}$ la courbe paramétrée de l'énoncé, on a $\mathscr{C} \subset \mathscr{D}$, puisque si $(x, y) \in \mathbf{R}$ est de la forme $(X(t), Y(t))$ pour un certain $t \in \mathbf{R}$, alors $x = t$ et $y = at + b = ax + b$, donc (x, y) satisfait l'équation « $y = ax + b$ », donc $(x, y) \in \mathscr{D}$. Réciproquement, si $(x, y) \in \mathscr{D}$, alors $y = ax + b$ et en posant $t := x$, on a $(x, y) = (t, at + b)$ qui appartient bien à $\mathscr{C}$. Par double inclusion, on a montré que $\mathscr{C} = \mathscr{D}$.

Proposition II.4.2.2. *Soit $x_0 \in \mathbf{R}$. On note $\mathscr{D} \subset \mathbf{R}^2$ la droite d'équation $x = x_0$. Alors $\mathscr{D}$ admet le paramétrage suivant :*

$$\left\{ \begin{array}{rcl} X(t) & = & x_0 \\ Y(t) & = & t \end{array} \right. \qquad \textit{pour } t \in \mathbf{R}\,.$$

Définition II.4.2.3. *Un paramétrage affine d'une droite est un paramétrage de la forme suivante*

$$\left\{ \begin{array}{rcl} X(t) & = & x_0 + v_1 t \\ Y(t) & = & y_0 + v_2 t \end{array} \right. \qquad \textit{pour } t \in \mathbf{R}\,,$$

où x_0, y_0, v_1 et v_2 sont réels, tels que $v_1 \neq 0$ ou $v_2 \neq 0$ (c'est-à-dire que $(v_1, v_2) \neq (0, 0)$).

(On remarque que dans la définition ci-dessus, si on avait $(v_1, v_2) = (0, 0)$, alors la « courbe » paramétrée ne contiendrait que le point (x_0, y_0), donc ce ne serait certainement pas une droite.)

Remarque II.4.2.4. *On dira qu'une droite est de dimension 1 : un paramètre t est nécessaire (et suffisant) pour en obtenir un paramétrage. On dira que $\mathbf{R}^2$ est de dimension 2. Un singleton est de dimension 0.*

Théorème II.4.2.5. *Si une courbe est définie par un paramétrage affine (cf. définition II.4.2.3), alors c'est une droite (au sens de la définition II.3.1), c'est-à-dire que c'est un ensemble défini par une équation de la forme $ax + by + c = 0$ avec a, b et c des réels tels que $(a, b) \neq (0, 0)$.*

Donnons-nous des coefficients x_0, y_0, v_1 et v_2 comme dans la définition II.4.2.3 et considérons la courbe paramétrée $\mathscr{C}$ définie par $X(t) = x_0 + v_1 t$ et $Y(t) = y_0 + v_2 t$. Tentons d'éliminer le paramètre t de ces équations. Pour cela, calculons $v_2 X(t) - v_1 Y(t)$:

$$v_2 X(t) - v_1 Y(t) = v_2 x_0 + v_2 v_1 t - v_1 y_0 - v_1 v_2 t = v_2 x_0 - v_1 y_0 \,.$$

On remarque que le membre de droite $v_2 x_0 - v_1 y_0$ est indépendant de t. Posons $a := v_2$, $b := -v_1$ et $c := -v_2 x_0 + v_1 y_0$. On a montré que pour tout $t \in \mathbf{R}$, on avait

$$aX(t) + bY(t) + c = 0\,.$$

On a bien sûr $(a, b) \neq (0, 0)$. On peut donc noter $\mathscr{D}$ la droite d'équation $ax + by + c = 0$. On vient de montrer l'inclusion $\mathscr{C} \subset \mathscr{D}$.

Pour conclure, il reste à montrer l'inclusion réciproque $\mathscr{D} \subset \mathscr{C}$. Plaçons-nous dans le cas particulier où $v_1 \neq 0$, et donc $b \neq 0$ (le cas où $v_1 = 0$ et $v_2 \neq 0$ se démontre de manière similaire). Soit $(x, y) \in \mathbf{R}$ tel que $ax + by + c = 0$. On cherche $t \in \mathbf{R}$ tel que les deux équations suivantes soient satisfaites :

$$\left\{\begin{array}{lcl} x & = & x_0 + v_1 t \\ y & = & y_0 + v_2 t \end{array}\right.$$

Si $t \in \mathbf{R}$ est solution de ce système d'équations, alors on a nécessairement $v_1 t = x - x_0$, donc $t = \frac{x - x_0}{v_1}$. On a alors bien sûr $x = x_0 + v_1 t$, donc la première équation ci-dessus est vraie. On a aussi :

$$y_0 + v_2 t = \frac{v_1 y_0 + v_2 x - v_2 x_0}{v_1} = -\frac{(c + ax)}{b} = \frac{by}{b} = y$$

La deuxième équation ci-dessus donc vraie également. Ainsi, pour le paramètre $t := \frac{x - x_0}{v_1}$, le point $(X(t), Y(t))$ est égal à (x, y). Ceci montre l'inclusion $\mathscr{D} \subset \mathscr{C}$. On a donc montré que $\mathscr{C} = \mathscr{D}$.

Proposition II.4.2.6. *Supposons que $A_1 = (x_1, y_1)$ et $A_2 = (x_2, y_2)$ soient deux éléments distincts de $\mathbf{R}^2$. Alors, il existe une unique droite $\mathscr{D}$ passant par A_1 et A_2 : c'est la droite $\mathscr{D}$ donnée par le paramétrage suivant :*

$$\left\{\begin{array}{lcl} X(t) & = & (1 - t)x_1 + t x_2 \\ Y(t) & = & (1 - t)y_1 + t y_2 \end{array}\right. \qquad \textit{pour } t \in \mathbf{R}\,.$$

(Cette droite $\mathscr{D}$ sera aussi notée $(A_1 A_2)$.)

Si on note $M(t) := (X(t), Y(t))$ pour tout $t \in \mathbf{R}$, on remarque que $M(0) = A_1$ et $M(1) = A_2$. Par ailleurs, il s'agit bien d'un paramétrage affine, donc $\mathscr{D}$ est bien une droite, qui contient A_1 et A_2.

Réciproquement, soit Δ une droite, d'équation cartésienne « $ax + by + c = 0$ », et contenant A_1 et A_2. Les identités suivantes sont vérifiées :

$$\begin{array}{rcl} ax_1 + by_1 + c & = & 0 \\ ax_2 + by_2 + c & = & 0 \end{array}$$

Soit $t \in \mathbf{R}$. On obtient :

$$\begin{array}{rcl} 0 & = & (1 - t)(ax_1 + by_1 + c) + t(ax_2 + by_2 + c) \\ & = & a((1 - t)x_1 + t x_2) + b((1 - t)y_1 + t y_2) + c \\ & = & aX(t) + bY(t) + c \end{array}$$

Ceci montre que le point $M(t)$ appartient à Δ. Ceci étant vrai pour tout $t \in \mathbf{R}$, on a montré que $\mathscr{D} \subset \Delta$. D'après la remarque II.3.2.3, cela implique $\mathscr{D} = \Delta$. Il existe donc bien une unique droite contenant les points A_1 et A_2.

Remarque II.4.2.7. *Dans la démonstration ci-dessus, le point $M(t)$ pour $t \in \mathbf{R}$ est caractérisé par l'identité vectorielle $\overrightarrow{A_1M(t)} = t\overrightarrow{A_1A_2}$.*

II.5. Différentes descriptions des droites

Plusieurs manières de décrire une droite ont été données plus haut :

— comme ensemble des points vérifiant une équation cartésienne « $ax + by + c = 0$ » avec a, b et c des réels tels que $(a, b) \neq (0, 0)$;

— comme ensemble des points vérifiant une équation « réduite » « $y = px + q$ » (avec $(p, q) \in \mathbf{R}^2$) ou bien « $x = x_0$ » (avec $x_0 \in \mathbf{R}$) ;

— comme courbe donnée par un paramétrage affine

$$\begin{cases} X(t) &= x_0 + v_1 t \\ Y(t) &= y_0 + v_2 t \end{cases} \qquad \text{pour } t \in \mathbf{R} \,;$$

— comme droite passant par deux points distincts.

Si une droite est décrite d'une des quatre manières ci-dessus, il sera important de savoir donner assez rapidement une de ses descriptions alternatives.

II.6. Intersection de deux droites

Théorème II.6.1. *Si $\mathscr{D}_1$ et $\mathscr{D}_2$ sont deux droites dans $\mathbf{R}^2$, alors l'intersection $\mathscr{D}_1 \cap \mathscr{D}_2$ est soit :*

— *vide ;*

— *un singleton (c'est-à-dire un ensemble contenant un seul élément) ;*

— *une droite.*

II.6.2.

La façon la plus directe de démontrer le théorème II.6.1 consiste à supposer que l'on a défini $\mathscr{D}_1$ par un paramétrage affine $M(t) := (X(t), Y(t))$ et que $\mathscr{D}_2$ est définie par une équation cartésienne « $ax + by + c = 0$ ». L'ensemble des points d'intersection est l'ensemble des points $M(t)$ tels que $M(t)$ appartienne à $\mathscr{D}_2$. Déterminer les points d'intersections revient essentiellement à déterminer les paramètres $t \in \mathbf{R}$ tels que $aX(t) + bY(t) + c = 0$. Comme $X(t)$ et $Y(t)$ sont des expressions affines de t, la quantité $aX(t) + bY(t) + c$ s'exprime sous la forme $\lambda t + \mu$. Déterminer les points d'intersection revient à chercher les $t \in \mathbf{R}$ tel que $\lambda t + \mu = 0$.

Plusieurs cas sont possibles. Soit $\lambda \neq 0$ et alors l'unique solution est $t := -\frac{\mu}{\lambda}$, de sorte que l'unique point d'intersection est $M(-\frac{\mu}{\lambda})$. Si $\lambda = 0$, deux sous-cas sont possibles :

— soit $\mu \neq 0$, et alors il n'y a pas de solution : $\mathscr{D}_1 \cap \mathscr{D}_2$ est vide ;

— soit $\mu = 0$, et alors tous les $t \in \mathbf{R}$ conviennent, ce qui revient à dire que $\mathscr{D}_1 \subset \mathscr{D}_2$. On sait alors d'après la remarque II.3.2.3 que $\mathscr{D}_1 = \mathscr{D}_2$, c'est-à-dire que les deux droites sont confondues, donc l'intersection est une droite.

II.6.3.

Le problème énoncé dans le théorème II.6.1 peut être posé différemment si $\mathscr{D}_1$ et $\mathscr{D}_2$ sont définies par deux équations cartésiennes : $a_1x + b_1y = c_1$ pour $\mathscr{D}_1$ et $a_2x + b_2y = c_2$ pour $\mathscr{D}_2$. Les points appartenant à l'intersection $\mathscr{D}_1 \cap \mathscr{D}_2$ sont les couples $(x, y) \in \mathbf{R}^2$ qui sont solutions du système d'équations suivant :

$$\begin{cases} a_1x + b_1y & = & c_1 & \qquad (E_1) \\ a_2x + b_2y & = & c_2 & \qquad (E_2) \end{cases}$$

Considérons le cas où $a_1 \neq 0$. Ceci implique que si on le souhaitait, on pourrait récrire l'équation (E_1) comme exprimant x en fonction de y. On va tenter de récrire ce système en gardant la première équation inchangée, mais avec une deuxième équation qui ne fasse plus intervenir que y. Pour cela on note $\lambda := \frac{a_2}{a_1}$, et on considère l'équation (E_2') obtenue en « retranchant à l'équation (E_2) l'équation (E_1) multipliée par λ ». Plus précisément, en notant $a_2' := a_2 - \lambda a_1 = 0$, $b_2' := b_2 - \lambda b_1$, $c_2' := c_2 - \lambda c_1$. On obtient un nouveau système :

$$\begin{cases} a_1x + b_1y & = & c_1 & \qquad (E_1) \\ b_2'y & = & c_2' & \qquad (E_2') \end{cases}$$

Ce système est équivalent au précédent : les deux systèmes ont les mêmes solutions (l'argument est détaillé plus bas à la remarque II.6.6). Ce nouveau système est plus facile à résoudre que le précédent. En effet, en considérant l'équation (E_2'), on détermine facilement les y qui peuvent convenir, et comme l'équation (E_1) équivaut à $x = \frac{c_1 - b_1y}{a_1}$, il suffit, pour chaque solution y à (E_2') de calculer $x := \frac{c_1 - b_1y}{a_1}$, et on obtient alors exactement tous les couples (x, y) solutions au système initial. Pour le détail de la résolution, on remarque que $b_2' = \frac{a_1b_2 - a_2b_1}{a_1}$ (Si $a_1 = 0$, alors on sait que $b_1 \neq 0$, et on peut procéder de même en échangeant les rôles de x et de y.)

Remarque II.6.4. *Dans tous les cas, on remarque que l'intersection est un singleton si et seulement si $a_1b_2 - a_2b_1 \neq 0$.*

Proposition II.6.5. *Supposons que $\lambda \in \mathbf{R}$. Soit $(X_1, s_1, X_2, s_2) \in \mathbf{R}^4$. Les conditions suivantes sont équivalentes :*
(i) les deux équations $X_1 = s_1$ et $X_2 = s_2$ sont vraies;
(ii) les deux équations $X_1 = s_1$ et $X_2 - \lambda X_1 = s_2 - \lambda s_1$ sont vraies;

Si (i) est vrai, on a $X_1 = s_1$ et $X_2 = s_2$, donc bien sûr $X_1 = s_1$ et $X_2 - \lambda X_1 = s_2 - \lambda s_1$, donc (i) est vrai.

Réciproquement, si (ii) est vrai, on a $X_1 = s_1$ et $X_2 - \lambda X_1 = s_2 - \lambda s_1$, donc $X_2 = s_2 + \lambda(X_1 - s_1) = s_2$, donc (ii) est vrai.

Remarque II.6.6. *Le principe général à retenir de la proposition précédente est que si on dispose de deux équations (E_1) et (E_2), et d'un coefficient λ, alors le système formé par ces deux équations est équivalent au système formé par (E_1) et (E_2') où E_2' est l'« équation obtenue en retranchant l'équation (E_1) multipliée par le coefficient λ ».*

Exemple II.6.7. *Dans ce système d'équations*

$$\begin{cases} 2x & + & 3y & = 7 \\ 4x & + & 7y & = 15 \end{cases}$$

on peut retrancher à la deuxième équation deux fois la première pour obtenir un système équivalent :

$$\begin{cases} 2x & + & 3y & = 7 \\ & & y & = 1 \end{cases}$$

L'unique solution de ces systèmes est $(2, 1)$.

II.7. Droites parallèles, vecteurs directeurs

Définition II.7.1. *Soit $\mathscr{D}$ une droite d'équation cartésienne « $ax + by + c = 0$ ». On appelle « droite vectorielle » $\overrightarrow{\mathscr{D}}$ associée à $\mathscr{D}$ la droite d'équation « $ax + by = 0$ ».*

La droite vectorielle associée ne dépend effectivement que de la droite $\mathscr{D}$. En effet, d'après le théorème II.3.2.2, si on avait une autre équation cartésienne « $a'x + b'y + c' = 0$ » de cette même droite, cette équation serait nécessairement obtenue en multipliant la première équation par un coefficient $\lambda \in \mathbf{R}^{\times}$. La droite d'équation $a'x + b'y = 0$ serait alors la même que la droite d'équation $ax + by = 0$.

Proposition II.7.2. *Soit $\mathscr{D}$ une droite et $\overrightarrow{\mathscr{D}}$ la droite vectorielle associée. Soit $(v_1, v_2) \in \overrightarrow{\mathscr{D}} - \{(0, 0)\}$. Alors, pour tout $(w_1, w_2) \in \overrightarrow{\mathscr{D}}$, il existe un unique $t \in \mathbf{R}$, tel que $w_1 = tv_1$ et $w_2 = tv_2$. On dit que (v_1, v_2) est un vecteur directeur de $\mathscr{D}$.*

Considérons $(v_1, v_2) \in \overrightarrow{\mathscr{D}} - \{(0, 0\}$. On peut définir $X(t) := tv_1$ et $Y(t) = tv_2$: c'est le paramétrage d'une droite contenant les deux points distincts $(0, 0)$ et (v_1, v_2), donc c'est la droite $\overrightarrow{\mathscr{D}}$. Ceci démontre au moins l'existence d'un $t \in \mathbf{R}$ tel que $w_1 = tv_1$ et $w_2 = tv_2$, et l'unicité est évidente.

Remarque II.7.3. *Par exemple, si $\mathscr{D}$ est d'équation « $ax + by + c = 0$ », alors $(b, -a)$ est un vecteur directeur de $\mathscr{D}$. Si $\mathscr{D}$ admet pour équation réduite « $y = px + q$ », alors $(1, p)$ est un vecteur directeur de $\mathscr{D}$.*

Proposition II.7.4. *Soit $\mathscr{D}$ une droite d'équation $ax + by + c = 0$. Soit $(x_0, y_0) \in \mathscr{D}$. Pour tout $(x, y) \in \mathbf{R}^2$, notons $x' := x - x_0$ et $y' = y - y_0$. Les conditions suivantes sont équivalentes :*
(i) $ax + by + c = 0$;
(ii) $ax' + by' = 0$;

En effet, on a toujours $ax' + by' = ax + by - (ax_0 + by_0) = ax + by + c$, donc $ax' + by' = 0$ si et seulement si $ax + by + c = 0$.

Remarque II.7.5. *Le principe à retenir de la proposition précédente est que pour déterminer les solutions à une équation du type (i), il suffit de connaître une « solution particulière » (x_0, y_0) et de savoir déterminer les solutions à l'équation (ii).*

Remarque II.7.6. *Géométriquement, si A_1 et A_2 sont deux points distincts d'une droite $\mathscr{D}$, le vecteur $\overrightarrow{A_1A_2}$ est un vecteur directeur de la droite $\mathscr{D}$. En effet, on peut considérer les coordonnées du point A_1 comme une solution particulière de l'équation $ax + by + c = 0$, et alors l'équivalence énoncée dans la proposition ci-dessus s'interprète en disant que A_2 appartient à la droite $\mathscr{D}$ si et seulement si le vecteur $\overrightarrow{A_1A_2}$ appartient à $\overrightarrow{\mathscr{D}}$.*

Définition II.7.7. *On dit de deux droites $\mathscr{D}_1$ et $\mathscr{D}_2$ qu'elles sont parallèles si elles ont la même droite vectorielle associée.*

Proposition II.7.8. *Soit $\mathscr{D}$ une droite d'équation $ax + by + c = 0$. Les droites parallèles à $\mathscr{D}$ sont les droites d'équation $ax + by + c' = 0$ pour $c' \in \mathbf{R}$.*

Cela résulte directement des définitions.

Proposition II.7.9. *Soit $\mathscr{D}$ une droite d'équation $ax+by+c = 0$. Soit $\mathscr{D}'$ une droite d'équation $a'x + b'y + c' = 0$. Alors, les droites $\mathscr{D}$ et $\mathscr{D}'$ sont parallèles si et seulement si $ab' - ba' = 0$.*

La droite vectorielle $\overrightarrow{\mathscr{D}}$ a pour équation $ax+by = 0$ tandis que $\overrightarrow{\mathscr{D}'}$ a pour équation $a'x+b'y = 0$. Il s'agit de montrer que la deuxième équation peut être obtenue en multipliant la première par un coefficient $\lambda \in \mathbf{R}^\times$ si et seulement si $ab' = ba'$.

Si $\mathscr{D}$ et $\mathscr{D}'$ sont parallèles, il existe $\lambda \in \mathbf{R}^\times$ tel que $a' = \lambda a$ et $b' = \lambda b$, donc $ab' = a(\lambda b) = \lambda ab$ et $ba' = b(\lambda a) = \lambda ab$, donc $ab' = ba'$. Réciproquement, supposons que $ab' = ba'$. Supposons dans un premier temps que $a \neq 0$, auquel cas on peut poser $\lambda := \frac{a'}{a}$ et on a bien $a' = \lambda a$ et $b' = \frac{ba'}{a} = \lambda b$, donc $\mathscr{D}$ et $\mathscr{D}'$ sont bien parallèles. Si $a = 0$, alors $b \neq 0$, et $ba' = ab' = 0$, donc $ba' = 0$. Comme $b \neq 0$, on en déduit $a' = 0$, puis $b' \neq 0$. Finalement, avec $\lambda := \frac{b'}{b}$, on obtient $b' = \lambda b$ et $a' = \lambda a = 0$, donc $\mathscr{D}$ et $\mathscr{D}'$ sont parallèles.

Corollaire II.7.10. *Supposons que $\mathscr{D}_1$ et $\mathscr{D}_2$ soient deux droites. Si $\mathscr{D}_1$ et $\mathscr{D}_2$ sont parallèles, alors $\mathscr{D}_1 \cap \mathscr{D}_2$ est soit vide soit une droite. Sinon, $\mathscr{D}_1$ et $\mathscr{D}_2$ sont sécantes, c'est-à-dire qu'elles possèdent un unique point d'intersection.*

Cela résulte de la proposition II.7.9 et de la remarque II.6.4.

III. Ensembles de points dans l'espace

Sommaire

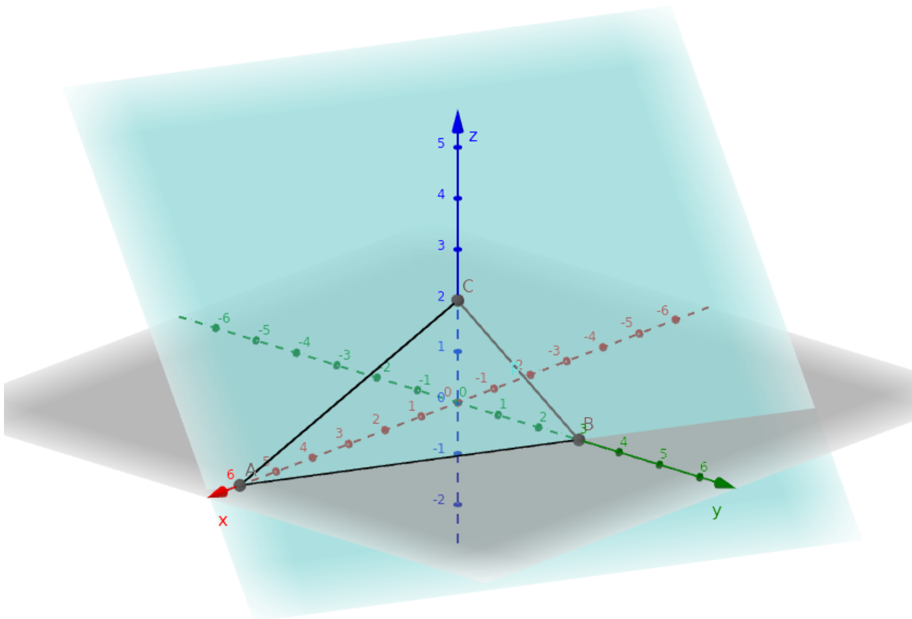
Dans ce chapitre, nous allons étudier des ensembles de points dans $\mathbf{R}^3$.

III.1. Plans et surfaces

Si a, b, c et d sont des réels, on peut considérer l'équation cartésienne en trois variables « $ax + by + cz = d$ », et donc considérer l'ensemble des points $(x, y, z) \in \mathbf{R}^3$ qui satisfont cette équation.

Définition III.1.1. *Dans $\mathbf{R}^3$, un plan est un sous-ensemble $\mathscr{P}$ qui peut être défini par une équation cartésienne de la forme « $ax + by + cz = d$ » avec $(a, b, c) \neq (0, 0, 0)$. (On dira que $\mathscr{P}$ est de dimension 2.)*

Si on autorisait ici le cas $a = b = c = 0$, on obtiendrait soit l'ensemble vide si $d \neq 0$ et $\mathbf{R}^3$ tout entier si $d = 0$.

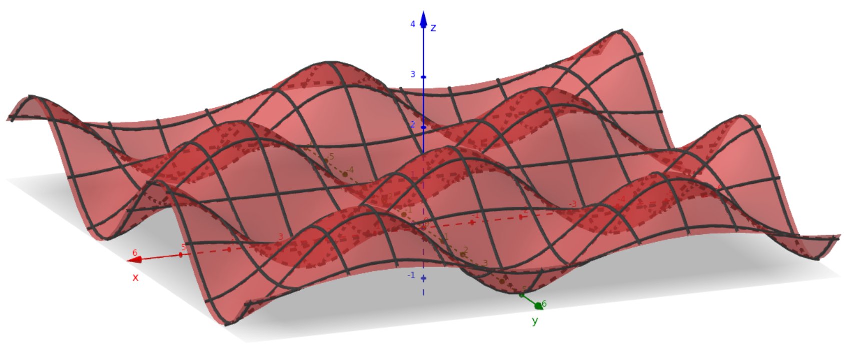


https://www.geogebra.org/m/btkxpgbq

Plan $\mathscr{P}$ d'équation $ax + by + cz = d$ (cas particulier $x + 2y + 3z = 6$)

Sur la figure ci-dessus, on a représenté les points d'intersection du plan $\mathscr{P}$ avec des axes : $A(\frac{d}{a}, 0, 0)$, $B(0, \frac{d}{b}, 0)$, $C(0, 0, \frac{d}{c})$. (Suivant que a, b ou c soient nuls ou non, ces points peuvent ne pas être bien définis.)

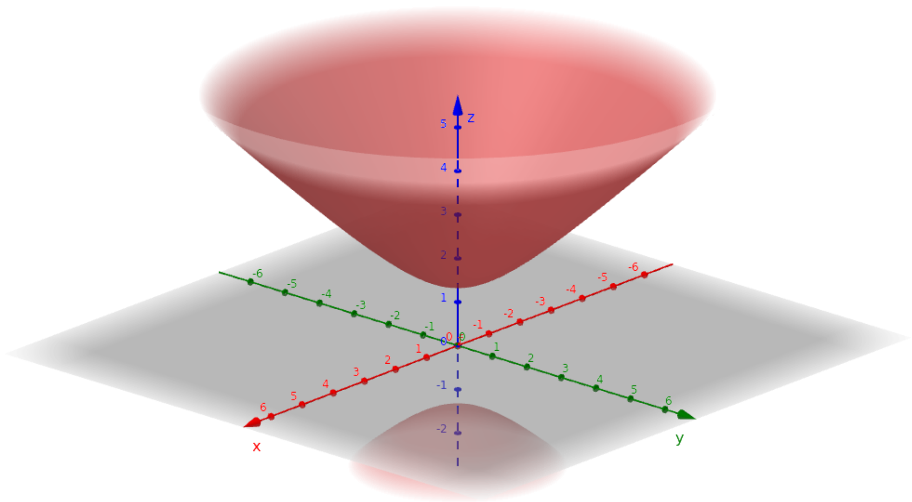
Remarque III.1.2. *On remarque que si $c \neq 0$, l'équation $ax + by + cz = d$ peut se récrire $z = \alpha x + \beta y + \gamma$ avec des coefficients réels α, β et γ bien choisis. En quelque sorte, on peut exprimer z en fonction de x et y : on peut alors définir une application $Z\colon \mathbf{R}^2 \to \mathbf{R}$ par la formule $Z(x, y) = \alpha x + \beta y + \gamma$, et on peut se représenter le plan $\mathscr{P}$ en disant que pour chaque $(x, y) \in \mathbf{R}^2$, il existe un unique $z \in \mathbf{R}$ (qui est $Z(x, y)$) tel que (x, y, z) appartienne au plan $\mathscr{P}$. On peut concevoir $\mathscr{P}$ comme une représentation graphique de la fonction de deux variables $Z\colon \mathbf{R}^2 \to \mathbf{R}$. (Il s'agit d'une variante en deux variables de la courbe représentative d'une fonction, cf. II.2.2.) En posant par exemple $Z(x, y) := \cos x \sin y + 1$, on obtiendrait la* surface *suivante :*



`https://www.geogebra.org/m/egcndzve`

Surface d'équation $z = \cos x \sin y + 1$

Remarque III.1.3. *En considérant d'autres équations cartésiennes, on peut obtenir d'autres surfaces :*



`https://www.geogebra.org/m/f9x2ctwv`

Hyperboloïde (à deux nappes) d'équation cartésienne $z^2 - x^2 - y^2 = 2$.

Théorème III.1.4. . *Supposons que $\mathscr{P}$ soit un plan d'équation $ax + by + cz = d$ et que $\mathscr{P}'$ soit un plan d'équation $a'x + b'y + c'z = d'$. Alors, les conditions suivantes sont équivalentes :*
(i) $\mathscr{P} = \mathscr{P}'$ *;*
(ii) il existe un $\lambda \in \mathbf{R}^\star$ (c'est-à-dire un nombre réel non nul), tel que $a' = \lambda a$, $b' = \lambda b$, $c' = \lambda c$ et $d' = \lambda d$.

Il s'agit d'un énoncé homologue au théorème II.3.2.2. L'implication $(ii) \implies (i)$ est évidente. Montrons $(i) \implies (ii)$. On suppose donc que le plan $\mathscr{P} = \mathscr{P}'$. Commençons par supposer que $c \neq 0$. Il est alors évident que $\mathscr{P}$ peut être décrit par l'équation cartésienne $z = Ax + By + D$ avec $A = -\frac{a}{c}$, $B = -\frac{b}{c}$, $D = -\frac{d}{c}$. On remarque donc que si $(x, y) \in \mathbf{R}^2$ est arbitraire, il existe un unique z (qui est $Ax + By + D$) tel que (x, y, z) appartienne à $\mathscr{P} = \mathscr{P}'$. On en déduit aussitôt que dans l'équation « $a'x + b'y + c'z = d'$ » de $\mathscr{P}'$, il n'est pas possible que $c' = 0$: en effet, sinon, si on pose, par exemple $x = y = 0$, l'ensemble des z tels que $(0, 0, z)$ appartienne à $\mathscr{P}$ serait soit vide soit $\mathbf{R}$ tout entier, alors que c'est un singleton. Ainsi, $c' \neq 0$, donc comme ci-dessus, $\mathscr{P}'$ peut être décrit par l'équation cartésienne $z = A'x + B'y + D'$ avec $A' = -\frac{a'}{c'}$, $B' = -\frac{b'}{c'}$, $D = -\frac{d'}{c'}$. Pour tout $(x, y) \in \mathbf{R}^2$, l'unique z tel que (x, y, z) appartienne à $\mathscr{P}'$ est donc $A'x + B'y + D'$. On a donc montré que pour tout $(x, y) \in \mathbf{R}^2$, on a $Ax + By + D = A'x + B'y + D'$, autrement dit :

$$(A' - A)x + (B' - B)y + (D' - D) = 0$$

En considérant le cas $(x, y) = (0, 0)$, on obtient $D' - D = 0$, donc $D' = D$. Si un des deux coefficients $\lambda := A' - A$ ou $\mu := B' - B$ était non nul, l'ensemble $\mathbf{R}^2$ serait contenu dans une droite (de $\mathbf{R}^2$) d'équation $\lambda x + \mu y = 0$, ce qui est bien sûr absurde d'après les résultats du chapitre précédent. Bref, $A' = A$, $B' = B$, $D' = D$. En posant $k := \frac{c'}{c}$, on en déduit aussitôt que $a' = ka$, $b' = kb$, $c' = kc$ et $d' = kd$.

Comme $(a, b, c) \neq (0, 0, 0)$, si $c = 0$, alors $a \neq 0$ ou $b \neq 0$. Si $a \neq 0$, on peut procéder de façon similaire à ci-dessus en permutant les rôles des variables x et z. Si $b \neq 0$, on peut permuter les rôles des variables y et z. (Alternativement, on pourrait montrer que si $c = 0$, alors $c' = 0$, de sorte que seules les variables x et y interviennent dans les équations, ce qui permet de se ramener à l'énoncé du théorème II.3.2.2.)

III.2. Droites

Définition III.2.1. *Dans l'espace $\mathbf{R}^3$, une droite est un ensemble de points $\mathscr{D}$ donné par un paramétrage $(X(t), Y(t), Z(t))$ pour $t \in \mathbf{R}$, où $X : \mathbf{R}$, $Y : \mathbf{R} \to \mathbf{R}$ et $Z : \mathbf{R} \to \mathbf{R}$ sont des applications de la forme $X(t) = x_0 + v_1 t$, $Y(t) = y_0 + v_2 t$, $Z(t) = z_0 + v_3 t$ avec des coefficients réels x_0, y_0, z_0, v_1, v_2 et v_3, tels que $(v_1, v_2, v_3) \neq (0, 0, 0)$. Plus précisément, un point $(x, y, z) \in \mathbf{R}^3$ appartient à $\mathscr{D}$ si et seulement s'il existe $t \in \mathbf{R}$ tel que $(x, y, z) = (X(t), Y(t), Z(t))$, c'est-à-dire qu'il existe $t \in \mathbf{R}$ tel que les équations suivantes soient vraies :*

$$\begin{cases} x &= x_0 + v_1 t \\ y &= y_0 + v_2 t \\ z &= z_0 + v_3 t \end{cases}$$

Avec ces notations, on remarque que (x_0, y_0, z_0) appartient à $\mathscr{D}$. On dit que (v_1, v_2, v_3) est un vecteur directeur de $\mathscr{D}$. On peut montrer, comme dans le cas des droites dans le plan, qu'un vecteur qui relie deux points de $\mathscr{D}$ est nécessairement de la forme $(\lambda v_1, \lambda v_2, \lambda v_3)$ pour $\lambda \in \mathbf{R}$.

III.2.2.

Avec les notations ci-dessus, si on fait l'hypothèse que $v_1 \neq 0$, on remarque que l'équation $x = x_0 + v_1 t$ équivaut à $t = \frac{x-x_0}{v_1}$. Pour tout $x \in \mathbf{R}$, notons $\tau(x) := \frac{x-x_0}{v_1}$. Ainsi, si un point (x, y, z) appartient à $\mathscr{D}$, c'est-à-dire est de la forme $(X(t), Y(t), Z(t))$, la seule possibilité est que $t = \tau(x)$ auquel cas la première des équations ci-dessus est satisfaite. Il en résulte que si $(x, y, z) \in \mathbf{R}^3$, alors (x, y, z) appartient à $\mathscr{D}$ si et seulement si (x, y, z) est solution du système de deux équations :

$$\left\{ \begin{array}{lcl} y & = & y_0 + v_2\tau(x) \\ z & = & z_0 + v_3\tau(x) \end{array} \right.$$

Si on remplace $\tau(x)$ par son expression, on obtient que $\mathscr{D}$ est précisément l'ensemble des solutions à un système de la forme suivante :

$$\left\{ \begin{array}{lcl} y & = & \alpha x + \beta \\ z & = & \alpha' x + \beta' \end{array} \right.$$

où les coefficients α, α', β et β' sont *a priori* arbitraires.

Chacune des deux équations peut être considérée comme l'équation d'un plan, donc on a montré que la droite $\mathscr{D}$ pouvait s'écrire comme intersection de deux plans. (C'est un fait général puisque si on avait $v_1 = 0$, alors $v_2 \neq 0$ ou $v_3 \neq 0$, et le même argument pourrait être fait en échangeant le rôle de la variable x avec celui de y ou de z.)

Théorème III.2.2.1. *Supposons que $\mathscr{P}$ et $\mathscr{P}'$ soient deux plans dans $\mathbf{R}^3$ définis respectivement par les équations « $ax + by + cz = d$ » et « $a'x + b'y + c'z = d'$ ». L'intersection $\mathscr{P} \cap \mathscr{P}'$ est soit*
- *vide ;*
- *une droite ;*
- *un plan.*

De la même façon que pour les droites dans le plan, le plan vectoriel $\overrightarrow{\mathscr{P}}$ associé à $\mathscr{P}$ est défini l'équation $ax + by + cz = 0$ et le plan vectoriel $\overrightarrow{\mathscr{P}'}$ associé à $\mathscr{P}'$ est défini par l'équation $a'x + b'y + c'z = 0$. On dit que $\mathscr{P}$ et $\mathscr{P}'$ sont parallèles si $\overrightarrow{\mathscr{P}} = \overrightarrow{\mathscr{P}'}$.

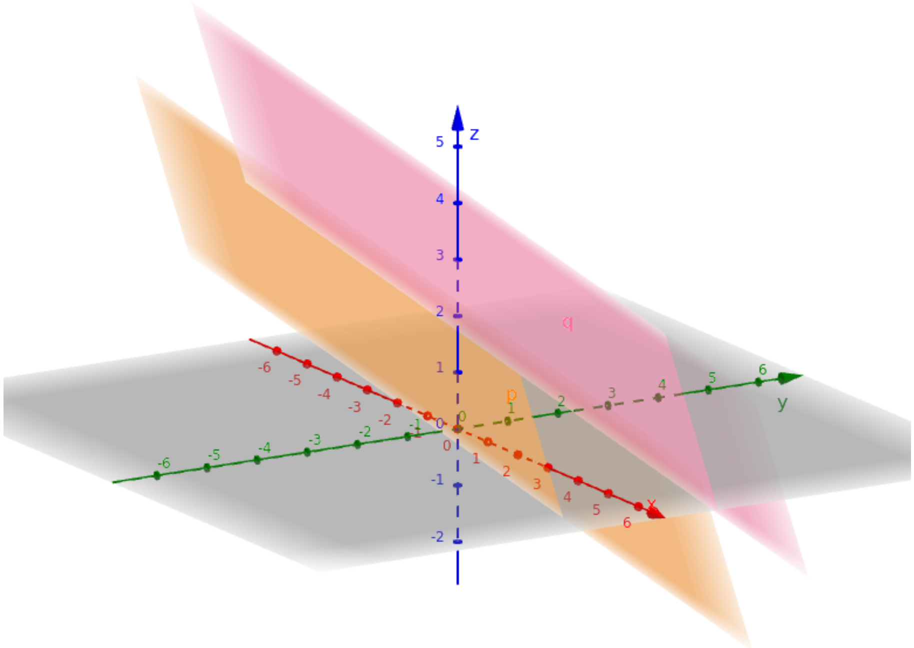
L'intersection $\mathscr{P} \cap \mathscr{P}'$ est formé des solutions du système suivant :

$$\left\{ \begin{array}{ccccccc} ax & + & by & + & cz & = & d \\ a'x & + & b'y & + & c'z & = & d' \end{array} \right. \qquad \begin{array}{c} (E_1) \\ (E_2) \end{array}$$

Si $\mathscr{P}$ et $\mathscr{P}'$ sont parallèles, alors il existe un coefficient $\lambda \in \mathbf{R}^\times$ tel que l'expression apparaissant dans le membre de gauche de (E_2) soit celui de la première équation multiplié par λ. En remplaçant (E_2) par l'équation (E_2') obtenue en « retranchant à (E_2) l'équation (E_1) multipliée par λ », on obtient le système

$$\left\{ \begin{array}{rcl} ax + by + cz & = & d \\ 0 & = & d' - \lambda d \end{array} \right. \qquad \begin{array}{c} (E_1) \\ (E_2') \end{array}$$

Suivant que $d' - \lambda d$ soit nul ou non, c'est-à-dire que $\mathscr{P}'$ soit égal à $\mathscr{P}$ ou non, on obtient que l'intersection $\mathscr{P} \cap \mathscr{P}'$ est soit le plan $\mathscr{P}$, soit l'ensemble vide.



https://www.geogebra.org/m/h8j4zmca
Deux plans parallèles (en rose et orange)

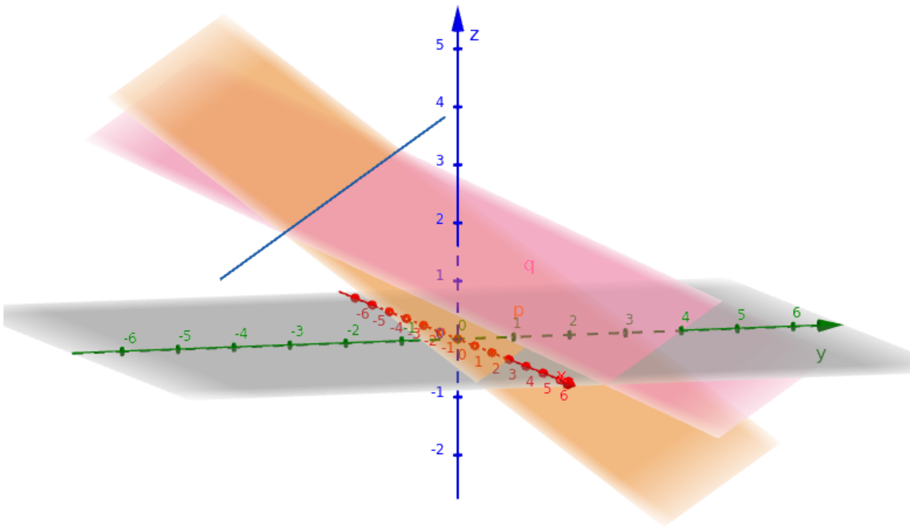
Supposons maintenant que $\mathscr{P}$ et $\mathscr{P}'$ ne sont pas parallèles. Il s'agit toujours de déterminer les solutions du système suivant :

$$\left\{\begin{array}{ccccccc} ax & + & by & + & cz & = & d \\ a'x & + & b'y & + & c'z & = & d' \end{array}\right. \qquad \begin{array}{c}(E_1)\\(E_2)\end{array}$$

On sait que $(a, b, c) \neq (0, 0, 0)$. Quitte à permuter le rôle des variables, on va supposer que $a \neq 0$. Notons $\lambda := \frac{a'}{a}$. Dans le système, on peut remplacer (E_2) par l'équation (E_2') obtenue en « retranchant à (E_2) l'équation (E_1) multipliée par λ » :

$$\left\{\begin{array}{ccccccc} ax & + & by & + & cz & = & d \\ & & b''y & + & c''z & = & d'' \end{array}\right. \qquad \begin{array}{c}(E_1)\\(E_2')\end{array}$$

Il n'est pas possible que l'on ait $b'' = c'' = 0$, puisque sinon $\mathscr{P}$ et $\mathscr{P}'$ seraient parallèles (cas précédent). Ainsi, si on considère « y » et « z » comme les coordonnées d'éléments $(y, z) \in \mathbf{R}^2$, on peut penser à l'équation (E_2') comme l'équation d'une droite dans $\mathbf{R}^2$, donc elle possède un paramétrage affine $(Y(t), Z(t))$ pour $t \in \mathbf{R}$. Autrement dit, si $(y, z) \in \mathbf{R}^2$, (y, z) vérifie (E_2') si et seulement s'il existe $t \in \mathbf{R}$ tel que $y = Y(t)$ et $z = Z(t)$. En posant $X(t) := \frac{d - bY(t) - cZ(t)}{a}$, on obtient que $\mathscr{P} \cap \mathscr{P}'$ est formé des (x, y, z) qui sont de la forme $(X(t), Y(t), Z(t))$ pour un certain $t \in \mathbf{R}$, bref, on a montré que $\mathscr{P} \cap \mathscr{P}'$ était une droite.



Deux plans sécants

III.3. Solutions de systèmes d'équations

Théorème III.3.1. *Soit $m \in \mathbf{N}$. On suppose donnés des coefficients réels arbitraires a_i, b_i, c_i, d_i pour $i \in \{1, \ldots, n\}$. On note (E_i) l'équation en trois variables « $a_i x + b_i y + c_i z = d_i$ ». On considère le système d'équations suivant :*

$$\left\{\begin{array}{ccccccclc} a_1 x & + & b_1 y & + & c_1 z & = & d_1 & & (E_1) \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_i x & + & b_i y & + & c_i z & = & d_i & & (E_i) \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_m x & + & b_m y & + & c_m z & = & d_m & & (E_m) \end{array}\right.$$

L'ensemble des solutions à ce système est soit :
- *l'espace $\mathbf{R}^3$ tout entier ;*
- *un plan ;*
- *une droite ;*
- *un singleton ;*
- *vide.*

Si l'ensemble des solutions est non vide, on vide qu'il est de dimension 3 si c'est $\mathbf{R}^3$, de dimension 2 si c'est un plan, de dimension 1 si c'est une droite et de dimension 0 si c'est un singleton.

On commence par remarquer que si une des équations (E_i) est telle que $a_i = b_i = c_i = 0$, alors soit $d_i \neq 0$ et alors l'ensemble des solutions est vide, soit $d_i = 0$, auquel cas on peut retirer cette équation du système sans changer l'ensemble des solutions. Ainsi, on se ramène au cas où chacune des équations (E_i) est l'équation d'un plan $\mathscr{P}_i$. On est ainsi ramené à montrer qu'une intersection de plans est de la forme voulue.

On procède par récurrence sur le nombre m de plans. Si $m = 0$, l'ensemble des solutions est bien sûr $\mathbf{R}^3$ tout entier. Si $m = 1$, l'intersection est le plan $\mathscr{P}_1$. Si $m = 2$, le théorème III.2.2.1 permet de conclure. Sinon, $m \geq 3$, et on peut noter $V = \mathscr{P}_1 \cap \cdots \cap \mathscr{P}_{m-1}$ l'ensemble des solutions du système suivant :

$$\left\{\begin{array}{ccccccclc} a_1x & + & b_1y & + & c_1z & = & d_1 & & (E_1) \\ \vdots & & \vdots & & \vdots & \vdots & & & \vdots \\ a_{m-1}x & + & b_{m-1}y & + & c_{m-1}z & = & d_{m-1} & & (E_{m-1}) \end{array}\right.$$

Comme ce système fait intervenir $m - 1$ équations, l'hypothèse de récurrence fait que V est de l'un des cinq types apparaissant dans l'énoncé du théorème. L'ensemble des solutions du système initial est $V \cap \mathscr{P}_m$. Examinons quelques-uns des différents cas :
- si $V = \mathbf{R}^3$, alors $V \cap \mathscr{P}_m = \mathscr{P}_m$ est un plan ;
- si V est plan, alors $V \cap \mathscr{P}_m$ est l'intersection de deux plans, donc d'après le theorème III.2.2.1 est soit vide, soit une droite, soit un plan.
- si V est un singleton $\{A\}$, $V \cap \mathscr{P}_m$ est contenu dans $V = \{A\}$, donc est soit le singleton $\{A\}$ si $A \in \mathscr{P}_m$ soit vide sinon.
- si V est vide, alors $V \cap \mathscr{P}_m$ est vide.

Il reste à traiter le cas où V est une droite. Supposons que V admette pour paramétrage affine $X(t) = x_0 + v_1t$, $Y(t) = y_0 + v_2t$, $Z(t) = z_0 + v_3t$. L'intersection $V \cap \mathscr{P}_m$ est formée des points $M(t) = (X(t), Y(t), Z(t))$ pour $t \in \mathbf{R}$ tel que $M(t)$ appartienne à $\mathscr{P}_m$ c'est-à-dire tel que :

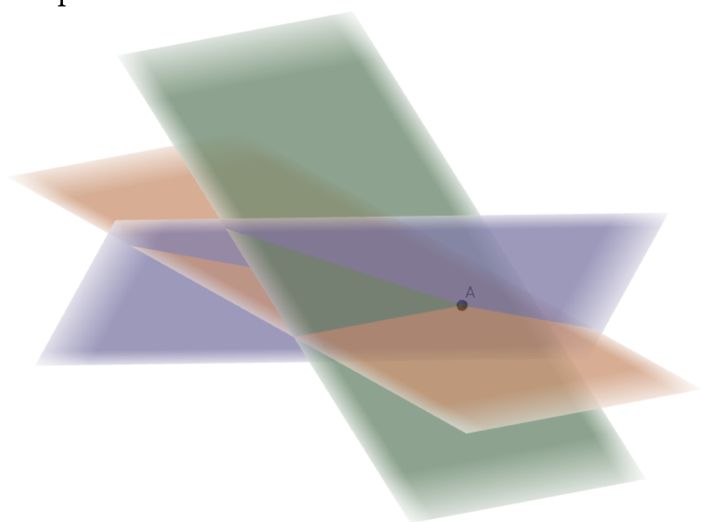
$$a_mX(t) + b_mY(t) + c_mZ(t) = d_m$$

Comme $X(t)$, $Y(t)$ et $Z(t)$ ont des expressions affines en t, cette condition peut se récrire :

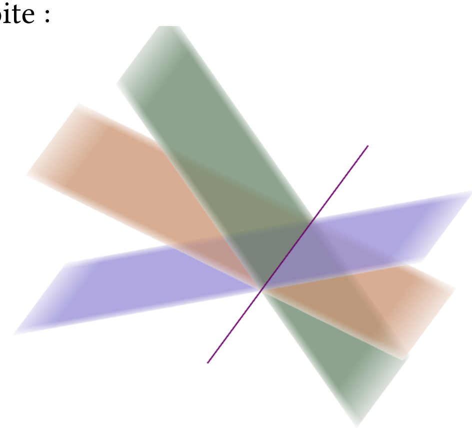
$$\lambda t = \mu$$

Si $\lambda \neq 0$, l'unique paramètre t qui convienne est $\frac{\mu}{\lambda}$ et donc l'intersection $V \cap \mathscr{P}_m$ est le singleton $\{M(\frac{\mu}{\lambda})\}$. Si $\lambda = 0$, soit $\mu \neq 0$ auquel cas l'intersection $V \cap \mathscr{P}_m$ est vide, soit $\mu = 0$ auquel cas $V \subset \mathscr{P}_m$ et alors $V \cap \mathscr{P}_m = V$ est une droite.

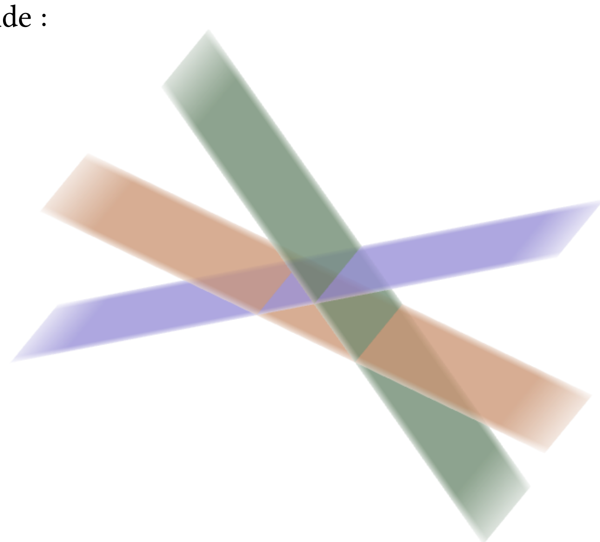
Voici quelques configurations possibles pour l'intersection de trois plans :
- L'intersection est un point :



— L'intersection est une droite :



— L'intersection est vide :



III.4. Initiation à l'algorithme du pivot de Gauss

On introduit ici les bases de l'algorithme du pivot de Gauss qui sera étudié de façon plus détaillée et dans un cadre plus général au chapitre V.

Considérons un système de m équations en trois varibles x, y et z comme dans l'énoncé du théorème III.3.1 :

$$\left\{\begin{array}{ccccccccc} a_1x & + & b_1y & + & c_1z & = & d_1 & \qquad & (E_1) \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_ix & + & b_iy & + & c_iz & = & d_i & & (E_i) \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_mx & + & b_my & + & c_mz & = & d_m & & (E_m) \end{array}\right.$$

Si tous les coefficients $a_1, \ldots, a_m$ sont nuls, la variable x n'intervient pas dans les équations, ce qui n'est pas très intéressant. On suppose que x intervient véritablement dans au moins une équation. Si $a_1 = 0$, quitte à échanger l'équation (E_1) avec une des autres équations, on peut se ramener au cas où $a_1 \neq 0$.

L'étape cruciale de l'algorithme du pivot de Gauss consiste à garder l'équation (E_1), mais à faire disparaître la variable x des autres équations. Ceci se fait en remplaçant l'équation (E_i) pour tout $i \in \{2, \ldots, m\}$, par « $(E_i) - \lambda_i(E_1)$ » où $\lambda_i = -\frac{a_i}{a_1}$. On obtient alors un nouveau système, équivalent au premier :

$$\left\{\begin{array}{ccccccccc} a_1x & + & b_1y & + & c_1z & = & d_1 & & (E_1) \\ & & b'_2y & + & c'_2z & = & d'_2 & & (E'_2) \\ & & \vdots & & \vdots & & \vdots & & \vdots \\ & & b'_my & + & c'_mz & = & d'_m & & (E'_m) \end{array}\right.$$

Dans les équations (E'_2) à (E'_m), il peut arriver que tous les coefficients devant y soient nuls : $b'_2 = \cdots ='_m = 0$, auquel cas on détermine facilement les éventuelles valeurs possibles pour z, et en tenant compte du fait que y peut être arbitraire, on détermine toutes les solutions en récrivant (E_1) sous la forme $x = \frac{d_1 - b_1y - c_1z}{a_1}$.

Supposons donc que la variable y apparaisse véritablement dans au moins des équations (E'_2) à (E'_m). Quitte à échanger si nécessaire (E'_2) avec une autre équation, on peut supposer que $b'_2 \neq 0$. On procède alors comme précédemment pour éliminer la variable y des équations (E'_3) à (E'_m) : pour $i \in \{3, \ldots, m\}$, on remplace (E'_i) par « $(E'_i) - \frac{b'_i}{b'_2}(E'_2)$ ».

$$\left\{\begin{array}{ccccccccc} a_1x & + & b_1y & + & c_1z & = & d_1 & & (E_1) \\ & & b'_2y & + & c'_2z & = & d'_2 & & (E'_2) \\ & & & & c''_3z & = & d'_3 & & (E''_3) \\ & & & & \vdots & & \vdots & & \vdots \\ & & & & c''_mz & = & d'_m & & (E''_m) \end{array}\right.$$

On examine alors les équations (E'_3) à (E'_m) qui ne font plus intervenir que la variable z. On remarque que l'ensemble des valeurs possibles pour z est soit vide, soit un singleton, soit $\mathbf{R}$ tout entier. Dans le cas où cet ensemble est non vide, on se ramène facilement à deux cas intéressants : celui où $m = 3$ et $c''_3 \neq 0$, et celui où $m = 2$.

Dans le cas le plus courant, on obtiendra un système de la forme suivante, avec $a_1 \neq 0$, $b'_2 \neq 0$ et $c''_3 \neq 0$. On dira qu'un tel système est triangulaire.

$$\left\{\begin{array}{ccccccccc} a_1x & + & b_1y & + & c_1z & = & d_1 & & (E_1) \\ & & b'_2y & + & c'_2z & = & d'_2 & & (E'_2) \\ & & & & c''_3z & = & d'_3 & & (E''_3) \end{array}\right.$$

Pour un tel système, il n'y a une unique solution (x, y, z). La valeur de z est déterminée par la troisième équation. Connaissant z, la deuxième équation ne laisse qu'une seule possibilité pour y. Enfin, connaissant y et z, il n'y a qu'une possibilité pour x.

Dans l'autre cas, le système est de la forme suivante, avec $a_1 \neq 0$ et $b'_2 \neq 0$:

$$\left\{\begin{array}{ccccccccc} a_1x & + & b_1y & + & c_1z & = & d_1 & & (E_1) \\ & & b'_2y & + & c'_2z & = & d'_2 & & (E'_2) \end{array}\right.$$

La deuxième équation permet d'exprimer y en fonction de z. En utilisant cette expression, la première équation permet finalement d'exprimer x en fonction de z. L'ensemble des solutions est alors une droite dont on obtient un paramétrage $(\alpha t + \beta, \alpha' t + \beta', t)$ pour $t \in \mathbf{R}$ avec des coefficients $\alpha, \alpha', \beta, \beta'$ appropriés.

IV. Matrices

Sommaire

IV.1. Définitions

Définition IV.1.1. *Si m et n sont deux entiers naturels, une matrice de taille $m \times n$ (on dit aussi matrice à m lignes et n colonnes) est un tableau de nombres réels $A = (a_{i,j})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$, que l'on représente ainsi :*

$$A = \begin{pmatrix} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} \\ a_{2,1} & \dots & a_{2,j} & \dots & a_{2,n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{i,1} & \dots & a_{i,j} & \dots & a_{i,n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ a_{m,1} & \dots & a_{m,j} & \dots & a_{m,n} \end{pmatrix}$$

Il est important de respecter la convention selon laquelle le premier indice i désigne le numéro de la ligne et le deuxième indice j celui de la colonne. Le nombre $a_{i,j}$ est le coefficient en position (i,j) de la matrice A. On note $M_{m,n}(\mathbf{R})$ l'ensemble des matrices de taille $m \times n$.

Définition IV.1.2. *Si $m \in \mathbf{N}$, Un vecteur-colonne de taille m est une matrice appartenant à $M_{m,1}(\mathbf{R})$, c'est-à-dire possédant une unique colonne :*

$$C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \in M_{m,1}(\mathbf{R})$$

Définition IV.1.3. *Si $n \in \mathbf{N}$, Un vecteur-ligne de taille n est une matrice appartenant à $M_{1,n}(\mathbf{R})$, c'est-à-dire possédant une unique ligne :*

$$L = \begin{pmatrix} l_1 & l_2 & \dots & l_n \end{pmatrix} \in M_{1,n}(\mathbf{R})$$

IV.2. Premières opérations

IV.2.1. Addition

Définition IV.2.1.1. *Soit $(m,n) \in \mathbf{N}^2$. Supposons que $A = (a_{i,j})_{i,j}$ et $B = (b_{i,j})_{i,j}$ soient deux matrices* ***de même taille*** *$m \times n$. Alors, on définit $A + B \in M_{m,n}(\mathbf{R})$ comme étant la matrice dont le coefficient en position (i,j) soit $a_{i,j} + b_{i,j}$ (pour tous $i \in \{1,\ldots,m\}$ et $j \in \{1,\ldots,n\}$) :*

$$A + B = \begin{pmatrix} a_{1,1}+b_{1,1} & \ldots & a_{1,j}+b_{1,j} & \ldots & a_{1,n}+b_{1,n} \\ a_{2,1}+b_{2,1} & \ldots & a_{2,j}+b_{2,j} & \ldots & a_{2,n}+b_{2,n} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ a_{i,1}+b_{i,1} & \ldots & a_{i,j}+b_{i,j} & \ldots & a_{i,n}+b_{i,n} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ a_{m,1}+b_{m,1} & \ldots & a_{m,j}+b_{m,j} & \ldots & a_{m,n}+b_{m,n} \end{pmatrix}$$

Comme on peut considérer l'opération $+$ définie ci-dessus comme une application $+ \colon M_{m,n}(\mathbf{R}) \times M_{m,n}(\mathbf{R}) \to M_{m,n}(\mathbf{R})$, on dit que $+$ est une loi de composition interne sur $M_{m,n}(\mathbf{R})$.

Définition IV.2.1.2. *Soit $(m,n) \in \mathbf{N}^2$. La matrice nulle de taille $m \times n$ est la matrice dont tous les coefficients sont nuls. On peut la noter 0 ou $\mathbf{0}$ à condition que le nombre de lignes et de colonnes soit clair dans le contexte, on encore $\mathbf{0}_{m,n}$:*

$$\mathbf{0}_{m,n} = \begin{pmatrix} 0 & \ldots & 0 \\ \vdots & \ldots & \vdots \\ 0 & \ldots & 0 \end{pmatrix}$$

Proposition IV.2.1.3. *Soit $(m,n) \in \mathbf{N}^2$. La loi de composition interne $+$ sur $M_{m,n}(\mathbf{R})$ vérifie les propriétés suivantes :*

(i) Pour tous A, B et C, éléments de $M_{m,n}(\mathbf{R})$, on a :

$$(A + B) + C = A + (B + C)$$

(On dit que l'opération $+$ sur $M_{m,n}(\mathbf{R})$ est associative. Puisque l'ordre des opérations n'est pas important ici, on pourra écrire simplement $A + B + C$.)

(ii) Pour tout $A \in M_{m,n}(\mathbf{R})$, on a :

$$A + \mathbf{0}_{m,n} = \mathbf{0}_{m,n} + A = A$$

(On dit que $\mathbf{0}_{m,n}$ est l'élément neutre de la loi $+$.)

(iii) Pour tout $A \in M_{m,n}(\mathbf{R})$, il existe une (unique) matrice $A' \in M_{m,n}(\mathbf{R})$ telle que $A + A' = A' + A = \mathbf{0}_{m,n}$. (Cette matrice A' sera notée $-A$.)

(iv) Pour tous A et B, éléments de $M_{m,n}(\mathbf{R})$, on a :

$$A + B = B + A$$

(On dit que l'opération $+$ sur $M_{m,n}(\mathbf{R})$ est commutative.)

Montrons (i). Si on note $a_{i,j}$, $b_{i,j}$ et $c_{i,j}$ les coefficients de matrices A, B et C, alors on vérifie immédiatement que le coefficient en position (i,j) de $(A + B) + C$ est $(a_{i,j} + b_{i,j}) + c_{i,j}$, qui est égal à $a_{i,j} + (b_{i,j} + c_{i,j})$ (propriété connue pour l'addition de nombres réels), qui est le coefficient en position (i, j) de la matrice $A + (B + C)$. Les deux matrices $(A + B) + C$ et $A + (B + C)$ ont les mêmes coefficients, donc sont égales. Les propriétés (ii) et (iv) ne démontrent essentiellement de la même façon.

Il reste à démontrer (iii). Si $A = (a_{i,j})_{i,j}$, pour tout (i, j), posons $a'_{i,j} := -a_{i,j}$. On a alors $a_{i,j} + a'_{i,j} = a'_{i,j} + a_{i,j} = 0$ pour tout (i, j), ce qui revient à dire que $A + A' = A' + A = \mathbf{0}_{m,n}$.

Remarque IV.2.1.4. *Les propriétés (i), (ii) et (iii) de la proposition précédente signifient que $M_{m,n}(\mathbf{R})$ muni de la loi de composition interne $+$ est un* ***groupe****, et la propriété (iv) signifie que ce groupe est un* ***groupe commutatif*** *(on dit aussi « groupe abélien »).*

IV.2.2. Multiplication par un scalaire

(Pour nous, un scalaire est simplement un nombre réel.)

Définition IV.2.2.1. *Soit $(m, n) \in \mathbf{N}^2$. Supposons que $A = (a_{i,j})_{i,j} \in M_{m,n}(\mathbf{R})$ et que $\lambda \in \mathbf{R}$. On note λA (ou $\lambda \cdot A$) la matrice appartenant à $M_{m,n}(\mathbf{R}$ dont le coefficient en position (i, j) est $\lambda a_{i,j}$:*

$$\lambda A = \begin{pmatrix} \lambda a_{1,1} & \dots & \lambda a_{1,j} & \dots & \lambda a_{1,n} \\ \lambda a_{2,1} & \dots & \lambda a_{2,j} & \dots & \lambda a_{2,n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ \lambda a_{i,1} & \dots & \lambda a_{i,j} & \dots & \lambda a_{i,n} \\ \vdots & \dots & \vdots & \dots & \vdots \\ \lambda a_{m,1} & \dots & \lambda a_{m,j} & \dots & \lambda a_{m,n} \end{pmatrix}$$

Proposition IV.2.2.2. *Soit $(m, n) \in \mathbf{N}^2$. Pour tous $A \in M_{m,n}(\mathbf{R})$, $B \in M_{m,n}(\mathbf{R})$, $\lambda \in \mathbf{R}$ et $\mu \in \mathbf{R}$, on a les formules suivantes :*
(i) $\lambda \cdot (A + B) = \lambda \cdot A + \lambda \cdot B$;
(ii) $\lambda \cdot \mathbf{0}_{m,n} = \mathbf{0}_{m,n}$;
(iii) $(\lambda + \mu) \cdot A = \lambda \cdot A + \mu \cdot A$.
(iv) $0 \cdot A = \mathbf{0}_{m,n}$;
(v) $1 \cdot A = A$;
(vi) $\lambda \cdot (\mu \cdot A) = (\lambda\mu) \cdot A$;

IV.2.3. Structure d'espace vectoriel

L'ensemble $M_{m,n}(\mathbf{R})$ a été muni d'une loi de composition interne $+$ vérifiant les propriétés de la proposition IV.2.1.3 et d'une loi de multiplication par un scalaire $\mathbf{R} \times M_{m,n}(\mathbf{R}) \to M_{m,n}(\mathbf{R})$ qui à (λ, A) associe $\lambda \cdot A$ vérifiant les propriétés de la proposition IV.2.2.2. On dit alors que l'ensemble $M_{m,n}(\mathbf{R})$ a été muni d'**une structure d'espace vectoriel (réel)**.

En particulier, si $n = 1$ et que m est quelconque, $M_{m,1}(\mathbf{R})$ est l'ensemble des vecteurs-colonnes de taille m, que l'on peut identifier à $\mathbf{R}^m$. L'ensemble $M_{m,1}(\mathbf{R})$ est alors un espace vectoriel réel. On peut ainsi considérer qu'un vecteur est un élément de $M_{m,1}(\mathbf{R})$, on a défini une loi d'addition des vecteurs et une loi de multiplication d'un vecteur par un scalaire, et ces lois vérifient les règles de calcul qui ont été établies plus haut.

Définition IV.2.4. *Soit $m \in \mathbf{R}$. On suppose donnés $C_1, \dots, C_n$ n vecteurs-colonnes de taille m, c'est-à-dire des éléments de $M_{m,1}(\mathbf{R})$. Soient $\lambda_1, \dots, \lambda_n$ des nombres réels. La combinaison linéaire des vecteurs(-colonnes) $C_1, \dots, C_n$ affectés des coefficients $\lambda_1, \dots, \lambda_n$ est le vecteur-colonne suivant :*

$$\lambda_1 C_1 + \cdots + \lambda_n C_n \in M_{m,1}(\mathbf{R})$$

IV.3. Multiplication des matrices

IV.3.1. Multiplication d'une matrice et d'un vecteur-colonne

Définition IV.3.1.1. *Supposons que $A \in M_{m,n}(\mathbf{R})$ et $X \in M_{n,1}(\mathbf{R})$. Nous allons définir le produit $AX \in M_{m,1}(\mathbf{R})$. Pour cela, on décompose A en ses n vecteurs-colonnes (notés $C_1, \dots, C_n$) :*

$$A = \begin{pmatrix} C_1 & C_2 & \dots & C_n \end{pmatrix}$$

On note $x_1, \dots, x_n$ les coefficients de X :

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

Le produit AX est alors défini comme étant la combinaison linéaire suivante :

$$AX = x_1 C_1 + \cdots + x_n C_n \in M_{m,1}(\mathbf{R})$$

Avec les notations de la définition ci-dessus, on peut aussi noter $a_{i,j}$ les coefficients de A, de façon à ce que pour tout $j \in \{1, \dots, n\}$, on ait :

$$C_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{m,j} \end{pmatrix}$$

On a alors :

$$AX = \begin{pmatrix} a_{1,1}x_1 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{i,1}x_1 + \cdots + a_{i,n}x_n \\ \vdots \\ a_{m,1}x_1 + \cdots + a_{m,n}x_n \end{pmatrix}$$

Remarque IV.3.1.2. *Alternativement, si on note $L_1, \dots, L_m$ les m vecteurs-lignes de $A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix}$, alors pour tout $i \in \mathbf{N}$, $L_i = \begin{pmatrix} a_{i,1} & a_{i,2} & \dots & a_{i,n} \end{pmatrix}$ et le produit $L_i X$ a un sens* *et est une matrice 1×1 que l'on peut identifier au nombre réel $a_{i,1}x_1 + \cdots + a_{i,n}x_n$. On peut ainsi récrire chacun des coefficients de la matrice AX comme le produit d'un vecteur-ligne avec un vecteur-colonne :*

$$MX = \begin{pmatrix} L_1X \\ L_2X \\ \vdots \\ L_mX \end{pmatrix}$$

IV.3.2. Multiplication de deux matrices

La définition suivante ne sera utilisée qu'à partir du chapitre VIII. Pour le moment, il est plus important de bien comprendre la définition IV.3.1.1.

Définition IV.3.2.1. *Soit $(m, n, p) \in \mathbf{N}^3$. Supposons que $A \in M_{m,n}(\mathbf{R})$ et $B \in M_{n,p}(\mathbf{R})$. On décompose B en ses p vecteurs-colonnes (notés $B_1, \ldots, B_p$) :*

$$B = \begin{pmatrix} B_1 & B_2 & \ldots & B_p \end{pmatrix}$$

Pour tout $j \in \{1, \ldots, p\}$, $B_p \in M_{n,1}(\mathbf{R})$ est un vecteur-colonne et la définition IV.3.1.1 donne un vecteur-colonne $AB_j \in M_{m,1}(\mathbf{R})$. La matrice AB est définie comme étant la matrice dont les vecteurs-colonnes sont ces éléments AB_j :

$$AB = \begin{pmatrix} AB_1 & AB_2 & \ldots & AB_p \end{pmatrix} \in M_{m,p}(\mathbf{R})$$

Proposition IV.3.2.2. *Soit $(m, n, p) \in \mathbf{N}^3$. Supposons que A et A' sont deux matrices de taille $m \times n$, que B et B' sont matrices de taille $n \times p$ et que $\lambda \in \mathbf{R}$. Les formules suivantes sont vérifiées :*

(i) $A(B + B') = AB + AB'$ *;*
(ii) $(A + A')B = AB + A'B$ *;*
(iii) $(\lambda A)B = A(\lambda B) = \lambda(AB)$.

IV.4. Interprétation matricielle des systèmes linéaires

Fixons $(m, n) \in \mathbf{N}^2$. Fixons $A \in M_{m,n}(\mathbf{R})$ et $B = M_{m,1}(\mathbf{R})$. Introduisons des notations pour les coefficients de ces matrices :

$$A = \begin{pmatrix} a_{1,1} & \ldots & a_{1,j} & \ldots & a_{1,n} \\ a_{2,1} & \ldots & a_{2,j} & \ldots & a_{2,n} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ a_{i,1} & \ldots & a_{i,j} & \ldots & a_{i,n} \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ a_{m,1} & \ldots & a_{m,j} & \ldots & a_{m,n} \end{pmatrix} \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{pmatrix}$$

La donnée d'un n-uplet $(x_1, \ldots, x_n) \in \mathbf{R}^n$ équivaut à la donnée d'un vecteur-colonne $X \in M_{n,1}(\mathbf{R})$:

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

Avec ces notations, la matrice $AX \in M_{m,1}(\mathbf{R})$ est le vecteur-colonne suivant :

$$AX = \begin{pmatrix} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n \\ \vdots \\ a_{i,1}x_1 + a_{i,2}x_2 + \cdots + a_{i,n}x_n \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n \end{pmatrix}$$

On obtient ainsi que si $(x_1, \ldots, x_n) \in \mathbf{R}^n$, alors $(x_1, \ldots, x_n)$ est solution du système d'équations linéaires

$$\left\{ \begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \ldots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \ldots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & = & \vdots \\ a_{i,1}x_1 & + & a_{i,2}x_2 & + & \ldots & + & a_{i,n}x_n & = & b_i \\ \vdots & & \vdots & & \vdots & & \vdots & = & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \ldots & + & a_{m,n}x_n & = & b_m \end{array} \right.$$

si et seulement si le vecteur-colonne X défini ci-dessus vérifie l'équation matricielle $AX = B$.

V. Algorithme du pivot de Gauss

Sommaire

V.1. Matrice augmentée d'un système d'équations linéaires

On a vu au §IV.4 qu'un système d'équations de m équations linéaires en n variables pouvait se récrire sous la forme d'une équation matricielle $AX = B$ où $A \in M_{m,n}(\mathbf{R})$ et $B \in M_{m,1}(\mathbf{R})$, l'inconnue étant $X \in M_{n,1}(\mathbf{R})$.

Précisément, si $A = (a_{i,j})_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$, et que l'on introduit les notations suivantes :

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Alors l'équation matricielle $AX = B$ équivaut au système suivant :

$$\left\{\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & = & \vdots \\ a_{i,1}x_1 & + & a_{i,2}x_2 & + & \dots & + & a_{i,n}x_n & = & b_i \\ \vdots & & \vdots & & \vdots & & \vdots & = & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \dots & + & a_{m,n}x_n & = & b_m \end{array}\right.$$

Définition V.1.1. *La matrice augmentée du système d'équations linéaire ci-dessus est la matrice (que l'on notera ici $\widehat{A}$) de taille $m \times (n+1)$ suivante :*

$$\widehat{A} = \left(\begin{array}{ccccc|c} a_{1,1} & \dots & a_{1,j} & \dots & a_{1,n} & b_1 \\ a_{2,1} & \dots & a_{2,j} & \dots & a_{2,n} & b_2 \\ \vdots & \dots & \vdots & \dots & \vdots & \vdots \\ a_{i,1} & \dots & a_{i,j} & \dots & a_{i,n} & b_i \\ \vdots & \dots & \vdots & \dots & \vdots & \vdots \\ a_{m,1} & \dots & a_{m,j} & \dots & a_{m,n} & b_m \end{array}\right) \in M_{m,n+1}(\mathbf{R})$$

Cette matrice est obtenue en ajoutant à A une colonne supplémentaire qui est formée du vecteur-colonne B. On écrit aussi $\widehat{A} = \left(\begin{array}{c|c} A & B \end{array} \right)$, qui correspond à ce qu'on appelle une représentation de la matrice $\widehat{A}$ « par blocs ». (Certains auteurs appellent cette matrice la « matrice complète » du système. La ligne verticale qui sert à séparer visuellement la dernière colonne est optionnelle dans l'écriture, mais je conseille de la garder pour bien distinguer les rôles des différents coefficients.)

Remarque V.1.2. *Si on note* $\widehat{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ \hline -1 \end{pmatrix} = \begin{pmatrix} X \\ \hline -1 \end{pmatrix} \in M_{n+1,1}(\mathbf{R})$, *alors on peut remarquer que* $\widehat{A}\widehat{X} = AX - B$, *de sorte que* $AX = B$ *si et seulement si* $\widehat{A}\widehat{X} = 0$.

Exemple V.1.3. *Considérons le système d'équations linéaires suivant en les variables x, y et z :*

$$\left\{ \begin{array}{rcrcrcr} x & + & 2y & - & 3y & = & 3 \\ -2x & - & 3y & + & z & = & -2 \\ x & & & + & 8z & = & 0 \end{array} \right.$$

La matrice augmentée de ce système est

$$\widehat{A} = \left(\begin{array}{rrr|r} 1 & 2 & -3 & 3 \\ -2 & -3 & 1 & -2 \\ 1 & 0 & 8 & 0 \end{array} \right)$$

Dans ce système d'équations, les variables s'appellent x, y et z au lieu de x_1, x_2 et x_3, mais cela n'a pas d'importance. (Néanmoins, il convient d'avoir fixé un ordre sur les variables.)

V.2. Opérations élémentaires sur les lignes

Définition V.2.1. *Si $\widehat{A}$ est la matrice augmentée d'un système de m équations linéaires en n variables, on notera ici* $\mathrm{Sol}(\widehat{A}) \subset \mathbf{R}^n$ *l'ensemble des* $(x_1, \ldots, x_n) \in \mathbf{R}^n$ *qui sont solutions du système.*

Définition V.2.2. *Supposons que $\widehat{A}$ et $\widehat{A}'$ soient les matrices augmentées associés à deux systèmes de m équations linéaires en n variables. On dira que les matrices augmentées $\widehat{A}$ et $\widehat{A}'$ sont équivalentes (ce qui sera noté $\widehat{A} \sim \widehat{A}'$) si* $\mathrm{Sol}(\widehat{A}) = \mathrm{Sol}(\widehat{A}')$).

(Il existe plusieurs notions d'« équivalences » sur les matrices. Nous n'utiliserons que celle-ci dans ce cours.)

Définition V.2.3. *Une opération élémentaire sur les lignes d'une matrice M à m lignes est décrite en termes des vecteurs-lignes $L_1, \ldots, L_m$ de M. Elles sont de trois types :*

(i) Échanger deux lignes, ce qui est noté $L_i \longleftrightarrow L_j$ (pour $i \neq j$);

(ii) Multiplier une ligne par un scalaire non nul, ce qui est noté $L_i \longleftarrow \lambda L_i$ pour $\lambda \in \mathbf{R}^\times$ et $i \in \{1, \ldots, m\}$;

(iii) Retrancher à une ligne un certain multiple d'une autre ligne, ce qui est noté $L_i \longleftarrow L_i - \lambda L_j$ pour $i \neq j$ et $\lambda \in \mathbf{R}^\times$.

Proposition V.2.4. *Soit $\widehat{A}$ la matrice augmentée d'un système d'équations linéaires. Notons $\widehat{A}'$ une matrice obtenue à partir de $\widehat{A}$ en effectuant une opération élémentaire sur les lignes (cf. définition V.2.3). Alors $\widehat{A} \sim \widehat{A}'$.*

Si $(x_1, \ldots, x_n) \in \mathrm{Sol}(\widehat{A})$, il est évident que $(x_1, \ldots, x_n) \in \mathrm{Sol}(\widehat{A}')$, ce qui montre l'inclusion $\mathrm{Sol}(\widehat{A}) \subset \mathrm{Sol}(\widehat{A}')$.

Pour montrer l'inclusion inverse, il suffit de montrer que réciproquement $\widehat{A}$ peut être obtenue à partir de $\widehat{A}'$ en effectuant une opération élémentaire sur les lignes. C'est évident pour les opérations de type (i) ou (ii). Concernant (iii), si $\widehat{A}'$ est obtenue en effectuant l'opération $L_i \longleftarrow L_i - \lambda L_j$ sur la matrice $\widehat{A}$, alors $\widehat{A}$ peut être réobtenue en effectuant l'opération $L_i \longleftarrow L_i + \lambda L_j$. En effet, notons plus précisément $L_1, \ldots, L_m$ les vecteurs-lignes de $\widehat{A}$ et En effet, notons plus précisément $L'_1, \ldots, L'_m$ ceux de $\widehat{A}'$. On a bien sûr $L'_k = L_k$ pour tout k différent de i, et on a par définition $L'_i = L_i - \lambda L_j$, et on a bien $L_i = L'_i + \lambda L_j = L'_i + \lambda L'_j$.

V.3. Mise sous forme échelonnée

Définition V.3.1. *Soit $A = (a_{i,j}) \in M_{m,n}(\mathbf{R})$ une matrice dont on note $L_1, \ldots, L_m$ les vecteurs-lignes. On dira que A est échelonnée (ou échelonnée en lignes) s'il existe un entier $r \in \{0, \ldots, m\}$ tel que les lignes $L_1, \ldots, L_r$ soient non nulles et que $L_{r+1} = \cdots = L_m = \mathbf{0}$, et que si pour tout $i \in \{1, \ldots, r\}$ on note c_i le plus petit entier tel que $a_{i,x_i} \neq 0$, alors on a $c_1 < c_2 < \cdots < c_r$. Les cases (i, c_i) de la matrice A pour $1 \leq i \leq r$ seront appelées* pivots *(ou* coefficients principaux*).*

Exemple V.3.2. *Les matrices suivantes sont échelonnées :*

$$\begin{pmatrix} 2 & 3 & 0 \\ 0 & 1 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Les pivots ont été indiqués en rouge.

Théorème V.3.3. *Soit $A \in M_{m,n}(\mathbf{R})$. En partant de A et en effectuant des opérations élémentaires sur les lignes, il est possible d'obtenir une matrice échelonnée.*

Définition V.3.4 (Algorithme du pivot de Gauss : mise sous forme échelonnée)**.**

Soit $A \in M_{m,n}(\mathbf{R})$. (La matrice A va être modifiée au cours de l'avancement de l'algorithme.)

Première étape *Si A est la matrice nulle ou si $m = 1$, on a terminé ;*

Deuxième étape *On note c le plus petit indice tel que la c-ième colonne de A soit non nulle ;*

Troisième étape *On choisit un indice i tel que $a_{i,c} \neq 0$, et si $i \neq 1$ on effectue l'opération $L_1 \longleftrightarrow L_i$;*

Quatrième étape *La matrice courante A vérifie maintenant $a_{1,c} \neq 0$ (ce coefficient est appelé* pivot*), et pour tout $i \geq 2$, on effectue les opérations $L_i \longleftarrow L_i - \frac{a_{i,c}}{a_{1,c}} L_1$.*

Cinquième étape *À partir de maintenant, on ne touche plus à la première ligne, et on applique l'algorithme depuis le début à la matrice constituée des vecteurs-lignes $L_2, \ldots, L_m$ de la matrice A.*

Il est clair qu'à la fin de la quatrième étape, la matrice est de la forme suivante :

$$\begin{pmatrix} 0 & \dots & 0 & \blacksquare & \star & \dots & \star \\ \hline 0 & \dots & 0 & 0 & \star & \dots & \star \\ \vdots & \vdots & \vdots & \vdots & \star & \dots & \star \\ 0 & \dots & 0 & 0 & \star & \dots & \star \end{pmatrix}$$

La convention que nous utiliserons est que $\star$ peut être remplacé par n'importe quel coefficient, tandis que $\blacksquare$ désigne un coefficient qui est non nul. Ce coefficient $\blacksquare$ se trouve ici sur la c-ième colonne. La première ligne ne changera pas jusqu'à la fin. Les seules opérations élémentaires qui seront effectuées concerneront les lignes $L_2, \dots, L_m$, et donc, il est évident que l'on conservera la propriété selon laquelle les lignes $L_2, \dots, L_m$ ont c zéros sur la gauche.

On peut démontrer que l'algorithme termine et renvoie un résultat correct par récurrence sur le nombre de lignes m de la matrice. Si $m \in \{0, 1\}$, c'est évident. Pour démontrer le résultat pour une certaine matrice ayant $m \geq 2$, on peut supposer que l'algorithme termine et renvoie un résultat correct pour des matrices ayant $m - 1$ lignes. On peut donc supposer que la *cinquième étape* va effectivement mettre sous forme échelonnée la matrice obtenue en ignorant la première ligne. Comme la matrice obtenue conserve la forme générale discutée plus haut, il est évident que si à la fin la matrice formée des lignes $L_2, \dots, L_m$ est échelonnée, alors la matrice formée des lignes $L_1, \dots, L_m$ est échelonnée.

Théorème V.3.5. *Soit $\widehat{A} \in M_{m,n+1}(\mathbf{R})$ la matrice augmentée d'un système de m équations linéaires en n variables. Si on note $A \in M_{m,n}(\mathbf{R})$ et $B \in M_{m,1}(\mathbf{R})$ les matrices telles que $\widehat{A} = \left(\begin{array}{c|c} A & B \end{array} \right)$, il est possible d'effectuer des opérations élémentaires sur les lignes de $\widehat{A}$ de façon à obtenir $\widehat{A} \sim \widehat{A}'$ où $\widehat{A}' = \left(\begin{array}{c|c} A' & B' \end{array} \right)$ avec $A' \in M_{m,n}(\mathbf{R})$* ***échelonnée*** *et $B' \in M_{m,1}(\mathbf{R})$. (On dira que le système d'équations linéaires associé à $\widehat{A}'$ est échelonné.)*

Il suffit en effet d'appliquer aux lignes de $\widehat{A}$ les opérations élémentaires données par l'algorithme du pivot de Gauss permettant de mettre sous forme échelonnée la matrice A.

Exemple V.3.6. *Dans le cas de l'exemple V.1.3, l'algorithme du pivot de Gauss donne ceci :*

$$\begin{aligned} \widehat{A} &= \left(\begin{array}{rrr|r} 1 & 2 & -3 & 3 \\ -2 & -3 & 1 & -2 \\ 1 & 0 & 8 & 0 \end{array} \right) \\ \underset{\sim}{\left[\begin{array}{l} L_2 \longleftarrow L_2 + 2L_1 \\ L_3 \longleftarrow L_3 - L_1 \end{array} \right]} & \left(\begin{array}{rrr|r} 1 & 2 & -3 & 3 \\ 0 & 1 & -5 & 4 \\ 0 & -2 & 11 & -3 \end{array} \right) \\ \underset{\sim}{\left[L_3 \longleftarrow L_3 + 2L_1 \right]} & \left(\begin{array}{rrr|r} 1 & 2 & -3 & 3 \\ 0 & 1 & -5 & 4 \\ 0 & 0 & 1 & 5 \end{array} \right) \end{aligned}$$

Ceci implique que le système d'équation associé à $\widehat{A}$ est équivalent au système suivant :

$$\left\{ \begin{array}{rcrcrcr} x & + & 2y & - & 3y & = & 3 \\ & & y & - & 5z & = & 4 \\ & & & & z & = & 5 \end{array} \right.$$

En partant de la dernière équation puis en remontant, on peut ainsi déterminer z, puis y et enfin x. Ceci sera interprété dans la section suivante par la mise sous forme échelonnée réduite.

V.4. Mise sous forme échelonnée réduite

Définition V.4.1. *Soit $A = (a_{i,j}) \in M_{m,n}(\mathbf{R})$ une matrice échelonnée. On dit que A est échelonnée réduite si les pivots de A sont égaux à 1 et que tous les coefficients placés au-dessus d'un pivot sont nuls. Autrement dit, avec les notations de la définition V.3.1, pour tout $i \in \{1, \dots, r\}$, on a $a_{i,c_i} = 1$ et pour tout $i' \in \{1, \dots, i-1\}$, on a $a_{i',c_i} = 0$.*

Théorème V.4.2. *Soit $A \in M_{m,n}(\mathbf{R})$. En partant de A et en effectuant des opérations élémentaires sur les lignes, il est possible d'obtenir une matrice échelonnée réduite A'.*

Pour démontrer ce théorème, on peut supposer que A est échelonnée : en effet, si ce n'est pas le cas, on peut commencer par mettre A sous forme échelonnée en utilisant l'algorithme de la définition V.3.4.

Exemple V.4.3. *Si la matrice échelonnée A est de la forme suivante,*

$$\begin{pmatrix} 0 & \blacksquare & \star & \star & \star & \star & \star \\ 0 & 0 & 0 & \blacksquare & \star & \star & \star \\ 0 & 0 & 0 & 0 & \blacksquare & \star & \star \\ 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare \end{pmatrix}$$

on souhaite, par des opérations élémentaires, obtenir une matrice de la forme :

$$\begin{pmatrix} 0 & 1 & \star & 0 & 0 & \star & 0 \\ 0 & 0 & 0 & 1 & 0 & \star & 0 \\ 0 & 0 & 0 & 0 & 1 & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Pour réaliser cela, il suffit d'appliquer l'algorithme de la définition suivante :

Définition V.4.4 (Algorithme du pivot de Gauss : mise sous forme échelonnée réduite).
Soit $A \in M_{m,n}(\mathbf{R})$ une matrice échelonnée. (On utilise les notations de la définition V.3.1.)
Première étape *Pour tout $i \in \{1, \dots, r\}$, on effectue l'opération élémentaire $L_i \longleftarrow \frac{1}{a_{i,c_i}} L_i$.*
Deuxième étape *(Maintenant, $a_{i,c_i} = 1$ pour tout $i \in \{1, \dots, r\}$.) On effectue les opérations suivantes :*
- *Pour i allant de r à 2 (en décroissant)*
  - *Pour j allant de 1 à $i-1$, on fait $L_j \longleftarrow L_j - a_{j,c_i} L_i$.*

Les opérations élémentaires qui sont faites dans cet algorithme ne changent pas le caractère échelonné de la matrice et ne changent pas non plus la position des pivots. C'est évident pour les opérations de la *première étape* qui font que les pivots deviennent tous égaux à 1. Ensuite, il est important de noter que les opérations qui sont faites sont de la forme $L_j \longleftarrow L_j - \lambda L_i$ avec $i > j$, ce qui permet de conserver le caractère échelonné, sans changer les pivots.

Ensuite, concernant la boucle « Pour i allant de r à 2 (en décroissant) », on peut introduire l'invariant de boucle suivant :

(H_i) Pour tout entier k tel que $i < k \leq r$, les coefficients a_{j,c_k} sont nuls pour tout $j \in \{1, \dots, k-1\}$.

Cette propriété (H_i) signifie que les coefficients au-dessus des pivots a_{k,c_k} sont nuls pour $k \in \{i+1, \dots, r\}$. Cette condition est vérifiée en entrant dans la boucle, puisque (H_r) est évidemment vérifiée. On vérifie facilement que si au début de l'itération correspondant à i, la propriété (H_i) est vraie, alors la propriété (H_{i-1}) sera vraie à l'itération suivante. À la sortie de la boucle la propriété (H_1) sera donc vérifiée, et cela signifie précisément que la matrice (dont les pivots vallent 1) est échelonnée réduite.

Remarque V.4.5. *Le théorème V.4.2 énonce l'existence d'une matrice échelonnée réduite A' qu'il est possible d'obtenir à partir d'une matrice quelconque A en effectuant des opérations élémentaires sur les lignes. Il est en fait possible de montrer que A' est unique, mais c'est un théorème un peu plus difficile, et dont on n'aura pas à faire usage dans la suite de ce cours.*

Théorème V.4.6. *Soit $\widehat{A} \in M_{m,n+1}(\mathbf{R})$ la matrice augmentée d'un système de m équations linéaires en n variables. Si on note $A \in M_{m,n}(\mathbf{R})$ et $B \in M_{m,1}(\mathbf{R})$ les matrices telles que $\widehat{A} = \left(\begin{array}{c|c} A & B \end{array} \right)$, il est possible d'effectuer des opérations élémentaires sur les lignes de $\widehat{A}$ de façon à obtenir $\widehat{A} \sim \widehat{A}'$ où $\widehat{A}' = \left(\begin{array}{c|c} A' & B' \end{array} \right)$ avec $A' \in M_{m,n}(\mathbf{R})$ **échelonnée réduite** et $B' \in M_{m,1}(\mathbf{R})$.*

D'après le théorème V.3.5, on sait déjà que l'on peut s'arranger pour que A soit échelonnée, et alors en appliquant les mêmes opérations à $\widehat{A}$ que celles qui permettent de mettre A sous forme échelonnée réduite, on obtient le résultat.

Exemple V.4.7. *Dans le cas de l'exemple V.1.3, on sait depuis V.3.6 que la matrice $\widehat{A}$ est équivalente à :*

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 3 \\ 0 & 1 & -5 & 4 \\ 0 & 0 & 1 & 5 \end{array}\right) \underset{\sim}{\left[\begin{array}{c} L_1 \longleftarrow L_1+3L_3 \\ L_2 \longleftarrow L_2+5L_3 \end{array}\right]} \left(\begin{array}{ccc|c} 1 & 2 & 0 & 18 \\ 0 & 1 & 0 & 29 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

$$\underset{\sim}{\left[L_1 \longleftarrow L_1-2L_2 \right]} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -40 \\ 0 & 1 & 0 & 29 \\ 0 & 0 & 1 & 5 \end{array}\right)$$

Ceci implique que le système d'équation associé à $\widehat{A}$ est équivalent au système suivant :

$$\left\{ \begin{array}{rcr} x & = & -40 \\ y & = & 29 \\ z & = & 5 \end{array} \right.$$

L'unique solution au système d'équations linéaires de l'exemple V.1.3 est donc $(-40, 29, 5)$.

V.5. Description des solutions

Dans ce paragraphe, on suppose que $\widehat{A} = \left(\begin{array}{c|c} A & B \end{array} \right)$ est la matrice augmentée (de taille $m \times (n+1)$) d'un système d'équations linéaires telle que $A \in M_{m,n}(\mathbf{R})$ soit **échelonnée** et $B \in M_{n,1}(\mathbf{R})$.

On reprend les notations de la définition V.3.1. En particulier, on note $L_1, \ldots, L_m$ les lignes de la matrice A : chacune des r premières lignes contient un pivot (égal à 1) à la position (i, c_i), tandis que les lignes $L_{r+1}, \ldots, L_m$ sont nulles.

Définition V.5.1. *On dit que le système d'équations linéaires associé à $\widehat{A}$ est incompatible s'il existe $i \in \{r+1, \ldots, m\}$ tel que $b_i \neq 0$. On dit qu'il est compatible sinon.*

Si le système est incompatible, il existe un indice i tel que la i-ième équation soit de la forme « $0 = b_i$ ». Comme $b_i \neq 0$, cette équation n'est jamais satisfaite. L'ensemble des solutions d'un système d'équations incompatible est donc vide.

Exemple V.5.2. *Le système d'équations linéaires associé à la matrice augmentée suivante est incompatible :*

$$\widehat{A} = \left(\begin{array}{ccc|c} 1 & 2 & 0 & 5 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 2 \end{array}\right)$$

Nous allons maintenant décrire les solutions dans le cas où le système est compatible, ce qui nécessite d'introduire une certaine terminologie sur les variables :

Définition V.5.3. *Supposons que le système d'équations considéré soit compatible. Les* ***variables principales*** *sont les variables $x_{c_1}, x_{c_2}, \dots, x_{c_r}$, c'est-à-dire les variables associés aux colonnes de A sur lesquelles se trouvent les pivots. Les autres variables sont appelées* ***variables secondaires****.*

Exemple V.5.4. *Pour la matrice augmentée suivante, les variables principales sont x_1 et x_2, et x_3 est la seule variable secondaire :*

$$\widehat{A} = \left(\begin{array}{ccc|c} 1 & 0 & 2 & 5 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

On peut alors récrire le système sous la forme :

$$\begin{cases} x_1 & = & 5 - 2x_3 \\ x_2 & = & 4 - 3x_3 \end{cases}$$

Dans l'exemple précédent, le système se reformule sous la forme d'une équation par variable principale, cette variable principale devant être égale à une certaine expression en les variables secondaires. C'est un fait général dans le cas où la matrice A est échelonnée réduite :

Proposition V.5.5. *On conserve les notations précédentes. On suppose que le système est compatible et que A est* **échelonnée réduite***. On introduit J l'ensemble des indices j tels que x_j soit une variable secondaire, tandis que les $x_{c_1}, \dots, x_{c_r}$ sont les variables principales. Le système est équivalent à un système de r équations, où pour tout $i \in \{1, \dots, r\}$, la i-ième équation est*

$$x_{c_i} = b_i - \sum_{j \in J} a_{i,j} x_j$$

Ainsi, une solution $(x_1, \dots, x_n)$ au système d'équations est uniquement déterminée par les valeurs x_j pour $j \in J$, et celles-ci peuvent être choisies arbitrairement.

En effet, en tenant compte du caractère échelonné réduit de A, la i-ième équation est « $x_{c_i} + \sum_{j \in J} a_{i,j} x_j = b_i$ ».

Remarque V.5.6. *On dira qu'on a résolu un système d'équations linéaires (avec la méthode du pivot de Gauss) si on a déterminé s'il admet ou non des solutions, et dans le cas où il admet des solutions, on a récrit le système sous la forme d'équations comme dans la proposition précédente. Dans la rédaction, il conviendra de préciser les opérations sur les lignes qui auront été réalisées sur la matrice augmentée.*

Remarque V.5.7. *On peut penser aux variables secondaires comme des paramètres en fonction desquels les autres variables (les variables principales) s'expriment. Ainsi, dans l'exemple V.5.4, les solutions sont les triplets de la forme* $(5-2x_3, 4-3x_3, x_3)$ *pour* $x_3 \in \mathbf{R}$*. De façon équivalente, si pour tout* $t \in \mathbf{R}$*, on pose* $M(t) = (t-2t, 4-2t, t)$*, alors* $M(t)$ *pour* $t \in \mathbf{R}$ *est un paramétrage des solutions (qui est une droite dans* $\mathbf{R}^3$*).*

La remarque précédente suggère d'introduire la définition suivante :

Définition V.5.8. *Quand le système est compatible, l'ensemble des solutions sera dit de dimension* d*, où* d *est le nombre de variables secondaires.*

VI. Sous-espaces vectoriels de $\mathbf{R}^n$

Sommaire

VI.1. Définition

Définition VI.1.1. *Soit $n \in \mathbf{N}$. Un sous-espace vectoriel de $M_{n,1}(\mathbf{R})$ (que l'on peut identifier à $\mathbf{R}^n$) est une partie $F \subset M_{n,1}(\mathbf{R})$ telle que*
 (i) $\mathbf{0}_{n,1} \in F$ *;*
 (ii) pour tous $X_1 \in F$ et $X_2 \in F$, $X_1 + X_2 \in F$;
 (iii) pour tout $\lambda \in \mathbf{R}$ et tout $X \in F$, on a $\lambda X \in F$.

Exemple VI.1.2. *$\{\mathbf{0}_{n,1}\}$ et $M_{n,1}(\mathbf{R})$ sont des sous-espaces vectoriels de $M_{n,1}(\mathbf{R})$. On dit que $\{\mathbf{0}_{n,1}\}$ est le sous-espace vectoriel nul de $M_{n,1}(\mathbf{R})$.*

Proposition VI.1.3. *Soit $A \in M_{m,n}(\mathbf{R})$. Notons F l'ensemble des $X \in M_{n,1}(\mathbf{R})$ tels que $AX = \mathbf{0}_{m,1}$. Alors, F est un sous-espace vectoriel de $M_{n,1}(\mathbf{R})$.*

On a bien sûr, $A\mathbf{0}_{n,1} = \mathbf{0}_{m,1}$. Si X_1 et X_2 sont deux éléments de F, c'est-à-dire deux éléments de $M_{n,1}(\mathbf{R})$ tels que $AX_1 = \mathbf{0}_{m,1}$ et $AX_2 = \mathbf{0}_{m,1}$, alors $A(X_1 + X_2) = AX_1 + AX_2 = \mathbf{0}_{m,1} + \mathbf{0}_{m,1} = \mathbf{0}_{m,1}$, donc $X_1 + X_2 \in F$. Si $X \in F$, on obtient de même que pour tout $\lambda \in \mathbf{R}$, on a $A(\lambda X) = \lambda(AX) = \lambda \cdot \mathbf{0}_{m,1} = \mathbf{0}_{m,1}$, donc $\lambda X \in F$.

Définition VI.1.4. *Un système d'équations linéaires est dit homogène si les membres de droite de toutes les équations sont nuls :*

$$\left\{\begin{array}{ccccccccc}
a_{1,1}x_1 & + & a_{1,2}x_2 & + & \dots & + & a_{1,n}x_n & = & 0 \\
a_{2,1}x_1 & + & a_{2,2}x_2 & + & \dots & + & a_{2,n}x_n & = & 0 \\
\vdots & & \vdots & & \vdots & & \vdots & = & \vdots \\
a_{i,1}x_1 & + & a_{i,2}x_2 & + & \dots & + & a_{i,n}x_n & = & 0 \\
\vdots & & \vdots & & \vdots & & \vdots & = & \vdots \\
a_{n,1}x_1 & + & a_{n,2}x_2 & + & \dots & + & a_{n,n}x_n & = & 0
\end{array}\right.$$

Corollaire VI.1.5. *L'ensemble des solutions d'un système d'équations linéaires* homogène *de m équations en n variables est un sous-espace vectoriel de $\mathbf{R}^n$.*

En effet, un système homogène se reformule par une équation matricielle $AX = \mathbf{0}_{m,1}$.

Remarque VI.1.6. *Il est possible de montrer que tous les sous-espaces vectoriels de $\mathbf{R}^n$ sont obtenus de cette façon, mais la démonstration ne sera pas donnée dans ce cours.*

Remarque VI.1.7.
- *Dans le plan $\mathbf{R}^2$, une droite d'équation $ax+by=0$ (pour $(a,b) \neq (0,0)$) est un sous-espace vectoriel de $\mathbf{R}^2$. Autrement dit, une droite passant par $(0,0)$ est un sous-espace vectoriel de $\mathbf{R}^2$.*
- *Dans l'espace $\mathbf{R}^3$, un plan d'équation $ax+by+cz=0$ est un sous-espace vectoriel de $\mathbf{R}^3$; autrement dit, un plan contenant l'origine $(0,0,0)$ est un sous-espace vectoriel de $\mathbf{R}^3$.*
- *Dans l'espace $\mathbf{R}^3$, une droite $\mathscr{D}$ passant par l'origine est aussi un sous-espace vectoriel de $\mathbf{R}^3$: en effet, on peut écrire $\mathscr{D} = \mathscr{P}_1 \cap \mathscr{P}_2$ avec $\mathscr{P}_1$ et $\mathscr{P}_2$ deux plans contenant l'origine, donc $\mathscr{D}$ peut être décrite comme l'ensemble des solutions à un système linéaire homogène de deux équations en trois variables.*

VI.2. Interprétation vectorielle des systèmes linéaires

Supposons que $C_1, \ldots, C_n$ soient n vecteurs-colonnes de taille $m \times 1$. Soit $B \in M_{m,1}(\mathbf{R})$. On se pose la question suivante :

Existe-t-il $(x_1, \ldots, x_n) \in \mathbf{R}^n$ tel que $x_1C_1 + \cdots + x_nC_n = B$?

D'après la définition IV.2.4, les vecteurs de la forme $x_1C_1 + \cdots + x_nC_n$ sont appelés combinaisons linéaires des vecteurs $C_1, \ldots, C_n$ (et les $x_1, \ldots, x_n$ sont appelés coefficients de ces combinaisons linéaires). La question se reformule ainsi :

Le vecteur B est-il combinaison linéaire des vecteurs $C_1, \ldots, C_n$?

Pour répondre à cette question, on introduit la matrice $A = \begin{pmatrix} C_1 & \ldots & C_n \end{pmatrix}$ dont les vecteurs-colonnes sont les vecteurs C. Si on note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(\mathbf{R})$ pour tout n-uplet $(x_1, \ldots, x_n) \in \mathbf{R}^n$, on a par définition du produit AX :

$$AX = x_1C_1 + \cdots + x_nC_n$$

Avec ces notations, la question initiale se reformule en :

L'équation matricielle $AX = B$ admet-elle une solution ?

On peut répondre à cette question en utilisant l'algorithme du pivot de Gauss. Dans le cas où le système d'équations linéaires possède au moins une solution, la méthode de résolution du système permet de déterminer les coefficients $(x_1, \ldots, x_n) \in \mathbf{R}^n$ des combinaisons linéaires telle que $x_1C_1 + \cdots + x_nC_n = B$.

Exemple VI.2.1. *Si on a :*

$$C_1 = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \qquad C_2 = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 5 \\ 9 \end{pmatrix}$$

on peut introduire la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$$

Déterminer les x_1 et x_2 tels que $x_1C_1 + x_2C_2 = B$ revient à déterminer les vecteurs-colonnes X tels que $AX = B$, ce qui peut se faire en appliquant l'algorithme du pivot de Gauss à la matrice augmentée suivante :

$$\begin{aligned}
M &= \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 5 \\ 5 & 6 & 9 \end{array}\right) \\
&\underset{\left[\begin{smallmatrix} L_2 \longleftarrow L_2 - 3L_1 \\ L_3 \longleftarrow L_3 - 5L_1 \end{smallmatrix}\right]}{\sim} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & -2 & 2 \\ 0 & -4 & 4 \end{array}\right) \\
&\underset{\left[\begin{smallmatrix} L_2 \longleftarrow -\frac{1}{2}L_2 \\ L_3 \longleftarrow -\frac{1}{4}L_3 \end{smallmatrix}\right]}{\sim} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{array}\right) \\
&\underset{[L_3 \longleftarrow L_3 - L_2]}{\sim} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right) \quad \textit{(on note que le système associé est compatible)} \\
&\underset{[L_1 \longleftarrow L_1 - 2L_2]}{\sim} \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right)
\end{aligned}$$

L'unique solution au système est donnée par $x_1 = 3$, $x_2 = -1$, donc $B = 3C_1 - C_2$.

VI.3. Le sous-espace vectoriel engendré par une famille de vecteurs

Définition VI.3.1. *Supposons que $C_1, \ldots, C_n$ soient des éléments de l'ensemble $M_{m,1}(\mathbf{R})$. On note $\mathrm{Vect}(C_1, \ldots, C_n)$ l'ensemble des combinaisons linéaires des vecteurs $C_1, \ldots, C_n$, c'est-à-dire l'ensemble des vecteurs de la forme $x_1C_1 + \cdots + x_nC_n \in M_{m,1}(\mathbf{R})$ pour $(x_1, \ldots, x_n) \in \mathbf{R}^n$.*

Autrement dit, un élément $B \in M_{m,1}(\mathbf{R})$ appartient à $\mathrm{Vect}(C_1, \ldots, C_n)$ si et seulement s'il existe $(x_1, \ldots, x_n) \in \mathbf{R}^n$ tel que $B = x_1C_1 + \cdots + x_nC_n$.

Si on introduit la matrice $A := \begin{pmatrix} C_1 & \ldots & C_n \end{pmatrix} \in M_{m,n}(\mathbf{R})$, $\mathrm{Vect}(C_1, \ldots, C_n)$ peut aussi être décrit comme l'ensemble des vecteurs-colonnes de la forme AX où $X \in M_{n,1}(\mathbf{R})$.

Proposition VI.3.2. *Si $C_1, \dots, C_n$ sont des éléments de $M_{m,1}(\mathbf{R})$, alors $\mathrm{Vect}(C_1, \dots, C_n)$ est un sous-espace vectoriel de $M_{m,1}(\mathbf{R})$.*

Si F est un sous-espace vectoriel de $M_{m,1}(\mathbf{R})$ contenant les vecteurs $C_1, \dots, C_n$, alors on a automatiquement l'inclusion $\mathrm{Vect}(C_1, \dots, C_n) \subset F$: on dit que $\mathrm{Vect}(C_1, \dots, C_n)$ est le sous-espace vectoriel de $M_{m,1}(\mathbf{R})$ engendré par les vecteurs $C_1, \dots, C_n$.

En utilisant la matrice A définie ci-dessus, $\mathrm{Vect}(C_1, \dots, C_n)$ est l'ensemble des vecteurs de la forme AX pour $X \in M_{n,1}(\mathbf{R})$.

Avec $X := \mathbf{0}$, on obtient que le vecteur nul appartient à $\mathrm{Vect}(C_1, \dots, C_n)$.

Deux éléments V_1 et V_2 dans $\mathrm{Vect}(C_1, \dots, C_n)$ peuvent s'écrire $V_1 = AX_1$ et $V_2 = AX_2$, donc $V_1 + V_2 = AX_1 + AX_2 = A(X_1 + X_2)$.

Si $\lambda \in \mathbf{R}$ et $X \in M_{n,1}(\mathbf{R})$, le calcul $\lambda \cdot (AX) = A(\lambda X)$ montre que $\mathrm{Vect}(C_1, \dots, C_n)$ est stable par multiplication par λ.

Exemple VI.3.3. *Si $C \in M_{m,1}(\mathbf{R})$, $\mathrm{Vect}(C)$ est l'ensemble des vecteurs de la forme λC pour $\lambda \in \mathbf{R}$. Soit C est le vecteur nul, auquel cas $\mathrm{Vect}(C)$ est le sous-espace vectoriel nul, soit C est non nul, et alors on dit que $\mathrm{Vect}(C)$ est la droite vectorielle engendrée par C.*

Définition VI.3.4. *Si C_1 et C_2 sont deux vecteurs dans $M_{m,1}(\mathbf{R})$, on dit que C_1 et C_2 sont colinéaires si $C_2 = \mathbf{0}$ ou s'il existe $\lambda \in \mathbf{R}$ tel que $C_1 = \lambda C_2$. (En particulier, si C_1 et C_2 sont tous les deux non nuls, ils sont colinéaires si et seulement s'il existe $\lambda \in \mathbf{R}^\times$, $C_1 = \lambda C_2$.)*

Si C_1 et C_2 sont colinéaires, alors $\mathrm{Vect}(C_1, C_2)$ est égal à $\mathrm{Vect}(C_1)$ si C_2 est nul, et à $\mathrm{Vect}(C_2)$ si C_1 est de la forme $C_1 = \lambda C_2$. Le cas intéressant est celui où C_1 et C_2 ne sont pas colinéaires :

Définition VI.3.5. *Si C_1 et C_2 sont deux vecteurs non colinéaires de $M_{m,1}(\mathbf{R})$, on appelle plan vectoriel engendré par les vecteurs C_1 et C_2 l'ensemble $\mathrm{Vect}(C_1, C_2)$.*

VI.4. Équations définissant un sous-espace vectoriel engendré par des vecteurs

Dans l'exemple VI.2.1,

$$C_1 = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} \qquad C_2 = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$$

on a déterminé si un certain vecteur particulier $B \in M_{3,1}(\mathbf{R})$ était combinaison linéaire de C_1 et C_2. La même méthode peut être utilisée pour un vecteur B quelconque :

$$B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Ici, b_1, b_2 et b_3 sont des nombres réels arbitraires.

Pour résoudre la question générale de savoir si $B \in \mathrm{Vect}(C_1, C_2)$, on introduit cette fois-ci une matrice augmentée M dans laquelle les coefficients b_1, b_2 et b_3 apparaissent :

$$M = \left(\begin{array}{cc|c} 1 & 2 & b_1 \\ 3 & 4 & b_2 \\ 5 & 6 & b_3 \end{array}\right)$$

On applique la méthode du pivot de Gauss à cette matrice de façon à mettre sous forme échelonnée la « partie gauche » de la matrice :

$$\begin{aligned}
M &\underset{\left[\substack{L_2\longleftarrow L_2-3L_1\\ L_3\longleftarrow L_3-5L_1}\right]}{\sim} \left(\begin{array}{cc|l} 1 & 2 & b_1 \\ 0 & -2 & b_2-3b_1 \\ 0 & -4 & b_3-5b_1 \end{array}\right) \\
&\underset{[L_3\longleftarrow L_3-2L_2]}{\sim} \left(\begin{array}{cc|l} 1 & 2 & b_1 \\ 0 & -2 & b_2-3b_1 \\ 0 & 0 & b_3-5b_1-2(b_2-3b_1) \end{array}\right) \\
&= \left(\begin{array}{cc|l} 1 & 2 & b_1 \\ 0 & -2 & b_2-3b_1 \\ 0 & 0 & b_1-2b_2+b_3 \end{array}\right)
\end{aligned}$$

On peut conclure que le système est compatible si et seulement si $b_1 - 2b_2 + b_3 = 0$. (Et on peut remarquer que si le système est compatible, il admet ici une unique solution (x_1, x_2).)

Ainsi, $\mathrm{Vect}(C_1, C_2)$ est formé des vecteurs-colonnes $B = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ tels que $b_1 - 2b_2 + b_3 = 0$. Autrement dit, si au lieu d'utiliser les variables b_1, b_2, b_3, on utilise les variables x, y et z, $\mathrm{Vect}(C_1, C_2)$ est plan vectoriel d'équation $x - 2y + z = 0$.

Remarque VI.4.1. *La méthode utilisée ci-dessus permet en général de décrire un sous-espace vectoriel de la forme* $\mathrm{Vect}(C_1, \dots, C_n)$ *comme l'ensemble des solutions à un système d'équations linéaires homogène.*

VI.5. Familles libres

VI.5.1. Définition

Définition VI.5.1.1. *Supposons que $C_1, \dots, C_n$ est une famille de n éléments de $M_{m,1}(\mathbf{R})$. On dit que $(C_1, \dots, C_n)$ est une famille libre si pour tout $(x_1, \dots, x_n) \in \mathbf{R}^n$, on a l'implication*

$$x_1C_1 + \dots + x_nC_n = \mathbf{0}_{m,1} \quad \Longrightarrow \quad x_1 = \dots = x_n = 0$$

Une famille libre $(C_1, \dots, C_n)$ est aussi qualifiée de famille de vecteurs linéairement indépendants.

On dit que $(C_1, \dots, C_n)$ est une famille liée si elle n'est pas libre, ce qui signifie qu'il existe des réels $x_1, \dots, x_n$ non tous nuls tels que $x_1C_1 + \dots + x_nC_n = \mathbf{0}_{m,1}$.

En utilisant la terminologie des combinaisons linéaires, une famille de vecteurs est libre si et seulement si la seule combinaison linéaire de ces vecteurs qui est nulle est celle dont les coefficients sont nuls. Une famille de vecteurs est liée si le vecteur nul peut être obtenu comme combinaison linéaire de ces vecteurs avec au moins un coefficient non nul.

VI.5.2. Famille à un ou deux vecteurs

Proposition VI.5.2.1. *Soit $C \in M_{m,1}(\mathbf{R})$. La famille à un élément C est libre si et seulement si $C \neq \mathbf{0}_{m,1}$. Si c'est le cas, un élément de* $\mathrm{Vect}(C)$ *s'écrit de manière unique λC pour $\lambda \in \mathbf{R}$.*

Proposition VI.5.2.2. *Soit (C_1, C_2) un couple d'éléments de $M_{m,1}(\mathbf{R})$. La famille à deux éléments (C_1, C_2) est libre si et seulement si C_1 et C_2 ne sont pas colinéaires.*

Si C_1 et C_2 sont colinéaires, alors soit C_2 est nul auquel cas $0 \cdot C_1 + 1 \cdot C_2 = \mathbf{0}$, soit C_2 est non nul et alors il existe $\lambda \in \mathbf{R}$, tel que $C_1 = \lambda C_2$ et donc $1 \cdot C_1 - \lambda C_2 = \mathbf{0}$. Dans tous les cas, on a déterminé une combinaison linéaire de C_1 et C_2 égale au vecteur nul mais dont les coefficients ne sont pas tous nuls, donc (C_1, C_2) est liée, c'est-à-dire non libre.

Réciproquement, si (C_1, C_2) est liée, il existe $(\lambda, \mu) \in \mathbf{R}^2 - \{(0,0)\}$ tels que $\lambda C_1 + \mu C_2 = \mathbf{0}$. Si $\lambda = 0$, et donc $\mu \neq 0$, on a $C_2 = \mathbf{0}$. Sinon, $\lambda \neq 0$ et alors $C_1 = \frac{-\lambda}{\mu} C_2$. Dans tous les cas, on a montré que C_1 et C_2 étaient colinéaires.

Proposition VI.5.2.3. *Si $C_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ et $C_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$, on peut introduire la matrice $A = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \in M_{2,2}(\mathbf{R})$. Le déterminant des vecteurs (C_1, C_2), ou le déterminant de A, est noté et défini ainsi :*

$$\det(C_1, C_2) = \det A = \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} := x_1 y_2 - y_1 x_2$$

Les vecteurs (C_1, C_2) forment une famille libre (c'est-à-dire sont non colinéaires) si et seulement si $\det(C_1, C_2) \neq 0$.

Si C_1 et C_2 sont colinéaires, il existe $C \in \{C_1, C_2\}$ et des réels λ et μ tels que $C_1 = \lambda C$ et $C_2 = \mu C$. Si on note x et y les coordonnées de C, on a alors :

$$\det(C_1, C_2) = \begin{vmatrix} \lambda x & \mu x \\ \lambda y & \mu y \end{vmatrix} = \lambda \mu x y - \lambda \mu x y = 0$$

Réciproquement, si $\det(C_1, C_2) = 0$, montrons que C_1 et C_2 sont colinéaires. Si $C_2 = \mathbf{0}$, le résultat est évident. Supposons donc $C_2 \neq \mathbf{0}$. Pour $(x, y) \in \mathbf{R}^2$ arbitraire, considérons

$$\begin{vmatrix} x & x_2 \\ y & y_2 \end{vmatrix} = y_2 x - x_2 y$$

L'équation $y_2 x - x_2 y = 0$ est l'équation d'une droite $\mathscr{D}$ passant par les deux point $(0,0)$ et (x_2, y_2), donc admet le paramétrage $M(t) = (t x_2, t y_2)$. Par hypothèse, $(x_1, y_1) \in \mathscr{D}$, donc il existe t tel que $x_1 = t x_2$ et $y_1 = t y_2$, autrement dit $C_1 = t C_2$, ce qui montre que C_1 et C_2 sont colinéaires.

VI.5.3. Caractérisation matricielle

Proposition VI.5.3.1. *Supposons que $C_1, \ldots, C_n$ sont des éléments de $M_{m,1}(\mathbf{R})$. Notons $A \in M_{m,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont les vecteurs $C_1, \ldots, C_n$. La famille $C_1, \ldots, C_n$ est libre si et seulement si l'équation matricielle $AX = \mathbf{0}_{m,1}$ pour $X \in M_{n,1}(\mathbf{R})$ admet pour solution unique $X = \mathbf{0}_{n,1}$.*

Si on introduit les coefficients $x_1, \ldots, x_n$ de $X \in M_{n,1}(\mathbf{R})$, alors $AX = x_1 C_1 + \cdots + x_n C_n$. Il en résulte que la donnée d'une solution à l'équation matricielle $AX = \mathbf{0}_{m,1}$ revient à celle de coefficients $(x_1, \ldots, x_n) \in \mathbf{R}^n$ tels que $x_1 C_1 + \cdots + x_n C_n = \mathbf{0}_{m,1}$. L'équivalence voulue en résulte immédiatement.

Proposition VI.5.3.2. *Supposons que $C_1, \dots, C_n$ sont des éléments de $M_{m,1}(\mathbf{R})$. Notons $A \in M_{m,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont les vecteurs $C_1, \dots, C_n$. La famille $C_1, \dots, C_n$ est libre si et seulement si pour tout $B \in M_{m,1}(\mathbf{R})$, l'équation matricielle $AX = B$ admet au plus une solution $X \in M_{n,1}(\mathbf{R})$.*

Supposons que $C_1, \dots, C_n$ soit libre. Soit $B \in M_{m,1}(\mathbf{R})$. Montrons que $AX = B$ admet au plus une solution dans $M_{n,1}(\mathbf{R})$. Supposons que X et X' soient deux solutions, c'est-à-dire que $AX = B$ et $AX' = B$. Considérons $Y := X' - X \in M_{n,1}(\mathbf{R})$. On a $AY = A(X' - X) = AX' - AX = B - B = \mathbf{0}_{m,1}$. D'après la proposition VI.5.3.1, on a $Y = \mathbf{0}_{n,1}$, c'est-à-dire $X' - X = \mathbf{0}_{n,1}$, donc $X' = X$. L'équation $AX = B$ admet donc au plus une solution dans $M_{n,1}(\mathbf{R})$.

Proposition VI.5.3.3. *Soit $C_1, \dots, C_n$ une famille de vecteurs-colonnes appartenant à $M_{m,1}(\mathbf{R})$. On note $A \in M_{m,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont $C_1, \dots, C_n$. Notons $A' \in M_{m,n}(\mathbf{R})$ soit une matrice échelonnée obtenue à partir de A en effectuant des opérations élémentaires sur les lignes. La famille $(C_1, \dots, C_n)$ est libre si et seulement si la matrice A' a n pivots (autrement dit, il y a un pivot sur chaque colonne).*

En effet, si on considère $\widehat{A} = \left(\begin{array}{c|c} A & \mathbf{0} \end{array}\right) \in M_{m,n+1}(\mathbf{R})$ et $\widehat{A'} = \left(\begin{array}{c|c} A' & \mathbf{0} \end{array}\right) \in M_{m,n+1}(\mathbf{R})$ les matrices augmentées associées aux systèmes formulés matriciellement $AX = \mathbf{0}$ et $A'X = \mathbf{0}$ respectivement, alors $\widehat{A'}$ peut être obtenu à partir de $\widehat{A}$ en effectuant des opérations élémentaires sur les lignes.

Le système associé à la matrice augmentée $\widehat{A'}$ est évidemment compatible (puisque le vecteur nul est solution). La famille $C_1, \dots, C_n$ est libre si et seulement si ce système possède $\mathbf{0}_{n,1}$ comme unique solution, c'est-à-dire s'il n'y ait aucune variable secondaire, autrement dit que toutes les variables soient principales, ce qui signifie qu'un pivot se trouve sur chacune des n colonnes de A'.

Corollaire VI.5.3.4. *Soit $C_1, \dots, C_n$ une famille de vecteurs-colonnes appartenant à $M_{m,1}(\mathbf{R})$. Si $n \geq m + 1$, alors $C_1, \dots, C_n$ est une famille liée : il existe des coefficients $(x_1, \dots, x_n) \in \mathbf{R}^n - \{(0, \dots, 0)\}$ tels que $x_1C_1 + \cdots + x_nC_n = \mathbf{0}$.*

On utilise le critère de la proposition VI.5.3.3. Si on note $A' \in M_{m,n}(\mathbf{R})$ une matrice échelonnée obtenue par opérations élémentaires sur les lignes à partir de la matrice A dont les colonnes sont $C_1, \dots, C_n$, alors A' a au plus m pivots puisque A' a m lignes. Comme $n \geq m + 1$, la matrice échelonnée A' ne peut pas avoir n pivots, donc ne peut pas en avoir un sur chaque colonne : les vecteurs $C_1, \dots, C_n$ ne forment pas une famille libre.

VI.6. Familles génératrices de $M_{m,1}(\mathbf{R})$

Définition VI.6.1. *Supposons que $C_1, \dots, C_n$ est une famille de n éléments de $M_{m,1}(\mathbf{R})$. On dit que $(C_1, \dots, C_n)$ est une famille génératrice (de $M_{m,1}(\mathbf{R})$) si tout élément $B \in M_{m,1}(\mathbf{R})$ peut s'écrire sous la forme $B = x_1C_1 + \cdots + x_nC_n$ pour des coefficients $(x_1, \dots, x_n) \in \mathbf{R}^n$ bien choisis.*

Autrement dit, $C_1, \dots, C_n$ est une famille génératrice (de $M_{m,1}(\mathbf{R})$) si tout élément de $M_{m,1}(\mathbf{R})$ est combinaison linéaire des vecteurs $C_1, \dots, C_n$. Ceci se reformule aussi en énonçant l'égalité $\mathrm{Vect}(C_1, \dots, C_n) = M_{m,1}(\mathbf{R})$.

Exemple VI.6.2. *Dans $M_{2,1}(\mathbf{R})$, les vecteurs $C_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $C_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ forment une famille génératrice.*

Proposition VI.6.3. *Supposons que $C_1, \dots, C_n$ soient des éléments de $M_{m,1}(\mathbf{R})$. Notons $A = \begin{pmatrix} C_1 & \dots & C_n \end{pmatrix} \in M_{m,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont les vecteurs $C_1, \dots, C_n$. Soit $A' \in M_{m,n}(\mathbf{R})$ une matrice obtenue à partir de A en effectuant une opération élémentaire sur les lignes. Notons $C'_1, \dots, C'_n$ les vecteurs-colonnes de A'.*

Alors, $C_1, \dots, C_n$ est une famille génératrice de $M_{m,1}(\mathbf{R})$ si et seulement si $C'_1, \dots, C'_n$ est une famille génératrice de $M_{m,1}(\mathbf{R})$.

Pour démontrer cette proposition, on commence par le lemme suivant :

Lemme VI.6.4. *On conserve les notations de la proposition précédente. Soit $B \in M_{m,1}(\mathbf{R})$. On note $B' \in M_{m,1}(\mathbf{R})$ le vecteur-colonne obtenu en appliquant à B la même opération élémentaire sur les lignes que celle effectuée sur la matrice A pour obtenir A'.*

Alors, $B \in \mathrm{Vect}(C_1, \dots, C_n)$ si et seulement si $B' \in \mathrm{Vect}(C'_1, \dots, C'_n)$.

Notons $M := \left(A \mid B \right) = \left(C_1 \ \dots \ C_n \mid B \right) \in M_{m,n+1}(\mathbf{R})$ et $M' := \left(A' \mid B' \right) = \left(C'_1 \ \dots \ C'_n \mid B' \right) \in M_{m,n+1}(\mathbf{R})$.

Par hypothèse, les matrices augmentées M et M' sont équivalentes. Les systèmes d'équations linéaires associés à M et M' ont donc les mêmes solutions. Par conséquent, le système d'équation associé à M possède (au moins) une solution (c'est-à-dire $B \in \mathrm{Vect}(C_1, \dots, C_n)$) si et seulement si le système d'équation associé à M' possède (au moins) une solution (c'est-à-dire $B' \in \mathrm{Vect}(C'_1, \dots, C'_n)$).

Revenons à la démonstration de la proposition VI.6.3. Supposons que $C'_1, \dots, C'_n$ soit une famille génératrice de $M_{m,1}(\mathbf{R})$, montrons que $C_1, \dots, C_n$ est aussi génératrice. Soit $B \in M_{m,1}(\mathbf{R})$. Notons $B' \in M_{m,1}(\mathbf{R})$ le vecteur défini dans l'énoncé du lemme ci-dessus. Par hypothèse, $B' \in \mathrm{Vect}(C'_1, \dots, C'_n)$, donc l'équivalence donnée dans le lemme montre que $B \in \mathrm{Vect}(C_1, \dots, C_n)$. Ceci étant vrai pour tout $B \in M_{m,1}(\mathbf{R})$, on a montré que $C_1, \dots, C_n$ était une famille génératrice de $M_{m,1}(\mathbf{R})$.

On a montré un des deux sens d'implication dans l'énoncé de la proposition VI.6.3. Pour démontrer la réciproque, il suffit de remarquer que dans les hypothèses, si on peut obtenir A' à partir de A en effectuant une opération élémentaire sur les lignes de A, alors d'après l'argument vu dans la démonstration de proposition V.2.4, on peut aussi obtenir A à partir de A' en effectuant une opération élémentaire sur les lignes. En utilisant l'implication déjà démontrée en échangeant les rôles des familles $C_1, \dots, C_n$ et $C'_1, \dots, C'_n$, on obtient l'équivalence voulue.

Théorème VI.6.5. *Supposons que $C_1, \dots, C_n$ soient des éléments de $M_{m,1}(\mathbf{R})$. Notons $A = \begin{pmatrix} C_1 & \dots & C_n \end{pmatrix} \in M_{m,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont $C_1, \dots, C_n$. Notons A' une matrice échelonnée obtenue en effectuant des opérations élémentaires sur A.*

Alors, $C_1, \dots, C_n$ est une famille génératrice si et seulement si A' contient un pivot sur chaque ligne (autrement dit, si A' n'a pas de ligne formée uniquement de zéros).

D'après la proposition VI.6.3, le fait que les vecteurs-colonnes d'une matrice engendrent $M_{m,1}(\mathbf{R})$ est invariant par opérations élémentaires. On peut donc supposer que $A' = A$. Par ailleurs, la position des pivots est inchangée quand on met une matrice échelonnée sous forme échelonnée réduite. On peut donc supposer que $A = A'$ est une matrice échelonnée réduite.

Si A possède une ligne de zéros (par exemple sur la dernière ligne), alors la dernière coordonnée de tous les vecteurs C_i est nulle, donc il est évident que le vecteur $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ n'appartient pas à $\mathrm{Vect}(C_1, \ldots, C_n)$ donc cette famille n'est pas génératrice.

Si au contraire toutes les lignes de la matrice échelonnée réduite A sont non nulles, c'est-à-dire que pour tout $i \in \{1, \ldots, m\}$, la i-ième ligne contient un pivot sur la c_i-ième colonne, alors le caractère échelonné réduit de A implique que le c_i-ième vecteur-colonne de A est le vecteur E_i :

$$E_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in M_{m,1}(\mathbf{R})$$

où le coefficient égal à 1 est sur la i-ème ligne. Par conséquent, tous les vecteurs $E_1, \ldots, E_m$ appartiennent à $\mathrm{Vect}(C_1, \ldots, C_n)$ qui est donc égal à $M_{m,1}(\mathbf{R})$ tout entier.

Corollaire VI.6.6. *Supposons que $C_1, \ldots, C_n$ soient une famille de vecteurs-colonnes appartenant à $M_{n,1}(\mathbf{R})$. (Si on note $A \in M_{n,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont $C_1, \ldots, C_n$, alors la matrice A est* carrée *: elle possède le même nombre de lignes que de colonnes, et on note $A \in M_n(\mathbf{R}) := M_{n,n}(\mathbf{R})$.)*

Les conditions suivantes sont équivalentes :

(i) $C_1, \ldots, C_n$ est une famille libre;

(ii) $C_1, \ldots, C_n$ est une famille génératrice de $M_{n,1}(\mathbf{R})$.

Si on note $A' \in M_n(\mathbf{R})$ une matrice échelonnée obtenue en partant de A et en effectuant des opérations élémentaires sur les lignes. La condition (ii) se traduit en disant que A' possède un pivot sur chaque ligne. Si (ii) est vrai, on peut noter $c_1 < c_2 < \cdots < c_n$ les numéros des colonnes sur lesquels se trouvent les pivots. Comme $1 \leq c_1 < \cdots < c_n \leq n$, on a forcément $c_i = i$ pour tout $i \in \{1, \ldots, n\}$, donc chaque colonne de A' possède un pivot, donc (i) est vrai d'après la proposition VI.5.3.1.

Réciproquement, si (i) est vrai, la proposition VI.5.3.1, montre que A' possède un pivot sur chaque colonne, donc possède n pivots, ce qui implique que toutes les lignes de A' sont non nulles, donc A' possède un pivot sur chaque ligne, donc (ii) est vrai d'après le théorème VI.6.5.

Corollaire VI.6.7. *Si C_1 et C_2 sont deux vecteurs non-colinéaires dans $M_{2,1}(\mathbf{R})$, alors ils forment une famille génératrice de $M_{2,1}(\mathbf{R})$.*

Corollaire VI.6.8. *Soit $C_1, \ldots, C_n$ une famille de vecteurs-colonnes appartenant à $M_{m,1}(\mathbf{R})$. Si $n < m$, alors $C_1, \ldots, C_n$ n'est pas une famille génératrice de $M_{m,1}(\mathbf{R})$.*

En effet, si $A' \in M_{m,n}(\mathbf{R})$ est une matrice échelonnée obtenue en effectuant des opérations élémentaires sur la matrice A dont les colonnes sont les vecteurs $C_1, \ldots, C_n$, alors comme A' possède n colonnes, elle possède au plus n pivots. Comme $n < m$, A' ne peut pas avoir

m pivots, donc d'après le critère du théorème VI.6.5, la famille $C_1, \ldots, C_n$ n'engendre pas $M_{m,1}(\mathbf{R})$.

VI.7. Dimension

Proposition VI.7.1. *Supposons que $C_1, \ldots, C_n$ soient des éléments de $M_{m,1}(\mathbf{R})$. Introduisons le sous-espace vectoriel $F := \mathrm{Vect}(C_1, \ldots, C_n) \subset M_{m,1}(\mathbf{R})$ de $M_{m,1}(\mathbf{R})$.*

Soit $(V_1, \ldots, V_{n'})$ une famille de d'éléments de F. Si $n' > n$, alors $(V_1, \ldots, V_{n'})$ n'est pas libre.

Notons $A = \begin{pmatrix} C_1 & \ldots & C_n \end{pmatrix} \in M_{m,n}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont les C_i.

Pour tout $i \in \{1, \ldots, n'\}$, le fait que V_i appartienne à F signifie que V_i est combinaison linéaire de $C_1, \ldots, C_n$, autrement dit qu'il existe $W_i \in M_{n,1}(\mathbf{R})$ tel que $V_i = AW_i$.

On dispose ainsi de n' vecteurs $W_1, \ldots, W_{n'}$ dans $M_{n,1}(\mathbf{R})$. D'après le corollaire VI.5.3.4, le fait que $n' > n$ implique qu'il existe des coefficients $(x_1, \ldots, x_{n'}) \in \mathbf{R}^{n+1} - \{(0, \ldots, 0)\}$ tels que

$$x_1 W_1 + \cdots + x_{n'} W_{n'} = \mathbf{0}$$

En multipliant à gauche par la matrice A, on obtient :

$$x_1 V_1 + \cdots + x_{n'} V_{n'} = x_1 A W_1 + \cdots + x_{n'} A W_{n'} = A(x_1 W_1 + \cdots + x_{n'} W_{n'}) = \mathbf{0}$$

On a bien montré que la famille $(V_1, \ldots, V_{n'})$ était liée.

Définition VI.7.2. *Soit F un sous-espace vectoriel de $M_{m,1}(\mathbf{R})$. La dimension de F est le maximum des cardinaux des familles libres d'éléments de F. On note* $\dim F$ *cette dimension.*

On sait déjà que le cardinal d'une famille libre dans $M_{m,1}(\mathbf{R})$ est majoré par m. Les axiomes fondamentaux portant sur les entiers naturels justifient l'existence de l'entier $d := \dim F$.

Si $d := \dim F$, il est évident qu'il existe une famille libre $(V_1, \ldots, V_d)$ d'éléments de F. J'affirme qu'on a alors $F = \mathrm{Vect}(V_1, \ldots, V_d)$. En effet, sinon, compte tenu du fait que l'on a $\mathrm{Vect}(V_1, \ldots, V_d) \subset F$, on obtiendrait l'existence d'un $W \in F$ tel que $W \notin \mathrm{Vect}(V_1, \ldots, V_d)$. On montrerait alors facilement que la famille $(V_1, \ldots, V_d, W)$ serait libre, ce qui contredirait la maximalité de d. On a donc $F = \mathrm{Vect}(V_1, \ldots, V_d)$.

Inversement, si on dispose d'une famille libre $(W_1, \ldots, W_e)$ de vecteurs de F telle que $F = \mathrm{Vect}(W_1, \ldots, W_e)$, alors $\dim F = e$. En effet, par définition de $\dim F$, on a bien sûr $\dim F \geq e$, mais la proposition VI.7.1 donne l'autre inégalité $\dim F \leq e$. On a ainsi montré :

Proposition VI.7.3. *Soit F un sous-espace vectoriel de $M_{m,1}(\mathbf{R})$. Si $(W_1, \ldots, W_e)$ est une famille libre de vecteurs de F telle que $F = \mathrm{Vect}(W_1, \ldots, W_e)$, alors $\dim F = e$.*

VII. Représentation vectorielle des solutions

Sommaire

VII.1. Un exemple

Considérons un système de trois équations linéaires en quatre variables :

$$\left\{\begin{array}{rcrcrcrcr} x_1 & + & 2x_2 & + & 3x_3 & + & 4x_4 & = & 7 \\ 5x_1 & + & 6x_2 & + & 7x_3 & + & 8x_4 & = & 11 \\ 9x_1 & + & 10x_2 & + & 11x_3 & + & 12x_4 & = & 15 \end{array}\right.$$

La matrice augmentée associée est

$$M = \left(\begin{array}{rrrr|r} 1 & 2 & 3 & 4 & 7 \\ 5 & 6 & 7 & 8 & 11 \\ 9 & 10 & 11 & 12 & 15 \end{array}\right)$$

Quand on applique l'algorithme du pivot de Gauss à ce système, on obtient finalement la matrice échelonnée réduite ci-dessous :

$$M \sim M' = \left(\begin{array}{rrrr|r} 1 & 0 & -1 & -2 & -5 \\ 0 & 1 & 2 & 3 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Ce système est compatible, les variables principales sont x_1 et x_2, tandis que les variables secondaires sont x_3 et x_4. Le système associé à la matrice augmentée M' se reformule en :

$$\left\{\begin{array}{rcl} x_1 & = & x_3 + 2x_4 - 5 \\ x_2 & = & -2x_3 - 3x_4 + 6 \end{array}\right.$$

Les solutions (identifiées à des vecteurs) sont donc les vecteurs-colonnres de la forme suivante, pour $(x_3, x_4) \in \mathbf{R}^2$:

$$\begin{pmatrix} x_3 + 2x_4 - 5 \\ -2x_3 - 3x_4 + 6 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} -5 \\ 6 \\ 0 \\ 0 \end{pmatrix}$$

Si on introduit les vecteurs-colonnes suivants

$$V_1 := \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \end{pmatrix} \qquad V_2 := \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix} \qquad X_0 := \begin{pmatrix} -5 \\ 6 \\ 0 \\ 0 \end{pmatrix}$$

les solutions sont les vecteurs de la forme $X_0 + t_1V_1 + t_2V_2$ pour $(t_1, t_2) \in \mathbf{R}^2$, et chaque solution s'écrit de façon unique sous cette forme. (Et il se trouve que dans le cas présent, t_1 et t_2 correspondent respectivement aux coordonnées x_3 et x_4.)

VII.2. Solutions aux systèmes homogènes

Proposition VII.2.1. *Soit $F \subset M_{n,1}(\mathbf{R})$ l'ensemble des solutions à un système de m équations linéaires* homogène. *Il existe alors une famille libre $V_1, \dots, V_d$ d'éléments de $M_{n,1}(\mathbf{R})$ telle que $F = \mathrm{Vect}(V_1, \dots, V_d)$.*

Plus précisément, tout élément de F s'écrit de manière unique sous la forme $t_1V_1 + \cdots + t_dV_d$ pour $(t_1, \dots, t_d) \in \mathbf{R}^d$.

En outre, on a l'inégalité $d \geq n - m$.

Supposons que $F = \{X \in M_{n,1}(\mathbf{R}),\ AX = \mathbf{0}\}$ où $A \in M_{m,n}(\mathbf{R})$.

On peut appliquer l'algorithme du pivot de Gauss à A pour obtenir $A' \in M_{m,n}(\mathbf{R})$ échelonnée réduite. On a alors :

$$F = \{X \in M_{n,1}(\mathbf{R}),\ A'X = \mathbf{0}\}$$

Puisqu'il est homogène, le système est forcément compatible. En procédant comme dans l'exemple VII.1, on peut noter $I \subset \{1, \dots, n\}$ l'ensemble des indices correspondant aux variables principales et $J \subset \{1, \dots, n\}$ l'ensemble des indices associés aux variables secondaires. (Dans l'exemple, on avait $I = \{1, 2\}$ et $J = \{3, 4\}$.)

Comme la matrice échelonnée A' a m lignes, elle possède au plus m pivots, donc le cardinal de I est $\leq m$. Par conséquent, le cardinal d de J vérifie $d \geq n - m$.

Comme A' est échelonnée réduite, l'équation matricielle $A'X = \mathbf{0}$ se reformule en un système tel que pour tout $i \in I$, on a une équation exprimant x_i en fonction des variables secondaires x_j pour $j \in J$.

De façon similaire à l'exemple VII.1, pour chaque $j \in J$, c'est-à-dire pour chaque variable secondaire x_j, on peut déterminer un vecteur-colonne $V_j \in M_{n,1}(\mathbf{R})$ tel que les solutions s'écrivent de manière unique :

$$\sum_{j \in J} x_j W_j$$

Si on choisit une énumération $j_1, \dots, j_d$ des éléments de J, on peut poser $V_k := W_{j_k}$ pour $k \in \{1, \dots, d\}$, et alors les solutions, c'est-à-dire les éléments de F sont exactement les combinaisons linéaires $t_1V_1 + \cdots + t_dV_d$. Le fait que l'écriture soit unique fait que la famille de vecteurs $V_1, \dots, V_d$ est libre.

Remarque VII.2.2. *L'algorithme du pivot de Gauss utilisé dans la démonstration permet d'obtenir une famille de vecteurs $V_1, \dots, V_d$ qui convienne, mais il n'y a pas unicité. Cependant, l'entier d est unique : en effet, $d = \dim F$.*

VII.3. Cas général

Proposition VII.3.1. *Notons $G \subset M_{n,1}(\mathbf{R})$ l'ensemble des solutions à une équation matricielle $AX = B$ d'inconnue $X \in M_{n,1}(\mathbf{R})$ pour $A \in M_{m,n}(\mathbf{R})$ et $B \in M_m(\mathbf{R})$.*

Notons $F \subset M_{m,1}(\mathbf{R})$ l'ensemble des solutions à l'équation matricielle homogène $AX = \mathbf{0}$.

Si G est non vide, choisissons $X_0 \in G$. Alors, les éléments de G sont exactement ceux de la forme $X_0 + U$ avec $U \in F$. Si on note $(V_1, \dots, V_d)$ une famille libre telle que $\mathrm{Vect}(V_1, \dots, V_d) = F$ *(cf. proposition VII.2.1), alors les éléments de G sont exactement les éléments de la forme*

$$X_0 + t_1 V_1 + \cdots + t_d V_d$$

pour $(t_1, \dots, t_d) \in \mathbf{R}^d$, et chaque solution peut s'écrire de façon unique sous cette forme.

Fixons $X_0 \in G$. Si $X \in M_{n,1}(\mathbf{R})$, alors l'équation $AX = B$ équivaut à $AX = AX_0$, c'est-à-dire $A(X - X_0) = \mathbf{0}$, autrement dit $X - X_0 \in F$, c'est-à-dire qu'il existe $U \in F$ tel que $X = X_0 + U$. La proposition résulte alors aussitôt de la description des solutions dans le cas homogène.

Remarque VII.3.2. *Avec les notations ci-dessus, on dit que X_0 est une* solution particulière *à l'équation $AX = B$. On a montré qu'une solution à l'équation $AX = B$ s'écrit de façon unique $X_0 + U$ où U est une solution à l'équation matricielle homogène $AX = \mathbf{0}$.*

Remarque VII.3.3. *Les points abordés dans les chapitres VI et VII seront approfondis au deuxième semestre.*

VIII. Applications linéaires et multiplication des matrices

Sommaire

VIII.1. Application linéaires

Définition VIII.1.1. *Soit $A \in M_{m,n}(\mathbf{R})$. Si $X \in M_{n,1}(\mathbf{R})$, on a défini le produit $AX \in M_{m,1}(\mathbf{R})$ (cf. définition IV.3.1.1).*

On peut noter $\alpha\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ l'application qui à un vecteur-colonne $X \in M_{n,1}(\mathbf{R})$ associe $\alpha(X) := AX \in M_{m,1}(\mathbf{R})$.

Exemple VIII.1.2. *Si $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}$, alors pour tout $X = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in M_{2,1}(\mathbf{R})$, on a :*

$$AX = x_1 \cdot \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 5x_1 + 6x_2 \end{pmatrix}$$

Ainsi, l'application $\alpha\colon M_{2,1}(\mathbf{R}) \to M_{3,1}(\mathbf{R})$ est caractérisée par le fait que pour tout $(x_1, x_2) \in \mathbf{R}^2$, on a :

$$\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 5x_1 + 6x_2 \end{pmatrix}$$

Définition VIII.1.3. *Supposons que m et n soient deux entiers. Soit $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ une application. On dit que f est une application linéaire si pour tous $X \in M_{n,1}(\mathbf{R})$, $Y \in M_{n,1}(\mathbf{R})$ et $\lambda \in \mathbf{R}$, on a les identités suivantes dans $M_{m,1}(\mathbf{R})$:*

(i) $f(X + Y) = f(X) + f(Y)$;

(ii) $f(\lambda X) = \lambda f(X)$;

Proposition VIII.1.4. *Soit $A \in M_{m,n}(\mathbf{R})$. Notons $\alpha\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ l'application associée à A par la définition VIII.1.1 : $\alpha(X) := AX$.*

Alors, $\alpha\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ est une application linéaire.

Avec les notation de la définition VIII.1.3, cela provient des règles de calculs déjà vues :

$$\begin{array}{rcl}\alpha(X+Y) & = & A\cdot(X+Y) = AX+AY = \alpha(X)+\alpha(Y)\\ \alpha(\lambda X) & = & A\cdot(\lambda X) = \lambda\cdot(AX) = \lambda\alpha(X)\end{array}$$

Proposition VIII.1.5. *Si $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ est une application linéaire. Il existe une unique matrice $A \in M_{m,n}(\mathbf{R})$ telle que pour tout $X \in M_{n,1}(\mathbf{R})$, on ait $f(X) = AX$.*

On dit que $A \in M_{m,n}(\mathbf{R})$ est la matrice de l'application linéaire $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$, et on la note $A := \mathrm{Mat}(f)$.

Pour tout $i \in \{1,\dots,n\}$, notons $E_i := \begin{pmatrix} 0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix} \in M_{n,1}(\mathbf{R})$ le vecteur-colonne où le coefficient 1 est sur le case $(1,i)$.

Soit f une application linéaire comme dans l'énoncé. Notons $V_i := f(E_i) \in M_{m,1}(\mathbf{R})$. Pour tout $(x_1,\dots,x_n) \in \mathbf{R}^n$, notons

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

On a bien sûr $X = x_1E_1 + \cdots + x_nE_n$. En utilisant la linéarité, on obtient :

$$\begin{array}{rcl} f(X) & = & f(x_1E_1+\cdots+x_nE_n)\\ & = & f(x_1E_1)+\cdots+f(x_nE_n)\\ & = & x_1f(E_1)+\cdots+x_nf(E_n)\\ & = & x_1V_1+\cdots+x_nV_n\end{array}$$

Si on note $A \in M_{m,1}(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont $V_1,\dots,V_n$, on vient de montrer que

$$f(X) = x_1V_1+\cdots+x_nV_n = AX$$

Ceci montre l'existence de A. Pour l'unicité, il suffit d'observer que si A convient, alors pour tout $i \in \{1,\dots,n\}$, la i-ième colonne de A est $AE_i = f(E_i) = V_i$.

Proposition VIII.1.6. *Supposons que $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ et $g\colon M_{m,1}(\mathbf{R}) \to M_{p,1}(\mathbf{R})$ soient deux applications linéaires. Alors, l'application composée $g\circ f\colon M_{n,1}(\mathbf{R}) \to M_{p,1}(\mathbf{R})$ (qui est l'application qui à $X \in M_{n,1}(\mathbf{R})$ associe $g(f(X))$) est une application linéaire.*

Si X et Y sont des éléments de $M_{n,1}(\mathbf{R})$ et $\lambda \in \mathbf{R}$, la linéarité de f et g impliquent

$$\begin{array}{rcl}(g\circ f)(X+Y) & = & g(f(X+Y)) = g(f(X)+f(Y))\\ & = & g(f(X))+g(f(Y))\\ & = & (g\circ f)(X)+(g\circ f)(Y)\\ (g\circ f)(\lambda X) & = & g(f(\lambda X)) = g(\lambda f(X))\\ & = & \lambda g(f(X)) = \lambda((g\circ f)(X))\end{array}$$

Définition VIII.1.7. *Soit $n \in \mathbf{N}$. L'application identité $M_{n,1}(\mathbf{R}) \to M_{n,1}(\mathbf{R})$ qui à X associe X a pour matrice la matrice identité $\mathbf{I}_n$:*

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \in M_n(\mathbf{R})$$

VIII.2. Interprétation géométrique

Supposons que l'on dispose d'une matrice $A \in M_{m,n}(\mathbf{R})$ qui corresponde à une application linéaire $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$.

Notons $E_1, \dots, E_n$ les vecteurs introduits précédemment, et notons pour tout $i \in \{1, \dots, n\}$, $V_i := AE_i = f(E_i)$. Les V_i sont les vecteurs-colonnes de A.

Si $M = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n,1}(\mathbf{R})$, on a alors :

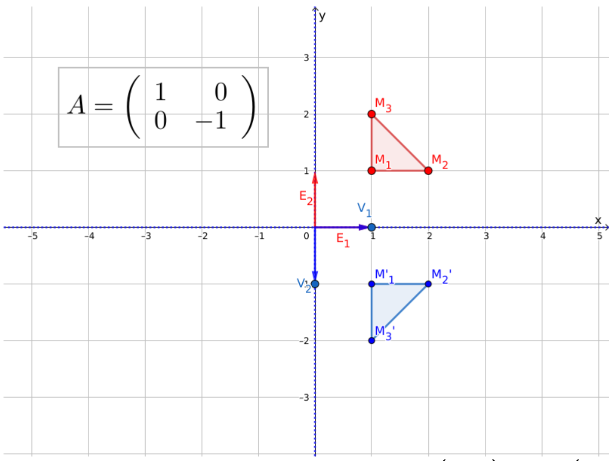
$$M = x_1 E_1 + \dots + x_n E_n \in M_{n,1}(\mathbf{R})$$

et donc si on note $M' := f(M)$, on a :

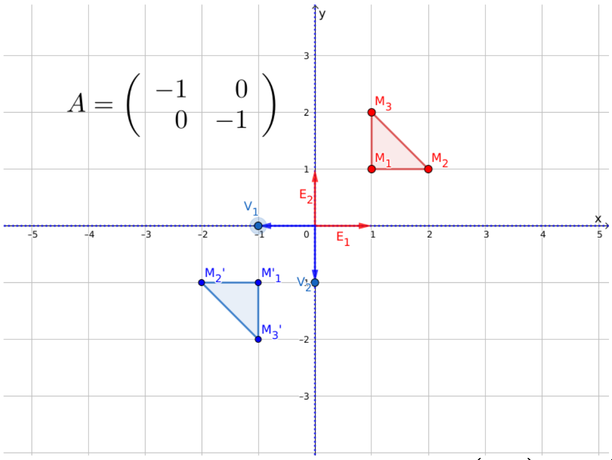
$$M' = x_1 V_1 + \dots + x_n V_n \in M_{m,1}(\mathbf{R})$$

Dans le cas $m = n = 2$, l'application f va de $M_{2,1}(\mathbf{R})$ dans lui-même, et on peut l'identifier à une application $\mathbf{R}^2 \to \mathbf{R}^2$, c'est-à-dire du plan dans lui-même.

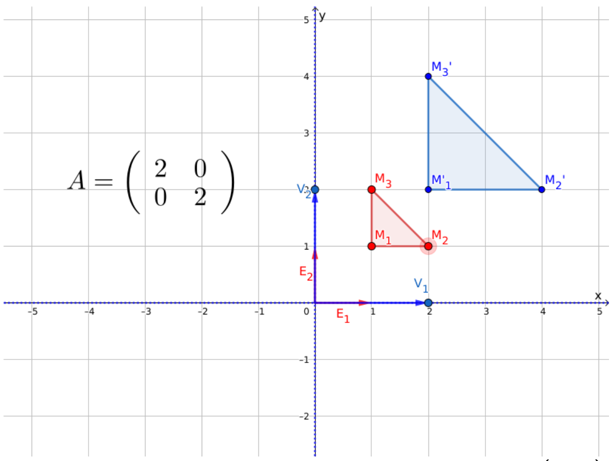
Dans les exemples suivants, on a représenté en rouge les vecteurs $E_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $E_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, et trois points points M_1, M_2, M_3 fixés. On a représenté en bleu deux vecteurs $V_1 = f(E_1)$ et $V_2 = f(E_2)$, ainsi que les points $M_1' := f(M_1)$, $M_2' := f(M_2)$ et $M_3' := f(M_3)$, où f est l'application linéaire associée à la matrice $A \in M_{2,2}(\mathbf{R})$.



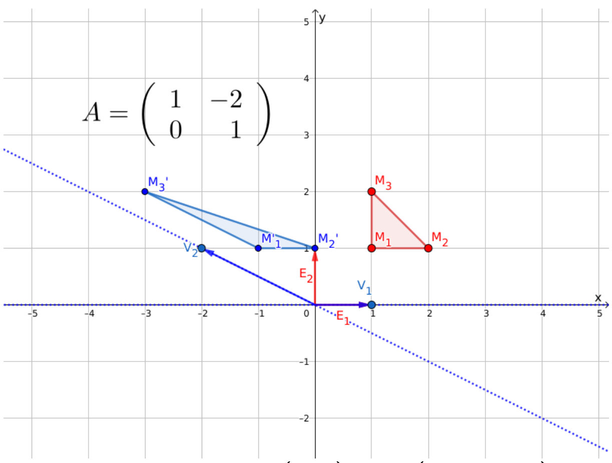
Symétrie orthogonale d'axe (Ox) : $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ -y \end{pmatrix}$



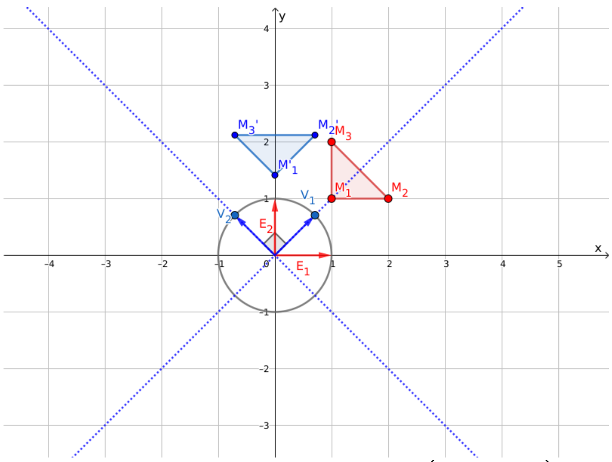
Symétrie centrale de centre l'origine O : $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}$



Homothétie de centre l'origine O et de rapport 2 : $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$



Transvection : $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - 2y \\ y \end{pmatrix}$



Rotation d'angle $\frac{\pi}{2}$: $A = \frac{\sqrt{2}}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

VIII.3. Multiplication des matrices

Pour le moment, comme produits matriciels, nous n'avons utilisé que le cas de produits AX où X est un vecteur-colonne qui a autant de cases que A possède de colonnes.

Définition VIII.3.1. *Supposons que $B \in M_{p,m}(\mathbf{R})$ et $A \in M_{m,n}(\mathbf{R})$ soient deux matrices (telles que le nombre de colonnes de B soit égal au nombre de lignes de A).*

On note BA ou $B \cdot A \in M_{p,n}(\mathbf{R})$ la matrice de l'application linéaire $\varphi\colon M_{n,1}(\mathbf{R}) \to M_{p,1}(\mathbf{R})$ définie par $\varphi(X) := B(AX)$.

Autrement dit, *par définition,* la proposition suivante est vraie :

Proposition VIII.3.2. *Supposons que $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ et $g\colon M_{m,1}(\mathbf{R}) \to M_{p,1}(\mathbf{R})$ soient deux applications linéaires, alors*

$$\mathrm{Mat}(g \circ f) = \mathrm{Mat}(g) \cdot \mathrm{Mat}(f)$$

Exemple VIII.3.3. *Si $B := \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ et $A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, on peut noter $f\colon M_{2,1}(\mathbf{R}) \to M_{2,1}(\mathbf{R})$ l'application linéaire associée à A et $g\colon M_{2,1}(\mathbf{R}) \to M_{3,1}(\mathbf{R})$ celle qui est associée à B. On a les formules suivantes pour tous réels x_1, x_2, y_1, y_2 :*

$$\begin{aligned} f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{pmatrix} \\ g\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} &= B\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_1 + y_2 \end{pmatrix} \end{aligned}$$

Pour calculer $(g \circ f)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = g\left(f\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right)$, il s'agit de faire ci-dessus les substitutions

$y_1 := x_1 + 2x_2$ *et* $y_2 := 3x_1 + 4x_2$, *ce qui donne :*

$$
\begin{aligned}
(g \circ f) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ (x_1 + 2x_2) + (3x_1 + 4x_2) \end{pmatrix} \\
&= \begin{pmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \\ 4x_1 + 6x_2 \end{pmatrix}
\end{aligned}
$$

On reconnaît alors que la matrice de l'application linéaire $g \circ f$ *est :*

$$\mathrm{Mat}(g \circ f) = C := \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 4 & 6 \end{pmatrix}$$

En effet, pour tout $(x_1, x_2) \in \mathbf{R}^2$, *on a :*

$$(g \circ f) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = C \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Par définition, on en déduit que $BA = C$, *c'est-à-dire :*

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

La façon de procéder ci-dessus est très compliquée. Il est possible de calculer un produit matriciel de façon beaucoup plus simple, en se ramenant au cas déjà étudié du produit d'une matrice avec un vecteur-colonne :

Proposition VIII.3.4. *Supposons que* $B \in M_{p,m}(\mathbf{R})$ *et* $A \in M_{m,n}(\mathbf{R})$ *soient deux matrices (telles que le nombre de lignes de* A *soit égal au nombre de colonnes de* B*). Notons* $A_1, \ldots, A_n$ *les vecteurs-colonnes de* A.

Alors, la matrice $BA \in M_{p,n}(\mathbf{R})$ *est la matrice dont les vecteurs-colonnes sont* $BA_1, \ldots, BA_n$.

En effet, on a vu que pour tout $i \in \{1, \ldots, n\}$, on avait $A_i = AE_i$ où $E_i \in M_{n,1}(\mathbf{R})$ est le vecteur précédemment introduit. Si on note f et g les applications linéaires associées respectivement à A et B, on a ainsi $f(E_i) = AE_i = A_i$, puis $(g \circ f)(E_i) = g(f(E_i)) = g(A_i) = BA_i$. Pour tout $i \in \{1, \ldots, n\}$. $(g \circ f)(E_i)$ est le i-ème vecteur-colonne de $\mathrm{Mat}(g \circ f) = BA$, et on vient de montrer que c'était BA_i.

Exemple VIII.3.5. *Avec cette méthode, pour calculer le produit*

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

il suffit de calculer :

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

et donc

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

Alternativement, on dispose d'une formule faisant intervenir trois indices :

Proposition VIII.3.6. *Si $B = (b_{ij})_{1\leq i\leq p, 1\leq j\leq m} \in M_{p,m}(\mathbf{R})$ et Si $A = (a_{ij})_{1\leq i\leq m, 1\leq j\leq n} \in M_{m,n}(\mathbf{R})$, alors le coefficient en position (i,j) de la matrice $BA \in M_{p,n}(\mathbf{R})$ est :*

$$\sum_{k=1}^{m} b_{i,k}a_{k,j} = b_{i,1}a_{1,j} + b_{i,2}a_{2,j} + \cdots + b_{i,m}a_{m,j}$$

(les coefficients qui apparaissent dans la formule sont ceux de la i-ième ligne de B et ceux de la j-ième colonne de A.)

Remarque VIII.3.7. *Si A et B sont deux matrices telles que le produit BA soit défini, suivant les dimensions de ces matrices, il se peut que AB n'ait pas de sens (par exemple $B \in M_{3,2}(\mathbf{R})$ et $A \in M_{2,2}(\mathbf{R})$).*

Si les deux produits AB et BA ont un sens et que ces deux matrices ont la même dimension, il arrive très souvent que $BA \neq AB$.

VIII.4. Règles de calcul

Un certain nombre de règles de calcul ont déjà été énoncées dans la proposition IV.3.2.2.

Proposition VIII.4.1 (Associativité)**.** *Supposons que l'on dispose de trois matrices,*
- $A \in M_{q,p}(\mathbf{R})$;
- $B \in M_{p,m}(\mathbf{R})$;
- $C \in M_{m,n}(\mathbf{R})$.

Il est alors possible de calculer $AB \in M_{q,m}(\mathbf{R})$, et donc $(AB)C \in M_{q,n}(\mathbf{R})$. Par ailleurs, on peut calculer $BC \in M_{p,n}(\mathbf{R})$, et donc $A(BC) \in M_{q,n}(\mathbf{R})$.

Alors, on a $(AB)C = A(BC) \in M_{q,n}(\mathbf{R})$. On peut simplement noter $ABC := (AB)C = A(BC)$.

Notons $h\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ l'application linéaire associée à C, $g\colon M_{m,1}(\mathbf{R}) \to M_{p,1}(\mathbf{R})$ l'application linéaire associée à B et $f\colon M_{p,1}(\mathbf{R}) \to M_{q,1}(\mathbf{R})$ l'application linéaire associée à A.

Par définition, on a :

$$\begin{aligned} AB &= \mathrm{Mat}(f \circ g) \\ BC &= \mathrm{Mat}(g \circ h) \\ (AB)C &= \mathrm{Mat}((f \circ g) \circ h) \\ A(BC) &= \mathrm{Mat}(f \circ (g \circ h)) \end{aligned}$$

Par définition, pour tout $X \in M_{n,1}(\mathbf{R})$, on a :

$$((f \circ g) \circ h)(X) = (f \circ g)(h(X)) = f(g(h(X)))$$
$$(f \circ (g \circ h))(X) = f((g \circ h)(X)) = f(g(h(X)))$$

Par conséquent, $(f \circ g) \circ h = f \circ (g \circ h)$: c'est un cas particulier de l'associativité de la composition des applications. Ce composé peut être noté $f \circ g \circ h$. On en déduit :

$$(AB)C = A(BC)$$

Proposition VIII.4.2. *Supposons que $A \in M_{m,n}(\mathbf{R})$. On a alors :*

$$\mathbf{I}_m \cdot A = A = A \cdot \mathbf{I}_n$$

En effet, si on compose l'application linéaire $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ associée à A avec l'identité de $M_{m,1}(\mathbf{R})$ d'un côté ou avec l'identité de $M_{n,1}(\mathbf{R})$ de l'autre côté, on obtient f.

VIII.5. Noyau, image

La terminologie introduite dans cette section n'est pas exigible à l'examen, mais ce sont des notions essentielles qui seront approfondies au deuxième semestre.

Définition VIII.5.1. *Soit $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ une application linéaire. On note* $\ker f$*, et on appelle « noyau de f », l'ensemble des vecteurs $X \in M_{m,1}(\mathbf{R})$ tels que $f(X) = \mathbf{0}$.*

Dans la notation, « ker » est l'abréviation de *kernel*.

Avec ces notations, si on note $A := \mathrm{Mat}(f) \in M_{m,n}(\mathbf{R})$, alors le noyau $\ker f$ est l'ensemble des vecteurs-colonnes X tels que $AX = \mathbf{0}$. Ainsi, $\ker f$ est l'ensemble des solutions d'un système d'équations linéaires *homogène* : c'est un sous-espace vectoriel de $M_{n,1}(\mathbf{R})$.

Définition VIII.5.2. *Soit $f\colon M_{n,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ une application linéaire. On note* $\mathrm{im}\, f$*, et on appelle « image de f », l'ensemble des vecteurs $Y \in M_{m,1}(\mathbf{R})$ pour lesquels il existe $X \in M_{n,1}(\mathbf{R})$ tel que $Y = f(X)$.*

Si on note encore $A := \mathrm{Mat}(f)$, et que l'on introduit $A_1, \ldots, A_n$ les vecteurs-colonnes de A, alors $\mathrm{im}\, f = \mathrm{Vect}(A_1, \ldots, A_n)$: c'est un sous-espace vectoriel de $M_{m,1}(\mathbf{R})$.

IX. Inverse d'une matrice

Sommaire

IX.1. Caractérisations

Définition IX.1.1. *Soit $A \in M_{m,n}(\mathbf{R})$. On dit que la matrice A est inversible s'il existe une matrice $A^{-1} \in M_{n,m}(\mathbf{R})$ telle que*

$$A \cdot A^{-1} = \mathbf{I}_m \quad \textit{et} \quad A^{-1} \cdot A = \mathbf{I}_n$$

On dit que A^{-1} est la matrice inverse de A.

Si A est inversible, la matrice A^{-1} est unique. En effet, si $B \in M_{n,m}(\mathbf{R})$ vérifie aussi $AB = \mathbf{I}_m$, on a :

$$\begin{aligned} A^{-1}AB &= (A^{-1}A)B = \mathbf{I}_n B = B \\ &= A^{-1}(AB) = A^{-1} \cdot \mathbf{I}_m = A^{-1} \end{aligned}$$

donc $B = A^{-1}$.

On rappelle qu'une matrice carrée est une matrice qui a autant de lignes que de colonnes, c'est-à-dire qu'elle appartient à $M_n(\mathbf{R}) := M_{n,n}(\mathbf{R})$ pour un certain entier n.

Proposition IX.1.2. *Si A est une matrice inversible, alors A est une matrice carrée.*

Notons m et n les entiers tels que $A \in M_{m,n}(\mathbf{R})$, et $A^{-1} \in M_{n,m}(\mathbf{R})$. Notons $A' \in M_{m,n}(\mathbf{R})$ une matrice échelonnée obtenue à partir de A en effectuant des opérations élémentaires sur les lignes.

Étudions la condition $A^{-1}A = \mathbf{I}_n$. Considérons un vecteur-colonne $X \in M_{n,1}(\mathbf{R})$. Si $AX = \mathbf{0}_m$, alors $X = \mathbf{I}_n X = A^{-1}AX = A^{-1}\mathbf{0}_m = \mathbf{0}_n$, donc $X = \mathbf{0}$. Ceci signifie que le système linéaire homogène formulé matriciellement sous la forme $AX = \mathbf{0}_m$ possède $\mathbf{0}_n$ comme unique solution. Ce système (compatible) ne possède donc pas de variable secondaire, c'est-à-dire que A' possède un pivot sur chaque colonne, autrement dit A' possède exactement n pivots.

Étudions maintenant la condition $AA^{-1} = \mathbf{I}_m$. Soit $Y \in M_{m,1}(\mathbf{R})$. Si on pose $X := A^{-1}Y \in M_{n,1}(\mathbf{R})$, on obtient $AX = AA^{-1}Y = \mathbf{I}_m Y = Y$. Ceci étant vrai pour tout vecteur-colonne Y, on a montré que si on note $A_1, \ldots, A_n$ les vecteurs-colonnes de A, alors $\mathrm{Vect}(A_1, \ldots, A_n)$ est égal à $M_{m,1}(\mathbf{R})$. D'après le théorème VI.6.5, ceci signifie que A' possède un pivot sur chaque ligne, c'est-à-dire que A' possède exactement m pivots.

En combinant les deux conditions obtenues, le nombre de pivots de A' est $m = n$, donc A est une matrice carrée.

Compte tenu de la proposition précédente, il n'est réellement intéressant de se demander si une matrice est inversible que s'il s'agit d'une matrice carrée.

Dans la proposition suivante, comme on l'a déjà fait précédemment, on note

$$E_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

le vecteur-colonne où le coefficient 1 est sur la i-ième ligne.

Proposition IX.1.3. *Soit $A \in M_n(\mathbf{R})$. Les conditions suivantes sont équivalentes :*
(i) A est inversible ;
(ii) pour tout $Y \in M_{n,1}(\mathbf{R})$, il existe $X \in M_{n,1}(\mathbf{R})$ tel que $AX = Y$;
(iii) pour tout $i \in \{1, \ldots, n\}$, il existe $X_i \in M_{n,1}(\mathbf{R})$ tel que $AX_i = E_i$.
(iv) il existe $W \in M_n(\mathbf{R})$ tel que $AW = \mathbf{I}_n$.

(i)$\Longrightarrow$(ii). Si A est inversible d'inverse A^{-1}, alors si $Y \in M_{n,1}(\mathbf{R})$, on peut poser $X := A^{-1}Y$, et on a alors $AX = (AA^{-1})Y = \mathbf{I}_n Y = Y$.

(ii)$\Longrightarrow$(iii). Supposons *(ii)*. Pour tout $i \in \{1, \ldots, n\}$, on peut appliquer la condition *(ii)* avec $Y := E_i$, on obtient alors un vecteur-colonne $X_i \in M_{n,1}(\mathbf{R})$ tel que $AX_i = E_i$.

(iii)$\Longrightarrow$(iv). Supposons *(iii)*. Notons $X_i \in M_{n,1}(\mathbf{R})$ des vecteurs-colonnes tels que $AX_i = E_i$ pour tout $i \in \{1, \ldots, n\}$. Notons $W = \begin{pmatrix} X_1 & \ldots & X_n \end{pmatrix} \in M_n(\mathbf{R})$ la matrice dont les vecteurs-colonnes sont les vecteurs X_i. Par une des caractérisations du produit matriciel, AW est la matrice dont les vecteurs colonnes sont $AX_i = E_i$, c'est-à-dire que $AW = \mathbf{I}_n$.

(iv)$\Longrightarrow$(i). Notons $W \in M_n(\mathbf{R})$ une matrice telle que $AW = \mathbf{I}_n$. D'après un des arguments donnés dans la démonstration de la proposition IX.1.2, si on note A' une matrice échelonnée obtenue à partir de A par opérations élémentaires sur les lignes, alors A' possède un pivot sur chaque ligne, et donc aussi sur chaque colonne. Si on note $A_1, \ldots, A_n$ les vecteurs-colonnes de A, alors $(A_1, \ldots, A_n)$ est une famille libre, c'est-à-dire que l'équation matricielle $AX = \mathbf{0}_n$ avec $X \in M_{n,1}(\mathbf{R})$ admet $X = \mathbf{0}_n$ comme unique solution. Notons $D := WA - \mathbf{I}_n$. On a $AD = AWA - A = (AW)A - A = \mathbf{I}_n A - A = A - A = \mathbf{0}_{n,n}$. On a montré $AD = \mathbf{0} \in M_n(\mathbf{R})$. Si on note $D_1, \ldots, D_n$ les vecteurs-colonnes de D, ceci signifie que $AD_i = \mathbf{0}$ pour tout $i \in \{1, \ldots, n\}$, et on a vu que cela impliquait $D_i = \mathbf{0}$. Par conséquent, $D = WA - \mathbf{I}_n$ est la matrice nulle, donc $WA = \mathbf{I}_n$. Finalement, on a $AW = WA = \mathbf{I}_n$, donc A est inversible et $A^{-1} = W$.

Remarque IX.1.4. *Si les conditions équivalentes de la proposition précédente sont vérifiées, alors dans les conditions (ii), (iii), (iv), les matrices dont l'existence est affirmée sont uniques.*

Proposition IX.1.5. *Soit $A \in M_n(\mathbf{R})$ une matrice inversible. Soit $Y \in M_{n,1}(\mathbf{R})$.*
Alors, le système d'équations formulé matriciellement sous la forme $AX = Y$ (avec $X \in M_{n,1}(\mathbf{R})$) possède une unique solution qui est $X = A^{-1}Y$.

En effet, si $AX = Y$, alors en multipliant à gauche par A^{-1}, on obtient $X = \mathbf{I}_n X = A^{-1}AX = A^{-1}Y$. Réciproquement, si on pose $X := A^{-1}Y$, on a bien $AX = A^{-1}AY = Y$.

IX.2. Cas de la dimension 2

Proposition IX.2.1. *Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbf{R})$. On rappelle que le déterminant* $\det A$ *est défini par :*

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

La matrice A est inversible si et seulement si $\det A \neq 0$.
Si $\det A \neq 0$*, alors :*

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Il a été vu au cours de la démonstration de la proposition IX.1.3 que si A est inversible, alors ses vecteurs-colonnes (ici, $\begin{pmatrix} a \\ c \end{pmatrix}$ et $\begin{pmatrix} b \\ d \end{pmatrix}$) forment une famille libre. D'après la proposition VI.5.2.3, ceci signifie que $\det A \neq 0$.
Il s'agit de montrer la réciproque. Commençons par observer que si on note

$$B := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

alors on a toujours :

$$AB = \begin{pmatrix} \det A & 0 \\ 0 & \det A \end{pmatrix}$$

Ainsi, si $\det A \neq 0$, on peut noter $W := \frac{1}{\det A} B$ et on a

$$AW = \mathbf{I}_2$$

D'après la condition *(iv)* de la proposition IX.1.3, on obtient que A est inversible, et son inverse est W.

IX.3. Calcul de l'inverse

Proposition IX.3.1. *Soit $A \in M_n(\mathbf{R})$. Soit A' une matrice échelonnée obtenue à partir de A en effectuant des opérations élémentaires sur les lignes.*
La matrice A est inversible si et seulement si la matrice A' possède n pivots (sur les cases (i, i) pour $i \in \{1, \dots, n\}$).
Si on suppose de plus que A' est échelonnée réduite, alors A est inversible si et seulement $A' = \mathbf{I}_n$.

La condition énoncée sur les pivots de A' (qui se traduit par $A' = \mathbf{I}_n$ dans le cas où A' est échelonnée réduite) signifie, d'après les résultats du chapitre VI que les vecteurs-colonnes de A forment une famille à la fois libre, et génératrice de $M_{n,1}(\mathbf{R})$, ou encore que pour tout $B \in M_{n,1}(\mathbf{R})$, l'équation matricielle $AX = B$ possède une unique solution $X \in M_{n,1}(\mathbf{R})$. D'après les caractérisations de la proposition IX.1.3, on obtient immédiatement l'équivalence avec le fait que A soit inversible.

Remarque IX.3.2. *Si $V_1, \ldots, V_n$ sont les vecteurs-colonnes de la matrice $A \in M_n(\mathbf{R})$, alors la proposition précédente permet de dire que la matrice A est inversible si et seulement si les vecteurs $V_1, \ldots, V_n$ forment une famille libre (puisque cette condition se traduit aussi par le fait que A' possède un pivot sur chaque colonne), et c'est aussi équivalent au fait que les vecteurs $V_1, \ldots, V_n$ forment une famille génératrice de $M_{n,1}(\mathbf{R})$ (présence d'un pivot sur chaque ligne de A'), cf. corollaire VI.6.6).*

On a vu à la proposition IX.1.3 que pour déterminer l'inverse d'une matrice inversible $A \in M_n(\mathbf{R})$, il s'agissait de déterminer pour tout $i \in \{1, \ldots, n\}$, l'unique vecteur-colonne $X_i \in M_{n,1}(\mathbf{R})$ qui vérifie $AX_i = E_i$. La matrice A^{-1} est alors la matrice dont les vecteurs-colonnes sont ces vecteurs X_i.

Une idée serait alors d'introduire pour tout $i \in \{1, \ldots, n\}$ la matrice augmentée

$$M_i := \left(\begin{array}{c|c} A & E_i \end{array} \right) \in M_{n,n+1}(\mathbf{R})$$

et d'appliquer l'algorithme du pivot de Gauss pour mettre A sous forme échelonnée réduite, on obtiendra alors finalement une matrice M_i' équivalente :

$$M_i' := \left(\begin{array}{c|c} \mathbf{I}_n & X_i \end{array} \right)$$

Par exemple, considérons la matrice

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 7 \end{pmatrix} \in M_2(\mathbf{R})$$

et appliquons l'algorithme du pivot de Gauss aux deux matrices augmentées suivantes :

$$M_1 = \left(\begin{array}{cc|c} 1 & 3 & 1 \\ 2 & 7 & 0 \end{array} \right) \qquad M_2 = \left(\begin{array}{cc|c} 1 & 3 & 0 \\ 2 & 7 & 1 \end{array} \right)$$

Effectuons l'opération $L_2 \longleftarrow L_2 - 2L_1$:

$$M_1 \sim \left(\begin{array}{cc|r} 1 & 3 & 1 \\ 0 & 1 & -2 \end{array} \right) \qquad M_2 \sim \left(\begin{array}{cc|c} 1 & 3 & 0 \\ 0 & 1 & 1 \end{array} \right)$$

Effectuons l'opération $L_1 \longleftarrow L_1 - 3L_2$:

$$M_1 \sim \left(\begin{array}{cc|r} 1 & 0 & 7 \\ 0 & 1 & -2 \end{array} \right) \qquad M_2 \sim \left(\begin{array}{cc|r} 1 & 0 & -3 \\ 0 & 1 & 1 \end{array} \right)$$

On en déduit que

$$X_1 = \begin{pmatrix} 7 \\ -2 \end{pmatrix} \qquad X_2 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

Par conséquent,

$$A^{-1} = \begin{pmatrix} X_1 & X_2 \end{pmatrix} = \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$$

Comme cela a déjà été remarqué, on peut faire les mêmes opérations élémentaires sur toutes les matrices $M_1, \ldots, M_n$: cela ne dépend que de la matrice A. Les méthodes de calcul décrites dans le paragraphe suivant tiennent compte de ce fait.

IX.3.3. Méthodes de calcul

Première méthode (recommandée)

Pour déterminer l'inverse d'une matrice $A \in M_n(\mathbf{R})$, une méthode consiste à considérer la matrice $M = \left(A \,\middle|\, \mathbf{I}_n \right) \in M_{n,2n}(\mathbf{R})$ obtenue en rajoutant n colonnes (les vecteurs $E_1, \ldots, E_n$) à la matrice A. Ceci permet de résoudre simultanément n systèmes linéaires (ayant les mêmes membres de gauche, mais des membres de droite différents).

Dans l'exemple précédent, on obtient ceci :

$$M = \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & 7 & 0 & 1 \end{array} \right) \in M_{2,4}(\mathbf{R})$$

Effectuons l'opération $L_2 \longleftarrow L_2 - 2L_1$:

$$M \sim \left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{array} \right)$$

Effectuons l'opération $L_1 \longleftarrow L_1 - 3L_2$:

$$M \sim \left(\begin{array}{cc|cc} 1 & 0 & 7 & -3 \\ 0 & 1 & -2 & 1 \end{array} \right)$$

On en déduit que

$$A^{-1} = \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$$

Deuxième méthode

Alternativement, si on note $f\colon M_{n,1}(\mathbf{R}) \to M_{n,1}(\mathbf{R})$ l'application linéaire associée à $A \in M_n(\mathbf{R})$, on peut chercher à déterminer l'application linéaire $f^{-1}\colon M_{n,1}(\mathbf{R}) \to M_{n,1}(\mathbf{R})$ associée à A^{-1}.

Pour tout $Y \in M_{n,1}(\mathbf{R})$, on cherche $X \in M_{n,1}(\mathbf{R})$ tel que $f(X) = Y$, c'est-à-dire $AX = Y$, de sorte que $f^{-1}(Y) = X$. Si on fixe $Y \in M_{n,1}(\mathbf{R})$, $f^{-1}(Y)$ est l'unique solution au système d'équation $AX = Y$ (en la variable vectorielle X), autrement dit l'unique solution au système associé à la matrice augmentée $\left(A \,\middle|\, Y \right) \in M_{n,n+1}(\mathbf{R})$.

Concrètement, de façon similaire à ce qui a été vu dans le paragraphe VI.4, il est possible de considérer des paramètres réels $y_1, \ldots, y_n$ arbitraires, et d'introduire la matrice augmentée :

$$\widehat{A} = \left(\begin{array}{ccccc|c} a_{1,1} & \ldots & a_{1,j} & \ldots & a_{1,n} & y_1 \\ a_{2,1} & \ldots & a_{2,j} & \ldots & a_{2,n} & y_2 \\ \vdots & \ldots & \vdots & \ldots & \vdots & \vdots \\ a_{i,1} & \ldots & a_{i,j} & \ldots & a_{i,n} & y_i \\ \vdots & \ldots & \vdots & \ldots & \vdots & \vdots \\ a_{n,1} & \ldots & a_{n,j} & \ldots & a_{n,n} & y_n \end{array} \right) \in M_{n,n+1}(\mathbf{R})$$

Si on note $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, l'unique solution de ce système d'équation sera $f^{-1}(Y)$. On peut ainsi déterminer l'expression de l'application linéaire $f^{-1}\colon M_{n,1}(\mathbf{R}) \to M_{n,1}(\mathbf{R})$, ce qui permet d'obtenir A^{-1}.

Dans le cas de l'exemple considéré précédent, il s'agit d'appliquer l'algorithme du pivot de Gauss à la matrice augmentée suivante :

$$\widehat{A} = \left(\begin{array}{cc|c} 1 & 3 & y_1 \\ 2 & 7 & y_2 \end{array}\right) \in M_{2,3}(\mathbf{R})$$

Effectuons l'opération $L_2 \longleftarrow L_2 - 2L_1$:

$$M \sim \left(\begin{array}{cc|c} 1 & 3 & y_1 \\ 0 & 1 & -2y_1 + y_2 \end{array}\right)$$

Effectuons l'opération $L_1 \longleftarrow L_1 - 3L_2$:

$$M \sim \left(\begin{array}{cc|c} 1 & 0 & 7y_1 - 3y_2 \\ 0 & 1 & -2y_1 + y_2 \end{array}\right)$$

On en déduit que l'unique couple $(x_1, x_2) \in \mathbf{R}^2$ qui soit solution du système d'équation considéré est donné par les formules suivantes :

$$\begin{cases} x_1 &= 7y_1 - 3y_2 \\ x_2 &= -2y_1 + y_2 \end{cases}$$

Autrement dit, on a :

$$f^{-1}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 7y_1 - 3y_2 \\ -2y_1 + y_2 \end{pmatrix}$$

On retrouve donc bien :

$$A^{-1} = \mathrm{Mat}(f^{-1}) = \begin{pmatrix} 7 & -3 \\ -2 & 1 \end{pmatrix}$$

Remarque IX.3.4. *Dans les méthodes qui ont été discutées ci-dessus, si on ne sait pas* a priori *que la matrice A est inversible, on s'en rendra forcément compte en appliquant l'algorithme du pivot de Gauss, puisqu'une fois que l'on a mis A sous forme échelonnée réduite, les n premières colonnes de la matrice obtenue par opérations élémentaires s'identifient à la matrice identité $\mathbf{I}_n$ si et seulement si A est inversible (cf. proposition IX.3.1).*

IX.4. Propriétés des matrices inverses

Proposition IX.4.1. *Soit $n \in \mathbf{N}$.*

- *Si $A \in M_n(\mathbf{R})$, est une matrice inversible, alors A^{-1} est aussi une matrice inversible et $(A^{-1})^{-1} = A$.*
- *Si $A \in M_n(\mathbf{R})$ et $B \in M_n(\mathbf{R})$ sont deux matrices inversibles, alors AB est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.*

La première assertion est évidente. Montrons la deuxième. Notons $C := B^{-1}A^{-1}$. On a :

$$(AB)C = (AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A\mathbf{I}_nA^{-1} = AA^{-1} = \mathbf{I}_n$$

$$C(AB) = (B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}\mathbf{I}_nB = B^{-1}B = \mathbf{I}_n$$

Ces identités montrent que AB est inversible et que son inverse est $C = B^{-1}A^{-1}$.

Définition IX.4.2 (Transposée). *Soit $A \in M_{m,n}(\mathbf{R})$ une matrice. La matrice transposée de A, notée ${}^tA \in M_{n,m}(\mathbf{R})$, est la matrice dont le coefficient en position (i,j) est le coefficient $a_{j,i}$. (On a bien sûr, ${}^t({}^tA) = A$.)*

Par exemple, la transposée de la matrice $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \in M_{2,3}(\mathbf{R})$ est :

$${}^tA = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in M_{3,2}(\mathbf{R})$$

(Dans certaines références, la transposée d'une matrice A est parfois notée A^T.)

Proposition IX.4.3. *Si $A \in M_{p,m}(\mathbf{R})$, $B \in M_{m,n}(\mathbf{R})$, alors on a l'égalité suivante dans $M_{n,p}(\mathbf{R})$:*

$${}^t(AB) = {}^tB\,{}^tA$$

Corollaire IX.4.4. *Soit $A \in M_n(\mathbf{R})$ une matrice inversible. Alors, ${}^tA$ est aussi inversible, et son inverse est ${}^t(A^{-1})$.*

Notons $B := A^{-1}$. On a donc ${}^t(AB) = {}^t\mathbf{I_n} = \mathbf{I}_n$. D'après la proposition précédente, ceci implique que ${}^tB\,{}^tA = \mathbf{I}_n$. De façon similaire, on montre que ${}^tA^tB = \mathbf{I}_n$. Ceci montre que ${}^tA$ est inversible d'inverse ${}^tB = {}^t(A^{-1})$.

On peut alors utiliser la notation ${}^tA^{-1}$ puisqu'il n'est pas nécessaire de spécifier si c'est l'inverse de la transposée ou bien la transposée de l'inverse.

IX.5. Opérations élémentaires

Considérons la matrice P suivante :

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

Si $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, on a alors :

$$PX = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 2x_1 + x_3 \end{pmatrix}$$

On remarque que PX est le vecteur-colonne obtenu à partir de X en ajoutant deux fois la première case à la troisième case.

Plus généralement, si $A \in M_{3,n}(\mathbf{R})$ est une matrice ayant un nombre arbitraire de colonnes, et que l'on note L_1, L_2, L_3 les vecteurs-lignes (dans $M_{1,n}(\mathbf{R})$), alors

$$PA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 + 2L_1 \end{pmatrix}$$

Autrement dit, PA est la matrice obtenue en effectuant l'opération élémentaire $L_3 \longleftarrow L_3 + 2L_1$ sur la matrice A.

Plus généralement, on a la proposition suivante :

Proposition IX.5.1. *Soit m un entier. Considérons une opération élémentaire de la forme $L_i \longleftarrow L_i - \lambda L_j$ pour $i \neq j$ deux entiers appartenant à $\{1, \ldots, m\}$. Notons $P \in M_m(\mathbf{R})$ la matrice dont les coefficients sont nuls, sauf sur les cases (k, k) pour $k \in \{1, \ldots, m\}$ où le coefficient vaut 1, et sur la case (i, j) où le coefficient vaut $-\lambda$.*

Alors, pour tout entier $n \in \mathbf{N}$, et toute matrice $A \in M_{m,n}(\mathbf{R})$, la matrice PA est le résultat de l'opération élémentaire sur les lignes $L_i \longleftarrow L_i - \lambda L_j$ appliquée à la matrice A.

(L'exemple étudié précédemment était le cas où $m = 3$, $i = 3$, $j = 1$ et $\lambda = -2$.)

Notons $f : M_{m,1}(\mathbf{R}) \to M_{m,1}(\mathbf{R})$ l'application linéaire associée à P, c'est-à-dire que $f(X) = PX$. On rappelle que l'on dispose des vecteurs-colonnes $E_1, \ldots, E_m$ appartenant à $M_{m,1}(\mathbf{R})$ et que si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$, alors

$$X = x_1 E_1 + \cdots + x_m E_m$$

Soit $k \in \{1, \ldots, m\}$. Calculons $f(E_k) = PE_k$: il s'agit de la k-ième colonne de la matrice P. On obtient immédiatement que

$$f(E_k) = PE_k = \begin{cases} E_k & \text{si } k \neq j \\ E_k - \lambda E_i & \text{si } k = j \end{cases}$$

Par linéarité, on en déduit en général :

$$\begin{aligned} PX = f(X) &= f(x_1 E_1) + \cdots + f(x_m) E_m \\ &= x_1 f(E_1) + \cdots + x_m f(E_m) \\ &= (x_1 E_1 + \cdots + x_m E_m) - \lambda x_j E_i \\ &= X - \lambda x_j E_i \end{aligned}$$

Le seul coefficient qui a peut-être changé entre X et PX est le i-ème : le i-ième coefficient de PX est $x_i - \lambda x_j$. Ceci montre le résultat voulu dans le cas où la matrice A de la proposition n'a qu'une seule colonne, donc est un vecteur-colonne que l'on a noté ici X. En raisonnant colonne par colonne, on obtient le résultat voulu.

Définition IX.5.2. *Les matrices P de la forme décrite dans la proposition précédente sont appelées « élémentaires ».*

On a ainsi montré qu'effectuer une opération élémentaire de la forme $L_i \longleftarrow L_i - \lambda L_j$ revenait à multiplier à gauche par une matrice élémentaire.

On pourrait démontrer de façon similaire que les autres opérations $L_i \longleftrightarrow L_j$ ($i \neq j$) et $L_i \longleftarrow \lambda L_i$ ($\lambda \neq 0$) peuvent aussi s'interpréter comme la multiplication à gauche par une matrice bien choisie, mais nous n'en aurons pas besoin.

Si, partant d'une matrice A, on effectue une première opération élémentaire correspondant à la multiplication à gauche par une matrice P_1, on obtient $P_1 A$. Si on continue et que l'on applique une deuxième opération élémentaire correspondant à la multiplication à gauche par P_2, on obtiendra la matrice $P_2(P_1 A) = P_2 P_1 A$.

En général, si on applique successivement à une matrice A une suite de ℓ opérations élémentaires sur les lignes dont la i-ième s'interprète comme la multiplication à gauche par la matrice élémentaire P_i, alors le résultat est le produit matriciel

$$A' := P_\ell \cdots P_2 P_1 A$$

Il a déjà été remarqué que si on effectue successivement une opération élémentaire $L_i \longleftarrow L_i - \lambda L_j$, puis l'opération $L_i \longleftarrow L_i + \lambda L_j$, alors on retombait sur la matrice de départ. Ceci permet de montrer que toutes les matrices P_i ci-dessus sont inversibles et que leur inverse P_i^{-1} est aussi une matrice élémentaire. On a donc :

$$(P_\ell \cdots P_2 P_1)^{-1} = P_1^{-1} P_2^{-1} \cdots P_\ell^{-1}$$

Ainsi, on a également :

$$A = P_1^{-1} P_2^{-1} \cdots P_\ell^{-1} A'$$

IX.6. Matrices triangulaires

Définition IX.6.1. *Soit $T = (t_{i,j})_{1 \leq i,j \leq n} \in M_n(\mathbf{R})$ une matrice carrée. On dit que T est triangulaire inférieure si $t_{i,j} = 0$ pour tout (i,j) tels que $i < j$.*

On dit que T est triangulaire supérieure si $t_{i,j} = 0$ pour tout (i,j) tel que $i > j$, c'est-à-dire si la transposée ${}^tT$ est triangulaire inférieure.

Par exemple, les matrices suivantes sont triangulaires inférieures :

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 4 & 5 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ -2 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ -2 & -1 & 0 \\ -5 & 4 & 1 \end{pmatrix}$$

Leurs transposées sont triangulaires supérieures :

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \\ 0 & 0 & 5 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 \\ 0 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 & -5 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix}$$

Proposition IX.6.2. *Si A et B sont deux matrices carrées de même taille, et que A et B sont triangulaires inférieures, alors AB aussi.*

Si A est une matrice triangulaire inférieure inversible, alors A^{-1} aussi.

Cette proposition est admise. Le même énoncé vaut pour les matrices triangulaires supérieures.

IX.7. Décomposition LU

Définition IX.7.1. *Soit $A \in M_{m,n}(\mathbf{R})$. Une décomposition LU de M consiste en la donnée deux matrices $L \in M_m(\mathbf{R})$ et $U \in M_{m,n}(\mathbf{R})$ telles que*

$$A = LU$$

avec L triangulaire inférieure avec des 1 sur la diagonale et U échelonnée.

(On remarque que si $m = n$, la matrice échelonnée U est nécessairement triangulaire supérieure.)

Dans l'algorithme du pivot de Gauss de mise sous forme échelonnée (cf. définition V.3.4), on trouve deux types d'opérations :
- des opérations de la forme $L_i \longleftarrow L_i - \lambda L_j$, avec $i > j$.
- des opérations d'échanges entre lignes $L_i \longleftrightarrow L_j$.

Proposition IX.7.2. *Soit $A \in M_{m,n}(\mathbf{R})$. Supposons que lors de la mise sous forme échelonnée de la matrice A avec l'algorithme du pivot de Gauss, il ne soit pas nécessaire d'échanger des lignes. Alors, la matrice A possède une décomposition LU.*

En outre, si A est inversible (et donc carrée), alors cette décomposition LU est unique (et alors, la matrice échelonnée U est aussi triangulaire supérieure).

L'algorithme du pivot de Gauss appliqué à la matrice A se déroule ainsi.

Si A est non nulle, on note c_1 l'indice de la première colonne non nulle de A. Comme on fait l'hypothèse qu'il n'est pas nécessaire d'échanger des lignes pendant l'application de l'algorithme du pivot, on a $a_{1,c_1} \neq 0$ et on choisit ce coefficient comme pivot. Pour i allant de 2 jusqu'à m, on effectue une opération élémentaire de la forme $L_i \longleftarrow L_i - \lambda_{i,1} L_1$ pour des coefficients $\lambda_{i,1}$ bien choisis.

Plus loin dans l'algorithme, lorsque l'on doit choisir le k-ième pivot (pour $k \in \{1, \dots, r\}$), on doit examiner certains coefficients de la c_k-ième colonne, et plus précisément ceux des cases (i, c_k) pour $k \leq i \leq m$. Notre hypothèse fait que l'on peut choisir le coefficient situé sur la case (k, c_k) comme pivot. On effectue alors des opérations élémentaires de la forme $L_i \longleftarrow L_i - \lambda_{i,k} L_k$ pour $i \in \{k+1, \dots, m\}$.

Notons U la matrice échelonnée obtenue à la fin de l'algorithme du pivot de Gauss. Si on note $P_1, P_2, \dots, P_\ell$ les matrices élémentaires associées aux opérations élémentaires sur les lignes effectuées lors de l'algorithme du pivot de Gauss, partant de la matrice A, on aura obtenu successivement les matrices $P_1 A$, $P_2 P_1 A$, etc, et finalement

$$U = P_\ell \dots P_2 P_1 A$$

Comme les opérations qui sont faites sont de la forme $L_i \longleftarrow L_i - \lambda_{i,k} L_k$ avec $i > k$, on remarque que toutes les matrices P_s pour $s \in \{1, \dots, \ell\}$ sont triangulaires inférieures avec des 1 sur le diagonale. Introduisons la matrice suivante :

$$L := (P_\ell \dots P_2 P_1)^{-1} = P_1^{-1} P_2^{-1} \dots P_\ell^{-1}$$

On a alors $A = LU$. Il reste à justifier que $L \in M_m(\mathbf{R})$ est bien triangulaire inférieure avec des 1 sur la diagonale. On pourrait montrer ce fait *a priori*, mais nous allons calculer explicitement cette matrice L.

La matrice L a la propriété que

$$P_\ell \dots P_2 P_1 L = \mathbf{I}_m$$

Autrement dit, si on applique à L les opérations élémentaires sur les lignes correspondant aux matrices P_i (en commençant par P_1, puis P_2, etc., jusqu'à P_ℓ), alors on obtient la matrice identité.

Vérifions que L est la matrice dont tous les coefficients sont nuls, sauf les coefficients diagonaux qui vallent 1, et les coefficients des cases (i,k) pour $i > k$, qui vallent $\lambda_{i,k}$. (Au cours de l'algorithme ci-dessus, ces coefficients ne sont définis que pour $k \in \{1,\dots,r\}$. On convient que ces coefficients sont nuls pour $k > r$.)

$$L = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ \lambda_{2,1} & 1 & 0 & \dots & 0 \\ \lambda_{3,1} & \lambda_{3,2} & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \lambda_{m,1} & \lambda_{m,2} & \dots & \lambda_{m,m-1} & 1 \end{pmatrix}$$

Si on applique la première série d'opérations élémentaires faites lors de l'algorithme du pivot de Gauss ci-dessus, il s'agit d'appliquer les opérations $L_i \longleftarrow L_i - \lambda_{i,1} L_1$ pour $i \in \{2,\dots,m\}$. On voit immédiatement que le résultat est d'annuler les coefficients de la première colonne à l'exception du coefficient 1 situé en haut. Notons $L^{(1)}$ la matrice obtenue :

$$L^{(1)} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_{3,2} & 1 & \ddots & 0 \\ 0 & \vdots & \ddots & \ddots & 0 \\ 0 & \lambda_{m,2} & \dots & \lambda_{m,m-1} & 1 \end{pmatrix}$$

Ensuite, il s'agit d'appliquer successivement pour k allant de 2 jusqu'à r, un certain ensemble d'opérations élémentaires, de la forme $L_i \longleftarrow L_i - \lambda_{i,k} L_k$ pour $i \in \{k+1,\dots,m\}$. Notons $L^{(k)}$ la matrice obtenue en appliquant ces opérations (à k fixé) à la matrice $L^{(k-1)}$.

Il est immédiat que $L^{(2)}$ est obtenue en remplaçant les coefficients de la deuxième colonne par des zéros (à l'exception du coefficient de la case $(2,2)$ qui est inchangé). Plus généralement, de proche en proche, la matrice $L^{(k)}$ aura la propriété que ses k premières colonnes sont les mêmes que celles de la matrice $\mathbf{I}_n$, tandis que les $m-k$ dernières colonnes seront les mêmes que celles de la matrice L.

Finalement, il est évident qu'à la fin on a : $L^{(r)} = \mathbf{I}_m$. Ceci montre que la matrice qui a été construite en utilisant les coefficients $\lambda_{i,j}$ est effectivement une matrice L triangulaire inférieure avec des 1 sur les diagonale qui vérifie $A = LU$.

Exemple IX.7.3. *Déterminons la décomposition LU de la matrice suivante :*

$$A = \begin{pmatrix} {\color{red}1} & -5 & 2 \\ 3 & -11 & -4 \\ 5 & -17 & 2 \end{pmatrix}$$

Effectuons les opérations $L_2 \longleftarrow L_2 - 3L_1$ et $L_3 \longleftarrow L_3 - 5L_1$:

$$A \sim \begin{pmatrix} 1 & -5 & 2 \\ 0 & 4 & -2 \\ 0 & 8 & -8 \end{pmatrix}$$

Enfin, faisons l'opération $L_3 \longleftarrow L_3 - 2L_2$:

$$A \sim U := \begin{pmatrix} 1 & -5 & 2 \\ 0 & 4 & -2 \\ 0 & 0 & -4 \end{pmatrix}$$

On a effectué successivement les opérations élémentaires suivantes :

$$L_2 \longleftarrow L_2 - \lambda_{2,1} L_1$$

$$L_3 \longleftarrow L_3 - \lambda_{3,1} L_1$$

$$L_3 \longleftarrow L_3 - \lambda_{3,2} L_2$$

avec $\lambda_{2,1} = 3$, $\lambda_{3,1} = 5$ et $\lambda_{3,2} = 2$.

D'après la méthode générale, on peut poser :

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \lambda_{2,1} & 1 & 0 \\ \lambda_{3,1} & \lambda_{3,2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 5 & 2 & 1 \end{pmatrix}$$

Et on a alors $A = LU$.

(On peut vérifier que si on applique les opérations élémentaires réalisées plus haut (dans le même ordre) à la matrice L, on obtient bien la matrice identité $\mathbf{I}_3$.)

Remarque IX.7.4. *Si $A \in M_n(\mathbf{R})$ admet une décomposition LU, alors pour déterminer les solutions X à une équation matricielle $AX = B$ avec B et X dans $M_{n,1}(\mathbf{R})$, il s'agit de déterminer les X tels que $LUX = B$. On peut alors commencer par déterminer l'unique $Y \in M_{n,1}(\mathbf{R})$ tels que $LY = B$, ce qui revient à résoudre un système plus ou moins triangulaire. Ensuite, il reste à déterminer X en résolvant l'équation $UX = Y$ qui correspond à un système triangulaire.*

Exemple IX.7.5. *Avec la matrice A de l'exemple IX.7.3, on peut considérer l'équation*

$$AX = \begin{pmatrix} 3 \\ 7 \\ 23 \end{pmatrix}$$

D'après la remarque précédente, on étudie d'abord le système correspondant à l'équation $LY = \begin{pmatrix} 3 \\ 7 \\ 23 \end{pmatrix}$:

$$\begin{cases} y_1 & & & = 3 \\ 3y_1 + y_2 & & & = 7 \\ 5y_1 + 2y_2 + y_3 & & & = 23 \end{cases}$$

L'unique solution est $Y = \begin{pmatrix} 3 \\ -2 \\ 12 \end{pmatrix}$.

On déterminer enfin les éventuels $X \in M_{3,1}(\mathbf{R})$ *vérifiant l'équation* $UX = Y$ *:*

$$\left\{ \begin{array}{rcrcrcr} x_1 & - & 5x_2 & + & 2x_3 & = & 3 \\ & & 4x_2 & - & 2x_3 & = & -2 \\ & & & & -4x_3 & = & 12 \end{array} \right.$$

Finalement, l'unique solution est $X := \begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix}$.