Algèbre linéaire : Fondements et applications

Nathan Allaire Sacha Benarroch-Lelong Théo Denorme
Mathieu Gervais-Dubé Andréa Gourion Alain Hertz
Sébastien Le Digabel Rémi Pédenon-Orlanducci Camille Pinçon

31 décembre 2025

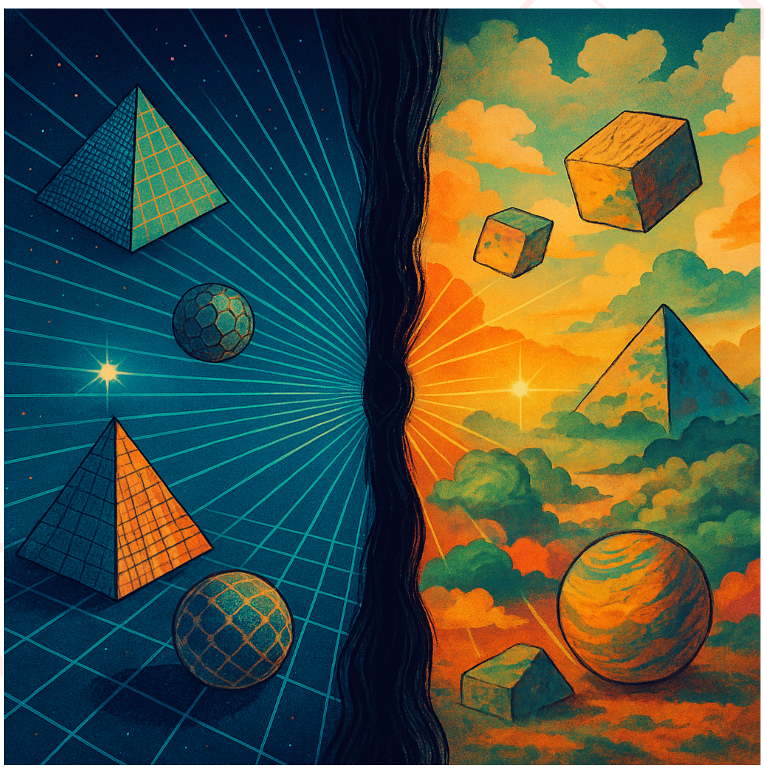

image générée par DALL·E.

- ▷ Ce document est un prototype de manuel pour des cours d’algèbre linéaire de premier cycle.
- ▷ À terme, il sera distribué en tant que RÉL (ressource éducative libre), en source LaTeX libre, utilisation libre, et adaptation libre.
- ▷ Cette version prototype est destinée au seul usage des étudiants du cours MTH1008 de Polytechnique Montréal, pour la session d’hiver 2026.
- ▷ Le document est en cours d’écriture, mais assez mature pour être testé dans un cadre réel. Il sera progressivement augmenté durant la session d’hiver 2026, en fonction de l’avancement de son écriture et des commentaires reçus de la part des étudiants.
- ▷ Il est donc normal que plusieurs liens ne fonctionnent pas encore.
- ▷ Merci de reporter toute erreur à `sebastien.le-digabel@polymtl.ca`.

Table des matières

Préface

Des systèmes dynamiques au traitement du signal, de l'optimisation à la commande automatique, des réseaux de communication à l'apprentissage machine, l'algèbre linéaire est, dans tous les domaines scientifiques, essentielle. Souvent perçue comme inutilement abstraite ou trop théorique, son utilité concrète n'est pas toujours perceptible de prime abord. Le but de cet ouvrage est de proposer une approche pragmatique de l'algèbre linéaire, structurée clairement entre fondements théoriques et applications pratiques. La partie Fondements pose un cadre théorique et rigoureux, et définit toutes les notions nécessaires aux principales utilisations de l'algèbre linéaire appliquée, présentées dans la partie Applications : projections, diagonalisation, inégalités matricielles, décomposition en valeurs singulières, etc.

Le présent ouvrage s'adresse en priorité aux étudiants de premier cycle universitaire en sciences appliquées, et plus particulièrement en ingénierie, ainsi qu'à toute personne souhaitant renforcer sa maîtrise de l'algèbre linéaire en vue d'applications techniques, en proposant une approche résolument tournée vers l'utilisateur plutôt que le théoricien, sans pour autant sacrifier la rigueur ni le formalisme mathématique. Il est conçu pour être lu en format électronique et non imprimé : la multitude de liens hypertextes et de références croisées a été pensée pour favoriser la navigation entre les différentes notions et idées.

Les auteurs sont des membres actifs de la communauté scientifique : professeurs et étudiants au doctorat ou à la maîtrise en mathématiques appliquées et ingénierie, tous utilisent l'algèbre linéaire à des fins d'applications et ont à cœur d'en transmettre une vision pragmatique (et une passion !) directement reliée aux problématiques de recherche actuelles. Tous les auteurs ont également une expérience d'enseignement universitaire, et sont donc particulièrement sensibles à la pédagogie et à la transmission, en particulier à des scientifiques non spécialistes des mathématiques.

Il ne s'agit pas d'une traduction mais d'une création originale, en langue française, mise à disposition de tous en tant que Ressource Éducative Libre (RÉL). L'écriture de la première version de ce document a été financée par une subvention des Fonds d'Actions Pédagogiques Stratégiques (FAPS) de Polytechnique Montréal.

Les auteurs.

Auteurs et contributions

Les auteurs sont triés par ordre alphabétique.

Directeurs de projet

Alain Hertz et **Sébastien Le Digabel**, professeurs titulaires au département de mathématiques et génie industriel de Polytechnique Montréal.

Co-auteurs

- **Nathan Allaire** : Contributions à venir ;
- **Sacha Benarroch-Lelong** : Contributions à venir ;
- **Théo Denorme** : Contributions à venir ;
- **Mathieu Gervais-Dubé** : Contributions à venir ;
- **Andréa Gourion** : Contributions à venir ;
- **Rémi Pédenon-Orlanducci** : Écriture du chapitre 4, de la préface, études de cas sur le binôme de Newton et le théorème de Cayley-Hamilton, propositions d'exercices, contributions *back-end*.
- **Camille Pinçon** : Contributions à venir ;

Remerciements

Les auteurs tiennent à remercier chaleureusement leurs relecteurs et étudiants qui se sont prêtés au jeu en travaillant avec les premières versions de cet ouvrage.

Notations et terminologie

Ce chapitre vise à fournir au lecteur les outils formels nécessaires à la bonne compréhension de cet ouvrage. Toutes les notations et conventions utilisées sont détaillées ici. La plupart sont standard et cohérentes avec le reste de la littérature en algèbre linéaire et en mathématique en général, mais certaines sont moins usuelles. Afin de faciliter la transition vers ou depuis d'autres ouvrages, des termes et notations alternatifs à ceux utilisés ici, ainsi que certaines traductions anglaises, sont proposés dans la dernière section de ce chapitre.

i Généralités

▷ *Ensembles de nombres* (voir figure 2) : ils sont supposés connus, on rappelle les inclusions $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

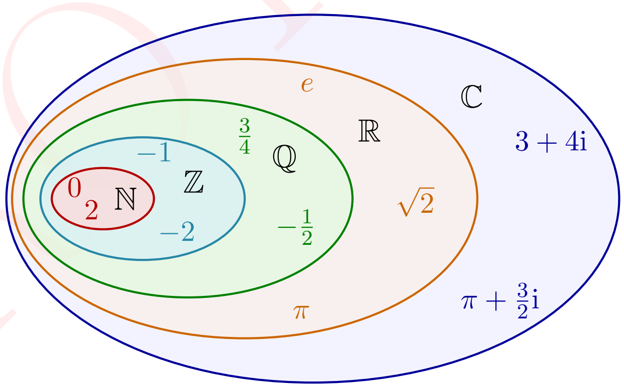

Figure 2 – Les ensembles de nombres.

- Un *astérisque* en exposant d'un ensemble de nombres signifie qu'on en exclut 0. Exemple : $\mathbb{R}^*$ est l'ensemble des nombres réels non nuls.
- Un *signe* $+$ *ou* $-$ en indice signifie qu'on ne retient que les nombres du signe prescrit. Exemples : $\mathbb{R}_+$ est l'ensemble des nombres réels positifs ou nuls, et $\mathbb{R}_+^*$ est l'ensemble des réels strictement positifs.
- On appelle *scalaire* un élément de $\mathbb{R}$ ou de $\mathbb{C}$.
- *Produit de scalaires :* si p et q sont deux scalaires, leur produit est noté pq sans le signe "$\times$". Ce signe est en revanche utilisé en présence de valeurs numériques, par exemple $2 \times 3 = 6$. On bannit totalement l'usage du point "$\cdot$" pour la multiplication, car il est trop discret à l'écrit et se confond trop facilement avec une décimale.
- *Approximations :* $\approx$ signifie "environ égal" et $\lessapprox$ signifie "inférieur, à une tolérance près".
- On adopte l'usage suivant, non usuel en français, pour le point et la virgule dans l'écriture des chiffres décimaux :
  - Les *décimales* d'un nombre sont séparées de sa partie entière par un point. Exemple : $1/3 \approx 0.33$.
  - La virgule est utilisée comme *séparateur de milliers*. Exemple : $3,250,125.22$.
- Les *intervalles* sont décrits avec des crochets ouverts ou fermés, et un point-virgule sépare les deux bornes. Exemples : $[1;2]$, $]-0.5;10]$.
- L'ensemble des *entiers consécutifs* $\{p, p+1, \dots, n\}$ avec $p < n$ est noté $[\![p;n]\!]$.

ii Ensembles, familles, espaces

- Bien qu'ils aient de nombreuses propriétés en commun, on distingue ici soigneusement *ensembles* et *familles*. Voir l'exercice 0.2 après avoir lu les deux définitions suivantes.
- Un *ensemble* est une collection non ordonnée d'éléments sans répétition. Ils sont notés avec des accolades. S'il existe un ordre naturel entre les éléments (un ensemble de réels, par exemple), ils sont rangés suivant l'ordre croissant. Exemples : $\{-2, 1/3, \pi\} = \{1/3, \pi, -2\}$ (présence d'un ordre, donc on retiendra la première écriture), $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \{\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_1\}$ (absence d'ordre, donc les deux écritures seraient valables).
- Une *famille* est une collection ordonnée d'éléments pouvant se répéter. Elles sont notées avec des parenthèses. Exemples : $(\pi, 1/3, -2, 1/3) \neq (-2, 1/3, 1/3, \pi)$ (ordre différent), $(\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_1) \neq (\mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_1, \mathbf{u}_3)$ (répétition d'un élément).
- Les ensembles et les familles sont notés avec des majuscules usuelles (sauf les bases, voir plus bas). Exemples : U, V.
- L'*appartenance* d'un élément $\mathbf{u}$ à un ensemble (ou une famille) U est notée $\mathbf{u} \in U$, ce qui signifie que $\mathbf{u}$ est l'un des éléments de U. Dans le cas où U est un espace vectoriel, $\mathbf{u}$ en est alors appelé un vecteur.

▷ On appelle *cardinalité* d'un ensemble ou d'une famille U le nombre d'éléments qu'il ou elle contient, et on la note $|U|$.

▷ *Langage des ensembles :* si A et B sont deux ensembles (ou familles), on rappelle les significations des symboles suivants :

- $A = B$: les ensembles A et B contiennent exactement les mêmes éléments (pour deux familles, cela signifie que répétitions et ordre des éléments sont également identiques) ;
- $A \subseteq B$: A est un *sous-ensemble* (ou une *sous-famille*) de B ;
- $A \subset B$: A est un *sous-ensemble strict* (ou une *sous-famille stricte*) de B, i.e. $A \subseteq B$ mais $A \neq B$;
- $A \setminus B = \{x \in A : x \notin B\}$: *différence* entre A et B (ou "A privé de B") ;
- $A \cup B = \{x : x \in A \text{ ou } x \in B\}$ est *l'union* des ensembles (ou familles) A et B ;
- $A \cap B = \{x : x \in A \text{ et } x \in B\}$ est *l'intersection* des ensembles (ou familles) A et B.

Ces opérateurs peuvent être utilisés indifféremment si A et B sont tous deux des ensembles ou des familles, et produisent un objet de même nature. Dans le cas où A et B ne sont pas de même nature, on s'autorise tout de même à utiliser ces opérateurs, en posant la convention que l'objet produit est systématiquement un ensemble. Exemple : $\{-2, 1/3, \pi\} \cup (-1/3, 4) = \{-2, -1/3, 1/3, \pi, 4\}$.

▷ Les termes "espace" et "sous-espace" sont utilisés comme raccourcis pour désigner un *espace vectoriel* et un *sous-espace vectoriel*.

▷ Les *bases* sont des familles particulières et sont notées avec des majuscules calligraphiques. Exemple : $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$.

▷ On appelle *ensemble vide* (respectivement *famille vide*) l'ensemble (respectivement la famille) ne contenant aucun élément. Ensemble et famille vides sont notés $\emptyset$ et vérifient $|\emptyset| = 0$. Le contexte dicte si $\emptyset$ désigne l'ensemble ou la famille vide.

▷ Si U est un espace vectoriel, on appelle *sous-espace triviaux de U* : l'espace $\{\mathbf{0}\}$ engendré par le vecteur nul, et l'espace U lui-même (voir les remarques suivant le théorème 5.1.1).

▷ Des ensembles (ou familles) de vecteurs peuvent *générer* des espaces vectoriels. On parle alors d'*ensembles générateurs* ou de *familles génératrices* de ces espaces vectoriels. Voir section 5.2.1.

▷ Les "droites", "plans", et "hyperplans" de $\mathbb{R}^n$ correspondent aux sous-espaces de dimensions 1, 2, et $n-1$, respectivement.

▷ Si $\mathbf{A}$ est une matrice, on appelle les *quatre sous-espaces fondamentaux de* $\mathbf{A}$ les espaces suivants :

- $\mathrm{Im}(\mathbf{A})$ est *l'image* de $\mathbf{A}$;

  - $\mathrm{Ker}(\mathbf{A})$ est *le noyau* de $\mathbf{A}$;
  - $\mathrm{Im}(\mathbf{A}^\top)$ est *l'espace des lignes* de $\mathbf{A}$;
  - $\mathrm{Ker}(\mathbf{A}^\top)$ est *le noyau à gauche* de $\mathbf{A}$.

  Le chapitre 6 leur est dédié, et ils sont visualisés aux figures 6.3 et 6.4.

- ▷ Si U est un espace vectoriel, sa *dimension* est notée $\dim(U)$. Voir section 5.2.4 et définition 5.2.11.

- ▷ *Orthogonalité* (voir la section 5.3) : $\mathbf{x} \perp \mathbf{y}$ signifie que les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont *orthogonaux*, alors que $U \perp V$ signifie que les espaces vectoriels U et V sont orthogonaux. $U^\perp$ est le *complément orthogonal* (ou simplement *l'orthogonal*) de l'espace vectoriel U. On dit aussi que U et $U^\perp$ sont des espaces complémentaires.

- ▷ Une *famille orthonormale* est une famille de vecteurs orthogonaux entre eux et tous unitaires. Si U est un espace vectoriel, une *base orthonormale* de U est une base de U constituée par une famille orthonormale (voir la section 5.3.2).

- ▷ *Opérateurs :* quelles que soient les natures des objets sur lesquels un opérateur agit, on utilise des parenthèses entre l'opérateur et son ou ses argument(s). Exemple : $\mathrm{Im}(\mathbf{A})$. Cependant, cette notation est parfois trop lourde et on s'autorise donc quelques écarts. Deux exemples de tels usages acceptables sont :

  - $\mathrm{Im}\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$ au lieu de $\mathrm{Im}\left(\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\right)$ (opérateur agissant sur une matrice donnée explicitement) ;
  - $\mathrm{Vect}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ au lieu de $\mathrm{Vect}((\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p))$ (opérateur agissant sur une famille).

iii Nombres complexes

Les *nombres complexes* sont le sujet du chapitre 3.

- ▷ L'*unité imaginaire* est notée i. Le i (en italique) désigne souvent les indices d'une énumération, ou ceux des lignes d'une matrice.

- ▷ Les *parties réelle* et *imaginaire* d'un nombre complexe $z \in \mathbb{C}$ sont notées respectivement $\mathcal{R}\mathrm{e}(z) \in \mathbb{R}$ et $\mathcal{I}\mathrm{m}(z) \in \mathbb{R}$. De même pour une matrice $\mathbf{A} \in \mathbb{C}^{m\times n}$, on écrit $\mathcal{R}\mathrm{e}(\mathbf{A}) \in \mathbb{R}^{m\times n}$ et $\mathcal{I}\mathrm{m}(\mathbf{A}) \in \mathbb{R}^{m\times n}$.

- ▷ Le *conjugué* de $z \in \mathbb{C}$ est $\overline{z}$.

- ▷ Le *module* de $z \in \mathbb{C}$ est $|z|$. Si $z \in \mathbb{R}$, on l'appelle sa *valeur absolue.*

- ▷ Un *argument* du nombre complexe non nul $z \neq 0$ est noté $\arg(z) \in \mathbb{R}$, et son *argument principal* est unique et noté $\mathrm{Arg}(z) \in]-\pi;\pi]$.

iv Vecteurs

Les *vecteurs* sont introduits à la section 1.1, puis généralisés avec une définition formelle au chapitre 5.

- ▷ $\mathbb{R}^n$ (respectivement $\mathbb{C}^n$) désigne l'ensemble des vecteurs colonne[1] constitués de n nombres réels (respectivement n nombres complexes). En utilisant cette notation, on suppose toujours que $n \in \mathbb{N}^*$, et on ne le précise donc jamais.
- ▷ On comprendra au chapitre 5 que le terme *vecteur* désigne en fait tout élément d'un espace vectoriel. Cependant, les éléments de $\mathbb{R}^n$ et $\mathbb{C}^n$ ($n \geq 2$) jouent un rôle particulier dans cet ouvrage. Sans plus de précision, le terme "vecteur" désigne donc un élément de $\mathbb{R}^n$, et le terme "vecteur complexe" désigne un élément de $\mathbb{C}^n$. On note de tels éléments avec des lettres minuscules grasses. Exemple : $\mathbf{x} \in \mathbb{R}^n$.
- ▷ On prend soin de distinguer les *vecteurs colonne* des *vecteurs ligne*, qu'on peut aussi appeler *colonnes* et *lignes*. Pour passer de l'un à l'autre, on utilise "$\top$" l'opérateur de *transposition* : si $\mathbf{x}$ est un vecteur colonne, $\mathbf{x}^\top$ est appelé son *transposé* et c'est un vecteur ligne, et *vice-versa*. Ainsi :
  - $\mathbf{c} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ représente un vecteur colonne ;
  - $\mathbf{l} = [1\ 2\ 3]$ représente un vecteur ligne ;
  - on a, pour ces exemples, $\mathbf{c}^\top = \mathbf{l}$;
  - pour une écriture plus compacte, on écrira $(1, 2, 3) = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^\top$ un vecteur colonne (voir la définition 1.1.5).
- ▷ Si $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ ou $\mathbb{C}^n$, les éléments $x_i = \mathbf{x}(i)$ ($i \in [\![1; n]\!]$) sont appelés les *composantes* du vecteur $\mathbf{x}$, et ce sont des scalaires (réels ou complexes). Cette notation permet également de considérer $\mathbf{x}$ comme une famille d'éléments de $\mathbb{R}$ ou de $\mathbb{C}$.
- ▷ Le nombre de composantes d'un vecteur est appelé sa *taille*. On n'emploie pas le terme "dimension", qui est réservé aux espaces vectoriels.
- ▷ Le *vecteur nul* de $\mathbb{R}^n$ est noté $\mathbf{0}$, sa taille est dictée par le contexte. Pour lever une éventuelle ambiguïté, on écrit parfois $\mathbf{0}_n$. Ce vecteur est aussi appelé l'*origine*.
- ▷ Si $\mathbf{x}$ est étudié comme élément d'un espace vectoriel et que $\mathcal{B}$ est une base de cet espace, alors $[\mathbf{x}]_\mathcal{B}$ désigne les *coordonnées* de $\mathbf{x}$ dans $\mathcal{B}$ (voir théorème 5.2.1).
- ▷ Pour simplifier définitions et calculs, on *assimile* certains objets de la façon suivante :
  - un vecteur (respectivement un vecteur complexe) à une seule composante est assimilé à un nombre réel (respectivement un nombre complexe), i.e. $\mathbb{R} = \mathbb{R}^1$ (respectivement $\mathbb{C} = \mathbb{C}^1$) ;

1. Ici "colonne" est bien au singulier, car on parle de vecteurs *sous forme de* colonne.

- une matrice réelle (respectivement complexe) à n lignes et une colonne est assimilée à un vecteur de $\mathbb{R}^n$ (respectivement $\mathbb{C}^n$), i.e. $\mathbb{R}^{n\times 1} = \mathbb{R}^n$ (respectivement $\mathbb{C}^{n\times 1} = \mathbb{C}^n$).

Il découle de ce qui précède qu'une matrice à 1 ligne et 1 colonne est assimilée à un scalaire, i.e. $\mathbb{R} = \mathbb{R}^1 = \mathbb{R}^{1\times 1}$ et $\mathbb{C} = \mathbb{C}^1 = \mathbb{C}^{1\times 1}$. Ainsi, pour un réel x, on considère que

$$x = (x) = [x] = [x]^\top \in \mathbb{R}$$

et pour un complexe z, on considère que

$$z = (z) = [z] = [z]^\top \in \mathbb{C}\ .$$

Exemples : on considère dans cet ouvrage que :

- $\underbrace{(1,2,3)}_{\in\mathbb{R}^3} = \underbrace{\begin{bmatrix}1\\2\\3\end{bmatrix}}_{\in\mathbb{R}^3} = \underbrace{\begin{bmatrix}1\\2\\3\end{bmatrix}}_{\in\mathbb{R}^{3\times 1}}$;
- $\underbrace{1-\mathrm{i}}_{\in\mathbb{C}} = \underbrace{(1-\mathrm{i})}_{\in\mathbb{C}^1} = \underbrace{[1-\mathrm{i}]}_{\in\mathbb{C}^{1\times 1}}$;
- pour $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, le produit scalaire $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^\top\mathbf{y}$ est bien défini comme le produit d'une matrice $1 \times n$ et d'une matrice $n \times 1$. Il produit une matrice 1×1, qui est assimilée à un scalaire.

Noter que si cette assimilation est pratique pour le calcul, elle n'est pas exacte du point de vue théorique. Elle est également considérée illicite par la plupart des langages de calcul scientifique, qui distinguent par exemple $\mathbb{R}^n$ de $\mathbb{R}^{n\times 1}$. Ceux-ci prévoient cependant des opérations implicites de conversion, et c'est dans cet esprit qu'on pose cette convention.

▷ Une *combinaison linéaire* (ou simplement *combinaison*) de vecteurs est une somme pondérée de ces vecteurs (voir la définition 1.1.9). On appelle *combinaison triviale* la combinaison dans laquelle tous les poids sont nuls, et son résultat est $\mathbf{0}$.

▷ La notion d'*indépendance linéaire* est abordée à la section 5.2.1. On s'attache à employer les adjectifs appropriés :

- une *famille* de vecteurs peut être *libre* ou *liée* ;
- un ensemble de vecteurs peut être *(linéairement) indépendant* ou *(linéairement) dépendant* ;
- des vecteurs eux-mêmes peuvent être *(linéairement) dépendants* ou *(linéairement) indépendants* ;
- deux vecteurs dépendants peuvent être appelés *colinéaires.*

▷ L'opérateur $\mathrm{Vect}(\cdot)$, formellement défini à la définition 1.1.10, s'applique sur une *famille* de vecteurs (et non un ensemble), et produit l'*ensemble* (et non la famille) des combinaisons linéaires possibles de ces vecteurs. Voir l'exercice 0.4 pour la maîtrise de la notation liée à cet opérateur.

▷ Le *conjugué* d'un vecteur $\mathbf{x} = (x_1, x_2, \dots, x_n)$ est $\overline{\mathbf{x}} = (\overline{x_1}, \overline{x_2}, \dots, \overline{x_n})$. On note l'équivalence $\mathbf{x} \in \mathbb{R}^n \iff \overline{\mathbf{x}} = \mathbf{x}$. Le conjugué d'une ligne est défini de façon analogue, par application du conjugué complexe sur chacune de ses composantes.

▷ Le *transconjugué* d'un vecteur $\mathbf{x} = (x_1, x_2, \dots, x_n)$ est $\mathbf{x}^* = \overline{\mathbf{x}}^\top = \overline{\mathbf{x}^\top}$. On note l'équivalence $\mathbf{x} \in \mathbb{R}^n \iff \mathbf{x}^* = \mathbf{x}^\top$.

▷ La notion générale de *produit scalaire* est introduite à la définition 5.3.1.

- Si $\mathbf{x}$ et $\mathbf{y}$ sont deux éléments d'un espace vectoriel, leur produit scalaire est noté $\langle \mathbf{x}, \mathbf{y} \rangle$.
- Le *produit scalaire canonique* entre deux vecteurs $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ est $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$ (voir définition 1.1.12). L'usage du point "$\cdot$" est banni pour écrire le produit scalaire dans $\mathbb{R}^n$, car il est trop discret à l'écrit.
- Le produit scalaire canonique entre deux vecteurs $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ est $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^* \mathbf{y}$ (voir définition 3.6.5).

▷ Le *produit vectoriel* agit sur deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ et produit un vecteur $\mathbf{x} \wedge \mathbf{y}$ qui leur est à tous les deux orthogonal. Voir la section 1.1.5.

▷ La *norme* (euclidienne) du vecteur $\mathbf{x}$ est notée $\|\mathbf{x}\| \in \mathbb{R}$. Si $\|\mathbf{x}\| = 1$, on dit que $\mathbf{x}$ est un vecteur *unitaire*, ou *normalisé*.

▷ La *distance* entre deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ est $\mathrm{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

▷ La *base canonique* de $\mathbb{R}^n$ est notée $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ où $\mathbf{e}_i$ désigne le vecteur de $\mathbb{R}^n$ dont toutes les composantes sont nulles, sauf la i-ième qui est un 1 (autrement dit, la i-ième colonne de la matrice identité de taille n).

▷ L'opérateur de *projection orthogonale sur un sous-espace* est défini à la section 9.1. Plusieurs notations sont utilisées pour désigner cette opération. Soit V un espace vectoriel, alors

- si W est un sous-espace vectoriel de V et que $\mathbf{x} \in V$, $\mathrm{Proj}_W(\mathbf{x})$ est la projection de $\mathbf{x}$ sur le sous-espace W ;
- si $\mathbf{x}, \mathbf{y} \in V$, $\mathrm{Proj}_{\mathbf{y}}(\mathbf{x}) = \mathrm{Proj}_{\mathrm{Vect}(\mathbf{y})}(\mathbf{x})$ est le projeté de $\mathbf{x}$ sur la droite vectorielle générée par $\mathbf{y}$;
- si $\mathbf{A} \in \mathbb{R}^{m \times n}$ est telle que $\mathrm{Im}(\mathbf{A}) \subseteq V$ et $\mathbf{x} \in V$, $\mathrm{Proj}_{\mathbf{A}}(\mathbf{x}) = \mathrm{Proj}_{\mathrm{Im}(\mathbf{A})}(\mathbf{x})$ est le projeté de $\mathbf{x}$ sur l'image de $\mathbf{A}$.

▷ La *moyenne* du vecteur $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ est $\widehat{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n x_i$. On s'en sert pour exprimer facilement la somme des coordonnées d'un vecteur avec $\sum_{i=1}^n x_i = n\widehat{\mathbf{x}}$.

v Matrices

Les matrices sont introduites à la section 1.2.

- ▷ $\mathbb{R}^{m\times n}$ et $\mathbb{C}^{m\times n}$ désignent respectivement l'ensemble des matrices à m lignes et n colonnes constituées de nombres réels et complexes. En utilisant cette notation, on suppose toujours que $m, n \in \mathbb{N}^*$, et on ne le précise donc jamais.
- ▷ Les matrices sont notées par des lettres majuscules grasses. Exemple : $\mathbf{A} \in \mathbb{R}^{m\times n}$.
- ▷ Les nombres de lignes et de colonnes d'une matrice sont appelés sa *taille*. On n'emploie pas le terme "dimension", qui est réservé aux espaces vectoriels.
- ▷ L'expression générique "une matrice $m \times n$" désigne une matrice quelconque de $\mathbb{R}^{m\times n}$ ou de $\mathbb{C}^{m\times n}$. Lorsqu'une précision est nécessaire, on parle d'une "matrice réelle" ou d'une "matrice complexe".
- ▷ La *transposée* d'une matrice $\mathbf{A} \in \mathbb{R}^{m\times n}$ (respectivement $\mathbf{A} \in \mathbb{C}^{m\times n}$) est notée $\mathbf{A}^\top \in \mathbb{R}^{n\times m}$ (respectivement $\mathbb{C}^{n\times m}$).
- ▷ Si $\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$, les éléments $a_{i,j} = \mathbf{A}(i,j)$ ($i \in [\![1;m]\!]$, $j \in [\![1;n]\!]$) sont appelés les *composantes* de $\mathbf{A}$, et ce sont des scalaires (réels ou complexes). i désigne l'indice de ligne, et j désigne l'indice de colonne.
- ▷ Les colonnes et les lignes d'une matrice peuvent être sujettes à des manipulations. Il faut pour cela les désigner correctement. Soit $\mathbf{A} = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 3}$.
  - La première colonne de $\mathbf{A}$ est $(1,2) \in \mathbb{R}^2$, c'est un vecteur (sous-entendu : un vecteur colonne).
  - La première ligne de $\mathbf{A}$ est $\begin{bmatrix} 1 & 3 & 5 \end{bmatrix}$, ce n'est donc pas un vecteur. Si on souhaite l'étudier comme un vecteur, on étudiera son transposé $\begin{bmatrix} 1 & 3 & 5 \end{bmatrix}^\top \in \mathbb{R}^3$.
- ▷ La *matrice identité* de taille n est notée $\mathbf{I}$, ou éventuellement $\mathbf{I}_n$ si le contexte nécessite un éclaircissement. Elle est toujours carrée.
- ▷ La *matrice nulle* de taille $m \times n$ est notée $\mathbf{O}$ (noter la différence avec le vecteur nul $\mathbf{0}$), ou éventuellement $\mathbf{O}_{m,n}$ si le contexte nécessite un éclaircissement. Elle n'est pas forcément carrée. Voir définition 1.2.3.
- ▷ Une matrice diagonale (donc nécessairement carrée) de taille n peut être décrite uniquement par sa diagonale grâce à l'opérateur $\mathrm{Diag}(\cdot)$. Exemple : $\mathrm{Diag}(2,3) = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$.
- ▷ Le *produit matriciel* est introduit à la définition 1.2.14. Si $\mathbf{A} \in \mathbb{R}^{m\times p}$ et $\mathbf{B} \in \mathbb{R}^{p\times n}$, le produit de $\mathbf{A}$ et $\mathbf{B}$ est noté $\mathbf{AB} \in \mathbb{R}^{m\times n}$. On n'utilise pas le signe "$\times$". Ceci est vrai également pour les matrices complexes.

▷ Le *produit de Kronecker* de deux matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ et $\mathbf{B} \in \mathbb{R}^{p\times q}$ est noté $\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{pm\times qn}$. Voir étude de cas de la section 1.5.

▷ Le *rang* de la matrice $\mathbf{A}$ est noté $\mathrm{rg}(\mathbf{A})$. Voir la définition 6.1.5. Souvent, pour raccourcir la notation, on pose $r = \mathrm{rg}(\mathbf{A})$. De plus, si $\mathbf{A} \in \mathbb{R}^{m\times n}$, on dit que :

- $\mathbf{A}$ est de *plein rang colonne* si $\mathrm{rg}(\mathbf{A}) = n$;
- $\mathbf{A}$ est de *plein rang ligne* si $\mathrm{rg}(\mathbf{A}) = m$;
- $\mathbf{A}$ est de *plein rang* si $\mathbf{A}$ est carrée et que $\mathrm{rg}(\mathbf{A}) = m = n$.

▷ Si $\mathbf{A}$ est une matrice carrée inversible, on note $\mathbf{A}^{-1}$ la *matrice inverse* de $\mathbf{A}$. On dit aussi que $\mathbf{A}$ est *non singulière*. Dans le cas contraire, on dit que $\mathbf{A}$ est *non inversible*, ou *singulière*.

▷ Si $\mathbf{A} \in \mathbb{R}^{m\times n}$ avec $m > n$ est de plein rang colonne, la matrice *pseudo-inverse* de $\mathbf{A}$ est notée $\mathbf{A}^\dagger = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top$. Voir la section 6.4.

▷ Si $\mathbf{A}$ est une matrice carrée symétrique, on note son signe (voir chapitre 12) :

- $\mathbf{A} \succ 0$ si $\mathbf{A}$ est *définie positive* ;
- $\mathbf{A} \succeq 0$ ou “$\mathbf{A}$ est SDP” si $\mathbf{A}$ est *semi-définie positive* ;
- $\mathbf{A} \prec 0$ si $\mathbf{A}$ est *définie négative* ;
- $\mathbf{A} \preceq 0$ (ou “$\mathbf{A}$ est SDN”) si $\mathbf{A}$ *est semi-définie négative*.

▷ L'opérateur $\mathrm{Col}(\cdot)$ retourne la famille des colonnes d'une matrice (voir définition 1.2.4). Il ne produit pas un ensemble, puisque les colonnes d'une matrice peuvent être identiques.

▷ Si $\mathbf{A} \in \mathbb{R}^{m\times n}$ ou $\mathbf{A} \in \mathbb{C}^{m\times n}$, sa *norme matricielle* est $\|\mathbf{A}\|$ (voir section 1.2.8).

▷ Si $\mathbf{A} \in \mathbb{R}^{n\times n}$ ou $\mathbf{A} \in \mathbb{C}^{n\times n}$, le *déterminant* de $\mathbf{A}$ est noté $\det(\mathbf{A})$ (voir la section 1.3). Si on étudie le déterminant d'une matrice donnée explicitement, il est écrit entre barres simples. Exemple : avec $\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, on a $\det(\mathbf{A}) = \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix}$.

▷ La *trace* d'une matrice carrée $\mathbf{A} \in \mathbb{R}^{n\times n}$ (respectivement $\mathbf{A} \in \mathbb{C}^{n\times n}$) est notée $\mathrm{tr}(\mathbf{A}) \in \mathbb{R}$ (respectivement $\mathrm{tr}(\mathbf{A}) \in \mathbb{C}$). Voir la définition 1.2.12.

▷ Les *parties réelle* et *imaginaire* d'une matrice $\mathbf{A} \in \mathbb{C}^{m\times n}$ sont respectivement notées $\mathcal{R}\mathrm{e}(\mathbf{A}) \in \mathbb{R}^{m\times n}$ et $\mathcal{I}\mathrm{m}(\mathbf{A}) \in \mathbb{R}^{m\times n}$. Elles sont définies telles que $\mathbf{A} = \mathcal{R}\mathrm{e}(\mathbf{A}) + \mathrm{i}\,\mathcal{I}\mathrm{m}(\mathbf{A})$.

▷ La *transconjuguée* d'une matrice $\mathbf{A} \in \mathbb{C}^{m\times n}$ est $\mathbf{A}^* = \overline{\mathbf{A}}^\top = \overline{\mathbf{A}^\top}$. On note l'équivalence $\mathbf{A} \in \mathbb{R}^{m\times n} \iff \mathbf{A}^* = \mathbf{A}^\top$.

▷ Si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases d'un espace vectoriel de dimension n, la *matrice de changement de base* de $\mathcal{B}$ à $\mathcal{C}$ est notée $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} \in \mathbb{R}^{n\times n}$.

▷ Deux matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ ou $\mathbb{C}^{m\times n}$ sont dites *équivalentes* s'il existe une matrice $\mathbf{P}$ inversible telle que $\mathbf{A} = \mathbf{P}\mathbf{B}$. Ceci est noté $\mathbf{A} \sim \mathbf{B}$. Cette écriture est particulièrement utilisée pour pour indiquer que $\mathbf{B}$ est obtenue à partir d'une ou plusieurs opération(s) d'élimination sur $\mathbf{A}$ (voir la section 2.1.1).

▷ Deux matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ (ou $\mathbb{C}^{m\times n}$) sont dites *semblables* s'il existe une matrice $\mathbf{P}$ inversible telle que $\mathbf{A} = \mathbf{PBP}^{-1}$.

▷ Il est souvent utile d'étudier une matrice par *blocs*, de différentes façons. Si $\mathbf{A}$ est de forme $m \times n$, on peut la voir :

- comme l'ensemble de ses colonnes, on note alors $\mathbf{A} = \begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n\end{bmatrix}$ où $\mathbf{c}_j \in \mathbb{R}^m$ ($j \in [\![1;n]\!]$) ;
- comme l'ensemble de ses lignes. On considère alors que la i-ième ligne de $\mathbf{A}$ est le transposé du vecteur $\mathbf{l}_i \in \mathbb{R}^n$ ($i \in [\![1;m]\!]$), et on note $\mathbf{A} = \begin{bmatrix}\mathbf{l}_1^\top \\ \mathbf{l}_2^\top \\ \vdots \\ \mathbf{l}_m^\top\end{bmatrix}$;
- comme un ensemble de blocs appelés *sous-matrices*, que l'on peut éventuellement délimiter par des séparations virtuelles. Par exemple, pour décomposer la matrice $\mathbf{A}$ en 2×2 sous-matrices, on écrit :

$$\mathbf{A} = \begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} = \left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right].$$

  où $\mathbf{A}_{11} \in \mathbb{R}^{m_1\times n_1}$, $\mathbf{A}_{12} \in \mathbb{R}^{m_1\times n_2}$, $\mathbf{A}_{21} \in \mathbb{R}^{m_2\times n_1}$ et $\mathbf{A}_{22} \in \mathbb{R}^{m_2\times n_2}$. Les tailles doivent être cohérentes avec celle de $\mathbf{A}$, i.e. $m_1 + m_2 = m$ et $n_1 + n_2 = n$.

▷ Une *matrice orthogonale* est une matrice carrée dont les colonnes sont orthonormales (voir la définition 5.3.5). Une matrice non carrée dont les colonnes sont orthonormales n'admet pas d'appellation particulière.

▷ Un vocabulaire varié est à acquérir pour désigner différents types de matrices. On parle ainsi de matrices *diagonales*, *triangulaires*, *symétriques*, *hermitiennes*, *échelonnées*, *échelonnées réduites*, *unipotentes*, *d'élimination*, *de permutation*, *de rotation*, *de réflexion*, etc. Une liste exhaustive est donnée à la section 1.2.3.

vi Systèmes

Les *systèmes d'équations linéaires* (ou simplement *systèmes*) sont le sujet du chapitre 8.

▷ Plusieurs éléments sont nécessaires pour définir un système. Ces éléments sont appelés les *données du système* et sont :

- une matrice $\mathbf{A} \in \mathbb{R}^{m\times n}$ appelée *la matrice des coefficients du système* ;
- un vecteur $\mathbf{b} \in \mathbb{R}^m$ appelé *le membre de droite* du système ;

On peut aussi considérer des systèmes avec des données complexes.

▷ *Résoudre* le système donné par $\mathbf{A}$ et $\mathbf{b}$ consiste à identifier un vecteur $\mathbf{x} \in \mathbb{R}^n$, ou un ensemble de tels vecteurs, satisfaisant les équations linéaires décrites par l'égalité $\mathbf{Ax} = \mathbf{b}$.

▷ Dans le système donné par $\mathbf{Ax} = \mathbf{b}$, le vecteur $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ est appelé vecteur des *variables* (ou des *inconnues*) et la composante x_i est appelée la i-ième variable (ou inconnue) du système ($i \in [\![1; n]\!]$).

▷ Si $\mathbf{b} = \mathbf{0}$, le système $\mathbf{Ax} = \mathbf{0}$ est un *système homogène*, et sert à définir le noyau de $\mathbf{A}$ (voir la définition 6.1.1).

▷ L'ensemble de toutes les solutions possibles d'un système se nomme sa *solution complète*, définie à la section 8.4. La solution complète s'écrit à partir d'une *solution particulière* et de *solutions spéciales*. Voir les définitions 8.4.1 et 8.3.1.

▷ La matrice $\mathbf{N}$ dont les colonnes sont les solutions spéciales d'une matrice $\mathbf{A}$ se nomme la *matrice noyau de* $\mathbf{A}$ définie à la définition 8.3.2. Il ne faut pas confondre "noyau" (un espace) et "matrice noyau" (une matrice).

▷ On comprendra dans la section 1.2.5 que dans le système $\mathbf{Ax} = \mathbf{b}$, chaque colonne de $\mathbf{A} \in \mathbb{R}^{m \times n}$ est associée à une variable tandis que chaque ligne de $\mathbf{A}$ est associée à une équation. On dit donc que le système $\mathbf{Ax} = \mathbf{b}$ est constitué de *n variables et m équations*.

▷ Un système admettant au moins une solution est appelé *réalisable* (ou *compatible*). Un système n'en admettant aucune est appelé *irréalisable* (ou *incompatible*). On préfère le terme de *réalisabilité* plutôt que *compatibilité* car le premier est incontournable dans le domaine de l'optimisation, tandis que le second est ambigu et est réservé dans ce manuel pour décrire des opérations matricielles valides selon les tailles de matrices. Voir la définition 8.1.5.

▷ Le chapitre 2 décrit des techniques pour résoudre des systèmes, appelées *élimination de Gauss* et *élimination de Gauss-Jordan*. Ces techniques utilisent un vocabulaire axé sur la notion de *pivot*.

- Cette notion est introduite à la définition 2.1.6, et précise qu'un pivot est toujours un scalaire *non nul*.
- On dit que (i, j) est une *position de pivot* de la matrice $\mathbf{A}$ si sa composante a_{ij} est un pivot.
- Toute colonne d'une matrice $\mathbf{A}$ contenant un pivot est appelée *colonne pivot*[2] de $\mathbf{A}$, et la variable associée dans un système est appelée *variable pivot*.
- Toute colonne d'une matrice $\mathbf{A}$ ne contenant pas de pivot est appelée *colonne libre*[3], ou *colonne non-pivot* de $\mathbf{A}$, et la variable associée dans un système est appelée *variable libre*.
- Attention : Les concepts de "ligne pivot" et de "ligne libre" ne sont pas définis.
- Attention : Les pivots ne se voient pas dans $\mathbf{A}$, mais dans les formes échelonnée et échelonnée réduite obtenues à partir de $\mathbf{A}$.

2. dont le pluriel est "colonnes pivot", car chaque colonne est *de type pivot*.
3. dont le pluriel est "colonnes libres", car ici ce sont les colonnes elles-mêmes qui sont libres.

▷ La matrice formée des blocs $\begin{bmatrix}\mathbf{A} & \mathbf{b}\end{bmatrix} \in \mathbb{R}^{m\times(n+1)}$ est appelée la *matrice augmentée* du système $\mathbf{Ax} = \mathbf{b}$. On dit également qu'elle *le représente sous forme augmentée*. Attention à la terminologie : c'est bien la représentation qui est sous forme augmentée, on ne parle pas de "système sous forme augmentée". Cette notion se généralise à une agrégation de p systèmes avec les mêmes coefficients dans $\mathbf{A}$ mais des membres de droite différents placés dans une matrice $\mathbf{B}$, qui donne la matrice augmentée $[\mathbf{A}\ \mathbf{B}]$. Voir la définition 8.1.6.

▷ On appelle *opérations élémentaires* les opérations sur les lignes d'une matrice effectuées dans la méthode d'élimination de Gauss :

- la *combinaison linéaire* de deux lignes consiste à ajouter à une ligne un multiple d'une autre. Par exemple, on peut remplacer la deuxième ligne d'une matrice par cette ligne à laquelle on retranche deux fois la première. Ceci est noté $L_2 \leftarrow L_1 - 2L_2$;
- la multiplication d'une ligne par un scalaire. Par exemple, on peut multiplier la première ligne d'une matrice par 2. Ceci est noté $L_1 \leftarrow 2L_1$;
- la permutation de deux lignes. Par exemple, on peut permuter les lignes 1 et 2 d'une matrice. Ceci est noté $L_1 \leftrightarrow L_2$.

▷ Une suite d'opérations élémentaires appliquées sur une matrice $\mathbf{A}$ produit une matrice équivalente à $\mathbf{A}$. On désigne ceci par le signe $\sim$ entre ces deux matrices, et on précise sous ce signe les opérations élémentaires qui ont été effectuées. Par exemple, l'opération suivante remplace la troisième ligne par elle-même à laquelle on retranche trois fois la deuxième :

$$\begin{bmatrix} \boxed{1} & 3 & 0 \\ 0 & \boxed{1} & 1 \\ 0 & 3 & 5 \end{bmatrix} \underset{L_3 \leftarrow L_3 - 3L_2}{\sim} \begin{bmatrix} \boxed{1} & 3 & 0 \\ 0 & \boxed{1} & 1 \\ 0 & 0 & \boxed{2} \end{bmatrix}.$$

Lorsque des composantes sont encadrées comme ci-dessus, elles désignent les pivots du système. Cet encadrement n'est pas systématique.

▷ Une *solution au sens des moindres carrés* d'un système est notée $\widehat{\mathbf{x}}$ (voir la section 9.1.3).

vii Valeurs propres et valeurs singulières

Les valeurs propres sont décrites au chapitre 7 et les valeurs singulières au chapitre 13.

▷ Le *spectre* d'une matrice $\mathbf{A} \in \mathbb{R}^{n\times n}$ est noté $\mathrm{Sp}(\mathbf{A})$. Il correspond à l'ensemble des valeurs propres de $\mathbf{A}$. En tant qu'ensemble, il n'admet aucune répétition et donc, ne contient que les valeurs propres distinctes de $\mathbf{A}$. Par exemple, si $\mathbf{A}$ admet pour seule valeur propre 2, de multiplicité deux, on écrit tout de même $\mathrm{Sp}(\mathbf{A}) = \{2\}$, et non $\mathrm{Sp}(\mathbf{A}) = \{2, 2\}$.

- ▷ Si $\lambda \in \mathrm{Sp}(\mathbf{A})$, la *multiplicité algébrique* de λ est notée $\mu_{\mathbf{A}}^{\mathrm{a}}(\lambda)$ et sa *multiplicité algébrique* est notée $\mu_{\mathbf{A}}^{\mathrm{g}}(\lambda)$.
- ▷ Le *sous-espace propre* associé à la valeur propre $\lambda \in \mathrm{Sp}(\mathbf{A})$ n'admet aucune notation particulière. Il est directement désigné par $\mathrm{Ker}(\mathbf{A} - \lambda\mathbf{I})$.
- ▷ Le *polynôme caractéristique* de $\mathbf{A} \in \mathbb{R}^{n\times n}$ est noté $p_{\mathbf{A}}$. Exprimé en fonction de l'inconnue λ, il s'écrit $p_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$.
- ▷ Les *valeurs singulières* d'une matrice sont notées avec la lettre σ. En particulier, la plus petite valeur singulière de $\mathbf{A}$ est notée $\sigma_{\min}(\mathbf{A})$, et la plus grande est notée $\sigma_{\max}(\mathbf{A})$.
- ▷ Le *conditionnement* d'une matrice $\mathbf{A}$ est $\kappa(\mathbf{A}) = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}$.

viii Fonctions

Le chapitre 4 est dédié aux fonctions.

- ▷ Une *application* est décrite par

$$\begin{array}{rccc} f\colon & X & \longrightarrow & Y \\ & \mathbf{x} & \longmapsto & f(\mathbf{x}) \end{array}$$

  où X est le *domaine* de f et Y est l'*espace d'arrivée* de f. Voir la définition 4.1.1.
- ▷ L'écriture raccourcie $f : X \to Y$ désigne une application quelconque de domaine X et d'espace d'arrivée Y.
- ▷ On désigne par *fonction réelle de plusieurs variables* une application ayant pour domaine $\mathbb{R}^n$ et pour espace d'arrivée $\mathbb{R}^n$.
- ▷ Les ensembles de *polynômes* sont définis comme suit :
  - $\mathcal{P}(\mathbb{R})$ (respectivement $\mathcal{P}(\mathbb{C})$) est l'ensemble des polynômes à coefficients réels (respectivement complexes).
  - $\mathcal{P}_n(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$ (respectivement $\mathcal{P}_n(\mathbb{C}) \subset \mathcal{P}(\mathbb{C})$) est l'ensemble des polynômes à coefficients réels (respectivement complexes) de degré inférieur ou égal à n.
- ▷ Le *degré* du polynôme $p \in \mathcal{P}(\mathbb{R})$ ou $p \in \mathcal{P}(\mathbb{C})$ est noté $\deg(p)$.
- ▷ Les polynômes considérés dans cet ouvrage sont des fonctions d'une seule variable, i.e. leur inconnue est un scalaire, et la valeur qu'ils produisent est un scalaire.
- ▷ Pour $f : \mathbb{R}^n \to \mathbb{R}$ et $\mathbf{x} \in \mathbb{R}^n$,
  - si f est différentiable en $\mathbf{x}$, $\nabla f(\mathbf{x}) \in \mathbb{R}^n$ est le *gradient* de f en $\mathbf{x}$;

- si f est deux fois différentiable en $\mathbf{x}$, $\nabla^2 f(\mathbf{x}) \in \mathbb{R}^{n\times n}$ est la *matrice hessienne* de f en $\mathbf{x}$.

▷ La notion de fonction *convexe* est définie à la section 12.5.3.

▷ Les notions de *minimum*/*maximum* et *minima*/*maxima* sont définies à l'étude de cas de la section 12.5. Y figurent aussi les notions d'*optimalité locale* et *optimalité globales*.

Correspondance de termes

Toutes les sections précédentes sont retracées ici pour présenter des termes et notations alternatifs. Lorsqu'elles peuvent s'avérer utiles, des traductions en anglais sont également proposées, et signalées par le symbole ⚑.

i Généralités

▷ L'ensemble $\mathbb{R}_+$ est parfois écrit $\mathbb{R}_{\geq 0}$, et l'ensemble $\mathbb{R}_+^*$ est parfois écrit $\mathbb{R}_{>0}$.

▷ Le point médian est parfois utilisé pour remplacer le signe "$\times$" dans la multiplication entre scalaires : $p \times q = p \cdot q$.

▷ Dans l'écriture des intervalles, des parenthèses peuvent être utilisées pour indiquer une borne exclue. Exemple : $]-0.5; 10] = (-0.5; 10]$.

ii Ensembles, familles, espaces

▷ La cardinalité d'un ensemble U est parfois notée $\text{card}(U)$ ou $\#U$.

⚑ La littérature anglophone ne fait pas la distinction entre ensemble et famille. Les deux sont généralement désignés par le terme *set*.

▷ Générer un espace vectoriel se dit également *engendrer* un espace vectoriel.

  ⚑ Engendrer un espace vectoriel = *to span a vector space*, ensemble générateur/famille génératrice = *spanning set*.

⚑ Image d'une matrice = *Range of a matrix* ou *Column space of a matrix*, noté $R(\mathbf{A})$ ou $C(\mathbf{A})$.

▷ Le noyau de $\mathbf{A}$ est parfois noté $N(\mathbf{A})$.

  ⚑ Noyau = *Nullspace*, noté parfois $\text{Nul}(\mathbf{A})$.

▷ L'espace des lignes est parfois noté $\text{Lgn}(\mathbf{A})$.

  ⚑ Espace des lignes = *Row space*.

▷ Famille/base orthonormale = famille/base orthonormée.

iii Nombres complexes

▷ L'unité imaginaire est parfois notée en italique i. En physique, elle est souvent notée j.

▷ Les parties réelle et imaginaire de $z \in \mathbb{C}$ sont parfois notées $\Re(z)$ et $\Im(z)$, respectivement.

iv Vecteurs

- ▷ Les vecteurs de $\mathbb{R}^n$ et $\mathbb{C}^n$ sont parfois notés avec de simples minuscules (exemple : x), ou avec une flèche (exemple : $\overrightarrow{x}$). La notation avec une flèche n'est plus acceptable en mathématique de niveau universitaire.
- ▷ Le transposé d'un vecteur est souvent noté ${}^t x$ dans la littérature française (qui n'utilise pas de lettres grasses pour les vecteurs).
- ▷ Composantes d'un vecteur = coordonnées, éléments, coefficients, entrées.

  ⚑ *Vector entries.*

- ⚑ Famille libre/ensemble linéairement indépendant = *linearly independent set.*
- ⚑ Famille liée/ensemble linéairement dépendant = *linearly dependent set.*
- ⚑ $\mathrm{Vect}(\cdot) = \mathrm{Span}(\cdot)$.
- ▷ Le transconjugué d'un vecteur $\mathbf{x}$ est parfois appelé son transposé conjugué, son adjoint ou son conjugué hermitien, et noté $\mathbf{x}^H$.
- ▷ Au sens général, un produit scalaire est parfois appelé produit intérieur. Il peut être noté $\langle \mathbf{x} \mid \mathbf{y} \rangle$ (notamment en physique quantique).

  ⚑ *Scalar product*, *Inner product.*

- ▷ Le produit scalaire canonique de $\mathbb{R}^n$ est parfois noté $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y}$. Cette notation n'est plus acceptable au niveau universitaire.

  ⚑ *Dot product.*

- ▷ Le produit vectoriel entre $\mathbf{x}$ et $\mathbf{y}$ est parfois noté $\mathbf{x} \times \mathbf{y}$.
- ▷ En géométrie dans le plan ou dans l'espace, les vecteurs de la base canonique sont parfois notés différemment. La base canonique de $\mathbb{R}^2$ est notée $(\mathbf{i}, \mathbf{j})$ avec $\mathbf{i} = \mathbf{e}_1 = (1,0)$ et $\mathbf{j} = \mathbf{e}_2 = (0,1)$. De même, la base canonique de $\mathbb{R}^3$ peut être notée $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ avec $\mathbf{i} = \mathbf{e}_1 = (1,0,0)$, $\mathbf{j} = \mathbf{e}_2 = (0,1,0)$, et $\mathbf{k} = (0,0,1)$.

v Matrices

- ▷ Les matrices sont parfois notées avec de simples majuscules non grasses. Exemple : $A \in \mathbb{R}^{m\times n}$.
- ▷ Les ensembles de matrices sont parfois notés différemment. C'est notamment le cas dans la littérature française.
  - $\mathbb{R}^{m\times n}$ et $\mathbb{C}^{m\times n}$ sont respectivement notés $\mathcal{M}_{m,n}(\mathbb{R})$ et $\mathcal{M}_{m,n}(\mathbb{C})$, ou $\mathscr{M}_{m,n}(\mathbb{R})$ et $\mathscr{M}_{m,n}(\mathbb{C})$.
  - $\mathbb{R}^{n\times n}$ et $\mathbb{C}^{n\times n}$ sont respectivement notés $\mathcal{M}_n(\mathbb{R})$ et $\mathcal{M}_n(\mathbb{C})$, ou $\mathscr{M}_n(\mathbb{R})$ et $\mathscr{M}_n(\mathbb{C})$.
  - Pour désigner l'ensemble des matrices réelles inversibles de taille $n \times n$, on utilise souvent la notation $\mathrm{GL}_n(\mathbb{R})$. On appelle cet ensemble le *groupe linéaire de degré/d'ordre/d'indice n.*

- ▷ La transposée d'une matrice est souvent notée ${}^tA$ dans la littérature française (qui n'utilise pas de lettres grasses pour les matrices).
- ▷ Les composantes d'une matrice sont parfois appelés ses *coefficients*, ses *termes* ou ses *entrées*.

  ⚑ *Matrix entries.*
- ▷ La pseudo-inverse de $\mathbf{A}$ est parfois notée $\mathbf{A}^+$.
- ▷ Le signe d'une matrice est parfois noté $\mathbf{A} > 0$ au lieu de $\mathbf{A} \succ 0$. De même pour $\mathbf{A} \geq 0$, $\mathbf{A} < 0$ et $\mathbf{A} \leq 0$, au lieu de $\mathbf{A} \succeq 0$, $\mathbf{A} \prec 0$ et $\mathbf{A} \preceq 0$.

  ⚑ Les traductions anglaises peuvent prêter à confusion :
  - ▷ définie-positive = *positive-definite*, parfois on précise *symmetric positive-definite* et on l'abrège *SPD*. "SPD" signifie donc : semi-définie positive en français, mais définie-positive en anglais ;
  - ▷ semi-définie positive = *positive-semidefinite*, souvent abrégé *PSD*. Celui-ci n'existe pas en français.

  Il faut garder en tête que même sans la mention de *symétrique* ou *symmetric*, la notion de signe d'une matrice ne s'applique qu'aux matrices symétriques.
- ▷ Pour parler d'une matrice de plein rang, on ne précise pas toujours "plein rang *ligne*" ou "plein rang *colonne*", même si elle n'est pas carrée. Si $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathrm{rg}(\mathbf{A}) \leq \min\{m,n\}$ donc "$\mathbf{A}$ est de plein rang" peut aussi signifier :
  - $\mathrm{rg}(\mathbf{A}) = m$ si $m < n$;
  - $\mathrm{rg}(\mathbf{A}) = n$ si $n < m$.
- ▷ Le déterminant de $\mathbf{A}$ est parfois directement noté $|\mathbf{A}|$.
- ▷ La transconjuguée de $\mathbf{A}$ est parfois appelée sa transposée conjuguée, son adjoint ou son conjugué hermitien, et noté $\mathbf{A}^H$.
- ▷ Une matrice hermitienne (voir la définition 3.6.7) est parfois appelée *matrice auto-adjointe*.

  ⚑ *Self-adjoint matrix.*
- ▷ Matrice de changement de base = matrice de passage.

  ⚑ *Change-of-coordinates matrix*, *Change-of-basis matrix.*

vi Systèmes

- ▷ Si deux matrices sont équivalentes parce que l'une a été obtenue après des opérations d'élimination sur l'autre, elles sont parfois dites *équivalentes selon les lignes*, puisque les opérations élémentaires ne concernent que les lignes.

  ⚑ *Row equivalent matrices.*

vii Valeurs propres et valeurs singulières

- ▷ Le spectre de $\mathbf{A}$ est parfois noté $\sigma(\mathbf{A})$.
- ▷ La multiplicité algébrique de $\lambda \in \mathrm{Sp}(\mathbf{A})$ est parfois notée $\mathrm{MA}(\lambda)$, et sa multiplicité géométrique $\mathrm{MG}(\lambda)$. Il arrive que l'on n'utilise pas le terme "multiplicité géométrique", mais que l'on parle directement de "dimension du sous-espace propre associé à λ".
- ▷ Le sous-espace propre associé à $\lambda \in \mathrm{Sp}(\mathbf{A})$ est parfois noté E_λ (cette notation ne fait pas mention de la matrice $\mathbf{A}$, qui est rendue claire par le contexte).
- ▷ Le polynôme caractéristique de $\mathbf{A}$ est parfois noté $\chi_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I})$.

viii Fonctions

- ▷ On confond ici les trois termes de *fonction*, *application* et *transformation*.
- Espace d'arrivée d'une fonction = *Codomain*. Ce terme a donné lieu à l'utilisation de *codomaine* en français, parfois.
- ▷ L'ensemble des polynômes à coefficients réels est parfois noté $\mathbb{R}[x]$, ou $\mathbb{R}[X]$. Cette seconde notation fait intervenir la "grande indéterminée" X et est surtout pertinente dans l'étude algébrique des polynômes, en dehors du cadre de cet ouvrage. De même, l'ensemble des polynômes à coefficients complexes est parfois noté $\mathbb{C}[x]$ ou $\mathbb{C}[X]$.
- ▷ Le gradient de f en $\mathbf{x}$ est parfois noté $\mathrm{grad} f(\mathbf{x})$, et sa hessienne peut être notée $\mathbf{H}_f(\mathbf{x})$, $\mathbf{H}(\mathbf{x})$ (quand f est claire par le contexte) ou $\mathrm{Hess} f(\mathbf{x})$.

ix Exercices sur les notations

Solutions disponibles à l'annexe A.

Série d'exercices 0: Notations

Exercice 0.1: valide ou invalide ? ($\star$)

Soient A un ensemble de vecteurs, F une famille de vecteurs, et $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs. Parmi les expressions suivantes, déterminer lesquelles sont valides.

1 $\mathbf{x} \in A$;
2 $\mathbf{x} \in F$;
3 $\mathbf{x} \subseteq A$;
4 $\mathbf{x} \subseteq F$;
5 $\{\mathbf{x}, \mathbf{y}\} = A$;
6 $\{\mathbf{x}, \mathbf{y}\} \in A$;
7 $\{\mathbf{x}, \mathbf{y}\} \subseteq A$;
8 $\mathbf{x} \in \mathbb{R}^2$;
9 $\mathbf{x} \subset \mathbb{R}^2$
10 $\mathbf{x} = \mathbb{R}^2$
11 $\mathbf{x} = (1\ 2)$;
12 $\mathbf{x} = (1\ 2)^\top$;
13 $\mathbf{x} = (1, 2)$;
14 $\mathbf{x} = (1, 2)^\top$;
15 $\mathbf{x} = [1, 2]$;
16 $\mathbf{x} = [1, 2]^\top$;
17 $\mathbf{x} = [1\ 2]$;
18 $\mathbf{x} = [1\ 2]^\top$;
19 $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$;
20 $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}^\top$.

Exercice 0.2: ensembles et familles : repérage et réécriture ($\star$)

1. Pour chacune des écritures suivantes, préciser si elle désigne un ensemble ou une famille, puis la réécrire en respectant les conventions :
   - ▷ $(5, 2, 5)$;
   - ▷ $\{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3\}$;
   - ▷ $\{p, p+1, \ldots, n\}$.
2. Noter sous forme compacte l'ensemble des entiers de -3 à 4 (bornes incluses) en utilisant la notation $[\![\cdot]\!]$.
3. Dire si la base canonique $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ est un ensemble ou une famille.

Exercice 0.3: opérations sur ensembles et familles ($\star$)

Soient $A = \{1, 3, 5\}$ et $B = (1, 3, 3, 4)$.

1. Préciser pour chacun s'il s'agit d'un ensemble ou d'une famille.
2. Déterminer :
   (a) $A \cup B$;

(b) $A \cap B$;
(c) $B \setminus A$;
(d) $A \setminus B$.

Exercice 0.4: bon usage de l'opérateur $\mathrm{Vect}(\cdot)$ **($\star$)**

Soient A et B deux ensembles, F et G deux familles, et $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs. Parmi les expressions suivantes, déterminer lesquelles sont valides.

1 $\mathrm{Vect}(B) = A$;

2 $\mathrm{Vect}(F) = A$;

3 $\mathrm{Vect}(F) = G$;

4 $\mathrm{Vect}(A) = F$;

5 $\mathrm{Vect}(\mathbf{x}, \mathbf{y}) \in A$;

6 $\mathrm{Vect}(\mathbf{x}, \mathbf{y}) \subset A$;

7 $\mathrm{Vect}(\mathbf{x}, \mathbf{y}) = A$;

9 $\mathrm{Vect}((\mathbf{x}, \mathbf{y})) = A$;

10 $\mathrm{Vect}\,\{\mathbf{x}, \mathbf{y}\} = A$.

Exercice 0.5: appartenance et inclusion – ensembles ($\star$)

Soient les ensembles $A = \{1, 2, 4, 6\}$ et $B = \{2, 4\}$. Indiquer à l'aide des opérateurs d'appartenance et d'inclusion si un lien existe entre chacune des paires d'éléments suivantes :

1. 2 et B ;
2. $\{2\}$ et A ;
3. 6 et B ;
4. $\{1, 6\}$ et A ;
5. $\{1, 3\}$ et A ;
6. A et B.

Exercice 0.6: appartenance et inclusion – familles ($\star$)

Soient les familles $F = (1, 2, 4, 6)$ et $G = (2, 4)$. Indiquer à l'aide des opérateurs d'appartenance et d'inclusion si un lien existe entre chacune des paires d'éléments suivantes.

1. 2 et G ;
2. (2) et F ;
3. $\{2\}$ et F ;
4. F et G ;
5. $(4, 2)$ et F ;
6. $(1, 2, 2, 4, 6)$ et F.

Exercice 0.7: appartenance et inclusion – familles de vecteurs ($\star\star$)
Soient les vecteurs $\mathbf{u}_1 = (-1, 0, 1)$, $\mathbf{u}_2 = (1, 0, -1)$, et $\mathbf{u}_3 = (2, 0, 0)$, et les deux familles $F = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$ et $G = (\mathbf{u}_1, \mathbf{u}_3)$.

Indiquer à l'aide des opérateurs d'appartenance et d'inclusion si un lien existe entre chacune des paires d'éléments suivantes.

1. $\mathbf{u}_3$ et G ;
2. -1 et $\mathbf{u}_2$;
3. $(\mathbf{u}_2)$ et F ;
4. $\{\mathbf{u}_2\}$ et F ;
5. F et G ;
6. $(\mathbf{u}_3, \mathbf{u}_2)$ et F ;
7. $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2, \mathbf{u}_3)$ et F.

Première partie

Algèbre linéaire : Fondements

Chapitre 1

Vecteurs et matrices

Ce chapitre est consacré à l'étude des *vecteurs et matrices*, qui sont des outils fondamentaux en algèbre linéaire Il sert d'introduction à leur manipulation, alors que leur vraie nature est formalisée et étudiée au chapitre 5.

Plus précisément, on manipule dans ce chapitre les vecteurs colonne[1] de $\mathbb{R}^n$, et les matrices réelles. Ces deux structures sont en fait des tableaux de réels desquels découlent des définitions, propriétés et méthodes fondamentales en algèbre linéaire. Les manipulations présentées ici sont simples, mais à la base du reste de l'ouvrage.

Les vecteurs colonne de $\mathbb{C}^n$ et les matrices complexes sont définis à la section 3.6. Cependant, tout au long de ce chapitre, il est mentionné lorsqu'il est possible de travailler avec $\mathbb{C}^n$. Dans ce cas, les subtilités inhérentes à leur nature sont explicitées.

Par ailleurs, le mot "scalaire" apparaîtra souvent : ce terme désigne, selon le contexte, soit un réel, soit un complexe. Si on choisit un $k \in \mathbb{C}$, c'est une généralisation du cas réel. En effet, rien n'empêche un complexe k d'être dans $\mathbb{R}$ (figure 2).

On précise enfin que les premières applications de ces concepts, telles que la résolution de systèmes et l'inversion matricielle, sont vus plus tard dans cet ouvrage, au chapitre 8. De façon générale, les applications des concepts fondamentaux de la partie I sont vues à la partie II.

1. On ne met pas de "s" à "colonne" car on manipule la notion de vecteurs *de type* colonne.

1.1 Vecteurs colonne

Cette section est consacrée à la présentation des vecteurs colonne. Ces vecteurs peuvent être additionnés et multipliés par un scalaire pour donner de nouveaux vecteurs colonne grâce à des combinaisons linéaires présentées à la définition 1.1.9. Enfin, la notion de produit scalaire est introduite afin de définir le concept de norme pour un vecteur colonne.

1.1.1 Définitions

Tel qu'indiqué dans l'introduction du chapitre, un vecteur est un tableau de scalaires. Ce tableau est, par convention, représenté sous forme d'une colonne[2], d'où le nom de "vecteur colonne".

Le nombre de lignes de ce tableau détermine la taille du vecteur, tel que vu dans l'exemple 1.1.1. Les éléments d'un vecteur colonne (ce qui apparaît sur chaque ligne) sont appelés les composantes. Ce sont ces composantes qui sont des scalaires. On utilise des parenthèses pour y accéder. L'exemple 1.1.1 clarifie cette notion de composante.

> **Définition 1.1.1** (*vecteur colonne*)
>
> Un *vecteur colonne* est un tableau constitué d'une colonne de scalaires. Les scalaires sur chacune des lignes sont appelés les *composantes* du vecteur.

Il convient de noter qu'au chapitre 5, consacré aux bases, le terme "coordonnées" est préféré à celui de "composantes". Le théorème 5.2.1 énonce l'emploi correct de ce terme.

Bien que les vecteurs colonne soient souvent représentés en 2D ou 3D par une flèche, comme dans la figure 1.1, leur notation avec une flèche au dessus de la minuscule n'est pas appropriée au niveau universitaire. En effet, pour de plus grandes dimensions, cette représentation perd son sens, et on préfère donc la notation en minuscules grasses.

Il existe trois façons d'écrire une colonne pour représenter un vecteur colonne. Celles-ci sont détaillées à la section iv. Tous les autres usages liés aux notations sont explicités dans cette section.

Les vecteurs colonne ayant le même nombre de lignes (composantes) forment un ensemble introduit à la définition 1.1.2.

> **Définition 1.1.2** ($\mathbb{R}^n$)
>
> $\mathbb{R}^n$ désigne l'ensemble des vecteurs colonne à n composantes réelles.

2. C'est la convention usuelle mais, on aurait pu les laisser sous forme de lignes. Cependant, cette configuration facilite la compréhension des opérations présentées par la suite.

Pour $\mathbb{C}^n$, la définition reste la même, mais, les composantes sont cette fois-ci des complexes. L'ensemble $\mathbb{C}^n$ est donc une extension de $\mathbb{R}^n$ puisque $\mathbb{R}^n \subset \mathbb{C}^n$. On utilise également le terme de "vecteur" (sans le terme "colonne") pour les éléments de $\mathbb{R}^n$ ou $\mathbb{C}^n$ ainsi que pour les éléments de certains sous-ensembles de $\mathbb{R}^n$ ou $\mathbb{C}^n$.

La définition formelle des vecteurs est donnée au chapitre 5, et l'ensemble V de ces vecteurs est alors appelé un *espace vectoriel*.

Remarque

Si $n = 1$, on *assimile* le vecteur à un scalaire : $\mathbb{R} = \mathbb{R}^1$ et $\mathbb{C} = \mathbb{C}^1$. Par exemple, $[3] = 3$.

Exemple 1.1.1

- ▷ $\mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ est un vecteur (de $\mathbb{R}^2$) à deux composantes réelles, et $\mathbf{u}(2) = 4$ est la seconde composante du vecteur $\mathbf{u}$.
- ▷ $\mathbf{v} = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 4 \\ \pi \end{bmatrix}$ est un vecteur de $\mathbb{R}^5$ à cinq composantes réelles, et $\mathbf{v}(5) = \pi$ est la cinquième composante du vecteur $\mathbf{v}$;
- ▷ Pour $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, un vecteur colonne à n composantes réelles ou complexes, $\mathbf{x}(i) = x_i$ est la i-ième composante de $\mathbf{x}$: c'est donc un scalaire, ce qui explique pourquoi cette composante n'est pas écrite en gras comme le vecteur.

Pour un n donné, les vecteurs les plus connus et utiles dans de nombreux cas sont les colonnes de la matrice identité, tel que décrit à la définition 1.1.3. Le concept de "matrice identité" sera expliqué de manière formelle à la définition 1.2.16.

Définition 1.1.3 *(j-ième colonnes de la matrice identité)*

Pour un n donné, on définit n vecteurs $\mathbf{e}_j \in \mathbb{R}^n$ pour $j \in [\![1; n]\!]$ tels que $\mathbf{e}_j(i) = 0$ si $i \neq j$ pour $i \in [\![1; n]\!]$ et $\mathbf{e}_j(j) = 1$.

Exemple 1.1.2

Dans $\mathbb{R}^3$, les trois colonnes de la matrice identité sont $\mathbf{e}_1 = (1, 0, 0)$ (souvent noté $\mathbf{i}$), $\mathbf{e}_2 = (0, 1, 0)$ (souvent noté $\mathbf{j}$), et $\mathbf{e}_3 = (0, 0, 1)$ (souvent noté $\mathbf{k}$).

Pour un vecteur colonne donné, il est utile de définir son transposé dans lequel une colonne (un vecteur) devient une ligne et donc une *matrice* (ce terme est défini à la section suivante). La transposée est notamment utile dans la notion de produit scalaire vue ultérieurement à la section 1.1.4.

Définition 1.1.4 (*transposé d'un vecteur*)
Le transposé d'un vecteur $\mathbf{x} \in \mathbb{R}^n$, noté $\mathbf{x}^\top$, est une matrice à une ligne et n colonnes :

$$\mathbf{x}^\top = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}^\top = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} .$$

Pour les vecteurs complexes, transposer n'est souvent pas suffisant. On fait alors appel à la transconjuguée sur les complexes (la définition 3.6.4 donne plus de détails).

Remarque

- ▷ Il est important de noter qu'une matrice composée d'une ligne n'est pas un vecteur mais une matrice $1 \times n$;
- ▷ On n'assimile pas $\mathbb{R}^{1\times n}$ à $\mathbb{R}^n$ pour $n > 1$. C'est $\mathbb{R}^{n\times 1}$ qui est assimilé à $\mathbb{R}^n$ (il en est de même pour les complexes).

On conclut cette sous-section en décrivant trois façons différentes d'écrire les vecteurs colonne.

Définition 1.1.5 (*les trois différentes écritures d'un vecteur colonne*)
Soit $\mathbf{x}$ un vecteur (colonne) de n composantes notées x_1, x_2, …, x_n. On utilise les trois façons suivantes d'écrire $\mathbf{x}$ et ses composantes :

$$\begin{aligned} \mathbf{x} &= (x_1, x_2, \dots, x_n) \\ &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^\top \\ &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} . \end{aligned}$$

1.1.2 Opérations vectorielles simples

Dans cette section, deux opérations simples sur les vecteurs sont abordées : l'addition et la multiplication par un scalaire. Ces deux types d'opérations sont combinés dans la section 1.1.3 portant sur les combinaisons linéaires. Ces opérations se font dans un ensemble $\mathbb{R}^n$, où n est un entier naturel supérieur à un, mais, pour des questions de visualisation, les exemples sont donnés avec des vecteurs de $\mathbb{R}^2$. Un vecteur de $\mathbb{R}^2$ a donc deux composantes que l'on note x et y par la suite. On peut donc représenter un vecteur $\mathbf{x} = (x, y) \in \mathbb{R}^2$ comme un *point* dans un repère avec deux axes perpendiculaires, chacun des axes représentant une dimension. C'est ce que l'on appelle le repère "2D". L'origine de ce repère représente le vecteur nul noté $\mathbf{0} = (0, 0)$. Les propriétés de ce vecteur sont vues au chapitre 5 et dans la section 1.1.3. Le repère 2D, le vecteur nul $\mathbf{0} = (0, 0)$ et le vecteur $\mathbf{v} = (v_1, v_2)$ sont représentés dans la figure 1.1. Bien que le vecteur $\mathbf{v}$ corresponde au point de coordonnées (v_1, v_2), la coutume est d'ajouter une (bleue dans la figure), de manière symbolique, pour expliciter ce vecteur. Ici, le mot vecteur prend une signification plus générale : Pour n quelconque, un vecteur n'a pas de "direction", ni d'"origine", et ni d'"extrémité", même si ces termes peuvent être utiles pour $n \leq 3$.

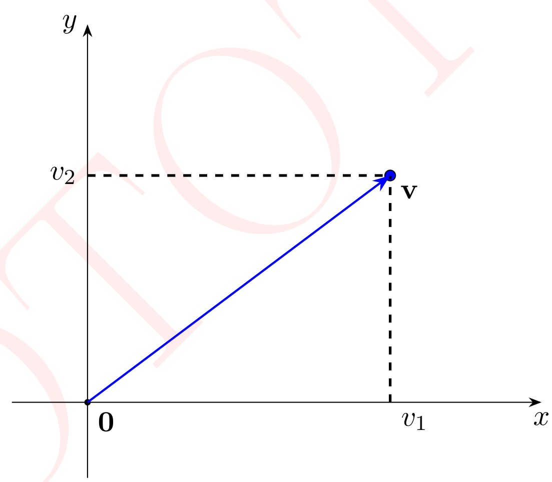

FIGURE 1.1 – Représentation d'un repère de $\mathbb{R}^2$, du vecteur nul $\mathbf{0}$ et du vecteur $\mathbf{v} = (v_1, v_2)$.

La première opération étudiée dans cette section est l'addition de deux vecteurs, soit l'opération qui consiste à additionner chacune des composantes une à une.

Définition 1.1.6 (*addition de deux vecteurs*)

Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Le vecteur $\mathbf{z} = \mathbf{x} + \mathbf{y}$ correspond à l'addition de $\mathbf{x}$ et $\mathbf{y}$. Il est calculé en additionnant les composantes des deux vecteurs. Plus précisément :

$$\mathbf{z} = \mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix} .$$

La soustraction de deux vecteurs, formalisée à la définition 1.1.7, se fait de manière similaire en soustrayant les composantes une à une. Il convient de souligner que l'ordre des vecteurs dans une soustraction est important.

Définition 1.1.7 (*soustraction de deux vecteurs*)

Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Le vecteur $\mathbf{z} = \mathbf{x} - \mathbf{y}$ correspond à la soustraction de $\mathbf{x}$ et $\mathbf{y}$. Il est calculé en soustrayant les composantes des deux vecteurs. Plus précisément :

$$\mathbf{z} = \mathbf{x} - \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} - \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_n - y_n \end{bmatrix} .$$

La figure 1.2 illustre ces notions d'addition et de soustraction dans $\mathbb{R}^2$. On peut remarquer que le vecteur $\mathbf{v} - \mathbf{w}$ est différent du vecteur $\mathbf{w} - \mathbf{v}$.

Le deuxième type d'opération vu dans cette section est la multiplication d'un vecteur par un scalaire. La définition 1.1.8 formalise cette opération.

Définition 1.1.8 (*multiplication d'un vecteur par un scalaire*)

Soit $\mathbf{x} \in \mathbb{R}^n$ et soit k un scalaire. Le vecteur $\mathbf{y} = k\mathbf{x}$ correspond à la multiplication de $\mathbf{x}$ par le scalaire k. Il est calculé en multipliant par k les composantes de $\mathbf{x}$. Plus précisément :

$$\mathbf{y} = k\mathbf{x} = k \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} kx_1 \\ kx_2 \\ \vdots \\ kx_n \end{bmatrix}.$$

Le scalaire k est également appelé *poids*.

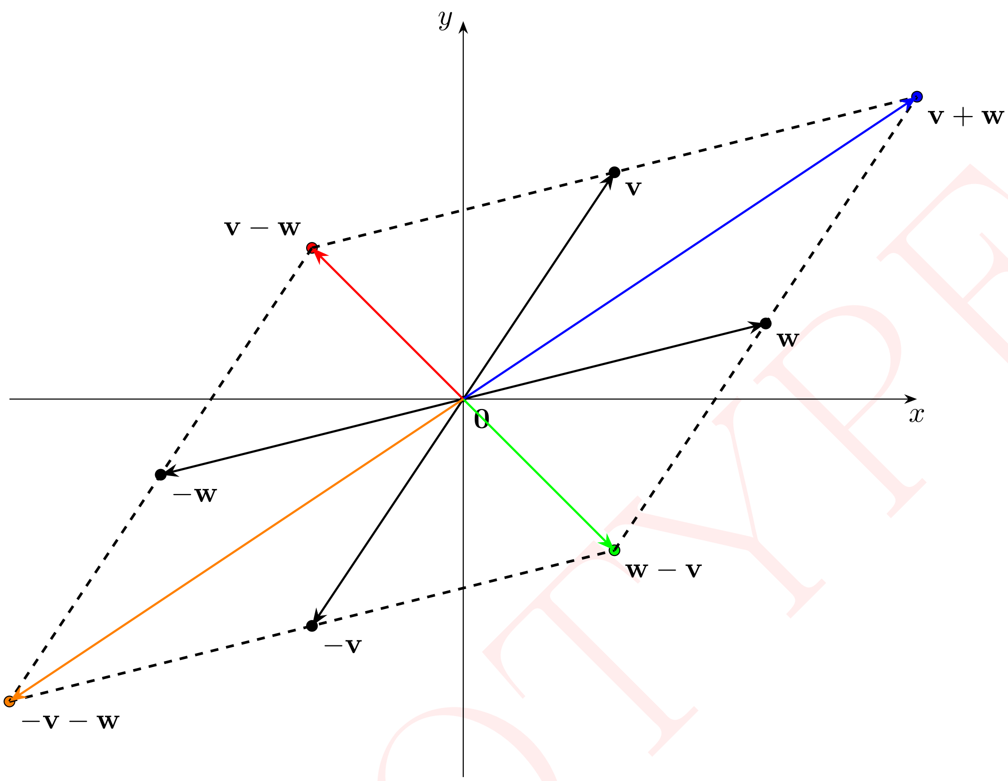

FIGURE 1.2 – Représentation de deux vecteurs $\mathbf{v}$ et $\mathbf{w}$ dans $\mathbb{R}^2$ ainsi que de leur addition et de leur soustraction.

Remarque

Selon les valeurs du scalaire k, le vecteur résultant a des propriétés :

- ▷ Si $k < 0$, les flèches qui relient $\mathbf{0}$ à $\mathbf{x}$ et à $k\mathbf{x}$ sont de sens opposé ;
- ▷ Si $0 < k < 1$, $k\mathbf{x}$ est “rétréci” par rapport à $\mathbf{x}$;
- ▷ Si $k > 1$, $k\mathbf{x}$ est “agrandi” par rapport à $\mathbf{x}$.

Ces notions de “rétréci” et “agrandi” prendront leur sens lorsque la norme d'un vecteur sera définie à la définition 1.1.13.

La figure 1.3 illustre cette opération de multiplication par un scalaire dans $\mathbb{R}^2$ pour différentes valeurs de k.

Si deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont tels que $\mathbf{x} = k\mathbf{y}$, où k est un scalaire[3], alors on dit

3. Ce scalaire peut être nul : Comme $\mathbf{0} = 0\mathbf{x}$ pour tout vecteur $\mathbf{x}$, le vecteur nul $\mathbf{0}$ est colinéaire à tous les vecteurs.

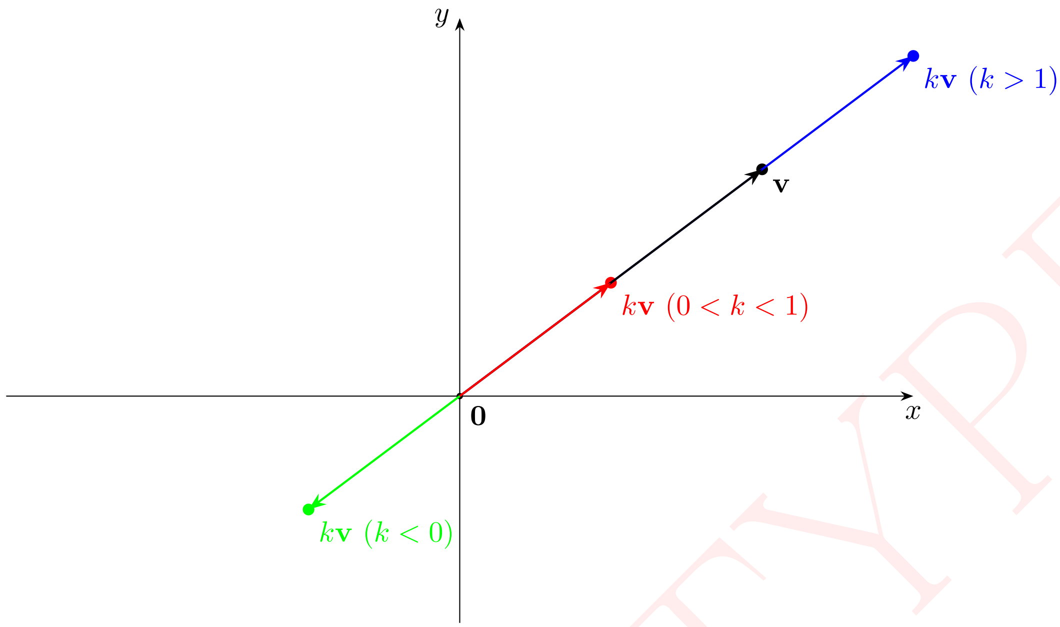

FIGURE 1.3 – Représentation de $\mathbf{x}$ et $k\mathbf{x}$ pour différentes valeurs de k dans $\mathbb{R}^2$.

que $\mathbf{x}$ et $\mathbf{y}$ sont *colinéaires*, ou *dépendants*. La notion d'*indépendance linéaire* est vue à la section 5.2.1. En combinant les opérations d'addition et de multiplication par un scalaire, on fait intervenir la notion de combinaison linéaire qui est introduite à la définition 1.1.9 de la section suivante.

1.1.3 Combinaisons linéaires

Une combinaison linéaire de vecteurs est une somme de ces vecteurs qui sont chacun pondérés par un scalaire. Le résultat est lui aussi un vecteur. La définition 1.1.9 formalise ceci.

Définition 1.1.9 (*combinaison linéaire*)

Soit V un ensemble de vecteurs. Une combinaison (linéaire) de vecteurs de V est une opération qui consiste à multiplier ces vecteurs par des scalaires (appelés des *poids*), puis à les additionner. On parle aussi de somme pondérée.
Plus précisément, avec les p vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ de V et avec les p poids $\alpha_1, \alpha_2, \ldots, \alpha_p$ associés à ces vecteurs, la combinaison linéaire de ces vecteurs pondérés s'écrit

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_p\mathbf{v}_p = \sum_{k=1}^{p} \alpha_k\mathbf{v}_k \ .$$

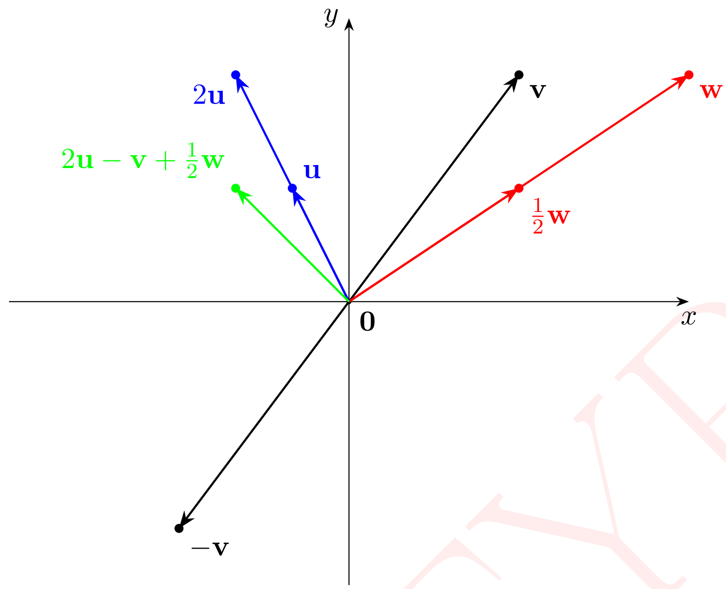

FIGURE 1.4 – Illustration de l'exemple 1.1.3.

Cette notion de combinaison linéaire s'applique non seulement à $V = \mathbb{R}^n$ ou $\mathbb{C}^n$, mais également à des sous-ensembles de ces ensembles.

Dans le chapitre 5, l'ensemble V est généralisé et appelé *espace vectoriel*. Grâce à certaines propriétés, on peut montrer qu'une combinaison linéaire de vecteurs d'un espace vectoriel appartient également à cet espace.

Dans l'exemple 1.1.3, la combinaison linéaire de trois vecteurs de $\mathbb{R}^2$ est étudiée.

Exemple 1.1.3

Soient $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ et $\mathbf{w} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$. Une combinaison linéaire de ces trois vecteurs est par exemple : $2\mathbf{u} - \mathbf{v} + \frac{1}{2}\mathbf{w}$. Dans ce cas,

$$2\mathbf{u} - \mathbf{v} + \frac{1}{2}\mathbf{w} = 2\begin{bmatrix} -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 4 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix} + \begin{bmatrix} -3 \\ -4 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}.$$

Le vecteur résultant est bien un vecteur de $\mathbb{R}^2$. Les différentes étapes de la somme sont représentées à la figure 1.4.

Ainsi, si on considère un ensemble $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$ de p vecteurs de V et un ensemble $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ de p scalaires, alors $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_p\mathbf{x}_p$ est une combinaison linéaire des p vecteurs, qui est elle-même un vecteur de V. Cette appartenance à V est formellement démontrée à la proposition 5.1.1.

Remarque (*combinaisons linéaires particulières de deux vecteurs*)

Soient $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs, et soit $\alpha\mathbf{x} + \beta\mathbf{y}$ leur combinaison linéaire avec deux scalaires α et β :

- ▷ Si $\alpha = \beta = 1$: la combinaison est la somme des deux vecteurs : $\mathbf{x} + \mathbf{y}$;
- ▷ Si $\alpha = 1$, $\beta = -1$: la combinaison est la soustraction des deux vecteurs : $\mathbf{x} - \mathbf{y}$;
- ▷ Si $\beta = 0$: la combinaison est la multiplication de $\mathbf{x}$ par le scalaire α ;
- ▷ Si $\alpha = \beta = 0$: cette combinaison donne $0\mathbf{x} + 0\mathbf{y} = \mathbf{0}$, c'est-à-dire le vecteur nul. Cette combinaison linéaire est appelée *triviale*.

Il existe une infinité de combinaisons linéaires d'un sous-ensemble de vecteurs avec des scalaires quelconque. L'opérateur $\text{Vect}(\cdot)$ traduit se phénomène. Il permet d'écrire d'une façon condensée l'ensemble des combinaisons linéaires d'une famille de vecteurs. La définition 1.1.10 explicite ceci. Pour maîtriser cette notation, il est recommandé de faire l'exercice 0.4.

Définition 1.1.10 (*opérateur* $\text{Vect}(\cdot)$)

Soit $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ une famille de p vecteurs de $\mathbb{R}^n$. L'opérateur $\text{Vect}(\cdot)$ est défini comme

$$\text{Vect}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_p\mathbf{v}_p \text{ avec } c_1, c_2, \ldots, c_p \in \mathbb{R}\} .$$

Remarque (*entrées et sorties de* $\text{Vect}(\cdot)$)

L'opérateur $\text{Vect}(\cdot)$:

- ▷ prend une famille (et non un ensemble) de vecteurs en entrée ;
- ▷ retourne l'ensemble (et non une famille) de toutes les combinaisons linéaires possibles de ces vecteurs.

Il est montré au corollaire 5.1.1 que la sortie de $\text{Vect}(\cdot)$ est un espace vectoriel. L'exemple 1.1.4 donne une intuition de ce résultat ainsi que plusieurs exemples d'ensembles retournés par l'opérateur $\text{Vect}(\cdot)$.

Exemple 1.1.4

- ▷ Pour $\mathbf{u} \in \mathbb{R}^n$ et $\mathbf{u} \neq \mathbf{0}$, l'ensemble $\text{Vect}(\mathbf{u})$ est une droite. En 2D ($n = 2$), on peut visualiser ceci sur la figure 1.3 ;

- ▷ Étant donnés deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ non nuls de $\mathbb{R}^3$ qui ne sont pas colinéaires, l'ensemble $\mathrm{Vect}(\mathbf{u}, \mathbf{v})$ est un plan.

Ces différents cas de figures sont étudiés à la section 5.2.4.

1.1.4 Produit scalaire réel

Cette section introduit la notion de produit scalaire, une opération qui permet d'effectuer la “multiplication” de deux vecteurs colonne afin, notamment, de définir dans la section 1.2.6 non seulement la norme d'un vecteur, mais aussi le produit de deux matrices.

Pour multiplier deux vecteurs colonne, on introduit la notion de produit ligne-colonne à la définition 1.1.11. Le terme ligne vient du fait que l'on transpose un vecteur colonne et qu'il devient donc une ligne.

Définition 1.1.11 *(produit ligne-colonne)*

Étant donnés $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, le *produit ligne-colonne* de $\mathbf{x}$ avec $\mathbf{y}$, noté $\mathbf{x}^\top\mathbf{y}$, est défini par :

$$\mathbf{x}^\top\mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Il s'agit de la somme des produits des composantes des deux vecteurs.

Remarque

- ▷ Cette opération est généralisée par le produit matriciel vu à la définition 1.2.14 ;
- ▷ Formellement, l'opération $\mathbf{x}^\top\mathbf{y}$ est un produit de matrices, car il a été vu plus tôt qu'un vecteur transposé est une matrice. Il n'est pas nécessaire ici de maîtriser ce concept car la formule donnée dans la définition 1.1.11 permet de facilement faire ce calcul ;
- ▷ $\mathbf{x}^\top\mathbf{y} = \mathbf{y}^\top\mathbf{x}$: peu importe le vecteur qui est transposé, le résultat est le même.

Ce produit ligne-colonne représente en fait un *produit scalaire* dans $\mathbb{R}^n$. Un produit scalaire est une fonction qui prend en entrée deux vecteurs et qui donne un scalaire en sortie.

La définition 1.1.12 donne la définition du produit scalaire dans $\mathbb{R}^n$ qui est utilisée dans ce document.

Définition 1.1.12 *(produit scalaire dans $\mathbb{R}^n$)*

Soient $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ et $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Le *produit scalaire* de $\mathbf{x}$

et $\mathbf{y}$, noté $\langle \mathbf{x}, \mathbf{y} \rangle$, est égal à

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i = \mathbf{x}^\top \mathbf{y} \ .$$

Remarque

- ▷ $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$;
- ▷ Dans certains ouvrages, le produit scalaire est noté avec un point entre les vecteurs à la place de la notation utilisée ici. Cela crée de la confusion et est donc proscrit dans cet ouvrage ;
- ▷ Dans la définition 1.1.12, il faut que les vecteurs soient dans $\mathbb{R}^n$. Pour $\mathbb{C}^n$, une autre définition est nécessaire : voir la définition 3.6.5 pour plus de détails ;
- ▷ La notion de produit scalaire est généralisée à la définition 5.3.1 ;
- ▷ Un produit scalaire peut être positif ou négatif.

Le produit scalaire permet de définir la *norme* d'un vecteur. Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$, c'est la “longueur” de la flèche qui représente symboliquement le vecteur.

Définition 1.1.13 (*norme d'un vecteur*)

Étant donné $\mathbf{x} \in \mathbb{R}^n$, la *norme* de $\mathbf{x}$, notée $\|\mathbf{x}\|$, est définie par

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2} \ .$$

Remarque

- ▷ La norme est bien définie car c'est la racine d'une somme d'éléments positifs ou nuls ;
- ▷ Cette définition étend la notion de valeur absolue. En effet, pour $n = 1$, $\mathbf{x} = (x_1)$ et on retrouve $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^{1} x_i^2} = \sqrt{x_1^2} = |x_1|$.

Dans les sections précédentes, la façon de multiplier un vecteur par un scalaire a été étudiée. Le résultat de cette opération est aussi un vecteur et possède donc une norme. Tel qu'indiqué dans la proposition 1.1.1, il n'est pas nécessaire de refaire le calcul de la norme de ce nouveau vecteur si l'on connaît la norme du vecteur initial. Cette propriété n'est présentée ici que pour les réels, mais elle est généralisée à la proposition 3.6.1 pour les complexes.

Proposition 1.1.1 *(norme et multiplication par un scalaire)*
Étant donnés $k \in \mathbb{R}$ et $\mathbf{x} \in \mathbb{R}^n$, on a $\|k\mathbf{x}\| = |k|\,\|\mathbf{x}\|$.

Preuve. Soit $k \in \mathbb{R}$ et soit $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. On a $k\mathbf{x} = (kx_1, kx_2, \ldots, kx_n)$ et donc $\|k\mathbf{x}\| = \sqrt{\langle k\mathbf{x}, k\mathbf{x}\rangle} = \sqrt{\sum_{i=1}^n (kx_i)^2} = \sqrt{k^2}\sqrt{\sum_{i=1}^n x_i^2} = |k|\,\|\mathbf{x}\|$. $\square$

Les vecteurs qui ont une norme égale à 1 (définition 1.1.14) sont très utilisés en pratique et notamment dans la section 5.3 et le chapitre 13.

Définition 1.1.14 *(vecteur normalisé)*
Un vecteur est *normalisé* (ou *unitaire*) si sa norme est égale à 1.

Remarque
Ceci est valide également pour les complexes avec la norme de la définition 3.6.6.

Une façon très simple d'obtenir un vecteur normalisé (non nul) est de le diviser par sa norme tel que vu à la proposition 1.1.2.

Proposition 1.1.2 *(normalisation d'un vecteur)*
Si $\mathbf{x} \in \mathbb{R}^n$ est un vecteur autre que le vecteur nul $\mathbf{0}$, alors $\dfrac{\mathbf{x}}{\|\mathbf{x}\|}$ est normalisé.

Preuve. D'après la proposition 1.1.1, on a $\left\|\dfrac{\mathbf{x}}{\|\mathbf{x}\|}\right\| = \left|\dfrac{1}{\|\mathbf{x}\|}\right| \|\mathbf{x}\|$. De plus, $\left|\dfrac{1}{\|\mathbf{x}\|}\right| \|\mathbf{x}\| = \dfrac{1}{\|\mathbf{x}\|}\|\mathbf{x}\|$ car $\|\mathbf{x}\| > 0$. Il s'ensuit que $\left\|\dfrac{\mathbf{x}}{\|\mathbf{x}\|}\right\| = \dfrac{\|\mathbf{x}\|}{\|\mathbf{x}\|} = 1$. $\square$

Remarque
Cette technique de normalisation fonctionne pour les complexes : voir la proposition 3.6.2.

Il existe une relation très connue qui relie le produit scalaire de deux vecteurs et leur norme : c'est l'inégalité de Cauchy-Schwarz. Elle est très utile dans de nombreuses démonstrations et notamment celle du théorème 1.1.2. Le théorème 1.1.1 est aussi valide pour les complexes grâce aux définitions de norme et de produit scalaire adéquates vues dans le chapitre 3.

Théorème 1.1.1 (*inégalité Cauchy-Schwarz*)
Étant donnés $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ou $\mathbb{C}^n$, on a

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| \; .$$

La compréhension de la preuve repose sur la connaissance des propriétés des fonctions du second degré (section 4.3.2) et des normes.

Preuve. Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ou $\mathbb{C}^n$ et soit f une fonction réelle à une variable définie par $f(\alpha) = \|\mathbf{x} + \alpha \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\alpha \langle \mathbf{x}, \mathbf{y} \rangle + \alpha^2 \|\mathbf{y}\|^2$. Si $\mathbf{y} = \mathbf{0}$, l'inégalité est vraie puisque $|\langle \mathbf{x}, \mathbf{y} \rangle| = 0 = \|\mathbf{x}\| \, \|\mathbf{y}\|$. Si $\mathbf{y} \neq \mathbf{0}$, sa norme est aussi différente de 0 (tel que vu dans le chapitre 5) et f est une fonction du second degré positive car elle est définie grâce à une norme au carré. Le discriminant associé à cette fonction est donc négatif ou nul car elle possède au plus une racine réelle (voir chapitre 4). On a donc

$$(2 \langle \mathbf{x}, \mathbf{y} \rangle)^2 - 4 \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \leq 0$$

c'est à dire :

$$\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \; .$$

En prenant la racine de chaque côté de l'inégalité, on déduit le résultat car une telle opération ne change pas le sens de l'inégalité (puisque la fonction racine est croissante et que les termes $|\langle \mathbf{x}, \mathbf{y} \rangle|$, $\|\mathbf{x}\|$ et $\|\mathbf{y}\|$ sont positifs). □

Comme la notion de norme étend la valeur absolue définie sur $\mathbb{R}$, l'inégalité triangulaire reste valide. Cette inégalité demeure vraie quelle que soit la norme considérée, qu'il s'agisse de celle définie ici ou de celle sur les complexes (définition 3.6.6) car les deux normes s'appuient sur les propriétés du produit scalaire.

Théorème 1.1.2 (*inégalité triangulaire*)
Étant donnés $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ou $\mathbb{C}^n$, on a $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Preuve. Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ou $\mathbb{C}^n$. En utilisant la distributivité du produit scalaire vue dans la définition 5.3.1, on obtient

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} + \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + 2 \langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \, . \end{aligned}$$

Il découle du théorème 1.1.1 que $\langle \mathbf{x}, \mathbf{y} \rangle \leq |\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|x\| \|y\|$. On a donc

$$\|\mathbf{x}+\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 = (\|\mathbf{x}\| + \|\mathbf{y}\|)^2.$$

En prenant la racine de chaque côté de l'inégalité, on déduit le résultat car une telle opération ne change pas le sens de l'inégalité (puisque la fonction racine est croissante et que les termes $\|\mathbf{x}+\mathbf{y}\|$, $\|\mathbf{x}\|$ et $\|\mathbf{y}\|$ sont positifs. □

Le cas d'égalité de l'inégalité triangulaire est traité dans l'exercice 1.7. Il est aussi possible de dériver d'autres égalités comme l'égalité du parallélogramme ou même le théorème de Pythagore vus dans l'exercice 1.6. Pour résoudre cet exercice, la notion d'orthogonalité doit être explicitée. Ici encore, cela étend une notion vue dans $\mathbb{R}^2$: la perpendicularité. On dit que deux vecteurs sont orthogonaux si leur produit scalaire est nul (définition 1.1.15). Si de plus, il sont normalisés, on dit qu'ils sont orthonormaux (définition 1.1.16).

Définition 1.1.15 *(vecteurs orthogonaux)*

Deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^n$ ou $\mathbb{C}^n$ sont dits orthogonaux si $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Remarque

La définition s'applique aussi au vecteur nul qui est donc orthogonal à tous les autres.

Définition 1.1.16 *(vecteurs orthonormaux)*

Deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^n$ ou $\mathbb{C}^n$ sont dits orthonormaux s'ils sont orthogonaux et normalisés, c'est-à-dire que $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ et $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

La notion d'orthogonalité est généralisée à la section 5.3.

En plus, de la "taille" d'un vecteur, on voudrait être capable d'évaluer à quel point des vecteurs sont "éloignés" l'un de l'autre, la notion de distance est introduite à la définition 1.1.17.

Définition 1.1.17 *(distance entre deux vecteurs)*

La *distance* entre deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^n$ est la norme du vecteur $\mathbf{x}-\mathbf{y}$, autrement dit $\|\mathbf{x}-\mathbf{y}\|$.

Remarque

- ▷ La norme d'un vecteur représente donc sa distance au vecteur nul ;
- ▷ Comme $\|\mathbf{x}-\mathbf{y}\| = \|\mathbf{y}-\mathbf{x}\|$, la distance entre deux vecteurs est unique ;

Dans $\mathbb{R}^2$, le signe du produit scalaire possède une interprétation géométrique. En effet, si l'angle entre deux vecteurs est inférieur à $\pi/2$, la produit scalaire est positif. Cela vient du fait que l'on peut exprimer le produit scalaire de $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^2$ grâce au cosinus de l'angle formé par ces deux vecteurs.

Proposition 1.1.3 (*produit scalaire dans* $\mathbb{R}^2$)
Soient $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs non nuls de $\mathbb{R}^2$ et soit $\theta \in [0;\pi]$ l'angle formé par les droites d'origine $\mathbf{0}$ et qui passent par $\mathbf{x}$ et $\mathbf{y}$, respectivement, dans le repère 2D. On a

$$\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos(\theta) \ .$$

Preuve. La distributivité du produit scalaire donne

$$\begin{aligned} \|\mathbf{x}-\mathbf{y}\|^2 &= \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x}-\mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x}-\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 \ . \end{aligned}$$

Considérons le triangle formé par les vecteurs $\mathbf{0}$, $\mathbf{x}$ et $\mathbf{y}$. La longueur de l'arête opposée à l'angle θ est égale à la distance $\|\mathbf{x}-\mathbf{y}\|$ (définition 1.1.17) et la loi des cosinus appliquée à ce triangle donne $\|\mathbf{x}-\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\,\|\mathbf{y}\|\cos(\theta)$.

On a donc $\|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\,\|\mathbf{y}\|\cos(\theta)$, ce qui implique $\langle \mathbf{x}, \mathbf{y} \rangle = \|x\|\,\|y\|\cos(\theta)$. □

Remarque
L'angle $\theta \in [0;\pi]$ entre deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ non nuls vérifie

$$\cos(\theta) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\,\|\mathbf{y}\|}$$

ce qui, par le théorème 1.1.1, donne bien un nombre dans l'intervalle $[-1;1]$. De plus,

- ▷ si $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, alors $\theta = \frac{\pi}{2}$ ($\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux) ;
- ▷ si $\langle \mathbf{x}, \mathbf{y} \rangle > 0$, alors $\theta \in [0; \frac{\pi}{2}[$;
- ▷ si $\langle \mathbf{x}, \mathbf{y} \rangle < 0$, alors $\theta \in]\frac{\pi}{2}; \pi]$.

L'exemple 1.1.5 illustre ce propos.

Exemple 1.1.5

Soient $\mathbf{u} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, et $\mathbf{w} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$:

- ▷ $\langle \mathbf{u}, \mathbf{v} \rangle = -1 \times 3 + 3 \times 4 = 9$;
- ▷ $\langle \mathbf{u}, \mathbf{w} \rangle = -1 \times 6 + 3 \times 1 = -3$.

Tel qu'illustré à la figure 1.5 l'angle θ_1 entre $\mathbf{u}$ et $\mathbf{w}$ est donc supérieur à $\frac{\pi}{2}$ (car $-3 < 0$) alors que l'angle θ_2 entre $\mathbf{u}$ et $\mathbf{v}$ est inférieur à $\frac{\pi}{2}$ (car $9 > 0$).

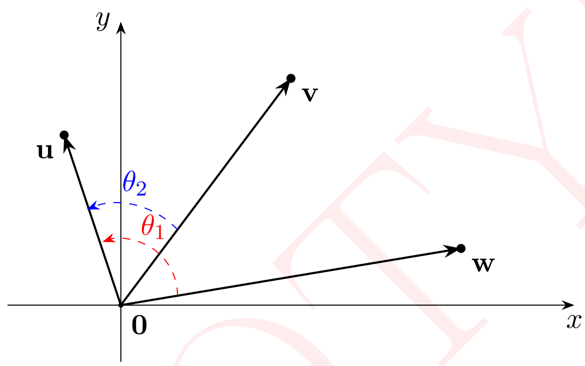

FIGURE 1.5 – Illustration de l'exemple 1.1.5.

1.1.5 Produit vectoriel

Comme le produit scalaire, le produit vectoriel prend en entrée deux vecteurs, mais y diffère en produisant un vecteur à la place d'un scalaire. Ce produit est défini pour les vecteurs à composantes réelles appartenant à $\mathbb{R}^3$. Il permet, pour deux vecteurs linéairement indépendants, d'en créer un troisième qui leur est orthogonal.

Définition 1.1.18 (*produit vectoriel dans* $\mathbb{R}^3$)

Soient $\mathbf{x} = (x_1, x_2, x_3)$ et $\mathbf{y} = (y_1, y_2, y_3)$ deux vecteurs de $\mathbb{R}^3$. Le produit vectoriel de $\mathbf{x}$ et $\mathbf{y}$ est

$$\mathbf{x} \wedge \mathbf{y} = \begin{bmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{bmatrix} .$$

Proposition 1.1.4 (*propriétés du produit vectoriel dans* $\mathbb{R}^3$)

Soient $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ trois vecteurs de $\mathbb{R}^3$, ainsi que α un réel. Le produit vectoriel respecte les propriétés suivantes.

1. Anti-commutativité : $\mathbf{x} \wedge \mathbf{y} = -(\mathbf{y} \wedge \mathbf{x})$;
2. Distributivité par rapport à l'addition : $\mathbf{x} \wedge (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \wedge \mathbf{y}) + (\mathbf{x} \wedge \mathbf{z})$;
3. Compatibilité avec la multiplication par un scalaire : $\alpha(\mathbf{x} \wedge \mathbf{y}) = (\alpha\mathbf{x}) \wedge \mathbf{y} = \mathbf{x} \wedge (\alpha\mathbf{y})$;
4. Respect de l'identité de Jacobi : $\mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}) + \mathbf{z} \wedge (\mathbf{x} \wedge \mathbf{y}) + \mathbf{y} \wedge (\mathbf{z} \wedge \mathbf{x}) = \mathbf{0}$;
5. Orthogonalité des vecteurs en entrée : $\langle (\mathbf{x} \wedge \mathbf{y}), \mathbf{x} \rangle = \langle (\mathbf{x} \wedge \mathbf{y}), \mathbf{y} \rangle = 0$;
6. Donne le vecteur nul pour deux vecteurs linéairement dépendants : $\mathbf{x} \wedge (\alpha\mathbf{x}) = \mathbf{0}$;
7. Règle de l'échange : $\langle (\mathbf{x} \wedge \mathbf{y}), \mathbf{z} \rangle = \langle \mathbf{x}, (\mathbf{y} \wedge \mathbf{z}) \rangle$;
8. Formule du double produit : $\mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}) = \langle \mathbf{x}, \mathbf{z} \rangle \, \mathbf{y} - \langle \mathbf{x}, \mathbf{y} \rangle \, \mathbf{z}$.

Les propriétés 7 et 8 sont démontrées ci-dessous. Le reste des propriétés est facilement démontrable et laissé en exercice.

Preuve de la règle de l'échange. Soit $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ et $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$. D'une part on a

$$\begin{aligned} \langle (\mathbf{x} \wedge \mathbf{y}), \mathbf{z} \rangle &= z_1(x_2y_3 - x_3y_2) + z_2(x_3y_1 - x_1y_3) + z_3(x_1y_2 - x_2y_1) \\ & x_1y_2z_3 - x_1y_3z_2 + x_2y_3z_1 - x_2y_1z_3 + x_3y_1z_2 - x_3y_2z_1 \ . \end{aligned}$$

D'une autre part, on a

$$\begin{aligned} \langle \mathbf{x}, (\mathbf{y} \wedge \mathbf{z}) \rangle &= x_1(y_2z_3 - y_3z_2) + x_2(y_3z_1 - y_1z_3) + x_3(y_1z_2 - y_2z_1) \\ &= x_1y_2z_3 - x_1y_3z_2 + x_2y_3z_1 - x_2y_1z_3 + x_3y_1z_2 - x_3y_2z_1 \ . \end{aligned}$$

D'où $\langle (\mathbf{x} \wedge \mathbf{y}), \mathbf{z} \rangle = \langle \mathbf{x}, (\mathbf{y} \wedge \mathbf{z}) \rangle$ pour tous $\mathbf{x}$ $\mathbf{y}$ et $\mathbf{z}$ de $\mathbb{R}^3$. $\square$

Preuve de la formule du double produit.

Soit $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ et $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$.

$$
\begin{aligned}
\mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}) &= \mathbf{x} \wedge \begin{bmatrix} y_2z_3 - y_3z_2 \\ y_3z_1 - y_1z_3 \\ y_1z_2 - y_2z_1 \end{bmatrix} \\
&= \begin{bmatrix} x_2(y_1z_2 - y_2z_1) - x_3(y_3z_1 - y_1z_3) \\ x_3(y_2z_3 - y_3z_2) - x_1(y_1z_2 - y_2z_1) \\ x_1(y_3z_1 - y_1z_3) - x_2(y_2z_3 - y_3z_2) \end{bmatrix} \\
&= \begin{bmatrix} y_1(x_3z_3 + x_2z_2) - z_1(x_2y_2 + x_3y_3) \\ y_2(x_3z_3 + x_1z_1) - z_2(x_3y_3 + x_1y_1) \\ y_3(x_1z_1 + x_2z_2) - z_3(x_1y_1 + x_2y_2) \end{bmatrix} \\
&= \begin{bmatrix} y_1(x_3z_3 + x_2z_2) + \underbrace{(y_1x_1z_1 - z_1x_1y_2)}_{0} - z_1(x_2y_2 + x_3y_3) \\ y_2(x_3z_3 + x_1z_1) + \underbrace{(y_2x_2z_2 - z_2x_2y_2)}_{0} - z_2(x_3y_3 + x_1y_1) \\ y_3(x_1z_1 + x_2z_2) + \underbrace{(y_3x_3z_3 - z_3x_3y_3)}_{0} - z_3(x_1y_1 + x_2y_2) \end{bmatrix} \\
&= \begin{bmatrix} y_1(x_3z_3 + x_2z_2 + x_1z_1) - z_1(x_2y_2 + x_3y_3 + x_1z_1) \\ y_2(x_3z_3 + x_1z_1 + x_2z_2) - z_2(x_3y_3 + x_1y_1 + x_2y_2) \\ y_3(x_1z_1 + x_2z_2 + x_3z_3) - z_3(x_1y_1 + x_2y_2 + x_3y_3) \end{bmatrix} \\
&= \langle \mathbf{x}, \mathbf{z} \rangle \, \mathbf{y} - \langle \mathbf{x}, \mathbf{y} \rangle \, \mathbf{z} \ .
\end{aligned}
$$

□

Il est à noter que le produit vectoriel de deux vecteurs n'est pas associatif. C'est-à-dire qu'en général $(\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z} \neq \mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z})$.

Exemple 1.1.6 (*non-associativité du produit vectoriel*)

Voici un exemple qui montre que le produit vectoriel n'est pas associatif en général.

Soit $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ et $\mathbf{z} = \begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix}$.

D'une part on a

$$(\mathbf{x}\wedge\mathbf{y})\wedge\mathbf{z} = \left(\begin{bmatrix}1\\2\\3\end{bmatrix}\wedge\begin{bmatrix}2\\1\\2\end{bmatrix}\right)\wedge\begin{bmatrix}3\\4\\4\end{bmatrix} = \begin{bmatrix}2\times 2-3\times 1\\3\times 2-1\times 2\\1\times 1-2\times 2\end{bmatrix}\wedge\begin{bmatrix}3\\4\\4\end{bmatrix}$$
$$= \begin{bmatrix}1\\4\\-3\end{bmatrix}\wedge\begin{bmatrix}3\\4\\4\end{bmatrix} = \begin{bmatrix}4\times 4-4(-3)\\-3\times 3-4\times 1\\1\times 4-4\times 3\end{bmatrix} = \begin{bmatrix}28\\-13\\-8\end{bmatrix}.$$

D'un autre part, on a

$$\mathbf{x}\wedge(\mathbf{y}\wedge\mathbf{z}) = \begin{bmatrix}1\\2\\3\end{bmatrix}\wedge\left(\begin{bmatrix}2\\1\\2\end{bmatrix}\wedge\begin{bmatrix}3\\4\\4\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\end{bmatrix}\wedge\begin{bmatrix}1\times 4-2\times 4\\2\times 3-2\times 4\\2\times 4-1\times 3\end{bmatrix}$$
$$= \begin{bmatrix}1\\2\\3\end{bmatrix}\wedge\begin{bmatrix}-4\\-2\\5\end{bmatrix} = \begin{bmatrix}2\times 5-3(-2)\\3(-4)-1\times 5\\1(-2)-2(-4)\end{bmatrix} = \begin{bmatrix}16\\-17\\6\end{bmatrix}.$$

D'où $(\mathbf{x}\wedge\mathbf{y})\wedge\mathbf{z}\neq\mathbf{x}\wedge(\mathbf{y}\wedge\mathbf{z})$.

La propriété du point 5 indique que le vecteur produit par deux autres vecteurs sera toujours orthogonal à ceux-ci. Ceci implique qu'il sera aussi orthogonal au plan généré par les deux vecteurs en entrée. En effet, les vecteurs générés par deux vecteurs linéairement indépendants $\mathbf{v}$ et $\mathbf{w}$ de $\mathbb{R}^3$ sont donnés par une combinaison linéaire de $\mathbf{v}$ et $\mathbf{w}$. C'est-à-dire qu'ils sont représentables par

$$T(\alpha,\beta) = \alpha\mathbf{v}+\beta\mathbf{w},\ \text{pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R}.$$

En imaginant les composantes de chaque vecteur $T(\alpha,\beta)$ comme étant des coordonnées de points (x,y,z) dans $\mathbb{R}^3$, on remarque que ces vecteurs se retrouvent dans un plan de $\mathbb{R}^3$ passant par l'origine.

Définition 1.1.19 *(représentation paramétrique d'un plan passant par l'origine dans $\mathbb{R}^3$)*
Un plan passant par l'origine dans $\mathbb{R}^3$ peut être représenté par

$$T(\alpha,\beta) = \alpha\mathbf{v}+\beta\mathbf{w},\ \text{pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R},$$

où $\mathbf{v}$ et $\mathbf{w}$ sont des vecteurs linéairement indépendants de $\mathbb{R}^3$. C'est une représentation paramétrique d'un plan passant par l'origine généré par $\mathbf{v}$ et $\mathbf{w}$, de paramètres α et β.

Pour $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ et $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, la représentation est

$$T(\alpha, \beta) = \alpha \mathbf{v} + \beta \mathbf{w} = \begin{bmatrix} \alpha v_1 + \beta w_1 \\ \alpha v_2 + \beta w_2 \\ \alpha v_3 + \beta w_3 \end{bmatrix}, \text{ pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R}.$$

Autrement dit, n'importe quel vecteur $\mathbf{b} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ qui appartient à ce plan est tel que

$$\begin{aligned} x &= \alpha v_1 + \beta w_1, \\ y &= \alpha v_2 + \beta w_2 \text{ et} \\ z &= \alpha v_3 + \beta w_3. \end{aligned}$$

pour certains scalaires réels α et β. C'est l'équation paramétrique du plan généré par $\mathbf{v}$ et $\mathbf{w}$.

L'ensemble des vecteurs qui font partie d'un plan généré par deux vecteurs linéairement indépendants $\mathbf{v}$ et $\mathbf{w}$ dans $\mathbb{R}^3$ sont tous orthogonaux à $\mathbf{v} \times \mathbf{w}$, car pour toutes valeurs de α et β dans $\mathbb{R}$,

$$\begin{aligned} \langle \mathbf{v} \times \mathbf{w}, T(\alpha, \beta) \rangle &= \langle \mathbf{v} \times \mathbf{w}, \alpha \mathbf{v} + \beta \mathbf{w} \rangle \\ &= \langle \mathbf{v} \times \mathbf{w}, \alpha \mathbf{v} \rangle + \langle \mathbf{v} \times \mathbf{w}, \beta \mathbf{w} \rangle, \text{ distributivité du produit scalaire} \\ &= \alpha \langle \mathbf{v} \times \mathbf{w}, \mathbf{v} \rangle + \beta \langle \mathbf{v} \times \mathbf{w}, \mathbf{w} \rangle, \text{ propriétés de multiplication par un scalaire} \\ &= \alpha 0 + \beta 0 = 0 . \end{aligned}$$

En trois dimensions, le vecteur $\mathbf{v} \wedge \mathbf{w}$ produit à partir de deux vecteurs linéairement indépendants d'un plan est un vecteur qu'on dit normal à ce plan. Pour chaque plan dans $\mathbb{R}^3$, il existe une infinité de vecteurs normaux qui sont tous colinéaires à $\mathbf{v} \wedge \mathbf{w}$. C'est-à-dire qu'ils s'expriment comme $\mathbf{v} \wedge \mathbf{w}$ multiplié par une valeur réelle.

On peut montrer qu'un plan dans $\mathbb{R}^3$ passant par l'origine qui contient deux vecteurs linéairement indépendants $\mathbf{v}$ et $\mathbf{w}$ admet exactement tout les vecteurs dans $\mathbb{R}^3$ qui sont orthogonaux à $\mathbf{v} \wedge \mathbf{w}$.

Définition 1.1.20 (*équation cartésienne d'un plan passant par l'origine*)

Le plan passant par l'origine et contenant deux vecteurs $\mathbf{v}$ et $\mathbf{w}$ linéairement indépendants peut s'écrire sous la forme

$$\langle \mathbf{n}, \mathbf{x} \rangle = 0$$

où $\mathbf{n}$ est un vecteur normal au plan généré par $\mathbf{v}$ et $\mathbf{w}$ et où $\mathbf{x} = (x, y, z)$ est un vecteur appartenant au plan. L'ensemble des vecteurs $\mathbf{x}$ qui respectent cette équation forme un plan qui passe par l'origine. Peut importe le vecteur normal $\mathbf{n}$ choisi, l'équation représentera toujours le même plan. Ces équations sont les équations cartésiennes d'un plan.

Exemple 1.1.7 (*équation cartésienne d'un plan*)

On cherche l'équation d'un plan passant par l'origine dans $\mathbb{R}^3$ contenant les vecteurs indépendants $\mathbf{v} = (1, -2, 4)$ et $\mathbf{w} = (3, 7, 5)$. Il est possible d'utiliser le produit vectoriel suivant pour trouver un vecteur normal qui permettra de décrire le plan comme une équation cartésienne :

$$\mathbf{n} = \mathbf{v} \wedge \mathbf{w} = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix} \wedge \begin{bmatrix} 3 \\ 7 \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \times 5 - 4 \times 7 \\ 4 \times 3 - 1 \times 5 \\ 1 \times 7 - (-2) \times 3 \end{bmatrix} = \begin{bmatrix} -38 \\ 7 \\ 13 \end{bmatrix}.$$

Donc $\mathbf{x} = (x, y, z) \in \mathrm{Vect}(\mathbf{v}, \mathbf{w})$ (le plan généré par $\mathbf{v}$ et $\mathbf{w}$) s'il respecte

$$\langle \mathbf{n}, \mathbf{x} \rangle = 0 \iff \begin{bmatrix} -38 & 7 & 13 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0 \iff -38x + 7y + 13z = 0$$

qui est une équation cartésienne du plan généré par les deux vecteurs $\mathbf{v}$ et $\mathbf{w}$.

Que faire lorsqu'on cherche à obtenir une équation d'un plan qui ne passe pas par l'origine ? Dans ce cas, les vecteurs appartenant à ce plan sont générés en partant d'un vecteur $\mathbf{r}$ appartenant au plan et en se déplaçant dans les directions de deux vecteurs $\mathbf{v}$, $\mathbf{w}$ tangents au plan. Les vecteurs qui appartiennent à ce plan sont donnés par

$$T(\alpha, \beta) = \mathbf{r} + \alpha\mathbf{v} + \beta\mathbf{w}, \text{ pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R}.$$

Définition 1.1.21 (*représentation paramétrique d'un plan passant par un point quelconque dans $\mathbb{R}^3$*)

Un plan passant par un vecteur quelconque $\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}$ dans $\mathbb{R}^3$, où $\mathbf{v}$ et $\mathbf{w}$ sont deux vecteurs linéairement indépendants tangents à ce plan, peut être représenté paramétriquement par

$$T(\alpha, \beta) = \mathbf{r} + \alpha\mathbf{v} + \beta\mathbf{w}, \text{ pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R}.$$

Les paramètres de cette représentation sont α et β.
Si $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ et $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$, la représentation paramétrique devient

$$T(\alpha, \beta) = \mathbf{r} + \alpha\mathbf{v} + \beta\mathbf{w} = \begin{bmatrix} r_1 + \alpha v_1 + \beta w_1 \\ r_2 + \alpha v_2 + \beta w_2 \\ r_3 + \alpha v_3 + \beta w_3 \end{bmatrix}, \text{ pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R}.$$

On remarque que pour n'importe quel vecteur $\mathbf{b} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ appartenant à ce plan, il existe un scalaire α et un scalaire β tel que

$$\begin{aligned} x &= r_1 + \alpha v_1 + \beta w_1, \\ y &= r_2 + \alpha v_2 + \beta w_2 \text{ et} \\ z &= r_3 + \alpha v_3 + \beta w_3. \end{aligned}$$

C'est l'équation paramétrique du plan passant par $\mathbf{r}$, généré par $\mathbf{v}$ et $\mathbf{w}$.

Alors, n'importe quel vecteur $\mathbf{v}$ appartenant à ce plan est aussi tel que $\mathbf{b} - \mathbf{r}$ appartient au plan passant par l'origine et généré par $\mathbf{v}$ et $\mathbf{w}$. Le prochain exemple montre comment utiliser ce fait pour obtenir une équation cartésienne d'un plan ne passant pas par l'origine.

Définition 1.1.22 *(équation cartésienne d'un plan passant par un point quelconque)*
Soit $\mathbf{v}$, $\mathbf{w}$, $\mathbf{n}$ et $\mathbf{r}$, quatre vecteurs de $\mathbb{R}^3$ où $\mathbf{n}$ est un vecteur normal au plan généré par $\mathbf{v}$ et $\mathbf{w}$. Si un plan de $\mathbb{R}^3$ est généré par

$$T(\alpha, \beta) = \mathbf{r} + \alpha\mathbf{v} + \beta\mathbf{w}, \text{ pour tout } \alpha \text{ et } \beta \text{ dans } \mathbb{R},$$

alors n'importe quel vecteur $\mathbf{b}$ de $\mathbb{R}^3$ appartenant à ce plan est tel que

$$\langle \mathbf{n}, \mathbf{b} - \mathbf{r} \rangle = 0,$$

ou de manière équivalente

$$\langle \mathbf{n}, \mathbf{b} \rangle = \langle \mathbf{n}, \mathbf{r} \rangle .$$

C'est une équation cartésienne du plan contenant l'ensemble des vecteurs générés par $T(\alpha, \beta)$.

Exemple 1.1.8 (*équation cartésienne d'un plan ne passant pas par l'origine*)

Soit un plan qui contient le vecteur $\mathbf{r} = \begin{bmatrix} -1 \\ -2 \\ 3 \end{bmatrix}$ dont deux vecteurs tangents à celui-ci sont $\mathbf{v} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix}$ et $\mathbf{w} = \begin{bmatrix} -1 \\ 5 \\ 6 \end{bmatrix}$. Pour trouver une équation cartésienne de ce plan, il faut d'abord calculer un vecteur normal au plan généré par $\mathbf{v}$ et $\mathbf{w}$. On utilise le produit vectoriel

$$\mathbf{n} = \mathbf{v} \wedge \mathbf{w} = \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix} \wedge \begin{bmatrix} -1 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \times 6 - 2 \times 5 \\ (-2 \times 1) - 4 \times 6 \\ 4 \times 5 - (3 \times (-1)) \end{bmatrix} = \begin{bmatrix} 8 \\ -26 \\ 23 \end{bmatrix}.$$

Ainsi, un vecteur $\mathbf{b} = (x, y, z)$ appartenant à ce plan doit être tel que

$$\begin{aligned} \langle \mathbf{n}, \mathbf{b} \rangle &= \langle \mathbf{n}, \mathbf{r} \rangle \\ \iff \begin{bmatrix} 8 & -26 & 23 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} 8 & -26 & 23 \end{bmatrix} \begin{bmatrix} -1 \\ -2 \\ 3 \end{bmatrix} \\ \iff 8x - 26y + 23z &= 113. \end{aligned}$$

Est-il possible de partir d'une équation cartésienne pour obtenir une représentation paramétrique d'un plan dans $\mathbb{R}^3$? C'est effectivement possible en trouvant deux vecteurs linéairement indépendants entre eux et orthogonaux à $\mathbf{n}$.

Exemple 1.1.9 (*d'une équation cartésienne d'un plan dans $\mathbb{R}^3$ à une représentation paramétrique*)

Prenons l'équation cartésienne d'un plan dans $\mathbb{R}^3$

$$2x + 5y + 8z = 8 \ .$$

Pour obtenir une représentation paramétrique, il faut d'abord trouver un vecteur qui appartient à ce plan. Puisque cette équation possède trois inconnues, il suffit d'en fixer deux pour obtenir la troisième. On pose $x = 0$ et $y = 0$ pour obtenir $8z = 8$ ce qui implique $z = 1$. Alors le vecteur $\mathbf{r} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ appartient au plan. Il faut maintenant obtenir deux vecteurs qui sont orthogonaux à $\mathbf{n} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix}$ et qui sont linéairement indépendants entre eux.

Soit un vecteur $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$, alors

$$\begin{aligned} & \langle \mathbf{n}, \mathbf{v} \rangle = 0 \\ \iff & 2v_1 + 5v_2 + 8v_3 = 0 \end{aligned}$$

En posant $v_1 = 1$, alors l'équation devient

$$2 + 5v_2 + 8v_3 = 0$$

et on peut prendre $v_2 = -2$ et $v_3 = 1$ pour obtenir un vecteur $\mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$ orthogonal à $\mathbf{n}$. De plus, on pourrait continuer à chercher un autre vecteur $\mathbf{w}$ orthogonal à $\mathbf{n}$ qui est linéairement indépendant à $\mathbf{v}$. Une astuce rapide est de prendre le produit vectoriel de $\mathbf{v}$ avec $\mathbf{n}$. On a

$$\begin{aligned} \mathbf{w} &= \mathbf{v} \wedge \mathbf{n} \\ &= \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \wedge \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} = \begin{bmatrix} -16-5 \\ 2-8 \\ 5+4 \end{bmatrix} = \begin{bmatrix} -21 \\ -6 \\ 9 \end{bmatrix}. \end{aligned}$$

Ainsi, une représentation paramétrique du plan d'équation cartésienne $2x+5y+8z = 8$ est

$$T(\alpha, \beta) = \mathbf{r} + \alpha\mathbf{v} + \beta\mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} -21 \\ -6 \\ 9 \end{bmatrix} = \begin{bmatrix} \alpha - 21\beta \\ -2\alpha - 6\beta \\ 1 + \alpha + 9\beta \end{bmatrix}, \quad \text{avec } \alpha, \beta \text{ dans } \mathbb{R}.$$

Proposition 1.1.5 (*formule équivalente du produit vectoriel*)

Soit $\theta \in [0, \pi]$, l'angle entre $\mathbf{x}$ et $\mathbf{y}$ dans le plan qui les contient, alors le produit vectoriel de deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ est aussi

$$\mathbf{x} \wedge \mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \sin\theta \mathbf{u}$$

où $\mathbf{u}$ est le vecteur unitaire orthogonal au plan généré par $\mathbf{x}$ et $\mathbf{y}$ tel que $\langle (\mathbf{x} \times \mathbf{y}), \mathbf{u} \rangle >$

0. Attention, ici il faut quand même utiliser la formule de la définition du produit vectoriel pour trouver le sens du vecteur normal unitaire $\mathbf{u}$.

Preuve. Montrons que pour deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^3$ et $\theta \in [0;\pi]$ l'angle entre ces deux vecteurs, alors

$$\|\mathbf{x} \wedge \mathbf{y}\| = \|\mathbf{x}\| \, \|\mathbf{y}\| \sin\theta.$$

En effet, on a

$$\begin{aligned}
\|\mathbf{x} \wedge \mathbf{y}\|^2 &= \langle \mathbf{x} \wedge \mathbf{y}, \mathbf{x} \wedge \mathbf{y} \rangle \\
&= \langle \mathbf{x}, \mathbf{y} \wedge (\mathbf{x} \wedge \mathbf{y}) \rangle, \text{ par la règle de l'échange} \\
&= \langle \mathbf{x}, \langle \mathbf{y}, \mathbf{y} \rangle \mathbf{x} - \langle \mathbf{x}, \mathbf{y} \rangle \mathbf{y} \rangle, \text{ par la formule du double produit sur } \mathbf{y} \wedge (\mathbf{x} \wedge \mathbf{y}) \\
&= \langle \mathbf{y}, \mathbf{y} \rangle \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle \langle \mathbf{x}, \mathbf{y} \rangle, \text{ par distributivité du produit scalaire} \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - \langle \mathbf{x}, \mathbf{y} \rangle^2, \text{ par définition du produit scalaire} \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - (\|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta)^2, \text{ par la proposition 1.1.3} \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 - \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \cos^2\theta \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 (1 - \cos^2\theta) \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2\theta, \text{ par l'identité trigonométrique} \\
\implies \|\mathbf{x} \wedge \mathbf{y}\| &= \sqrt{\|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \sin^2\theta} \\
&= \|\mathbf{x}\| \, \|\mathbf{y}\| \, |\sin\theta| \\
&= \|\mathbf{x}\| \, \|\mathbf{y}\| \sin\theta, \text{ car } \theta \in [0,\pi].
\end{aligned}$$

Ainsi, si $\mathbf{u}$ est le vecteur unitaire orthogonal à $\mathbf{x}$ et $\mathbf{y}$ dans le même sens que $\mathbf{x} \wedge \mathbf{y}$, alors

$$\mathbf{x} \wedge \mathbf{y} = \|\mathbf{x} \wedge \mathbf{y}\| \frac{\mathbf{x} \wedge \mathbf{y}}{\|\mathbf{x} \wedge \mathbf{y}\|} = \|\mathbf{x} \wedge \mathbf{y}\| \, \mathbf{u} = \|\mathbf{x}\| \, \|\mathbf{y}\| \sin\theta \mathbf{u}.$$

□

1.2 Matrices

Cette section est consacrée à la présentation des *matrices*. Elles peuvent être additionnées et multipliées par un scalaire pour donner de nouvelles matrices grâce à des combinaisons linéaires. Ces matrices peuvent aussi être multipliées entre elles. Enfin, un produit scalaire est introduit afin de permettre la définition rigoureuse d'une norme associée à une matrice.

1.2.1 Définitions

Dans la section précédente, un vecteur colonne était un tableau de scalaires à une colonne. Les matrices sont des extensions de ce concept, en considérant plusieurs colonnes. Les éléments du tableau qui définissent la matrice sont aussi appelés des composantes.

Une matrice est donc un tableau de scalaires qui possède m lignes et n colonnes. Le couple (m, n) indique la taille de la matrice que l'on note $m \times n$. Il convient de noter que l'on parle ici de la "taille" d'une matrice et surtout pas de sa dimension. Le terme "dimension" est réservé pour les espaces vectoriels. Voir la section 5.2.4 pour une définition plus précise.

Définition 1.2.1 (*matrice*)

Une matrice $\mathbf{A}$ est un tableau de taille $m \times n$. Ses éléments sont ses *composantes* qui sont des scalaires notés $\mathbf{A}(i, j)$, avec $i \in [\![1; m]\!]$ l'indice de ligne et $j \in [\![1; n]\!]$ l'indice de colonne. La forme générale d'une matrice est donc :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}(1,1) & \mathbf{A}(1,2) & \cdots & \mathbf{A}(1,n) \\ \mathbf{A}(2,1) & \mathbf{A}(2,2) & \cdots & \mathbf{A}(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}(m,1) & \mathbf{A}(m,2) & \cdots & \mathbf{A}(m,n) \end{bmatrix}.$$

Remarque

Utiliser les expressions "la ligne i de $\mathbf{A}$" et "la colonne j de $\mathbf{A}$" revient à mentionner la i-ième ligne de $\mathbf{A}$ et sa j-ième colonne.

Remarque

Les noms attribués aux indices revêtent une importance particulière, dans la mesure où ce sont systématiquement les mêmes qui sont repris, par convention, ce qui facilite la compréhension de la suite. Ansi,

- ▷ Les notations i, j, m, n sont réservées tout au long de l'ouvrage, avec la signifi-

cation suivante :
- ▷ i est l'indice d'une ligne ;
- ▷ j est l'indice d'une colonne ;
- ▷ m est le nombre de lignes ;
- ▷ n est le nombre de colonnes.

▷ Par convention, lorsque l'on accède à une composante, on donne toujours la ligne en premier, et la colonne en second.

Si on dit que deux matrices ont la même taille, cela signifie qu'elles ont le même nombre de lignes et le même nombre de colonnes. Les matrices qui sont d'une certaine taille et dont les scalaires sont de même nature forment également un ensemble usuel de l'algèbre tel que mentionné à la définition 1.2.2. Si toutes les composantes sont réelles alors $\mathbf{A} \in \mathbb{R}^{m\times n}$.

Définition 1.2.2 ($\mathbb{R}^{m\times n}$)
$\mathbb{R}^{m\times n}$ est un ensemble d'éléments, appelés *matrices*, écrits sous forme de m *lignes et* n *colonnes* et à composantes réelles.

Remarque

- ▷ Lorsque les composantes sont des complexes, les matrices appartiennent à $\mathbb{C}^{m\times n}$. On a bien évidement $\mathbb{R}^{m\times n} \subset \mathbb{C}^{m\times n}$.
- ▷ Si $\mathbf{A} \in \mathbb{R}^{m\times n}$, on dit que $\mathbf{A}$ est une *matrice réelle.*
- ▷ Si $\mathbf{A} \in \mathbb{C}^{m\times n}$, on dit que $\mathbf{A}$ est une *matrice complexe.*
- ▷ Une matrice $\mathbf{A} \in \mathbb{C}^{m\times n}$ peut éventuellement être une matrice réelle.

Il est expliqué au chapitre 5 que ces ensembles sont également des espaces vectoriels. Cette connaissance n'est pas nécessaire dans le présent chapitre.

Tout comme avec les vecteurs colonne, certains ensembles peuvent être assimilés : $\mathbb{R}^{m\times 1} = \mathbb{R}^m$ et $\mathbb{R} = \mathbb{R}^1 = \mathbb{R}^{1\times 1}$. (c'est également vrai avec les ensembles complexes similaires). En effet, un vecteur colonne à m composantes peut être vu comme une matrice à m lignes et $n = 1$ colonne.

Tel que déjà mentionné à la suite de la définition 1.1.4, il convient de noter qu'une ligne n'est pas un vecteur mais une matrice (sauf si elle a une seule composante et que c'est un scalaire). Tout ceci est illustré dans l'exemple 1.2.1.

Exemple 1.2.1 ($\mathbb{R}^{m\times 1} = \mathbb{R}^m$ *et* $\mathbb{R} = \mathbb{R}^1 = \mathbb{R}^{1\times 1}$)

- ▷ $\mathbf{x} = (1,2,3) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = [1\ 2\ 3]^\top \neq [1\ 2\ 3]$.
- ▷ $x = 1 = (1) = [1] = [1]^\top$.

La matrice nulle (définition 1.2.3), notée $\mathbf{O}$, est aussi un élément central de l'ensemble des matrices. Sa définition sert surtout à la justification que l'ensemble des matrices est un espace vectoriel (tel que vu à l'exercice 5.3). Il est essentiel de ne pas confondre la matrice nulle $\mathbf{O}$ avec le vecteur nul $\mathbf{0}$: leur taille est différente (sauf, bien sûr, si on considère la matrice nulle à une seule colonne).

Définition 1.2.3 (*matrice nulle*)
La *matrice nulle* est la matrice dont toutes les composantes sont égales au scalaire nul (réel ou complexe). Elle n'est pas forcément carrée et est notée $\mathbf{O}$ ou ($\mathbf{O}_{m,n}$ en cas d'ambiguïté).

Une autre matrice très utile en pratique est la matrice identité. Elle sert surtout pour les produits de matrices et est introduite dans la section 1.2.6.

Une manière courante de représenter une matrice consiste à la considérer comme un ensemble de vecteurs colonne. Cette représentation s'avère utile pour introduire la notion de produit matriciel. Il est ainsi intéressant d'introduire un nouvel opérateur Col(.) à la définition 1.2.4 qui permet d'accéder aux colonnes d'une matrice d'une façon plus aisée.

Définition 1.2.4 (*opérateur* Col(.))
Soit $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n]$ une matrice de taille $m \times n$ dont les colonnes $\mathbf{a}_j$ sont des vecteurs réels ou complexes de taille m, pour $j \in [\![1; n]\!]$. L'opération $\mathrm{Col}(\mathbf{A})$ donne la famille des colonnes de $\mathbf{A}$:

$$\mathrm{Col}(\mathbf{A}) = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)\ .$$

La section v donne d'autres détails de notations.

1.2.2 Transposée de matrices

La transposée d'une matrice, tel qu'indiqué dans la définition 1.2.5, est similaire à la transposée d'un vecteur. Le principe est le même, en considérant les colonnes d'une matrice comme des vecteurs colonne : les colonnes deviennent des lignes et les lignes deviennent des colonnes.

Définition 1.2.5 *(transposée d'une matrice)*
Soit $\mathbf{A}$ une matrice de taille $m \times n$. La transposée de $\mathbf{A}$, notée $\mathbf{A}^\top$, est la matrice de taille $n \times m$ telle que

$$\mathbf{A}^\top(i,j) = \mathbf{A}(j,i)$$

pour $i \in [\![1;m]\!]$ et $j \in [\![1;n]\!]$.

Les propriétés de la transposée sont regroupées dans la proposition 1.2.1. Les notions nécessaires pour comprendre certaines de ces propriétés sont vues plus loin dans l'ouvrage. Les liens vers les sections pertinentes sont donnés dans la preuve de cette proposition.

Proposition 1.2.1 *(propriétés de la transposée)*
Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices.

1. si $\mathbf{A}$ et $\mathbf{B}$ sont de même taille, alors $(\mathbf{A}+\mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$ (la somme des transposées est la transposée de la somme) ;
2. $(k\mathbf{A})^\top = k\mathbf{A}^\top$;
3. $\left(\mathbf{A}^\top\right)^\top = \mathbf{A}$;
4. Si $\mathbf{A}$ est *triangulaire inférieure (supérieure)*, alors $\mathbf{A}^\top$ est triangulaire supérieure (inférieure) ;
5. Si le *produit matriciel* $\mathbf{AB}$ est possible, alors $(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top$;
6. Si $\mathbf{A}$ est *inversible*, alors $\left(\mathbf{A}^\top\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\top$;
7. $\det(\mathbf{A}^\top) = \det(\mathbf{A})$.

Preuve.

1 Les composantes de $\mathbf{A}+\mathbf{B}$ sont $\mathbf{A}(i,j)+\mathbf{B}(i,j)$. Lorsqu'on transpose $\mathbf{A}+\mathbf{B}$ les composantes deviennent $\mathbf{A}(j,i)+\mathbf{B}(j,i)$, ce qui correspond à la somme des composantes de $\mathbf{A}^\top$ et $\mathbf{B}^\top$;

2 Les composantes restent toutes multipliées par k, d'où le résultat ;

3 Comme les lignes et les colonnes sont échangées deux fois, elles reviennent à leur place initiale ;

4 Voir la section 1.2.3 pour la définition des matrices triangulaires ; la preuve découle alors directement de cette dernière et de la définition de la transposée ;

5 Voir le point 1 du théorème 1.2.2 pour la propriété du produit matriciel ;

6 Voir la proposition 1.4.3 pour la transposée de l'inverse ;

7 Voir le théorème 1.3.1 pour le déterminant de la transposée.

□

La notion de transposée constitue le fondement des matrices symétriques qui sont définies à la section 1.2.3. Il convient également de noter que cette définition de la transposée est vraie pour les matrices réelles et complexes. Toutefois, pour faire des manipulations sur les matrices complexes, la plupart du temps cela ne suffit pas. Cette définition de transposée est donc étendue via la notion de transconjuguée (définition 3.6.4).

On note finalement que la section 4.4 du chapitre 4 montre que la transposée est une application linéaire.

1.2.3 Formes de matrices

Les matrices peuvent avoir une forme particulière desquelles découlent des propriétés intéressantes. Par exemple, il sera établi au chapitre 7 que les matrices symétriques ont toujours des valeurs propres réelles.

Formes usuelles

Les formes les plus usuelles des matrices sont les suivantes :

- ▷ *matrice rectangulaire* : une telle matrice a un nombre de lignes différent du nombre de colonnes, c'est-à-dire qu'une matrice $\mathbf{A}$ de taille $m \times n$ est rectangulaire si $m \neq n$;
- ▷ *matrice carrée* : une telle matrice a autant de lignes que de colonnes. On distingue plusieurs types particuliers de ces matrices. Ainsi, une matrice carrée peut être :
  - ▷ *Triangulaire* : cette notion est évoquée à la section 1.2.2. Il existe deux types de matrices triangulaires :
    - ▷ *Triangulaire supérieure* : la matrice est constituée de zéros sous sa diagonale. En d'autres termes, $\mathbf{A}$ est triangulaire supérieure si $\mathbf{A}(i,j) = 0$ pour tous $i > j$. Les autres composantes peuvent aussi être nulles ;
    - ▷ *Triangulaire inférieure* : la matrice est constituée de zéros au-dessus de sa diagonale. En d'autres termes, $\mathbf{A}$ est triangulaire inférieure si $\mathbf{A}(i,j) = 0$ pour tous $i < j$. Les autres composantes peuvent aussi être nulles.

    Grâce à cette définition, il est facile de prouver le point 4 de la proposition 1.2.1. En effet, si $\mathbf{A}(i,j) = 0$ pour tout $i < j$, alors $\mathbf{A}^\top(j,i) = 0$ pour tout $j > i$;
  - ▷ *Diagonale* : une matrice $\mathbf{A}$ est diagonale si $\mathbf{A}(i,j) = 0$ pour tout $i \neq j$. Les seuls éléments non nuls sont sur la diagonale. Il peut également y avoir des 0 sur la diagonale. Une matrice diagonale est donc à la fois triangulaire supérieure et inférieure ;[4]
  - ▷ *Unipotente* : une matrice $\mathbf{A}$ est unipotente si on peut trouver un entier k tel que $(\mathbf{I} - \mathbf{A})^k = \mathbf{O}$;

4. la notation $\mathrm{Diag}(a,b) = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ simplifie l'écriture d'une matrice diagonale.

- ▷ *Symétrique* : une matrice $\mathbf{A}$ est symétrique si elle est égale à sa transposée, c'est-à-dire $\mathbf{A} = \mathbf{A}^\top$. Ainsi, toute matrice diagonale est symétrique car les termes échangés lors de la transposition sont des zéros ;
- ▷ *Antisymétrique* : une matrice $\mathbf{A}$ est antisymétrique si $\mathbf{A}^\top = -\mathbf{A}$. Tel qu'indiqué à l'exercice 1.29, il s'ensuit que les termes sur la diagonale de telles matrices sont nuls ;
- ▷ *Hermitienne* : Voir définition 3.6.7 ;
- ▷ *Orthogonale* : Voir définition 5.3.5.

Il se peut, bien sûr, qu'une matrice soit de plusieurs des formes décrites ci-dessus. Les exemples 1.2.2 et 1.2.3 illustrent cela.

Exemple 1.2.2

$\mathbf{O}$ est à la fois triangulaire supérieure, triangulaire inférieure, diagonale, unipotente, symétrique, antisymétrique et hermitienne.

Exemple 1.2.3

Toute matrice diagonale est triangulaire (inférieure et supérieure).

D'autres formes sont étudiées dans d'autres sections, par exemple les matrices échelonnées et échelonnées réduites (voir pour cela les définitions 2.1.4 et 2.1.5).

Matrices par blocs

Pour démontrer certaines propriétés ou pour faciliter certains calculs, on peut "découper" une matrice en blocs. Ce découpage permet également de paralléliser certaines tâches lors de calculs de grande taille. Dans ce qui suit, les matrices peuvent être réelles ou complexes.

Définition 1.2.6 *(sous-matrice)*

Une *sous-matrice* d'une matrice $\mathbf{A}$ de taille $m \times n$ est une matrice de taille $m' \times n'$ avec $m' \leq m$ et $n' \leq n$ dont on a gardé seulement certaines lignes ou colonnes de $\mathbf{A}$.

Une matrice définie par blocs est une matrice que l'on définit en fonction de certaines sous-matrices qu'on appelle "blocs" de $\mathbf{A}$. Chacun des blocs d'une telle matrice $\mathbf{A}$ est une sous-matrice de $\mathbf{A}$ pour laquelle on a gardé des lignes et des colonnes consécutives. Les tailles de ces blocs doivent être cohérentes avec la taille de $\mathbf{A}$, i.e., le nombre total de colonnes des sous-matrices, doit être égal au nombre total de colonnes de la matrice $\mathbf{A}$, et il en est de même pour les lignes.

Définition 1.2.7 (*matrice blocs*)

Soit $\mathbf{A}$ une matrice de taille $m \times n$. On peut définir $\mathbf{A}$ grâce à des sous-matrices appelées *blocs* de telle sorte que

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \cdots & \mathbf{A}_{1q} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \cdots & \mathbf{A}_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_{p1} & \mathbf{A}_{p2} & \cdots & \mathbf{A}_{pq} \end{bmatrix}$$

où $\mathbf{A}_{ij}$ est une sous-matrice de $\mathbf{A}$ pour tout $i \in [\![1;p]\!]$ et pour tout $j \in [\![1;q]\!]$, p étant le nombre de divisions des m lignes de A, et q le nombre de divisions des n colonnes de $\mathbf{A}$.

Les blocs peuvent, par exemple, être les colonnes d'une matrice, dans quel cas on a $\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_n]$ avec $\mathbf{a}_j \in \mathbb{R}^m$ ou $\mathbb{C}^m$ pour $j \in [\![1;n]\!]$ (pour rappel, l'indice j est réservé pour les colonnes).

Les blocs peuvent aussi être les lignes d'une matrice, dans quel cas, pour une matrice $\mathbf{B}$ dont les lignes sont $\mathbf{b}_1^\top, \mathbf{b}_2^\top, \dots, \mathbf{b}_m^\top$, on a

$$\mathbf{B} = \begin{bmatrix} \mathbf{b}_1^\top \\ \mathbf{b}_2^\top \\ \vdots \\ \mathbf{b}_m^\top \end{bmatrix}$$

avec $\mathbf{b}_i \in \mathbb{R}^n$ ou $\mathbb{C}^n$ pour $i \in [\![1;m]\!]$ (pour rappel, l'indice i est réservé pour les lignes). Il convient de noter l'utilisation du symbole de la transposée pour les lignes de $\mathbf{B}$, car chaque ligne de $\mathbf{B}$ est un vecteur colonne transposé.

On peut éventuellement délimiter les blocs par des séparations virtuelles pour des questions de visualisation, comme cela est fait dans l'exemple 1.2.4.

Exemple 1.2.4

- ▷ Lorsqu'une matrice $\mathbf{A} \in R^{4\times 4}$ est décomposée en blocs de taille 2×2, on écrira :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array}\right].$$

  où $\mathbf{A}_{11} \in \mathbb{R}^{2\times 2}$, $\mathbf{A}_{12} \in \mathbb{R}^{2\times 2}$, $\mathbf{A}_{21} \in \mathbb{R}^{2\times 2}$ et $\mathbf{A}_{22} \in \mathbb{R}^{2\times 2}$;
- ▷ Considérons le découpage ci-dessous de la matrice $\mathbf{B}$ en 4 blocs

$$\mathbf{B} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & 8 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array}\right].$$

> On a $\mathbf{B}_{11} = \begin{bmatrix}1 & 2 & 3\end{bmatrix} \in \mathbb{R}^{1\times 3}$, $\mathbf{B}_{12} = \begin{bmatrix}4\end{bmatrix} \in \mathbb{R}^{1\times 1}$, $\mathbf{B}_{21} = \begin{bmatrix}5 & 6 & 7\end{bmatrix} \in \mathbb{R}^{1\times 3}$ et $\mathbf{B}_{22} = \begin{bmatrix}8\end{bmatrix} \in \mathbb{R}^{1\times 1}$.

Certains découpages par blocs sont intéressants de par leur structure. C'est le cas, par exemple, des *matrices diagonales par blocs* et des matrices *triangulaires par blocs* décrites dans les définitions suivantes. L'exemple 1.2.5 illustre ces définitions alors que l'exercice 1.20 en démontre l'utilité.

Définition 1.2.8 (*matrice diagonale par blocs*)

Soit $\mathbf{A}$ une matrice définie par blocs selon la définition 1.2.7. Elle est *diagonale par blocs* si $p = q$, $\mathbf{A}_{ij} = 0$ pour tout $i \neq j$, et $\mathbf{A}_{ii}$ est quelconque (pas nécessairement une matrice diagonale classique) pour tout $i \in [\![1; p]\!]$.

Définition 1.2.9 (*matrice triangulaire par blocs*)

Soit $\mathbf{A}$ une matrice définie par blocs selon la définition 1.2.7. Elle est *triangulaire par blocs* si $p = q$ et

- ▷ $\mathbf{A}_{ij} = 0$ pour tout $i > j$ (triangulaire supérieure) ;
- ▷ $\mathbf{A}_{ij} = 0$ pour tout $i < j$ (triangulaire inférieure).

Exemple 1.2.5

- ▷ $\mathbf{A} = \begin{bmatrix}1 & 2 & 3 & 0 & 0 & 0\\ 0 & 0 & 0 & 8 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 & 1 & 1\end{bmatrix} = \left[\begin{array}{ccc|c|cc} 1 & 2 & 3 & 0 & 0 & 0\\ \hline 0 & 0 & 0 & 8 & 0 & 0\\ \hline 0 & 0 & 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0 & 1 & 1\end{array}\right]$ est diagonale par blocs.
- ▷ $\mathbf{B} = \begin{bmatrix}1 & 2 & 3 & 0\\ 5 & 6 & 7 & 8\end{bmatrix} = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 0\\ \hline 5 & 6 & 7 & 8\end{array}\right]$ est triangulaire inférieure par blocs.

Les matrices définies par blocs sont aussi utilisées à la section 8.1.4 pour la résolution de systèmes d'équations linéaires.

1.2.4 Opérations matricielles simples

Comme pour les vecteurs, il est possible de définir deux opérations matricielles, à savoir le produit par un scalaire et l'addition. Ces deux opérations sont utiles pour démontrer la structure d'espace vectoriel de $\mathbb{R}^{m\times n}$ (voir par exemple l'exercice 5.3). Dans les deux définitions qui suivent, les matrices peuvent être réelles ou complexes.

Définition 1.2.10 (*addition matricielle*)

Étant données deux matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m\times n}$ (ou $\mathbb{C}^{m\times n}$), leur *addition matricielle* est la matrice $\mathbf{C}$ notée

$$\mathbf{C} = \mathbf{A} + \mathbf{B}$$

et ses composantes sont définies par

$$\mathbf{C}(i,j) = \mathbf{A}(i,j) + \mathbf{B}(i,j).$$

Remarque

$\mathbf{A}$ et $\mathbf{B}$ doivent être de même taille pour pouvoir être additionnées !

Définition 1.2.11 (*produit scalaire-matrice*)

Étant donnés une matrice $\mathbf{A} \in \mathbb{R}^{m\times n}$ (ou $\mathbb{C}^{m\times n}$) et un scalaire k, le *produit scalaire-matrice* est la matrice $\mathbf{C}$ notée

$$\mathbf{C} = k\mathbf{A}$$

et ses composantes sont celles de $\mathbf{A}$ multipliées par le scalaire k, c'est-à-dire

$$\mathbf{C}(i,j) = k\mathbf{A}(i,j).$$

Remarque

Il est possible de combiner ces différentes opérations pour faire des combinaisons linéaires de matrices.

La proposition 1.2.2 donne quelques propriétés de base de ces deux opérations matricielles.

Proposition 1.2.2 (*propriétés de l'addition matricielle et du produit scalaire-matrice*)

1. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$;
2. $k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}$;
3. $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$.

Preuve. Pour prouver ces trois propriétés, il suffit d'étudier les composantes de chacune de ces matrices, en tenant compte des définitions précédentes. Les propriétés découlent alors directement du fait qu'elles sont vraies pour des scalaires alors que les composantes de ces matrices sont toutes des scalaires. $\square$

Les exercices 1.9 et 1.11 permettent de se familiariser avec ces opérations et la notion de transposée.

Chaque matrice carrée (réelle ou complexe) possède une trace qui, telle que définie ci-dessous, est la somme de ses éléments diagonaux.

Définition 1.2.12 (*trace d'une matrice*)

Si $\mathbf{A}$ est une matrice carrée $n \times n$ réelle ou complexe, alors sa *trace* est définie comme la somme de ses éléments diagonaux. Elle est notée $\mathrm{tr}(\mathbf{A})$, ce qui donne

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} \mathbf{A}(i,i) \ .$$

La proposition 1.2.3 regroupe les propriétés importantes portant sur la trace des matrices.

Proposition 1.2.3 (*propriétés de la trace*)

Soient $\mathbf{A}, \mathbf{B}$ deux matrices réelles ou complexes et soit k un scalaire. On a

1. $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$, où $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices carrées ;
2. $\mathrm{tr}(k\mathbf{A}) = k\,\mathrm{tr}(\mathbf{A})$, où $\mathbf{A}$ est une matrice carrée ;
3. $\mathrm{tr}(\mathbf{A}^\top) = \mathrm{tr}(\mathbf{A})$, où $\mathbf{A}$ est une matrice carrée ;
4. $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, où $\mathbf{AB}$ et $\mathbf{BA}$ sont des matrices carrées, alors que $\mathbf{A}$ et $\mathbf{B}$ ne le sont pas forcément ;

Preuve.

1 En supposant que $\mathbf{A}$ et $\mathbf{B}$ sont de taille $n \times n$, étant donné que les coefficients diagonaux de $\mathbf{A}+\mathbf{B}$ sont $\mathbf{A}(i,i)+\mathbf{B}(i,i)$, il s'ensuit que

$$\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \sum_{i=1}^{n}(\mathbf{A}(i,i)+\mathbf{B}(i,i)) = \sum_{i=1}^{n}\mathbf{A}(i,i) + \sum_{i=1}^{n}\mathbf{B}(i,i) = tr(\mathbf{A}) + \mathrm{tr}(\mathbf{B});$$

2 En supposant que $\mathbf{A}$ est de taille $n \times n$, étant donné que les coefficients diagonaux de $k\mathbf{A}$ sont $k\mathbf{A}(i,i)$, il s'ensuit que

$$\mathrm{tr}(k\mathbf{A}) = \sum_{i=1}^{n}(k\mathbf{A}(i,i)) = k\sum_{i=1}^{n}\mathbf{A}(i,i) = \mathrm{tr}(\mathbf{A});$$

3 Cette propriété découle du fait que les coefficients diagonaux ne changent pas lors d'une transposition ;

4 La démonstration de ce point est faite à la section 1.2.6 (voir la proposition 1.2.6).

□

Il convient de noter qu'on n'a pas nécessairement $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{A})\,\mathrm{tr}(\mathbf{B})$, tel qu'illustré dans l'exemple 1.2.6. On peut également remarquer que $\mathrm{tr}(\mathbf{AB})$ est défini même si la trace de $\mathbf{A}$ et la trace de $\mathbf{B}$ n'existent pas, lorsque $\mathbf{A}$ et $\mathbf{B}$ ne sont pas des matrices carrées. La section 1.2.6 traite de ces problématiques de taille.

Exemple 1.2.6 *(contre-exemple pour la trace du produit de matrices)*

En général, on a $\mathrm{tr}(\mathbf{AB}) \neq \mathrm{tr}(\mathbf{A})\,\mathrm{tr}(\mathbf{B})$. Par exemple, si $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$, alors $\mathbf{AB} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathrm{tr}(\mathbf{A}) = \mathrm{tr}(\mathbf{B}) = 1$, $\mathrm{tr}(\mathbf{AB}) = 0$, et donc $0 \neq 1 \times 1$.

1.2.5 Produit matrice-vecteur

Avant de multiplier des matrice entre-elles, il est essentiel de maîtriser la notion de produit matrice-vecteur de la définition 1.2.13. Elle est basée sur une opération de type produit ligne-colonne (définition 1.1.11). Le produit matrice-vecteur est illustré à la figure 1.6.

Définition 1.2.13 *(produit matrice-vecteur)*

Soient

- ▷ $\mathbf{A}$ une matrice réelle ou complexe de taille $m \times n$ dont les lignes sont $\mathbf{a}_1^\top, \mathbf{a}_2^\top, \ldots, \mathbf{a}_m^\top$;
- ▷ $\mathbf{x} \in \mathbb{R}^n$ ou $\mathbb{C}^n$ un vecteur à n composantes réelles ou complexes.

Le *produit matrice-vecteur* qu'on prononce “$\mathbf{A}$ fois $\mathbf{x}$” est le vecteur $\mathbf{Ax} \in \mathbb{R}^m$ ou $\mathbb{C}^m$ et défini par :

$$(\mathbf{Ax})(i) = \mathbf{a}_i^\top \mathbf{x} \text{ pour } i \in [\![1; m]\!].$$

Remarque

- ▷ Il est important d'observer que le produit matrice-vecteur $\mathbf{Ax}$ n'est possible que si le vecteur $\mathbf{x}$ a autant de composantes que le nombre de colonnes de la matrice $\mathbf{A}$. L'exercice 1.9 s'attarde sur ce point ;
- ▷ Si $\mathbf{A} \in \mathbb{R}^{m\times n}$, alors $(\mathbf{Ax})(i) = \mathbf{a}_i^\top \mathbf{x} = \langle \mathbf{a}_i, \mathbf{x} \rangle$ pour $i \in [\![1; m]\!]$. C'est un produit scalaire. Pour les complexes, ce n'est pas le cas (voir la définition 3.6.5).

Le produit matrice-vecteur constitue le fondement des systèmes d'équations linéaires,

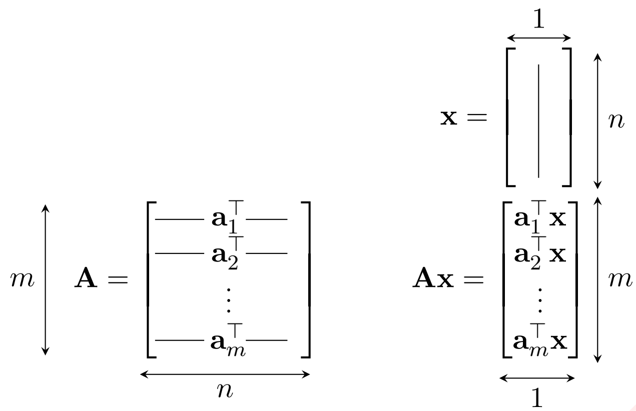

FIGURE 1.6 – Produit matrice-vecteur $\mathbf{Ax}$ de la matrice $\mathbf{A}$ et du vecteur $\mathbf{x}$ vu comme des produits ligne-colonne (définitions 1.1.11 et 1.2.13).

dans lesquels on cherche à déterminer un certain $\mathbf{x}$ tel que $\mathbf{Ax} = \mathbf{b}$ avec $\mathbf{A}$ une matrice et $\mathbf{b}$ un vecteur colonne. La définition formelle se trouve à la définition 8.1.2 et les méthodes de résolution sont vues dans le chapitre 8.

On montre à la définition 4.4.1 que le produit matrice-vecteur est la le résultat d'une *application linéaire*.

Le théorème 1.2.1 qui suit permet d'envisager le produit matrice-vecteur sous un angle différent : $\mathbf{Ax}$ est une combinaison linéaire des colonnes de $\mathbf{A}$. Ce point de vue est illustré à la figure 1.7.

Théorème 1.2.1 (*le produit matrice-vecteur est une combinaison de colonnes*)

Soient

- ▷ $\mathbf{A} = [\mathbf{c}_1 \ \mathbf{c}_2 \ \cdots \ \mathbf{c}_n]$ une matrice réelle ou complexe de taille $m \times n$ où chaque $\mathbf{c}_j$ est donc un vecteur colonne réel ou complexe à m composantes ;
- ▷ $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ un vecteur réel ou complexe à n composantes.

On a : $\mathbf{Ax} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \ldots + x_n\mathbf{c}_n$.

Preuve. Il suffit de vérifier que les composantes de $\mathbf{Ax}$ sont les mêmes que celles de $x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$. Par définition du produit matrice–vecteur (définition 1.2.13), la i-ième composante de $\mathbf{Ax}$ est

$$(\mathbf{Ax})_i = \sum_{j=1}^{n} \mathbf{A}(i,j) x_j \ .$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \\ | & | & & | \end{bmatrix} \qquad \mathbf{A}\mathbf{x} = x_1 \begin{bmatrix} | \\ \mathbf{c}_1 \\ | \end{bmatrix} + x_2 \begin{bmatrix} | \\ \mathbf{c}_2 \\ | \end{bmatrix} + \cdots + x_n \begin{bmatrix} | \\ \mathbf{c}_n \\ | \end{bmatrix}$$

FIGURE 1.7 – Produit matrice-vecteur $\mathbf{Ax}$ vu comme une combinaison des colonnes de $\mathbf{A}$. Lien avec le théorème 1.2.1.

D'autre part, la i-ième composante de la combinaison linéaire $x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n$ est

$$(x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n)_i = \sum_{j=1}^{n} x_j\,(\mathbf{c}_j)_i = \sum_{j=1}^{n} x_j\,\mathbf{A}(i,j)\ .$$

Les deux sommes étant égales pour tout $i \in [\![1;m]\!]$, il s'ensuit que $\mathbf{Ax} = \sum_{j=1}^{n} x_j\mathbf{c}_n$. □

En fait, le théorème 1.2.1 est central pour la compréhension de la majeure partie de cet ouvrage, ce qui justifie son statut de théorème plutôt que de simple proposition.

La définition 1.2.13 et le théorème 1.2.1 offrent donc deux points de vue du produit matrice-vecteur qui sont illustrés dans l'exemple 1.2.7.

Exemple 1.2.7 (*deux points de vue du produit matrice-vecteur*)

Soient $\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 1 & 3 \\ 3 & 0 & -2 & 0 \end{bmatrix}$ et $\mathbf{x} = (1,2,2,-1)$.

▷ Point de vue de la définition 1.2.13 :

$\mathbf{Ax} = \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{x}\rangle \\ \langle \mathbf{a}_2, \mathbf{x}\rangle \\ \langle \mathbf{a}_3, \mathbf{x}\rangle \end{bmatrix}$ avec $\mathbf{a}_1 = (1,0,-1,2)$, $\mathbf{a}_2 = (0,-1,1,3)$ et $\mathbf{a}_3 = (3,0,-2,0)$

d'où : $\mathbf{Ax} = \begin{bmatrix} 1\times 1 + 0\times 2 + (-1)\times 2 + 2\times(-1) \\ 0\times 1 + (-1)\times 2 + 1\times 2 + 3\times(-1) \\ 3\times 1 + 0\times 2 + (-2)\times 2 + 0\times(-1) \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \\ -1 \end{bmatrix}$.

▷ Point de vue du théorème 1.2.1 :

$$\mathbf{Ax} = 1\begin{bmatrix}1\\0\\3\end{bmatrix} + 2\begin{bmatrix}0\\-1\\0\end{bmatrix} + 2\begin{bmatrix}-1\\1\\-2\end{bmatrix} - 1\begin{bmatrix}2\\3\\0\end{bmatrix} = \begin{bmatrix}-3\\-3\\-1\end{bmatrix}.$$

Tel qu'énoncé dans la proposition 1.2.4, le cas particulier du produit d'une matrice $\mathbf{A}$ par une colonne de la matrice identité (définition 1.1.3) donne une colonne de la matrice $\mathbf{A}$.

Proposition 1.2.4 *(produit matrice-colonne de l'identité)*

Si $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n] \in \mathbb{R}^{m\times n}$ ou $\mathbb{C}^{m\times n}$ alors $\mathbf{Ae}_j = \mathbf{a}_j$ pour $j \in [\![1; n]\!]$.

Preuve. Une conséquence directe du théorème 1.2.1 et de la définition 1.1.3 est que

$$\mathbf{Ae}_j = \mathbf{e}_j(1)\mathbf{a}_1 + \mathbf{e}_j(2)\mathbf{a}_2 + \ldots + \mathbf{e}_j(n)\mathbf{a}_n = \mathbf{e}_j(j)\mathbf{a}_j = \mathbf{a}_j\ .$$

□

1.2.6 Produit matriciel

Maintenant que le produit matrice-vecteur est défini, on peut introduire le produit $\mathbf{AB}$ [5] de deux matrices $\mathbf{A}$ et $\mathbf{B}$.

Définition 1.2.14 *(produit matriciel)*

Soient

- ▷ $\mathbf{A}$ une matrice réelle ou complexe de taille $m \times p$ dont les lignes sont $\mathbf{a}_1^\top, \mathbf{a}_2^\top, \ldots, \mathbf{a}_m^\top$;
- ▷ $\mathbf{B}$ une matrice réelle ou complexe de taille $p \times n$ dont les colonnes sont $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$.

Le *produit matriciel* $\mathbf{AB}$ qu'on prononce "$\mathbf{A}$ fois $\mathbf{B}$" est une matrice de taille $m \times n$ définie par

$$\mathbf{AB}(i,j) = \mathbf{a}_i^\top \mathbf{b}_j \text{ pour } i \in [\![1; m]\!] \text{ et } j \in [\![1; n]\!].$$

Remarque

- ▷ Il est important de noter que le produit matriciel $\mathbf{AB}$ n'est possible que si le nombre de colonnes de $\mathbf{A}$ est égale au nombre de lignes de $\mathbf{B}$. Les exercices 1.9 et 1.11 traitent de cette problématique.

5. On n'utilise pas l'opérateur "$\times$" qu'on réserve pour décrire la taille des matrices comme "$m \times n$" et pour multiplier des valeurs numériques, comme "$2 \times 3 = 6$".

▷ Si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices réelles, les composantes du produit $\mathbf{AB}$ sont obtenues pas produit scalaire puisque $\mathbf{AB}(i,j) = \mathbf{a}_i^\top \mathbf{b}_j = \langle \mathbf{a}_i, \mathbf{b}_j \rangle$ pour $i \in [\![1;m]\!]$ et $j \in [\![1;n]\!]$. Ce n'est pas le cas pour les matrices complexes (voir la définition 3.6.5).

Les notions de *produit intérieur* et de *produit extérieur*, liése au produit matriciel, sont explicitées dans la remarque suivante.

Remarque

Étant donnés deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^n$:

▷ Le produit scalaire $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$ est le produit matriciel d'une matrice de taille $1 \times n$ (une ligne) par une matrice de taille $n \times 1$ (une colonne), et le résultat est donc une matrice de taille 1×1 qu'on assimile à un scalaire. Cette opération est également appelée *produit intérieur*.

▷ Le produit matriciel $\mathbf{x}\mathbf{y}^\top$, est le produit d'une matrice de taille $n \times 1$ (une colonne) par une matrice de taille $1 \times n$ (une ligne), et le résultat est donc une matrice de taille $n \times n$. Cette opération est également appelée *produit extérieur*. Les colonnes de $\mathbf{x}\mathbf{y}^\top$ sont toutes des multiples de $\mathbf{x}$, ce qui implique que $\mathbf{x}\mathbf{y}^\top$ est de *de rang 1* (voir section 6.1.5). Les matrices de rang 1 sont formellement introduites à la définition 6.1.6.

On peut étendre le produit à plus de deux matrices, tel qu'illustré dans l'exemple 1.2.8 où il est question de multiplier trois matrices entre elles.

Exemple 1.2.8 (*produit de plus de deux matrices*)

Soient $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ et $\mathbf{C} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 0 & 1 \end{bmatrix}$. On a

$$\mathbf{ABC} = (\mathbf{AB})\mathbf{C} = \begin{bmatrix} 1 & 2 & 4 \\ 5 & 4 & 5 \end{bmatrix} \mathbf{C} = \begin{bmatrix} 4 & 3 \\ 14 & 6 \end{bmatrix}.$$

Lorsque l'on multiplie une matrice par elle-même plusieurs fois, on parle de puissance de matrice. On ne peut pas diviser $\mathbf{A}$ par une matrice $\mathbf{B}$, mais on peut parfois multiplier $\mathbf{A}$ par l'inverse de $\mathbf{B}$, tel qu'indiqué à la section 1.4. Ces notions de puissance et d'inverse d'une matrice sont introduites dans la définition 1.2.15, et une illustration est donnée à l'exemple 1.2.9 avec une matrice carrée de taille 2×2.

Définition 1.2.15 (*puissances de matrices*)

Étant donnés une matrice carrée $\mathbf{A}$ et un entier $p \in \mathbb{N} \cup \{-1\}$, la matrice $\mathbf{A}^p$ qu'on prononce "$\mathbf{A}$ à la puissance p" est la matrice définie comme suit :

- ▷ Si $p = -1$, $\mathbf{A}^{-1}$ est l'inverse de $\mathbf{A}$, si elle existe (voir section 1.4).
- ▷ Si $p = 0$, par convention, $\mathbf{A}^0 = \mathbf{I}$, la *matrice identité* (voir définition 1.2.16).
- ▷ Si $p \geq 1$, $\mathbf{A}^p = \underbrace{\mathbf{A}\mathbf{A}\cdots\mathbf{A}}_{p \text{ termes}}$.

Exemple 1.2.9 (*produits de matrices*)

Avec $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, on a

$\mathbf{A}^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, $\mathbf{A}^2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ et $\mathbf{A}^3 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$.

Le produit matriciel constitue une notion essentielle de cet ouvrage. Il est donc important de maîtriser les quatre façons de le représenter, tel qu'indiqué à la proposition 1.2.5.

Proposition 1.2.5 (*les quatre points de vue du produit matriciel*)

Soient

- ▷ $\mathbf{A}$ une matrice de taille $m \times p$ dont les lignes sont $\mathbf{a}_1^\top, \mathbf{a}_2^\top, \ldots, \mathbf{a}_m^\top$ et les colonnes sont $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p$;
- ▷ $\mathbf{B}$ une matrice de taille $p \times n$ dont les colonnes sont $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ et les lignes sont $\mathbf{d}_1^\top, \mathbf{d}_2^\top, \ldots, \mathbf{d}_p^\top$.

Le produit $\mathbf{AB}$ est donc une matrice $m \times n$ qui peut être interprété des quatre manières suivantes :

1. *Produits intérieurs.* La composante $\mathbf{AB}(i,j)$ est le produit intérieur de la i-ième ligne de $\mathbf{A}$ avec la j-ième colonne de $\mathbf{B}$:

$$\mathbf{AB}(i,j) = \mathbf{a}_i^\top \mathbf{b}_j \text{ pour } i \in [\![1;m]\!] \text{ et } j \in [\![1;n]\!] \ .$$

2. *Matrice fois colonnes.* Chaque colonne de $\mathbf{AB}$ est le produit de $\mathbf{A}$ par une colonne de $\mathbf{B}$, c'est-à dire que la j-ième colonne de $\mathbf{AB}$ est $\mathbf{Ab}_j$:

$$\mathbf{AB} = [\mathbf{Ab}_1 \ \mathbf{Ab}_2 \ \cdots \ \mathbf{Ab}_n].$$

3. *Lignes fois matrice.* Chaque ligne de $\mathbf{AB}$ est le produit d'une ligne de $\mathbf{A}$ par $\mathbf{B}$, c'est-à-dire que la i-ième ligne de $\mathbf{AB}$ est $\mathbf{a}_i^\top \mathbf{B}$:

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^\top \mathbf{B} \\ \mathbf{a}_2^\top \mathbf{B} \\ \vdots \\ \mathbf{a}_m^\top \mathbf{B} \end{bmatrix}$$

4. *Produits extérieurs.* Le produit $\mathbf{AB}$ est une somme de p matrices $m \times n$, chaque terme de la somme étant un produit extérieur :

$$\mathbf{AB} = \mathbf{c}_1\mathbf{d}_1^\top + \mathbf{c}_2\mathbf{d}_2^\top + \ldots + \mathbf{c}_p\mathbf{d}_p^\top = \sum_{k=1}^{p} \mathbf{c}_k\mathbf{d}_k^\top .$$

Preuve.

1 C'est la définition 1.2.14.

2 Les composantes de la j-ième colonne de $\mathbf{AB}$ sont $\mathbf{AB}(i,j) = \mathbf{a}_i^\top \mathbf{b}_j$ avec $i \in [\![1;m]\!]$. Il découle de la définition 1.2.13 que la j-ième colonne de $\mathbf{AB}$ est $\mathbf{Ab}_j$.

3 Les composantes de la i-ième ligne de $\mathbf{AB}$ sont $\mathbf{AB}(i,j) = \mathbf{a}_i^\top \mathbf{b}_j$ avec $j \in [\![1;n]\!]$. La i-ième ligne de $\mathbf{AB}$ est donc la matrice $[\mathbf{a}_i^\top \mathbf{b}_1 \ \mathbf{a}_i^\top \mathbf{b}_2 \ \cdots \ \mathbf{a}_i^\top \mathbf{b}_n]$ de taille $1 \times n$, c'est-à-dire $\mathbf{a}_i^\top \mathbf{B}$.

4 Il suffit de démontrer l'égalité pour chacune des composantes. La i-ième composante de la j-ième colonne du produit extérieur $\mathbf{c}_k\mathbf{d}_k^\top$ est le scalaire $\mathbf{c}_k(i)\mathbf{d}_k(j) = \mathbf{A}(i,k)\mathbf{B}(k,j)$. La i-ième composante de la j-ième colonne de $\sum_{k=1}^{p} \mathbf{c}_k\mathbf{d}_k^\top$ est donc $\sum_{k=1}^{p} \mathbf{A}(i,k)\mathbf{B}(k,j) = \mathbf{AB}(i,j)$.

□

Les quatre points de vue de la proposition 1.2.5 sont illustrés dans les figures 1.8, 1.9, 1.10 et 1.11 ainsi que dans l'exemple 1.2.10 où $\mathbf{A}$ et $\mathbf{B}$ sont de taille 2×2. L'exercice 1.10 permet de s'entraîner au calcul de ce produit matriciel.

Remarque

- ▷ Le point de vue 1 est le plus pratique pour calculer un produit matriciel à la main. Il généralise aussi la définition 1.2.13 ;
- ▷ Les points de vue 2 et 3 sont essentiels pour de nombreuses preuves ;
- ▷ Le point de vue 2 permet de réaliser que chaque colonne de $\mathbf{AB}$ est une combinaison linéaire particulière des colonnes de $\mathbf{A}$;

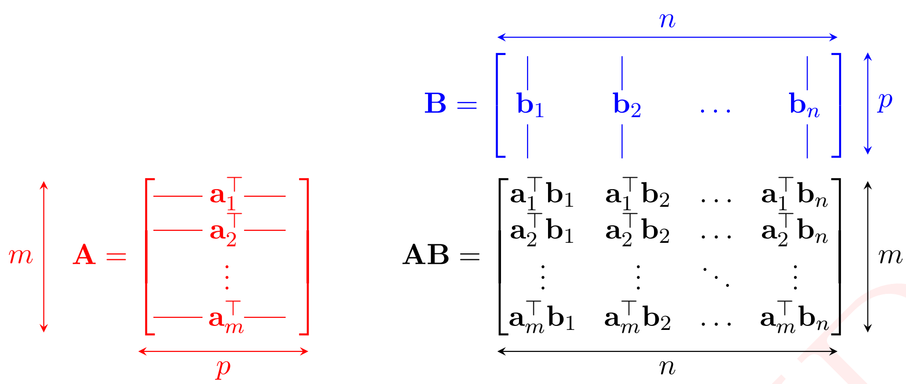

Figure 1.8 – Produits intérieurs. Point 1 de la proposition 1.2.5.

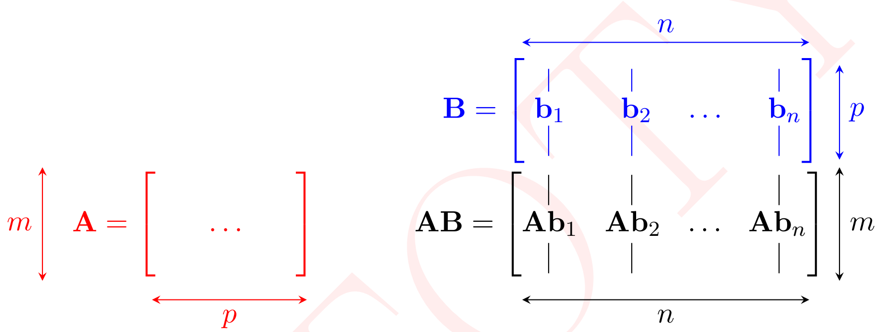

Figure 1.9 – Matrice fois colonnes. Point 2 de la proposition 1.2.5.

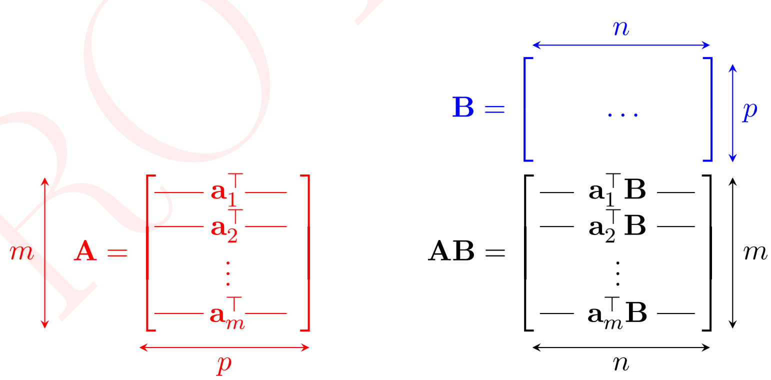

Figure 1.10 – Lignes fois matrice. Point 3 de la proposition 1.2.5.

$$\mathbf{B} = \begin{bmatrix} — \mathbf{d}_1^\top — \\ — \mathbf{d}_2^\top — \\ \vdots \\ — \mathbf{d}_p^\top — \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_p \\ | & | & & | \end{bmatrix} \qquad \mathbf{AB} = \begin{bmatrix} \\ \mathbf{c}_1\mathbf{d}_1^\top \\ \\ \end{bmatrix} + \begin{bmatrix} \\ \mathbf{c}_2\mathbf{d}_2^\top \\ \\ \end{bmatrix} + \dots + \begin{bmatrix} \\ \mathbf{c}_p\mathbf{d}_p^\top \\ \\ \end{bmatrix}$$

FIGURE 1.11 – Produits extérieurs. Point 4 de la proposition 1.2.5.

- ▷ Le point de vue 3 permet de réaliser que chaque ligne de $\mathbf{AB}$ est une combinaison linéaire particulière de lignes de $\mathbf{B}$;
- ▷ Le point de vue 4 peut paraître moins intuitif, mais il permet de mieux comprendre certaines décompositions vues dans des chapitres ultérieurs. Il étend le point de vue essentiel présenté dans le théorème 1.2.1.

Exemple 1.2.10 (*quatre points de vue du produit matriciel*)

Soients $\mathbf{A} = \begin{bmatrix} 1 & 4 \\ 1 & 5 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 3 & 2 \\ 1 & 0 \end{bmatrix}$. On a donc :

$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $\mathbf{a}_2 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\mathbf{c}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{c}_2 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\mathbf{d}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ et $\mathbf{d}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

- ▷ $\mathbf{AB}$ avec le point de vue 1 : $\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^\top\mathbf{b}_1 & \mathbf{a}_1^\top\mathbf{b}_2 \\ \mathbf{a}_2^\top\mathbf{b}_1 & \mathbf{a}_2^\top\mathbf{b}_2 \end{bmatrix}$. On a $\mathbf{a}_1^\top\mathbf{b}_1 = 1\times3+4\times1 = 7$, $\mathbf{a}_1^\top\mathbf{b}_2 = 1\times2+4\times0 = 2$, $\mathbf{a}_2^\top\mathbf{b}_1 = 1\times3+5\times1 = 8$ et $\mathbf{a}_2^\top\mathbf{b}_2 = 1\times2+5\times0 = 2$. D'où : $\mathbf{AB} = \begin{bmatrix} 7 & 2 \\ 8 & 2 \end{bmatrix}$.
- ▷ $\mathbf{AB}$ avec le point de vue 2 : $\mathbf{AB} = \begin{bmatrix} \mathbf{Ab}_1 & \mathbf{Ab}_2 \end{bmatrix}$. On a $\mathbf{Ab}_1 = \begin{bmatrix} \mathbf{a}_1^\top\mathbf{b}_1 \\ \mathbf{a}_2^\top\mathbf{b}_1 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$ et $\mathbf{Ab}_2 = \begin{bmatrix} \mathbf{a}_1^\top\mathbf{b}_2 \\ \mathbf{a}_2^\top\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$. D'où : $\mathbf{AB} = \begin{bmatrix} 7 & 2 \\ 8 & 2 \end{bmatrix}$.
- ▷ $\mathbf{AB}$ avec le point de vue 3 : $\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^\top\mathbf{B} \\ \mathbf{a}_2^\top\mathbf{B} \end{bmatrix}$. On a $\mathbf{a}_1^\top\mathbf{B} = \begin{bmatrix} \mathbf{a}_1^\top\mathbf{b}_1 & \mathbf{a}_1^\top\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 7 & 2 \end{bmatrix}$

et $\mathbf{a}_2^\top \mathbf{B} = \begin{bmatrix}\mathbf{a}_2^\top \mathbf{b}_1 & \mathbf{a}_2^\top \mathbf{b}_2\end{bmatrix} = \begin{bmatrix}8 & 2\end{bmatrix}$. D'où : $\mathbf{AB} = \begin{bmatrix}7 & 2\\ 8 & 2\end{bmatrix}$.

▷ $\mathbf{AB}$ avec le point de vue 4 : $\mathbf{AB} = \mathbf{c}_1\mathbf{d}_1^\top + \mathbf{c}_2\mathbf{d}_2^\top$. On a les produits extérieurs $\mathbf{c}_1\mathbf{d}_1^\top = \begin{bmatrix}3 & 2\\ 3 & 2\end{bmatrix}$ et $\mathbf{c}_2\mathbf{d}_2^\top = \begin{bmatrix}4 & 0\\ 5 & 0\end{bmatrix}$. D'où : $\mathbf{AB} = \begin{bmatrix}7 & 2\\ 8 & 2\end{bmatrix}$.

Le produit matriciel possède plusieurs propriétés énoncées au théorème 1.2.2.

Théorème 1.2.2 (*propriétés du produit matriciel*)
Soient $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ des matrices réelles ou complexes. On suppose que les tailles des matrices sont telles que les opérations existent.

1. $(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top$.
2. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (si les matrices $\mathbf{A}$ et $\mathbf{B}$ sont inversibles).
3. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$.
4. $(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA}+\mathbf{CA}$.
5. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$.
6. $(\mathbf{A}^p)(\mathbf{A}^q) = \mathbf{A}^{p+q}$ avec p et q des entiers naturels.
7. Le produit de matrices triangulaires supérieures (inférieures) donne une matrice triangulaire supérieure (inférieure).

Preuve.

1 Supposons $\mathbf{A}$ de taille $m \times p$ et $\mathbf{B}$ est de taille $p \times n$. Il suffit de vérifier que les matrices des deux côtés de l'égalité ont les mêmes composantes, c'est-à-dire que $((\mathbf{AB})^\top)(i,j) = (\mathbf{B}^\top\mathbf{A}^\top)(i,j)$ pour tout $i \in [\![1;m]\!]$ et $j \in [\![1;n]\!]$. On a

▷ $((\mathbf{AB})^\top)(i,j) = (\mathbf{AB})(j,i) = \sum_{k=1}^{p} \mathbf{A}(j,k)\,\mathbf{B}(k,i)$.

▷ $(\mathbf{B}^\top\mathbf{A}^\top)(i,j) = \sum_{k=1}^{p} (\mathbf{B}^\top)(i,k)(\mathbf{A}^\top)(k,j) = \sum_{k=1}^{p} \mathbf{B}(k,i)\mathbf{A}(j,k) = \sum_{k=1}^{p} \mathbf{A}(j,k)\mathbf{B}(k,i)$.

Les deux matrices ont donc bien les mêmes composantes.

2 Comme l'inverse d'une matrice n'a pas encore été introduite, ce résultat est prouvé à la section section 1.4 (voir la proposition 1.4.1).

3 Supposons $\mathbf{A}$ de taille $m \times p$ et $\mathbf{B}$ et $\mathbf{C}$ de taille $p \times n$. Il suffit de vérifier que les matrices des deux côtés de l'égalité ont les mêmes composantes, c'est-à-dire que

$(\mathbf{A}(\mathbf{B}+\mathbf{C}))(i,j) = (\mathbf{AB})(i,j) + (\mathbf{AC})(i,j)$. On a

$$\begin{aligned}
(\mathbf{A}(\mathbf{B}+\mathbf{C}))(i,j) &= \sum_{k=1}^{p} \mathbf{A}(i,k)(\mathbf{B}+\mathbf{C})(k,j) = \sum_{k=1}^{p} \mathbf{A}(i,k)(\mathbf{B}(k,j)+\mathbf{C})(k,j) \\
&= \sum_{k=1}^{p} \mathbf{A}(i,k)\mathbf{B}(k,j) + \sum_{k=1}^{p} \mathbf{A}(i,k)\mathbf{C}(k,j)) \\
&= (\mathbf{AB})(i,j) + (\mathbf{AC})(i,j).
\end{aligned}$$

4 On déduit des points 1 et 3 de ce théorème, ainsi que des points 1 et 3 de la proposition 1.2.1 que

$$\begin{aligned}
(\mathbf{B}+\mathbf{C})\mathbf{A} &= ((\mathbf{B}^\top)^\top + (\mathbf{C}^\top)^\top)(\mathbf{A}^\top)^\top = (\mathbf{B}^\top + \mathbf{C}^\top)^\top(\mathbf{A}^\top)^\top = (\mathbf{A}^\top(\mathbf{B}^\top + \mathbf{C}^\top))^\top \\
&= (\mathbf{A}^\top\mathbf{B}^\top + \mathbf{A}^\top\mathbf{C}^\top)^\top = ((\mathbf{BA})^\top + (\mathbf{CA})^\top)^\top = ((\mathbf{BA})^\top)^\top + ((\mathbf{CA})^\top)^\top \\
&= \mathbf{BA} + \mathbf{CA}.
\end{aligned}$$

5 Soient $\mathbf{A}$ de taille $m \times p$, $\mathbf{B}$ de taille $p \times q$ et $\mathbf{C}$ de taille $q \times n$. Les deux produits $\mathbf{A}(\mathbf{BC})$ et $(\mathbf{AB})\mathbf{C}$ sont bien définis et ont une taille $m \times n$. Il suffit de vérifier que les matrices des deux côtés de l'égalité ont les mêmes composantes, c'est-à-dire que $(\mathbf{A}(\mathbf{BC}))(i,j) = ((\mathbf{AB})\mathbf{C})(i,j)$ pour tout $i \in [\![1;m]\!]$ et $j \in [\![1;q]\!]$ Pour tout $i \in [\![1;m]\!]$ et $j \in [\![1;n]\!]$. Ceci découle de l'associativité et de la commutativité d'une somme finie. En effet, on a :

$$\begin{aligned}
(\mathbf{A}(\mathbf{BC}))(i,j) &= \sum_{k=1}^{p} \mathbf{A}(i,k)(\mathbf{BC})(k,j) = \sum_{k=1}^{p} \mathbf{A}(i,k) \left(\sum_{\ell=1}^{q} \mathbf{B}(k,\ell)\mathbf{C}(\ell,j) \right) \\
&= \sum_{k=1}^{p} \sum_{\ell=1}^{q} \mathbf{A}(i,k)\mathbf{B}(k,\ell)\mathbf{C}(\ell,j) = \sum_{\ell=1}^{q} \left(\sum_{k=1}^{p} \mathbf{A}(i,k)\mathbf{B}(k,\ell) \right) \mathbf{C}(\ell,j) \\
&= \sum_{\ell=1}^{q} (\mathbf{AB})(i,\ell)\mathbf{C}(\ell,j) = ((\mathbf{AB})\mathbf{C})(i,j).
\end{aligned}$$

6 Voici une preuve par induction sur p, qui s'appuie sur le point 5 de ce théorème.

- ▷ Si $p = 0$, on a $\mathbf{A}^p = \mathbf{I}$ et donc $(\mathbf{A}^p)(\mathbf{A}^q) = \mathbf{I}\mathbf{A}^q = \mathbf{A}^q = \mathbf{A}^{p+q}$.
- ▷ Soit $p > 0$, et supposons le résultat valide pour $p-1$. On a

$$\begin{aligned}
(\mathbf{A}^p)(\mathbf{A}^q) &= (\mathbf{A}^{p-1}\mathbf{A})(\mathbf{A}^q) = (\mathbf{A}^{p-1})(\mathbf{A}\mathbf{A}^q) = (\mathbf{A}^{p-1})(\mathbf{A}^{q+1}) = \mathbf{A}^{(p-1)+(q+1)} \\
&= \mathbf{A}^{p+q}.
\end{aligned}$$

7 La preuve est faite tout d'abord pour les matrices triangulaires supérieures. Soient donc $\mathbf{A}$ et $\mathbf{B}$ deux matrices de taille $n \times n$ et triangulaires supérieures. Par définition, on a $\mathbf{A}(i,j) = 0$ et $\mathbf{B}(i,j) = 0$ pour tout $i > j$. Soit $\mathbf{C} = \mathbf{AB}$. Il suffit de montrer que $\mathbf{C}(i,j) = 0$ pour tout $i > j$. Soit donc $i > j$. On a $\mathbf{C}(i,j) = \sum_{k=1}^{n} \mathbf{A}(i,k)\mathbf{B}(k,j)$:

- ▷ si $k < i$, $\mathbf{A}(i,k) = 0$;
- ▷ si $k \geq i$, alors $k > j$ (car $i > j$), et on a donc $\mathbf{B}(k,j) = 0$.

Donc, quel que soit $k \in [\![1;n]\!]$, on a $\mathbf{A}(i,k)\mathbf{B}(k,j) = 0$, ce qui prouve que $\mathbf{C}(i,j) = 0$.

Si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices triangulaires inférieures, alors $\mathbf{A}^\top$ et $\mathbf{B}^\top$ sont triangulaires supérieures, et on vient de voir que ceci implique que le produit $\mathbf{B}^\top\mathbf{A}^\top = (\mathbf{AB})^\top$ est également une matrice triangulaire supérieure, ce qui implique que $\mathbf{AB}$ est une matrice triangulaire inférieure.

□

Remarque

Concernant les tailles des matrices :

- ▷ Pour le point 1 : $\mathbf{B}$ doit avoir le même nombre de lignes que le nombre de colonnes de $\mathbf{A}$.
- ▷ Pour le point 3 : $\mathbf{B}$ et $\mathbf{C}$ doivent être de même taille et doivent avoir le même nombre de lignes que le nombre de colonnes de $\mathbf{A}$.
- ▷ Pour le point 4 : $\mathbf{B}$ et $\mathbf{C}$ doivent être de même taille et doivent avoir le même nombre de colonnes que le nombre de lignes de $\mathbf{A}$.
- ▷ Pour le point 5 : $\mathbf{B}$ doit avoir le même nombre de lignes que le nombre de colonnes de $\mathbf{A}$ et le même nombre de colonnes que le nombre de lignes de $\mathbf{C}$.
- ▷ Pour le point 6 : $\mathbf{A}$ doit être carrée.

Les propriétés du théorème 1.2.2 peuvent être utilisées pour l'exercice 1.11. L'exemple 1.2.11 traite du point 1 pour une matrice complexe.

Exemple 1.2.11 (*transposé d'un produit de matrices complexes*)

Soient $\mathbf{A} = \begin{bmatrix} 1+\mathrm{i} & 2 \\ 3\mathrm{i} & -1 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 2 & -\mathrm{i} \\ 1 & 4\mathrm{i} \end{bmatrix}$ deux matrices complexes. Montrons que $(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top$.

- ▷ D'une part, on a :

$$‘\mathbf{AB} = \begin{bmatrix} (1+\mathrm{i}) \times 2 + 2 \times 1 & (1+\mathrm{i}) \times (-\mathrm{i}) + 2 \times (4\mathrm{i}) \\ (3\mathrm{i}) \times 2 + (-1) \times 1 & (3\mathrm{i}) \times (-\mathrm{i}) + (-1) \times (4\mathrm{i}) \end{bmatrix}.$$

La simplification des quatre composantes donne :

- ▷ $(1+\mathrm{i}) \times 2 + 2 \times 1 = 2 + 2\mathrm{i} + 2 = 4 + 2\mathrm{i}$;
- ▷ $(1+\mathrm{i}) \times (-\mathrm{i}) + 2 \times (4\mathrm{i}) = -\mathrm{i} - \mathrm{i}^2 + 8\mathrm{i} = -\mathrm{i} + 1 + 8\mathrm{i} = 1 + 7\mathrm{i}$;

- ▷ $(3\mathrm{i}) \times 2 + (-1) \times 1 = -1 + 6\mathrm{i}$;
- ▷ $(3\mathrm{i}) \times (-\mathrm{i}) + (-1) \times (4\mathrm{i}) = -3\mathrm{i}^2 - 4\mathrm{i} = 3 - 4\mathrm{i}$.

Donc

$$\mathbf{AB} = \begin{bmatrix} 4+2\mathrm{i} & 1+7\mathrm{i} \\ -1+6\mathrm{i} & 3-4\mathrm{i} \end{bmatrix} \text{ et } (\mathbf{AB})^\top = \begin{bmatrix} 4+2\mathrm{i} & -1+6\mathrm{i} \\ 1+7\mathrm{i} & 3-4\mathrm{i} \end{bmatrix}.$$

▷ D'autre part, on a :

$$\mathbf{B}^\top = \begin{bmatrix} 2 & 1 \\ -\mathrm{i} & 4\mathrm{i} \end{bmatrix} \text{ et } \mathbf{A}^\top = \begin{bmatrix} 1+\mathrm{i} & 3\mathrm{i} \\ 2 & -1 \end{bmatrix}.$$

Leur produit donne :

$$\mathbf{B}^\top\mathbf{A}^\top = \begin{bmatrix} 2 \times (1+\mathrm{i}) + 1 \times 2 & 2 \times (3\mathrm{i}) + 1 \times (-1) \\ (-\mathrm{i}) \times (1+\mathrm{i}) + (4\mathrm{i}) \times 2 & (-\mathrm{i}) \times (3\mathrm{i}) + (4\mathrm{i}) \times (-1) \end{bmatrix}$$

La simplification des quatre composantes donne :

- ▷ $2 \times (1+\mathrm{i}) + 1 \times 2 = 2 + 2\mathrm{i} + 2 = 4 + 2\mathrm{i}$;
- ▷ $2 \times (3\mathrm{i}) + 1 \times (-1) = -1 + 6\mathrm{i}$;
- ▷ $(-\mathrm{i}) \times (1+\mathrm{i}) + (4\mathrm{i}) \times 2 = -\mathrm{i} - \mathrm{i}^2 + 8\mathrm{i} = -\mathrm{i} + 1 + 8\mathrm{i} = 1 + 7\mathrm{i}$;
- ▷ $(-\mathrm{i}) \times (3\mathrm{i}) + (4\mathrm{i}) \times (-1) = (-3\mathrm{i}^2) - 4\mathrm{i} = 3 - 4\mathrm{i}$.

Donc :

$$\mathbf{B}^\top\mathbf{A}^\top = \begin{bmatrix} 4+2\mathrm{i} & -1+6\mathrm{i} \\ 1+7\mathrm{i} & 3-4\mathrm{i} \end{bmatrix}.$$

On obtient bien

$$(\mathbf{AB})^\top = \mathbf{B}^\top\mathbf{A}^\top.$$

Il convient de noter que $\mathbf{AB}$ n'est en général pas égal à $\mathbf{BA}$. Si le nombre de colonnes de $\mathbf{A}$ est égal au nombre de lignes de $\mathbf{B}$ alors que le nombre de lignes de $\mathbf{A}$ n'est pas égal au nombre de colonnes de $\mathbf{B}$, le produit $\mathbf{AB}$ est défini alors que $\mathbf{BA}$ ne l'est pas. Mais même si les deux produits sont définis, ils peuvent être différents, tel qu'illustré dans l'exemple 1.2.12.

Exemple 1.2.12 *(le produit matriciel n'est pas commutatif)*

Si $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, on a $\mathbf{AB} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ alors que $\mathbf{BA} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$.

Il est intéressant cependant d'observer que si les deux produits $\mathbf{AB}$ et $\mathbf{BA}$ sont définis, alors leur trace est la même, tel qu'établi dans la proposition 1.2.6 (cette propriété ayant déjà été énoncée au point 4 de la proposition 1.2.3).

Proposition 1.2.6 (*trace du produit matriciel*)
Si $\mathbf{A}$ est une matrice de taille $m \times n$ et $\mathbf{B}$ une matrice de taille $n \times m$, alors

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA}).$$

Preuve. Étant donné que $(\mathbf{AB})(i,i) = \sum_{k=1}^{n} \mathbf{A}(i,k)\mathbf{B}(k,i)$, on a

$$\operatorname{tr}(\mathbf{AB}) = \sum_{i=1}^{m} (\mathbf{AB})_{ii} = \sum_{i=1}^{m}\sum_{k=1}^{n} \mathbf{A}(i,k)\mathbf{B}(k,i) \ .$$

On peut échanger l'ordre des sommes (car elles sont finies) ainsi que celui des produits, ce qui donne

$$\operatorname{tr}(\mathbf{AB}) = \sum_{k=1}^{n}\sum_{i=1}^{m} \mathbf{B}(k,i)\,\mathbf{A}(i,k) = \sum_{k=1}^{n} (\mathbf{BA})(k,k) = \operatorname{tr}(\mathbf{BA}) \ .$$

□

La *matrice identité* introduite à la définition 1.2.16 est ce qu'on appelle l'élément neutre de la multiplication, tel qu'expliqué à la proposition 1.2.7. Elle intervient, notamment, lors de l'inversion d'une matrice (section 1.4) ou dans la définition d'une base canonique (proposition 5.2.4).

Définition 1.2.16 (*matrice identité*)
La *matrice identité*, ou simplement l'*l'identité*, est une matrice carrée de taille $n \times n$, notée $\mathbf{I}$, telle que :

$$\mathbf{I} = \mathbf{I}_n = [\mathbf{e}_1 \ \mathbf{e}_2 \ \cdots \ \mathbf{e}_n] = \begin{bmatrix} \mathbf{e}_1^\top \\ \mathbf{e}_2^\top \\ \vdots \\ \mathbf{e}_n^\top \end{bmatrix} {}^{a} \ .$$

a. on utilise $\mathbf{I}_n$ plutôt que $\mathbf{I}$ pour éviter des ambiguïtés.

Remarque
On remarque que $\mathbf{I}$ est symétrique, la i-ième colonne étant égale à la transposée de la i-ième ligne, pour tout $i \in [\![1;n]\!]$.

Proposition 1.2.7 (*élément neutre du produit matriciel*)
L'identité $\mathbf{I}$ est l'élément neutre du produit matriciel, ce qui signifie que :

- ▷ $\mathbf{AI} = \mathbf{A}$ pour toute matrice $\mathbf{A}$ de taille $p \times n$;
- ▷ $\mathbf{IA} = \mathbf{A}$ pour toute matrice $\mathbf{A}$ de taille $n \times p$.

Preuve. Soit $\mathbf{A}$ une matrice de taille $p \times n$ dont les colonnes sont $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n$. En utilisant le point de vue 2 de la proposition 1.2.5 ainsi que la proposition 1.2.4, on obtient :

$$\mathbf{AI} = [\mathbf{Ae}_1 \ \mathbf{Ae}_2 \cdots \mathbf{Ae}_n] = [\mathbf{c}_1 \ \mathbf{c}_2 \cdots \mathbf{c}_n] = \mathbf{A} \ .$$

Si $\mathbf{A}$ est une matrice de taille $n \times p$, sa transposée est de taille $p \times n$, et tel qu'on vient de le voir, cela signifie que $\mathbf{A}^\top \mathbf{I} = \mathbf{A}^\top$. En utilisant le point 1 du théorème 1.2.2 et le point 3 de la proposition 1.2.1, on a donc :

$$\mathbf{IA} = (\mathbf{A}^\top \mathbf{I}^\top)^\top = (\mathbf{A}^\top \mathbf{I})^\top = (\mathbf{A}^\top)^\top = \mathbf{A}.$$

□

L'exemple 1.2.13 met en évidence certains pièges à éviter qui vont à l'encontre de l'intuition.

Exemple 1.2.13 (*pièges de la multiplication matricielle*)

- ▷ Si $\mathbf{AB} = \mathbf{O}$, cela n'implique pas nécessairement que $\mathbf{A} = \mathbf{O}$ ou $\mathbf{B} = \mathbf{O}$.
  Par exemple, si $\mathbf{A} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, on a $\mathbf{A} \neq \mathbf{O}$, $\mathbf{B} \neq \mathbf{O}$ et $\mathbf{AB} = \mathbf{O}$.
- ▷ Si $\mathbf{AB} = \mathbf{B}$ avec $\mathbf{B} \neq \mathbf{O}$, cela n'implique pas nécessairement que $\mathbf{A} = \mathbf{I}$.
  Par exemple, si $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, on a $\mathbf{AB} = \mathbf{B}$ et $\mathbf{A} \neq \mathbf{I}$.

1.2.7 Multiplication par blocs

Il est parfois utile d'effectuer la multiplication de matrices par blocs (telles que définies dans la section 1.2.3), comme dans l'exemple 1.2.14. La plupart des règles de la section précédente s'appliquent naturellement aux matrices par blocs .

Il faut prendre garde à ne pas confondre $\mathbf{A}_{ij}$ qui est un bloc de $\mathbf{A}$ et $\mathbf{A}(i,j)$ qui est une composante de la matrice $\mathbf{A}$. Notons que $\mathbf{A}_{11}(1,1) = \mathbf{A}(1,1)$ est la première composante de la première ligne (ou colonne) de $\mathbf{A}$.

Lorsqu'on effectue une multiplication par blocs, il est important de vérifier la compatibilité des tailles des blocs. Dans l'exemple 1.2.14, pour que la multiplication par blocs puisse être réalisée, il faut donc, entre autres, que le nombre de colonnes de $\mathbf{A}_{11}$ soit égal au nombre de lignes de $\mathbf{B}_{11}$ et que le nombre de colonnes de $\mathbf{A}_{12}$ soit égal au nombre de lignes de $\mathbf{B}_{21}$. Il faut aussi que les sommes soient bien définies ce qui implique, par exemple, que $\mathbf{A}_{11}$ et $\mathbf{A}_{12}$ doivent avoir le même nombre de lignes et que $\mathbf{B}_{11}$ et $\mathbf{B}_{21}$ doivent avoir le même nombre de colonnes.

Exemple 1.2.14

$$\begin{aligned}\mathbf{AB} &= \left[\begin{array}{c|c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]\left[\begin{array}{c|c|c}\mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23}\end{array}\right] \\ &= \left[\begin{array}{c|c|c}\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22} & \mathbf{A}_{11}\mathbf{B}_{13}+\mathbf{A}_{12}\mathbf{B}_{23} \\ \hline \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22} & \mathbf{A}_{21}\mathbf{B}_{13}+\mathbf{A}_{22}\mathbf{B}_{23}\end{array}\right].\end{aligned}$$

1.2.8 Produit scalaire et norme matricielle

Pour bien comprendre le contenu de cette section, il est préférable (mais pas nécessaire) de connaître les définitions des vecteurs et matrices complexes (section 3.6). Ainsi, par exemple, $\mathbf{A}^*$ est la transconjuguée de la matrice complexe $\mathbf{A}$, et on a $\mathbf{A}^* = \mathbf{A}^\top$ si $\mathbf{A}$ est réelle. Il est également utile de connaître la définition d'un produit scalaire (voir définition 5.3.1) ainsi que la définition d'un espace vectoriel (voir définition 5.1.1).

Tel qu'indiqué dans la définition 1.2.17, le produit scalaire matriciel considéré dans cet ouvrage est défini grâce à la trace.

Définition 1.2.17 *(produit scalaire matriciel)*

Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices complexes ou réelles de taille $m \times n$, . Le *produit scalaire matriciel*, noté $\langle \mathbf{A}, \mathbf{B}\rangle$, est défini par

$$\langle \mathbf{A}, \mathbf{B}\rangle = \operatorname{tr}(\mathbf{A}^*\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{A}^*).$$

Remarque

- ▷ Si les matrices $\mathbf{A}$ et $\mathbf{B}$ sont réelles, on a $\langle \mathbf{A}, \mathbf{B}\rangle = \operatorname{tr}(\mathbf{A}^\top\mathbf{B}) = \operatorname{tr}(\mathbf{B}\mathbf{A}^\top)$.
- ▷ Le produit scalaire matriciel n'est défini que pour deux matrices de même taille. En effet, si $\mathbf{A}$ et $\mathbf{B}$ sont toutes les deux de taille $m \times n$, alors $\mathbf{A}^*\mathbf{B}$ est de taille $n \times n$ et $\mathbf{B}\mathbf{A}^*$ est de taille $m \times m$. Dans les deux cas, la matrice résultante est carrée et on peut donc calculer sa trace.

Il conviendrait de vérifier toutes les propriétés de la définition 5.3.1 pour s'assurer que la définition 1.2.17 correspond bien à un produit scalaire. Une preuve formelle n'est pas donnée ici car elle repose sur des notions introduites ultérieurement dans cet l'ouvrage. Elle découle de la définition 1.2.14 du produit matriciel ainsi que de la définition 1.2.12 et des propriétés de la proposition 1.2.6 portant sur la trace.

Exemple 1.2.15 (*produit scalaire matriciel*)

Soients $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. De simples calculs donnent :

$$\mathbf{A}^\top \mathbf{B} = \begin{bmatrix} 4 & 3 \\ 4 & 2 \end{bmatrix}, \quad \mathbf{B}^\top \mathbf{A} = \begin{bmatrix} 4 & 4 \\ 3 & 2 \end{bmatrix}, \quad \mathbf{A}^\top \mathbf{A} = \begin{bmatrix} 10 & 8 \\ 8 & 8 \end{bmatrix} \quad \text{et} \quad \mathbf{B}^\top \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

On a donc bien $\langle \mathbf{A}, \mathbf{B} \rangle = 6 = \langle \mathbf{B}, \mathbf{A} \rangle$, $\langle \mathbf{A}, \mathbf{A} \rangle = 18 > 0$ et $\langle \mathbf{B}, \mathbf{B} \rangle = 3 > 0$.

On pourrait étendre le concept d'orthogonalité à autre chose que des vecteurs colonnes. Cette notion sera approfondie dans le chapitre 5. Il faut prendre garde cependant à ne pas confondre la notion de matrice orthogonale vue à la définition 5.3.5 avec le fait que deux matrices soient orthogonales entre elles.

L'introduction d'un produit scalaire sert souvent à définir une norme spécifique. On peut donc maintenant définir une norme pour les matrices : le produit scalaire de la définition 1.2.17 donne la *norme de Frobenius* tel qu'indiqué ci-dessous à la définition 1.2.18. L'exercice 1.29 traite de cette norme.

Définition 1.2.18 (*norme de Frobenius*)

La *norme de Frobenius* d'une matrice réelle ou complexe $\mathbf{A}$ de taille $m \times n$ est notée $\|\mathbf{A}\|_{\mathrm{F}}$ et est définie par

$$\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \sqrt{\mathrm{tr}(\mathbf{A}^* \mathbf{A})} = \sqrt{\mathrm{tr}(\mathbf{A} \mathbf{A}^*)} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |\mathbf{A}(i,j)|^2}\ .$$

Remarque

- ▷ Si $\mathbf{A}$ est une matrice réelle de taille $m \times n$, la formule se simplifie puisque $|\mathbf{A}(i,j)|^2 = \mathbf{A}(i,j)^2$. Dans ce cas, on a donc

$$\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\langle \mathbf{A}, \mathbf{A} \rangle} = \sqrt{\mathrm{tr}(\mathbf{A}^\top \mathbf{A})} = \sqrt{\mathrm{tr}(\mathbf{A} \mathbf{A}^\top)} = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} \mathbf{A}(i,j)^2}\ .$$

- ▷ La formule de la définition 1.2.18 est très intuitive. En effet, le carré de cette norme pour une matrice réelle est égal à la somme des carrés de ses composantes, exactement comme pour les vecteurs. De plus, Si $n = 1$ ($\mathbf{A}$ est donc un vecteur colonne réel ou complexe), on a la norme euclidienne définie en définition 1.1.13 pour les réels, et en définition 3.6.6 pour les complexes.

Exemple 1.2.16 (*normes de matrices réelles et complexes*)

- ▷ Si $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 1 & 3 \end{bmatrix}$, les calculs donnent $\langle \mathbf{A}, \mathbf{A} \rangle = \operatorname{tr}(\mathbf{A}^\top \mathbf{A}) = \operatorname{tr}\left(\begin{bmatrix} 11 & 7 \\ 7 & 9 \end{bmatrix}\right) = 20.$ On a donc $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{20} = 2\sqrt{5}$
- ▷ Si $\mathbf{B} = \begin{bmatrix} 1 & 2 \\ \mathrm{i} & 1 \end{bmatrix}$, on a $\mathbf{B}^* = \begin{bmatrix} 1 & -\mathrm{i} \\ 2 & 1 \end{bmatrix}$ et $\langle \mathbf{B}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{B}^* \mathbf{B}) = \operatorname{tr}\begin{bmatrix} 2 & 2-\mathrm{i} \\ 2+\mathrm{i} & 5 \end{bmatrix} = 7.$ On déduit que $\|\mathbf{B}\|_{\mathrm{F}} = \sqrt{7}$. C'est bien un réel !

D'autres produits scalaires et donc d'autres normes existent pour les matrices, mais cela sort du cadre de cet ouvrage. Chaque norme a son utilité pour prouver certaines propriétés utiles pour résoudre des problèmes. Un exemple d'utilisation de la norme de Frobenius se retrouve dans l'étude de cas de la section 9.3. L'intérêt de minimiser la norme de Frobenius de la hessienne d'un modèle quadratique est de pouvoir déterminer le modèle quadratique de courbure minimale.

1.3 Déterminant d'une matrice

Tout comme la trace, le déterminant d'une matrice est une des propriétés des matrices utiles dans de nombreuses applications, notamment pour montrer qu'une matrice est inversible (section 1.4). Le déterminant n'est défini que pour des matrices carrées.

Les matrices de cette section ne sont pas limitées aux réelles. Elle peuvent donc être complexes. Pour un formalisme complet du déterminant des matrices complexes, le lecteur est invité à consulter la section 3.6.4.

1.3.1 Définitions

Tel qu'indiqué dans la définition 1.3.1, le déterminant d'une matrice de taille 1×1 est la valeur de son unique composante. Cette notion de déterminant est étendue dans la définition 1.3.2 aux matrices de taille 2×2 . En 2D, le déterminant d'une matrice peut s'interpréter comme l'aire du parallélogramme engendré par les colonnes de la matrice. Plus précisément, si $\mathbf{A}$ est une matrice de taille 2×2 dont les colonnes sont $\mathbf{c}_1$ et $\mathbf{c}_2$, alors le déterminant de $\mathbf{A}$ est l'aire du parallélogramme dont les extrémités sont $\mathbf{0}, \mathbf{c}_1, \mathbf{c}_2$ et $\mathbf{c}_1 + \mathbf{c}_2$.

Le déterminant d'une matrice carrée quelconque peut être déterminé de manière récursive, tel qu'indiqué dans la définition 1.3.4.

Définition 1.3.1 *(déterminant d'une matrice de taille 1×1)*

Soit $\mathbf{A} = \begin{bmatrix} a \end{bmatrix}$ une matrice n'ayant qu'une composante qui peut être réelle ou complexe. Le *déterminant* $\det(\mathbf{A})$ de $\mathbf{A}$ est défini par $\det(\mathbf{A}) = a$.

La notion de déterminant d'une matrice de taille 2×2, telle que décrite dans la définition 1.3.2, a d'abord été introduite par Cardan pour résoudre des systèmes d'équations à deux inconnues et déterminer si le système considéré a une solution unique. La notion de résolution de systèmes d'équations est abordée dans le chapitre 8.

Définition 1.3.2 *(déterminant d'une matrice de taille 2×2)*

Soit $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ une matrice réelle ou complexe de taille 2×2.
Le *déterminant* $\det(\mathbf{A})$ de $\mathbf{A}$ est défini par $\det(\mathbf{A}) = ad - bc$.

Exemple 1.3.1 *(déterminant d'une matrice 2×2)*

Le déterminant $\det(\mathbf{A})$ de $\mathbf{A} = \begin{bmatrix} 2 & -3 \\ 6 & 1 \end{bmatrix}$ est égal à $2 \times 1 - (-3) \times 6 = -15$.

La définition précédente peut être généralisée à tout type de matrice en utilisant le concept de récursivité. Pour cette définition plus générale, il est nécessaire d'introduire d'abord la notion de cofacteur décrite dans la définition 1.3.3 et illustrée dans l'exemple 1.3.2.

Définition 1.3.3 (*cofacteur*)

Les *cofacteurs* d'une matrice $\mathbf{A}$ de taille $n \times n$ sont les termes

$$c_{i,j} = (-1)^{i+j}\mathbf{A}(i,j)\det(\mathbf{A}_{i,j})$$

avec $i \in [\![1;n]\!]$, $j \in [\![1;n]\!]$ et où $\mathbf{A}_{i,j}$ est la sous-matrice de $\mathbf{A}$ obtenue en supprimant la i-ième ligne et la j-ième colonne de $\mathbf{A}$.

Remarque

Une matrice possède autant de cofacteurs que de composantes.

Exemple 1.3.2

Les neuf cofacteurs de $\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}$ sont :

$$\begin{aligned}
c_{1,1} &= (-1)^{1+1} \times 2 \times \det\left(\begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix}\right) = -8 & c_{1,2} &= (-1)^{1+2} \times 3 \times \det\left(\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}\right) = 24 \\
c_{1,3} &= (-1)^{1+3} \times 1 \times \det\left(\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\right) = -4 & c_{2,1} &= (-1)^{2+1} \times 1 \times \det\left(\begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\right) = -1 \\
c_{2,2} &= (-1)^{2+2} \times 2 \times \det\left(\begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}\right) = -2 & c_{2,3} &= (-1)^{2+3} \times 3 \times \det\left(\begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}\right) = 15 \\
c_{3,1} &= (-1)^{3+1} \times 3 \times \det\left(\begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix}\right) = 21 & c_{3,2} &= (-1)^{3+2} \times 2 \times \det\left(\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}\right) = -10 \\
c_{3,3} &= (-1)^{3+3} \times 1 \times \det\left(\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}\right) = 1.
\end{aligned}$$

.

Tel que décrit dans la définition 1.3.4, on peut maintenant introduire la notion de déterminant d'une matrice de taille quelconque.

Définition 1.3.4 *(déterminant d'une matrice)*

Le *déterminant* $\det(\mathbf{A})$ d'une matrice $\mathbf{A}$ réelle ou complexe de taille $n \times n$ peut se calculer de manière récursive, à l'aide des cofacteurs $c_{i,j}$, de deux façons différentes :

- ▷ en développant selon la ligne i : $\det(\mathbf{A}) = \sum_{j=1}^{n} c_{i,j}$;
- ▷ en développant selon la colonne j : $\det(\mathbf{A}) = \sum_{i=1}^{n} c_{i,j}$.

Remarque

Les définitions 1.3.2 et 1.3.4 coïncident pour les matrices de taille 2×2. En effet, les sous-matrices sont ici les composantes de la matrice. Par exemple, $\mathbf{A}_{1,1} = \mathbf{A}(2,2)$ et donc $c_{1,1} = \mathbf{A}(1,1)\mathbf{A}(2,2)$.

L'exemple 1.3.3 illustre le calcul de déterminant à l'aide des cofacteurs pour deux matrices.

Exemple 1.3.3

- ▷ Soit $\mathbf{A} = \begin{bmatrix} -1 & 1 & 2 \\ 1 & 4 & 1 \\ 5 & 0 & 1 \end{bmatrix}$.

  Pour simplifier les calculs, puisque la deuxième colonne contient un zéro, on peut déterminer $\det(\mathbf{A})$ en développant selon cette colonne, ce qui donne :

$$\begin{aligned}
\det(\mathbf{A}) = &\sum_{i=1}^{3} (-1)^{i+2} \mathbf{A}(i,2) \det(\mathbf{A}_{i,2}) \\
= &(-1)^{1+2} \times 1 \times \det\left(\begin{bmatrix} 1 & 1 \\ 5 & 1 \end{bmatrix}\right) \\
&+ (-1)^{2+2} \times 4 \times \det\left(\begin{bmatrix} -1 & 2 \\ 5 & 1 \end{bmatrix}\right) \\
&+ (-1)^{3+2} \times 0 \times \det\left(\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}\right) \\
= &-(1 \times 1 - 5 \times 1) + 4 \times (-1 \times 1 - 5 \times 2) = -40.
\end{aligned}$$

  On peut également faire le calcul en développant selon la dernière ligne qui a aussi un zéro, ce qui donne :

$$\begin{aligned}\det(\mathbf{A}) =& \sum_{j=1}^{3}(-1)^{3+j}\mathbf{A}(3,j)\det(\mathbf{A}_{3,j})\\ =&(-1)^{3+1}\times 5\times\det\left(\begin{bmatrix}1 & 2\\ 4 & 1\end{bmatrix}\right)\\ &+(-1)^{3+2}\times 0\times\det\left(\begin{bmatrix}-1 & 2\\ 1 & 1\end{bmatrix}\right)\\ &+(-1)^{3+3}\times 1\times\det\left(\begin{bmatrix}-1 & 1\\ 1 & 4\end{bmatrix}\right)\\ =&5(1\times 1-4\times 2)+(-1\times 4-1\times 1)=-40.\end{aligned}$$

Les deux calculs donnent bien des résultats identiques !

▷ Pour la matrice de l'exemple 1.3.2, on remarque que peu importe le développement, le résultat est identique :

- ▷ selon la première ligne, on a $\det(\mathbf{A}) = c_{1,1}+c_{1,2}+c_{1,3} = -8+24-4 = 12$;
- ▷ selon la deuxième ligne, on a $\det(\mathbf{A}) = c_{2,1}+c_{2,2}+c_{2,3} = -1-2+15 = 12$;
- ▷ selon la troisième ligne, on a $\det(\mathbf{A}) = c_{3,1}+c_{3,2}+c_{3,3} = 21-10+1 = 12$;
- ▷ selon la première colonne, on a $\det(\mathbf{A}) = c_{1,1}+c_{2,1}+c_{3,1} = -8-1+21 = 12$;
- ▷ selon la deuxième colonne, on a $\det(\mathbf{A}) = c_{1,2}+c_{2,2}+c_{3,2} = 24-2-10 = 12$;
- ▷ selon la troisième colonne, on a $\det(\mathbf{A}) = c_{1,3}+c_{2,3}+c_{3,3} = -4+15+1 = 12$.

1.3.2 Propriétés du déterminant

La connaissance du déterminant d'une matrice peut être utile pour déterminer le déterminant d'autres matrices qui lui sont reliées d'une quelconque façon. Ceci fait l'objet de cette section.

Le théorème 1.3.1 énonce la propriété importante que le déterminant d'une matrice est invariant par transposition.

Théorème 1.3.1 (*déterminant de la transposée d'une matrice*)
Le déterminant d'une matrice carrée $\mathbf{A}$ est égal au déterminant de sa transposée, ce qui veut dire que

$$\det(\mathbf{A}^\top) = \det(\mathbf{A})\ .$$

Preuve. Soit $\mathbf{A}$ une matrice carrée de taille $n \times n$. La preuve se fait par induction sur n.

▷ **Initialisation.** Si $n = 1$, on a $\mathbf{A} = \mathbf{A}^\top = [a]$ et la définition 1.3.1) indique que

$\det(\mathbf{A}) = \det(\mathbf{A}^\top) = a$.

- ▷ **Induction.** Soit $n > 1$ et supposons que le théorème est valide pour les matrices carrées de taille $(n-1) \times (n-1)$. Le développement selon la dernière ligne donne :

$$\det(\mathbf{A}^\top) = \sum_{j=1}^{n} (-1)^{n+j} \mathbf{A}^\top(n,j) \det\left((\mathbf{A}^\top)_{n,j}\right).$$

Pour tout $j \in [\![1; n]\!]$, on a $(\mathbf{A}_{j,n})^\top = (\mathbf{A}^\top)_{n,j}$, et comme la sous-matrice $\mathbf{A}_{j,n}$ est de taille $(n-1) \times (n-1)$, on a $\det(\mathbf{A}_{j,n}) = \det\left((\mathbf{A}^\top)_{n,j}\right)$. De plus, $\mathbf{A}^\top(n,j) = \mathbf{A}(j,n)$, ce qui implique

$$\det(\mathbf{A}^\top) = \sum_{j=1}^{n} (-1)^{j+n} \mathbf{A}(j,n) \det(\mathbf{A}_{j,n}).$$

Le terme de droite de cette dernière égalité est la formule du déterminant de $\mathbf{A}$ développé selon sa dernière colonne. La propriété est donc démontrée.

□

Pour certaines matrices, comme les matrices triangulaires ou diagonales, le calcul du déterminant est simplifié. En effet, il suffit de considérer les éléments diagonaux. C'est l'objet de la proposition 1.3.1 pour les matrices triangulaires qui donne le corollaire 1.3.1 pour les matrices diagonales.

Proposition 1.3.1 (*déterminant d'une matrice triangulaire*)

Le déterminant d'une matrice triangulaire $\mathbf{A}$ de taille $n \times n$ est égal au produit des composantes sur sa diagonale :

$$\det(\mathbf{A}) = \prod_{i=1}^{n} \mathbf{A}(i,i) \ .$$

Preuve. Soit $\mathbf{A}$ une matrice triangulaire supérieure de taille $n \times n$. La preuve est pas induction sur n.

- ▷ **Initialisation.** Si $n = 1$, on a $\mathbf{A} = [a]$ et la définition 1.3.1 indique que $\det(\mathbf{A}) = a = \mathbf{A}(1,1)$.
- ▷ **Induction.** Soit $n > 1$ et supposons que la proposition est valide pour les matrices triangulaires supérieures de taille $(n-1) \times (n-1)$. Le développement selon la dernière ligne donne :

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{n+j} \mathbf{A}(n,j) \det(\mathbf{A}_{n,j}) \ .$$

Comme $\mathbf{A}$ est triangulaire supérieure, le seul indice de colonne j tel que $\mathbf{A}(n,j) \neq 0$ est $j = n$, et on a donc :

$$\det(\mathbf{A}) = \mathbf{A}(n,n)\det(\mathbf{A}_{n,n}).$$

Mais $\mathbf{A}_{n,n}$ est également triangulaire supérieure et de taille $(n-1)\times(n-1)$, ce qui implique que $\det(\mathbf{A}_{n,n}) = \prod_{i=1}^{n-1} \mathbf{A}_{n,n}(i,i) = \prod_{i=1}^{n-1} \mathbf{A}(i,i)$. On a donc

$$\det(\mathbf{A}) = \mathbf{A}(n,n)\prod_{i=1}^{n-1}\mathbf{A}(i,i) = \prod_{i=1}^{n}\mathbf{A}(i,i).$$

Si $\mathbf{A}$ st triangulaire inférieure, le théorème 1.3.1 indique que $\det(\mathbf{A}) = \det(\mathbf{A}^\top)$, et comme $\mathbf{A}^\top$ est triangulaire supérieure, on vient de démontrer que $\det(\mathbf{A}^\top) = \prod_{i=1}^{n} \mathbf{A}^\top(i,i)$. Pour conclure, il suffit d'observer que $\mathbf{A}^\top(i,i) = \mathbf{A}(i,i)$, ce qui donne

$$\det(\mathbf{A}) = \det(\mathbf{A}^\top) = \prod_{i=1}^{n}\mathbf{A}^\top(i,i) = \prod_{i=1}^{n}\mathbf{A}(i,i).$$

□

L'exemple 1.3.4 illustre l'utilité de la proposition 1.3.1 dans le cas d'une matrice triangulaire inférieure de taille 4×4.

Exemple 1.3.4

Le déterminant de la matrice triangulaire inférieure $\mathbf{A} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 78 & 1 & 0 & 0 \\ 34 & 22 & 1 & 0 \\ 12 & 14 & 3 & 1 \end{bmatrix}$ est $\det(\mathbf{A}) = 3\times 1\times 1\times 1 = 3$.

Remarque

La matrice identité étant une matrice triangulaire avec uniquement des 1 sur la diagonale, on a $\det(\mathbf{I}) = 1$.

Corollaire 1.3.1 *(déterminant d'une matrice diagonale)*

Le déterminant d'une matrice diagonale $\mathbf{A}$ de taille $n\times n$ est égal au produit des composantes sur sa diagonale :

$$\det(\mathbf{A}) = \prod_{i=1}^{n} \mathbf{A}(i,i) \ .$$

Preuve. Étant donné qu'une matrice diagonale est triangulaire, le résultat est une conséquence directe de la proposition 1.3.1. □

Une autre propriété fondamentale relie le déterminant à la multiplication d'une matrice par un scalaire : multiplier une matrice $n \times n$ par un facteur k revient à multiplier son déterminant par k dans chacune des n dimensions, ce qui conduit à la proposition 1.3.2 et à l'exemple 1.3.5

Proposition 1.3.2 (*déterminant du produit d'une matrice par un scalaire*)
Soit $\mathbf{A}$ une matrice carrée de taille $n \times n$ et soit k un scalaire. On a :

$$\det(k\mathbf{A}) = k^n \det(\mathbf{A}) \ .$$

Preuve. Soit $\mathbf{A}$ une matrice carrée de taille $n \times n$. Soient $c_{i,j}$ les cofacteurs de $\mathbf{A}$ et $d_{i,j}$ ceux de $k\mathbf{A}$, avec $i \in [\![1;n]\!]$ et $j \in [\![1;n]\!]$. La preuve est par induction sur n.

- ▷ **Initialisation.** Si $n = 1$, on a $\mathbf{A} = [a]$ et $k\mathbf{A} = [ka]$. La définition 1.3.1 indique que $\det(k\mathbf{A}) = ka = k\det(\mathbf{A})$.
- ▷ **Induction.** Soit $n > 1$ et supposons que la proposition est valide pour les matrices carrées de taille $(n-1) \times (n-1)$, La définition 1.3.3 indique que le cofacteur $d_{i,j}$ a la valeur

$$d_{i,j} = (-1)^{i+j}(k\mathbf{A})(i,j)\det\left((k\mathbf{A})_{i,j}\right)$$

Comme $(k\mathbf{A})_{i,j}$ est de taille $(n-1)\times(n-1)$, on a $\det\left((k\mathbf{A})_{i,j}\right) = k^{n-1}\det(\mathbf{A}_{i,j})$, ce qui implique

$$d_{i,j} = k^n\Big((-1)^{i+j}\mathbf{A}(i,j)\det(\mathbf{A}_{i,j})\Big) = k^n c_{i,j}.$$

On déduit le résultat souhaité puisque

$$\det(k\mathbf{A}) = \sum_{j=1}^{n} d_{1,j} = k^n \sum_{j=1}^{n} c_{1,j} = k^n \det(\mathbf{A}).$$

□

Exemple 1.3.5 *(multiplication d'une matrice par un scalaire)*
La matrice $\mathbf{B} = \begin{bmatrix} -2 & 2 & 4 \\ 2 & 8 & 2 \\ 10 & 0 & 2 \end{bmatrix}$ correspond à la matrice $\mathbf{A}$ de l'exemple 1.3.3 qui a été multipliée par deux. Comme c'est une matrice de taille 3×3 et qu'on a vu que $\det(\mathbf{A}) = -40$, on déduit que $\det(\mathbf{B}) = 2^3 \det(\mathbf{A}) = 8 \times -40 = -320$.

On conclut cette section en citant trois propriétés importantes sur les déterminants qui seront démontrées ultérieurement dans cet ouvrage.

- ▷ Théorème 2.4.3. Si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices carrées de même taille, alors $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$.
- ▷ Théorème 2.4.4. Si $\mathbf{A}$ est une matrice inversible, alors $\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$.
- ▷ Proposition 7.1.3 : Le déterminant d'une matrice carrée complexe est le produit de ses *valeurs propres*.

1.4 Matrices inverses

Tel que déjà mentionné, une matrice peut posséder une inverse, et on dit alors qu'elle est inversible. Cette section présente des conditions nécessaires et suffisantes à l'existence de l'inverse d'une matrice, ainsi que des propriétés de cette inverse. Par contre, cette section ne décrit aucune technique permettant d'inverser une matrice, ce point faisant l'objet de la section section 8.5.

Pour simplifier la présentation, on ne considère ici que l'inverse de matrices réelles. L'inversion de matrices complexes est traitée dans la section 3.6.4.

1.4.1 Caractérisation des matrices inversibles

On expose ici des conditions nécessaires et suffisantes pour qu'une matrice soit inversible. La définition 1.4.1 qui suit n'est qu'une des nombreuses caractérisations possibles des matrices inversibles. D'autres caractérisations sont données ultérieurement dans cette section.

Définition 1.4.1

Si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices carrées de même taille telles que $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$, alors $\mathbf{A}$ est dite *inversible*, ou *non-singulière*, et la *matrice inverse* de $\mathbf{A}$, notée $\mathbf{A}^{-1}$, est la matrice $\mathbf{B}$, c'est-à-dire que $\mathbf{A}^{-1} = \mathbf{B}$.

Remarque

- ▷ Par symétrie de la définition, $\mathbf{B}$ est aussi inversible et son inverse est $\mathbf{A}$.
- ▷ Si $\mathbf{A}$ est inversible, alors $\mathbf{A}^{-1}$ l'est aussi, et $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.
- ▷ Une matrice *non-singulière* est une matrice *inversible* et une matrice *singulière* est une matrice *non-inversible*.

Le théorème 1.4.1 énonce plusieurs caractérisations des matrices inversibles sous forme de conditions nécessaires et suffisantes. Certaines de ces caractérisations ainsi que certaines notions qui leur sont associées font appel à des concepts qui ne sont abordés qu'ultérieurement dans cet ouvrage. Mais par souci de regroupement, il a été décidé de toutes les réunir dans un même théorème. Toutefois, les termes qui n'ont pas encore été définis à ce stade sont indiqués en italique. Les renvois appropriés sont également fournis (en bleu) afin de permettre au lecteur de retrouver chacune des démonstrations.

Théorème 1.4.1 (*caractérisations des matrices inverses*)
La matrice carrée $\mathbf{A} \in \mathbb{R}^{n\times n}$ est inversible si et seulement si :

1. Il existe $\mathbf{B} \in \mathbb{R}^{n\times n}$ telle que $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ (et $\mathbf{A}^{-1}$ est alors égale à $\mathbf{B}$).
2. Il existe $\mathbf{B} \in \mathbb{R}^{n\times n}$ telle que $\mathbf{AB} = \mathbf{I}$ ou $\mathbf{BA} = \mathbf{I}$ (et $\mathbf{A}^{-1}$ est alors égale à $\mathbf{B}$).
3. Il existe $\mathbf{B} \in \mathbb{R}^{n\times n}$ inversible telle que $\mathbf{B}^{-1} = \mathbf{A}$ (et $\mathbf{A}^{-1}$ est alors égale à $\mathbf{B}$).
4. Le déterminant de $\mathbf{A}$ est différent de zéro : $\det(\mathbf{A}) \neq 0$.
5. Toutes les *valeurs propres de* $\mathbf{A}$ sont non nulles (chapitre 7).
6. Les colonnes de $\mathbf{A}$ sont linéairement indépendantes.
7. Un *algorithme d'élimination* permet d'obtenir n *pivots (non nuls)* (chapitre 2).
8. $\mathbf{A}$ est de *plein rang* : $\mathrm{rg}(\mathbf{A}) = n$ (section 6.1.5).
9. Le *noyau de* $\mathbf{A}$ est réduit au vecteur nul : $\mathrm{Ker}(\mathbf{A}) = \{\mathbf{0}\}$ (section 6.1.1).
10. L'*image de* $\mathbf{A}$ remplit tout l'espace : $\mathrm{Im}(\mathbf{A}) = \mathbb{R}^n$ (section 6.1.3).
11. L'*application linéaire* associée à $\mathbf{A}$ est bijective (section 4.4).
12. Le *système d'équations linéaires* $\mathbf{Ax} = \mathbf{b}$ possède une solution unique quel que soit le vecteur $\mathbf{b}$ (chapitre 8).

Preuve.

- ▷ Le point 1 est bien une condition nécessaire et suffisante pour que $\mathbf{A}$ soit inversible puisqu'il s'agit de la définition 1.4.1 de l'inverse d'une matrice.
- ▷ Il est montré au chapitre 6 que les points 6 à 12 sont équivalents et que, tel que prouvé dans le théorème 2.4.2, une matrice carrée $\mathbf{A}$ de taille $n \times n$ est inversible si et seulement si l'algorithme d'élimination produit n pivots (non nuls), ce qui correspond au point 7. Ces sept points sont donc des caractérisations équivalentes des matrices inversibles.
- ▷ Le point 7 indique que $\mathbf{A}$ est inversible si et seulement si les pivots générés par l'algorithme d'élimination sont non nuls, et il est montré à la proposition 2.4.1 que $\det(\mathbf{A})$ est égal au produit des pivots. On déduit donc que $\mathbf{A}$ est inversible si et seulement si son déterminant est non nul. Mais il est également montré au chapitre 7 que $\det(\mathbf{A})$ est égal au produit des valeurs propres de $\mathbf{A}$, ce qui implique que $\mathbf{A}$ est inversible si et seulement si ses valeurs propres sont non nulles. Les points 4 et 5 sont donc deux autres caractérisations des matrices inversibles.
- ▷ Le point 2 est également une condition nécessaire et suffisante pour que $\mathbf{A}$ soit inversible. En effet :
  - ▷ Si $\mathbf{A}$ est inversible, on déduit de la définition 1.4.1 que $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$, et le point 2 est donc valide avec $\mathbf{B} = \mathbf{A}^{-1}$.

- ▷ Si le point 2 est valide, alors il existe une matrice $\mathbf{B}$ telle que $\mathbf{AB} = \mathbf{I}$ ou $\mathbf{BA} = \mathbf{I}$. On a donc, d'après le théorème 2.4.3, que $\det(\mathbf{A})\det(\mathbf{B}) = \det(\mathbf{I}) = 1$, ce qui implique que $\det(\mathbf{A}) \neq 0$ et donc que $\mathbf{A}$ est inversible (selon le point 4).

▷ Finalement, le point 3 caractérise aussi les matrices inversibles car :

- ▷ Si $\mathbf{A}$ est inversible, alors la matrice $\mathbf{B} = \mathbf{A}^{-1}$ est également inversible, ce qui implique que $\mathbf{B}^{-1} = (\mathbf{A}^{-1})^{-1} = \mathbf{A}$, et donc que le point 3 est valide.
- ▷ Si le point 3 est valide, alors il existe une matrice $\mathbf{B}$ inversible telle que $\mathbf{B}^{-1} = \mathbf{A}$. Comme $\mathbf{B}$ est inversible, son inverse $\mathbf{A}$ l'est également, ce qui donne $\mathbf{A}^{-1} = (\mathbf{B}^{-1})^{-1} = \mathbf{B}$.

□

Remarque

Les conditions du théorème 1.4.1 étant équivalentes, on déduit que pour montrer que des vecteurs sont linéairement indépendants, on peut calculer le déterminant de la matrice dont les colonnes sont les vecteurs et s'assurer que celui-ci n'est pas nul. Cela découle de l'équivalence des points 4 et 6.

Le point 2 est une simplification de la définition d'une matrice inverse. Il stipule qu'il suffit de trouver une matrice $\mathbf{B}$ qui vérifie une seule des multiplications qui donne l'identité, tel qu'illustré dans l'exemple 1.4.1.

Exemple 1.4.1 (*stratégie intuitive pour trouver l'inverse d'une matrice*)

Pour savoir si $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$ est inversible et pour déterminer son inverse, si tel est le cas, on peut bien sûr utiliser la proposition 8.5.1 qui indique comment inverser des matrices de taille 2×2. En cas d'incertitude sur les calculs à effectuer, il est possible de s'appuyer sur l'intuition ou d'effectuer quelques opérations rapides sur un brouillon pour, par exemple, obtenir la matrice

$$\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 0 & 1/2 \end{bmatrix} .$$

Pour s'assurer qu'il s'agit bien de l'inverse de $\mathbf{A}$, il suffit de calculer $\mathbf{AB}$ et de vérifier que ce produit matriciel donne bien l'identité :

$$\mathbf{AB} = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 2 \times 0 & 1(-1) + 2 \times 1/2 \\ 0 \times 1 + 2 \times 0 & 0(-1) + 2 \times 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}.$$

Comme tel est bien le cas, on peut conclure que $\mathbf{A}$ est inversible et que $\mathbf{B}$ est son inverse. Le point 2 du théorème 1.4.1 indique qu'il est inutile de vérifier que $\mathbf{BA} = \mathbf{I}$.

Remarque

D'autres conditions que celles du théorème 1.4.1 permettent de caractériser une matrice inversible, Il s'agit de variantes des propriétés déjà énoncées. Par exemple, $\mathbf{A} \in \mathbb{R}^{n\times n}$ est inversible si et seulement si

- ▷ ses lignes sont linéairement indépendantes (variation du point 6) ;
- ▷ ses lignes génèrent $\mathbb{R}^n$ (variation du point 10) ;
- ▷ le système $\mathbf{Ax} = \mathbf{b}$ admet au moins une solution quel que soit le vecteur $\mathbf{b}$ (variation du point 12) ;
- ▷ elle peut se *réduire* à l'identité (voir la proposition 2.1.2 sur la *forme échelonnée réduite*) ;
- ▷ sa transposée est inversible (proposition 1.4.3).

Les preuves de ces points sont laissées en exercice.

Remarque

Si on s'intéresse à des matrices ayant certaines formes particulières, d'autres énoncés sont possibles. Par exemple :

- ▷ Une matrice triangulaire est inversible si et seulement si toutes les composantes sur sa diagonale sont non nulles. En effet, il découle de la proposition 1.4.5 ainsi que du point 4 du théorème 1.4.1 qu'une telle matrice est inversible si et seulement si son déterminant est non nul, et la proposition 1.3.1 indique que le déterminant d'une matrice triangulaire est égal au produit des composantes sur sa diagonale.
- ▷ Toutes les *matrices d'élimination* sont inversibles (voir la section 2.3.1).

1.4.2 Autres propriétés des matrices inverses

La section précédente contenait des conditions nécessaires et suffisantes pour qu'une matrice soit inversible. Cette section présente quelques propriétés de l'inverse d'une matrice.

La proposition 1.4.1 suivante, qui porte sur l'inverse d'un produit matriciel, a déjà été énoncée au point 2 du théorème 1.2.2. On dispose désormais des outils permettant de la prouver.

Proposition 1.4.1 (*inverse d'un produit de matrices*)

Si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices inversibles de même taille, leur produit $\mathbf{AB}$ l'est également et $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

Preuve. Soit $\mathbf{M} = \mathbf{B}^{-1}\mathbf{A}^{-1}$. On a $(\mathbf{AB})\mathbf{M} = \mathbf{ABB}^{-1}\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}$. On déduit du point 2 du théorème 1.4.1 que la matrice $\mathbf{AB}$ est inversible et que son inverse est $\mathbf{M}$. □

Le corollaire 1.4.1 est une généralisation de la proposition 1.4.1 et permet de déduire le corollaire 1.4.2.

Corollaire 1.4.1 (*inverse d'un produit de plusieurs matrices*)

Si $\mathbf{M}_1$, $\mathbf{M}_2$, …, $\mathbf{M}_p$ sont $p \geq 2$ matrices inversibles de même taille, leur produit est également inversible et $(\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_p)^{-1} = \mathbf{M}_p^{-1}\mathbf{M}_{p-1}^{-1}\cdots\mathbf{M}_1^{-1}$.

Preuve. La preuve est par induction sur p.

- ▷ **Initialisation.** Si $p = 2$, l résultat est valide puisqu'il s'agit de la proposition 1.4.1.
- ▷ **Induction.** Soit $p > 2$ et supposons que le résultat est vrai pour le produit de $p-1$ matrices, c'est-à-dire que $(\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_{p-1})^{-1} = \mathbf{M}_{p-1}^{-1}\mathbf{M}_{p-2}^{-1}\cdots\mathbf{M}_1^{-1}$. Il découle de la proposition 1.4.1 que :

$$\begin{aligned}(\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_p)^{-1} &= ((\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_{p-1})\mathbf{M}_p)^{-1} = \mathbf{M}_p^{-1}(\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_{p-1})^{-1}\\ &= \mathbf{M}_p^{-1}\mathbf{M}_{p-1}^{-1}\cdots\mathbf{M}_1^{-1}.\end{aligned}$$

□

Corollaire 1.4.2 (*inverse et puissance de matrice*)

Si $\mathbf{A}$ est une matrice inversible et si p est un entier, alors $\mathbf{A}^p$ est également inversible et $(\mathbf{A}^p)^{-1} = (\mathbf{A}^{-1})^p$.

Preuve. C'est un cas particulier du corollaire 1.4.1 avec $\mathbf{A} = \mathbf{M}_i$ pour tout $i \in [\![1;p]\!]$. □

La proposition 1.4.2 démontre que la multiplication une matrice carrée singulière par une matrice de même taille donne toujours une matrice singulière.

Proposition 1.4.2 (**A** *singulière implique* **AB** *singulière*)

Si $\mathbf{A}$ et $\mathbf{B}$ sont deux matrices carrées de même taille et si $\mathbf{A}$ est singulière, alors $\mathbf{AB}$ est également singulière.

Preuve. Il découle du théorème 2.4.3 que $\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$. Si $\mathbf{A}$ est singulière, on déduit du point 4 du théorème 1.4.1 que $\det(\mathbf{A}) = 0$ et donc que $\det(\mathbf{AB}) = 0$, ce qui signifie que $\mathbf{AB}$ est singulière. □

La proposition 1.4.3 indique que l'inverse de la transposée d'une matrice carrée $\mathbf{A}$ inversible est la transposée de l'inverse de $\mathbf{A}$, d'où on déduit le corollaire 1.4.3 qui stipule que l'inverse d'une matrice symétrique est symétrique.

Proposition 1.4.3 (*transposée de l'inverse*)

Si $\mathbf{A}$ est une matrice inversible, alors $\mathbf{A}^\top$ l'est également et $\left(\mathbf{A}^\top\right)^{-1} = \left(\mathbf{A}^{-1}\right)^\top$.

Preuve. La proposition 1.2.1 montre que $\mathbf{A}^\top\left(\mathbf{A}^{-1}\right)^\top = \left(\mathbf{A}^{-1}\mathbf{A}\right)^\top = \mathbf{I}^\top = \mathbf{I}$. On déduit donc du point 2 du théorème 1.4.1 que $\mathbf{A}^\top$ est inversible et que son inverse est $\left(\mathbf{A}^{-1}\right)^\top$. □

Corollaire 1.4.3 (*symétrie de l'inverse*)

Si $\mathbf{A}$ est une matrice inversible et symétrique, alors $\mathbf{A}^{-1}$ est également symétrique.

Preuve. Si $\mathbf{A}$ est une matrice inversible et symétrique, on déduit de la proposition 1.4.3 que $\left(\mathbf{A}^{-1}\right)^\top = \left(\mathbf{A}^\top\right)^{-1} = \mathbf{A}^{-1}$, ce qui signifie que $\mathbf{A}^{-1}$ est symétrique. □

La proposition 1.4.4 montre qu'une matrice triangulaire inversible et son inverse ont la même forme.

Proposition 1.4.4 (*inverse d'une matrice triangulaire inversible*)

L'inverse d'une matrice triangulaire inférieure (resp. supérieure) inversible est également triangulaire inférieure (resp. supérieure).

Preuve.

Il suffit de faire la démonstration pour les matrices triangulaires inférieures. En effet, si la propriété est vraie pour les matrices triangulaires inférieures et si $\mathbf{A}$ est triangulaire supérieure inversible, on déduit du théorème 1.3.1 que $\mathbf{A}^\top$ est inversible et que $(\mathbf{A}^{-1})^\top = (\mathbf{A}^\top)^{-1}$. Étant donné que $\mathbf{A}^\top$ est triangulaire inférieure, la matrice $(\mathbf{A}^\top)^{-1}$ est également triangulaire inférieure, ce qui signifie que $\mathbf{A}^{-1}$ est triangulaire supérieure.

Soit donc $\mathbf{A}$ une matrice triangulaire inférieure inversible de taille $n \times n$. On déduit du point 2 du théorème 1.4.1 qu'il existe une matrice $\mathbf{B}$ telle que $\mathbf{AB} = \mathbf{I}$, et il découle alors du théorème 2.4.3 que $\det(\mathbf{I}) = \det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) = 1$, ce qui implique $\det(\mathbf{A}) \neq 0$. La proposition 1.3.1 indique que le déterminant de $\mathbf{A}$ est égal au produit des composantes sur sa diagonale, ce qui implique que $\mathbf{A}(i,i) \neq 0$ pour tout $i \in [\![1;n]\!]$. Il reste à montrer que $\mathbf{B}$ est triangulaire inférieure. Il faut donc vérifier que quel que soit $i \in [\![1;n]\!]$, on a $\mathbf{B}(i,j) = 0$ pour tout $i < j \leq n$. La preuve est par induction sur i.

- ▷ **Initialisation.** Pour $i = 1$ et $j \in [\![2;n]\!]$, on a $0 = \mathbf{I}(1,j) = (\mathbf{AB})(1,j) = \mathbf{A}(1,1)\mathbf{B}(1,j)$ et donc $\mathbf{B}(1,j) = 0$.

- ▷ **Induction.** Soit $i > 1$ et supposons que $\mathbf{B}(k, j) = 0$ pour tout $k < i \leq j$. Si $j > i$, on a $0 = \mathbf{I}(i, j) = (\mathbf{AB})(i, j) = \sum_{k=1}^{i} \mathbf{A}(i, k)\mathbf{B}(k, j) = \mathbf{A}(i, i)\mathbf{B}(i, j)$ et donc $\mathbf{B}(i, j) = 0$.

□

Le point 4 du théorème 1.4.1 indique qu'une matrice carrée est inversible si et seulement si son déterminant est non nul. La proposition 1.4.5 donne une preuve alternative de ce résultat pour les matrices triangulaires.

Proposition 1.4.5 (*déterminant et inverse d'une matrice triangulaire inversible*)
Une matrice triangulaire est inversible si et seulement si son déterminant est non nul.

Preuve. Si $\mathbf{A}$ une une matrice triangulaire inversible, on déduit du point 2 du théorème 1.4.1 qu'il existe une matrice $\mathbf{B}$ telle que $\mathbf{AB} = \mathbf{I}$, et il découle alors du théorème 2.4.3 que $\det(\mathbf{I}) = \det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B}) = 1$, ce qui implique que $\det(\mathbf{A}) \neq 0$. Il reste donc à montrer que si $\mathbf{A}$ est une matrice triangulaire de déterminant non nul, alors $\mathbf{A}$ est inversible.

Il suffit de faire la démonstration pour les matrices triangulaires inférieures. En effet, si la propriété est vraie pour les matrices triangulaires inférieures et si $\mathbf{A}$ est triangulaire supérieure de déterminant non nul, alors il découle du théorème 1.3.1 que $\det(\mathbf{A}^\top) = \det(\mathbf{A}) \neq 0$, et donc que $\mathbf{A}^\top$ est inversible (puisque$\mathbf{A}^\top$ est triangulaire inférieure). On déduit alors de la proposition 1.4.3 que $\mathbf{A}$ est également inversible.

Soit donc $\mathbf{A}$ une matrice triangulaire inférieure de taille $n \times n$ telle que $\det(\mathbf{A}) \neq 0$. On a $\mathbf{A}(i, i) \neq 0$ pour tout $i \in [\![1; n]\!]$ car $\det(\mathbf{A}) = \prod_{i=1}^{n} \mathbf{A}(i, i)$ (selon la proposition 1.3.1).

On déduit du point 2 du théorème 1.4.1 que pour prouver que $\mathbf{A}$ est inversible, il suffit d'exhiber une matrice $\mathbf{B}$ telle que $\mathbf{AB} = \mathbf{I}$. Soit donc $\mathbf{B}$ une telle matrice. Il découle de la proposition 1.4.4 que $\mathbf{B}$ est également triangulaire inférieure, et donc que pour tout $j > i$, on a $\mathbf{B}(i, j) = 0$ et $(\mathbf{AB})(i, j) = \sum_{i=1}^{n} \mathbf{A}(i, k)\mathbf{B}(k, j) = 0 = \mathbf{I}(i, j)$. Il reste à montrer que pour tout $i \geq j$, en imposant $(\mathbf{AB})(i, j) = \mathbf{I}(i, j)$, les composantes $\mathbf{B}(i, j)$ peuvent être déterminées quelles que soient les composantes de $\mathbf{A}$.

- ▷ si $i = j$, on a $(\mathbf{AB})(i, i) = \mathbf{A}(i, i)\mathbf{B}(i, i)$, et pour avoir $(\mathbf{AB})(i, i) = \mathbf{I}(i, i) = 1$, on pose donc $\mathbf{B}(i, i) = \frac{1}{\mathbf{A}(i,i)}$ (qui est bien défini puisque $\mathbf{A}$ n'a pas de zéro sur la diagonale).
- ▷ Les composantes $\mathbf{B}(i, j)$ pour $i > j$ sont déterminées séquentiellement avec j allant de 1 à $n - 1$ et pour chacun de ces indices j de colonne, on considère les indices i allant de $j + 1$ à n. Ainsi, pour avoir $(\mathbf{AB})(i, j) = \mathbf{I}(i, j) = 0$, on a

$$\sum_{k=j}^{i} \mathbf{A}(i, k)\mathbf{B}(k, j) = 0 \qquad \Leftrightarrow \qquad \mathbf{B}(i, j) = -\frac{\sum\limits_{k=j}^{i-1} \mathbf{A}(i, k)\mathbf{B}(k, j)}{\mathbf{A}(i, i)}$$

ce qui montre que $\mathbf{B}(i, j)$ peut facilement être calculé puisque $\mathbf{A}(i, i) \neq 0$ et que tous les termes $\mathbf{B}(k, j)$ avec k allant de j à $i - 1$ sont connus lorsqu'on calcule $\mathbf{B}(i, j)$.

□

Tel qu'indiqué au point 12 du théorème 1.4.1, une matrice carrée $\mathbf{A}$ est inversible si et seulement si il existe un unique vecteur $\mathbf{x}$ qui vérifie $\mathbf{A} = \mathbf{b}$, quel que soit le vecteur $\mathbf{b}$. Le théorème 1.4.2 donne plus de précisions sur ce vecteur $\mathbf{x}$ lorsque $\mathbf{A}$ est inversible.

Théorème 1.4.2 (*solution unique d'un système*)

Si $\mathbf{A}$ est une matrice inversible de taille $n \times n$ et $\mathbf{b}$ un vecteur quelconque à n composantes, alors l'unique vecteur $\mathbf{x}$ qui vérifie $\mathbf{Ax} = \mathbf{b}$ est $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Preuve. Si $\mathbf{A}$ est une matrice inversible, on déduit du point 12 du théorème 1.4.1 qu'il existe un unique vecteur $\mathbf{x}$ qui vérifie $\mathbf{Ax} = \mathbf{b}$, et comme $\mathbf{A}(\mathbf{A}^{-1}\mathbf{b}) = (\mathbf{AA}^{-1})\mathbf{b} = \mathbf{Ib} = \mathbf{b}$, il s'ensuit que $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ est cet unique vecteur. □

Si $\mathbf{A}$ est une matrice inversible, le théorème 1.4.2 donne une formule directe pour déterminer l'unique vecteur qui vérifie $\mathbf{Ax} = \mathbf{b}$. Cependant, en pratique, il est extrêmement coûteux en termes de calculs d'inverser une matrice. Pour déterminer ce vecteur $\mathbf{x}$, on n'inverse donc jamais explicitement une matrice. On préfère utiliser les techniques vues au chapitre 8 basées sur le principe d'élimination qui est le sujet du chapitre 2.

Pour conclure cette section, notons que lorsqu'une matrice n'est pas inversible, on peut recourir à la notion de *pseudo-inverse*, qui est abordée à la section section 6.4.

1.5 Étude de cas : Produit de Kronecker

Le *produit de Kronecker* de deux matrices $\mathbf{A} \in \mathbb{R}^{m\times n}$ et $\mathbf{B} \in \mathbb{R}^{p\times q}$ est défini par blocs de la manière suivante :

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} \mathbf{A}(1,1)\mathbf{B} & \mathbf{A}(1,2)\mathbf{B} & \cdots & \mathbf{A}(1,n)\mathbf{B} \\ \mathbf{A}(2,1)\mathbf{B} & \mathbf{A}(2,2)\mathbf{B} & \cdots & \mathbf{A}(2,n)\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}(m,1)\mathbf{B} & \mathbf{A}(m,2)\mathbf{B} & \cdots & \mathbf{A}(m,n)\mathbf{B} \end{bmatrix} \in \mathbb{R}^{(mp)\times(nq)}.$$

Pour rappel, $\mathbf{A}(i,j) \in \mathbb{R}$ désigne l'entrée (i,j) de $\mathbf{A}$, pour $i \in [\![1;m]\!]$ et $j \in [\![1;n]\!]$.

Bien que très utile dans de nombreux domaines, le produit de Kronecker est peu étudié dans les premiers cours d'algèbre linéaire notamment parce que la plupart de ses propriétés sont souvent considérées comme trop techniques pour être démontrées à ce stade. L'idée de cette étude de cas n'est pas de prouver des propriétés mais plutôt de les énoncer et de les illustrer par des exemples simples afin de mettre en lumière la puissance du produit de Kronecker. Ensuite, certaines de ces propriétés seront appliquées au contexte de l'apprentissage automatique.

1.5.1 Propriétés utiles

Dans ce qui suit, les tailles des matrices sont supposées compatibles pour que les produits puissent être effectués.

Les propriétés suivantes sont vérifiées pour tout $\mathbf{A} \in \mathbb{R}^{m\times n}$ et $\mathbf{B} \in \mathbb{R}^{p\times q}$ compatibles :

$$\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{(mp)\times(nq)} \qquad \text{(tailles)} \tag{1.1}$$

$$(\mathbf{A}_1 + \mathbf{A}_2) \otimes \mathbf{B} = \mathbf{A}_1 \otimes \mathbf{B} + \mathbf{A}_2 \otimes \mathbf{B} \qquad \text{(linéarité à droite)} \tag{1.2}$$

$$\mathbf{A} \otimes (\mathbf{B}_1 + \mathbf{B}_2) = \mathbf{A} \otimes \mathbf{B}_1 + \mathbf{A} \otimes \mathbf{B}_2 \qquad \text{(linéarité à gauche)} \tag{1.3}$$

$$(c\mathbf{A}) \otimes \mathbf{B} = \mathbf{A} \otimes (c\mathbf{B}) = c(\mathbf{A} \otimes \mathbf{B}) \qquad \text{(scalaires)} \tag{1.4}$$

$$(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top \qquad \text{(transposition)} \tag{1.5}$$

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{M} \otimes \mathbf{D}) = (\mathbf{AM}) \otimes (\mathbf{BD}) \qquad \text{(produit mixte)} \tag{1.6}$$

$$(\mathbf{A} \otimes \mathbf{I}_r)(\mathbf{I}_r \otimes \mathbf{B}) = \mathbf{A} \otimes \mathbf{B} \qquad \text{(cas particulier du produit mixte)} \tag{1.7}$$

$$\mathrm{rg}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{rg}(\mathbf{A}) \times \mathrm{rg}(\mathbf{B}) \qquad \text{(rang)} \tag{1.8}$$

$$\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{tr}(\mathbf{A}) \times \mathrm{tr}(\mathbf{B}) \qquad \text{(trace)} \tag{1.9}$$

D'autres identités classiques incluent la suivante qui concerne les matrices carrées : pour $\mathbf{A} \in \mathbb{R}^{m\times m}$ et $\mathbf{B} \in \mathbb{R}^{p\times p}$ on a

$$\det(\mathbf{A} \otimes \mathbf{B}) = \det(\mathbf{A})^p \det(\mathbf{B})^m.$$

Aussi, pour une matrice $\mathbf{M} \in \mathbb{R}^{m\times n}$ quelconque, notons $\mathrm{vec}(\mathbf{M})$ le vecteur obtenu en empilant les colonnes de $\mathbf{M}$:

$$\mathrm{vec}(\mathbf{M}) = \begin{bmatrix} \mathbf{M}(1,1) \\ \mathbf{M}(2,1) \\ \vdots \\ \mathbf{M}(m,1) \\ \mathbf{M}(1,2) \\ \mathbf{M}(2,2) \\ \vdots \\ \mathbf{M}(m,n) \end{bmatrix} \in \mathbb{R}^{mn}.$$

Étant données $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{B} \in \mathbb{R}^{p\times q}$ et $\mathbf{M} \in \mathbb{R}^{n\times p}$, on a

$$\mathrm{vec}(\mathbf{AMB}) = (\mathbf{B}^\top \otimes \mathbf{A})\,\mathrm{vec}(\mathbf{M}),$$

Cette dernière identité est très utile pour vectoriser des égalités matricielles, par exemple dans le contexte de la régression linéaire.

1.5.2 Exemple

Dans cette partie, quelques propriétés du produit de Kronecker sont vérifiées. Considérons

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad \mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Par définition,

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} 1\,\mathbf{B} & 2\,\mathbf{B} \\ 0\,\mathbf{B} & 1\,\mathbf{B} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 4 & 0 \\ -1 & 3 & -2 & 6 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 3 \end{bmatrix}.$$

On retrouve bien la propriété (1.1) qui indique que $\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{4\times 4}$.

Aussi, on vérifie aisément la propriété (1.9) puisque $\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) = 10 = 2 \times 5 = \mathrm{tr}(\mathbf{A}) \times \mathrm{tr}(\mathbf{B})$.

Pour vérifier la propriété (1.7) avec la matrice $\mathbf{I}_2$, on peut calculer $\mathbf{A} \otimes \mathbf{I}_2$ et $\mathbf{I}_2 \otimes \mathbf{B}$, et ensuite multiplier ces deux matrices :

$$\mathbf{A}\otimes\mathbf{I}_2 = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{I}_2\otimes\mathbf{B} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 3 \end{bmatrix} \quad (\mathbf{A}\otimes\mathbf{I}_2)(\mathbf{I}_2\otimes\mathbf{B}) = \begin{bmatrix} 2 & 0 & 4 & 0 \\ -1 & 3 & -2 & 6 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & -1 & 3 \end{bmatrix}.$$

On constate que $(\mathbf{A} \otimes \mathbf{I}_2)(\mathbf{I}_2 \otimes \mathbf{B}) = \mathbf{A} \otimes \mathbf{B}$.

1.5.3 Applications

Imaginons, dans un cadre d'apprentissage automatique, que l'on souhaite utiliser sur son ordinateur personnel un modèle de traitement du langage naturel tel que GPT-2. Le problème est que le modèle est trop volumineux pour tenir dans la mémoire de l'ordinateur. On peut dans ce cas utiliser ce qu'on appelle la *factorisation tensorielle* pour décomposer le modèle en un produit de Kronecker de matrices plus petites, qui peuvent être chargées en mémoire. C'est précisément l'objet de cette section

De façon schématique, le modèle de traitement du langage se compose de *couches d'attention* qui constituent un mécanisme permettant de se concentrer sur les parties importantes d'un texte à analyser, de capturer les dépendances entre les mots, ce qui est essentiel pour comprendre le contexte global d'un texte et générer des réponses appropriées. Pour ceux qui souhaitent approfondir le sujet, cette architecture est à la base des modèles de langage actuels (Gemini, Claude, Mistral, LLaMA, GPT, entre autres) comme expliqué dans l'article Attention Is All You Need.

Les couches d'attention sont représentées par des matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_c$, avec $\mathbf{A}_i \in \mathbb{R}^{e_i \times s_i}$ pour $i \in [\![1; c]\!]$. Les éléments de ces matrices, appelés *paramètres*, sont des réels.

Le théorème suivant, à admettre sans démonstration, établit que chaque matrice peut être factorisée en une somme de produits de Kronecker de matrices plus petites, réduisant ainsi l'espace mémoire nécessaire au modèle

Théorème 1.5.1 (*factorisation de Kronecker d'une matrice*)

Quelle que soit la matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$, il existe $k \in \mathbb{N}^*$ ainsi que deux familles de matrices $\{\mathbf{A}_{1,i}\}_{i \in [\![1;k]\!]}$, $\{\mathbf{A}_{2,i}\}_{i \in [\![1;k]\!]}$ telles que, pour tout $i \in [\![1; k]\!]$, $\mathbf{A}_{1,i} \in \mathbb{R}^{m_1 \times n_1}$ et $\mathbf{A}_{2,i} \in \mathbb{R}^{m_2 \times n_2}$, avec $m_1 m_2 = m$ et $n_1 n_2 = n$, et

$$\mathbf{A} = \mathbf{A}_{1,1} \otimes \mathbf{A}_{2,1} + \mathbf{A}_{1,2} \otimes \mathbf{A}_{2,2} + \ldots + \mathbf{A}_{1,k} \otimes \mathbf{A}_{2,k}.$$

En d'autres mots, le Théorème 1.5.1 établit que toute matrice peut être représentée comme une somme finie de produits de Kronecker de matrices plus petites, ce qui réduit l'espace mémoire requis. Il convient à présent d'examiner le gain de mémoire offert par une telle factorisation.

Premier cas : GPT-2 base

Le modèle de base de GPT-2 est constitué de 48 couches d'attention, chacune de taille 768×768. Le nombre total de paramètres liés à ces matrices est donc $48 \times 768 \times 768 = 28,311,552$. La plupart du temps, ces paramètres sont stockés en 32 bits, soit 4 octets. Ces paramètres occupent donc environ 113.2 Mo de mémoire

Supposons que chacune de ces couches est factorisable en une somme de k produits de

Kronecker, avec $\mathbf{A}_{1,i} \in \mathbb{R}^{16\times 48}$, $\mathbf{A}_{2,i} \in \mathbb{R}^{48\times 16}$ pour tout i. Le nombre total de paramètres à stocker pour une seule factorisation est alors

$$k \times (16 \times 48 + 16 \times 48) = k \times 1536.$$

Pour 48 couches, le nombre total de paramètres devient

$$48 \times k \times 1536 = k \times 73,728.$$

Si l'on choisit $k = 10$ (ce qui est assez réaliste), le nombre total de paramètres à mémoriser est donc $737,280$, ce qui requiert environ 3 Mo. Le taux de compression est donc de

$$\tau = \frac{28,311,552}{737,280} \approx 38.4.$$

Second cas : GPT-2 XL

Le modèle de base de GPT-2 XL est constitué de 192 couches d'attention, chacune de taille 1600×1600. Le nombre total de paramètres liés à ces matrices est donc $192 \times 1600 \times 1600 = 491,520,000$, ce qui requiert environ 2 Go. de mémoire.

Supposons, que chacune de ces couches est factorisable en une somme de k produits de Kronecker, avec $\mathbf{A}_{1,i} \in \mathbb{R}^{32\times 50}$, $\mathbf{A}_{2,i} \in \mathbb{R}^{50\times 32}$ pour tout i. Le nombre total de paramètres à stocker pour une seule factorisation est alors

$$k \times (32 \times 50 + 32 \times 50) = k \times 3200.$$

Pour 192 couches, le nombre total de paramètres devient

$$192 \times k \times 3200 = k \times 614,400.$$

Si on choisit $k = 20$ (ce qui est à nouveau, assez réaliste), le nombre total de paramètres à mémoriser est $12,288,000$, ce qui requiert environ 49.2 Mo. Le taux de compression est donc de

$$\tau = \frac{491,520,000}{12,288,000} = 40.0.$$

Ainsi, que l'on considère le modèle de base ou le modèle XL, la factorisation de Kronecker permet de réduire significativement le nombre de paramètres à stocker. Évidemment, la structure du modèle a été simplifiée pour l'exemple ; les modèles actuels comptent souvent plusieurs centaines de milliards de paramètres. En pratique, le gain de compression sur l'ensemble du modèle peut atteindre l'ordre de la centaine, voire davantage, car les couches d'attention ne représentent qu'une fraction du total des paramètres.

1.8 Exercices sur les vecteurs et les matrices

Les exercices de ce chapitre se concentrent sur la manipulation de vecteurs et de matrices. La résolution de systèmes et l'inversion matricielle ne sont pas concernées car vues plus loin, au chapitre 8, même si ici on peut quand même manipuler des systèmes et des inverses de matrices. Les solutions sont disponibles à l'annexe A.

Série d'exercices 1: Matrices et vecteurs

Exercice 1.1: vrai ou faux ? ($\star$)
Les énoncés suivants sont-ils vrais ou faux ?

1. Soient $\mathbf{A}$ et $\mathbf{B}$ dans $\mathbb{R}^{m\times n}$. On a $(\mathbf{A}+\mathbf{B})^2 = \mathbf{A}^2 + 2\mathbf{A}\mathbf{B} + \mathbf{B}^2$.
2. Soient $\mathbf{A}$ et $\mathbf{B}$ dans $\mathbb{R}^{m\times n}$. Avec $\mathbf{B}$ inversible, on a $\operatorname{tr}(\mathbf{B}^{-1}\mathbf{A}\mathbf{B}) = \operatorname{tr}(\mathbf{A})$.
3. Toute matrice triangulaire est inversible.
4. Soient $\mathbf{A}$ et $\mathbf{B}$ dans $\mathbb{R}^{n\times n}$. Si $\mathbf{A}\mathbf{B}$ est inversible, alors $\mathbf{A}$ et $\mathbf{B}$ sont inversibles.
5. Soient $\mathbf{A}$ une matrice quelconque et $\mathbf{b}$ un vecteur quelconque. Si le système $\mathbf{A}\mathbf{x} = \mathbf{b}$ possède une solution unique, alors la matrice $\mathbf{A}$ est inversible.
6. Soient $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ et $\mathbf{D}$ des matrices inversibles de $\mathbb{R}^{n\times n}$. La matrice par blocs $\begin{bmatrix}\mathbf{A} & \mathbf{C}\\ \mathbf{B} & \mathbf{D}\end{bmatrix}$ est inversible et son inverse est donnée par $\begin{bmatrix}\mathbf{A}^{-1} & \mathbf{C}^{-1}\\ \mathbf{B}^{-1} & \mathbf{D}^{-1}\end{bmatrix}$.
7. Trois opérations élémentaires sont effectuées sur une matrice $\mathbf{A} \in \mathbb{R}^{n\times n}$ dans l'ordre suivant :
   (a) La deuxième ligne de $\mathbf{A}$ est multipliée par $\alpha \in \mathbb{R}^*$;
   (b) La première et la deuxième ligne de $\mathbf{A}$ sont permutées ;
   (c) La troisième ligne est ajoutée trois fois à la première ligne de $\mathbf{A}$.

   La matrice issue de ces trois transformations est notée $\mathbf{A}'$, et on a
   $$\det(\mathbf{A}') = -\alpha \det(\mathbf{A}) .$$

Exercice 1.2: ($\star$)
Soient $\mathbf{A} = \begin{bmatrix}1 & 2 & 3\\ 2 & 3 & 3\\ 1 & 4 & 1\end{bmatrix} \in \mathbb{R}^{3\times 3}$, $\mathbf{b} = \begin{bmatrix}1\\1\\1\end{bmatrix}$, et $\mathbf{c} = \begin{bmatrix}6\\8\\6\end{bmatrix} \in \mathbb{R}^3$.

1. Vérifier que $\mathbf{Ab} = \mathbf{c}$;
2. Écrire $\mathbf{c}$ comme une combinaison linéaire des colonnes de $\mathbf{A}$;
3. Sachant que $4\mathbf{b} = -\begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}2\\3\\4\end{bmatrix} + \begin{bmatrix}3\\3\\1\end{bmatrix}$, trouver une solution au système $\mathbf{Ax} = \mathbf{b}$.

Exercice 1.3: ($\star$)

Exprimer le produit $\mathbf{Ax}$ comme une combinaison linéaire des colonnes de $\mathbf{A}$, avec $\mathbf{A} = \begin{bmatrix}-1 & 1 & 2\\ 0 & 1 & -3\\ 1 & -2 & 1\\ 7 & 3 & -1\end{bmatrix}$ et $\mathbf{x} = (x_1, x_2, x_3)$.

Exercice 1.4: ($\star$)

Exprimer la combinaison linéaire suivante comme le produit d'une matrice $\mathbf{A}$ et d'un vecteur $\mathbf{x}$:

$$x_1 \begin{bmatrix}2\\1\end{bmatrix} + x_2 \begin{bmatrix}1\\1\end{bmatrix} + x_3 \begin{bmatrix}-1\\0\end{bmatrix} + x_4 \begin{bmatrix}6\\-4\end{bmatrix}$$

Exercice 1.5: ($\star$)

Soient $\mathbf{a} = \begin{bmatrix}2\\0\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}3\\3\end{bmatrix}$ et $\mathbf{c} = \begin{bmatrix}-\sqrt{3}\\1\end{bmatrix} \in \mathbb{R}^2$.

1. Calculer la norme de ces trois vecteurs.
2. Calculer $\langle \mathbf{a}, \mathbf{b} \rangle$ et $\langle \mathbf{a}, \mathbf{c} \rangle$ de deux façons différentes.
3. Sachant que $\mathbf{b} = \frac{3\sqrt{3}+3}{2}\mathbf{a} + 3\mathbf{c}$, en déduire $\langle \mathbf{b}, \mathbf{c} \rangle$.

Exercice 1.6: théorème de Pythagore ($\star\star$)

Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

1. Montrer que $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$.
2. Si, de plus, $\mathbf{x}$ et $\mathbf{y}$ sont orthogonaux, montrer que $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$.

Exercice 1.7: cas d'égalité de l'inégalité triangulaire ($\star\star\star$)

Soient $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ non nuls.

1. Montrer que si $\mathbf{y} = \alpha\mathbf{x}$ avec $\alpha > 0$ alors $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$.
2. Montrer que si $\mathbf{y} = \alpha\mathbf{x}$ avec $0 > \alpha > -1$ alors $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| - \|\mathbf{y}\|$.
3. Montrer que si $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\|\,\|\mathbf{y}\|$ alors $\mathbf{x}$ et $\mathbf{y}$ sont colinéaires en étudiant la norme de $\mathbf{y} - \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|^2}\mathbf{x}$.

4. Montrer que si $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\|$ alors $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \, \|\mathbf{y}\|$.
5. En déduire une condition nécessaire et suffisante sur la cas d'égalité de l'inégalité triangulaire.

Exercice 1.8: équation d'un plan généré par deux vecteurs ($\star$)
Soient $\mathbf{u} = (1, 2, 3)$ et $\mathbf{v} = (0, 1, 1)$ deux vecteurs de $\mathbb{R}^3$. Donner l'équation du plan $\mathrm{Vect}(\mathbf{u}, \mathbf{v})$.

Exercice 1.9: produits et additions compatibles ? ($\star$)
Soient les matrices et vecteurs suivants :

$$\mathbf{A} \in \mathbb{R}^{3\times 2}, \quad \mathbf{B} \in \mathbb{R}^{2\times 4}, \quad \mathbf{x} \in \mathbb{R}^2, \quad \mathbf{y} \in \mathbb{R}^3.$$

Pour chacune des expressions ci-dessous, indiquer si l'opération est bien définie. Si oui, préciser le type (matrice ou vecteur) et les dimensions du résultat. Sinon, expliquer pourquoi l'opération est impossible.

1. $\mathbf{Ax}$
2. $\mathbf{BA}$
3. $\mathbf{A} + \mathbf{y}$
4. $\mathbf{A}^\top \mathbf{y}$
5. $\mathbf{x}^\top \mathbf{A}$

Exercice 1.10: ($\star$)
Soient $\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 4 & 0 & 2 \end{bmatrix}$, et $\mathbf{B} = \begin{bmatrix} -2 & 3 \\ 1 & 2 \\ 6 & 0 \end{bmatrix}$.

1. Calculer $\mathbf{AB}$ selon les quatre points de vue du cours.
2. Calculer $\mathbf{BA}$ selon les quatre points de vue du cours.

Exercice 1.11: ($\star$)
Soient les matrices et vecteurs suivants :

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ -1 & 1 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \\ 1 & -2 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & 1 \end{bmatrix}$$

$$\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 3 \end{bmatrix}$$

Pour chacune des expressions suivantes, déterminer si l'opération est bien définie. Si oui, préciser les dimensions du résultat et le calculer. Si non, expliquer pourquoi.

1. $\mathbf{A} + \mathbf{B}^\top$
2. $(\mathbf{CB})^\top + \mathbf{A}$
3. $\mathbf{Au} + \mathbf{Bv}$
4. $\mathbf{C}(\mathbf{Au}) - 2\mathbf{v}$
5. $(\mathbf{A} + \mathbf{A}^\top)\mathbf{u}$

Exercice 1.12: $(\star)$

Soient $\mathbf{A} = \begin{bmatrix} \cos(a) & -\sin(a) \\ \sin(a) & \cos(a) \end{bmatrix}, \mathbf{B} = \begin{bmatrix} -\cos(b) & -\sin(b) \\ \sin(b) & \cos(b) \end{bmatrix}$, avec $a, b \in \mathbb{R}$.

1. Calculer $\mathbf{AB}$.
2. Calculer $\mathbf{A}^2$.
3. Calculer $\mathbf{A}^n$.

Exercice 1.13: $(\star\star)$

Soit $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}$. Trouver les matrices telles que $\mathbf{AB} = \mathbf{BA}$.

Exercice 1.14: $(\star\star)$

Soit $\mathbf{A} \in \mathbb{R}^{n\times n}$ avec des 1 sur la surdiagonale et des 0 ailleurs. Calculer $\mathbf{A}^k$ pour $k \in \mathbb{N}$.

Exercice 1.15: structure de la matrice noyau $(\star)$

Soit $\mathbf{R} = \begin{bmatrix} \mathbf{A} & \mathbf{F} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \in \mathbb{R}^{m\times n}$ et $\mathbf{N} = \begin{bmatrix} -\mathbf{F} \\ \mathbf{B} \end{bmatrix} \in \mathbb{R}^{n\times(n-r)}$ avec r un entier tel que $r \leq \min\{m, n\}$.

1. Si $\mathbf{A} \in \mathbb{R}^{r\times r}$, donner les tailles de $\mathbf{F}$, $\mathbf{C}$, $\mathbf{D}$, et aussi $\mathbf{B}$ si on veut que le produit $\mathbf{RN}$ soit défini.
2. Exprimer le produit $\mathbf{RN}$.
3. Si $\mathbf{A}$ et $\mathbf{B}$ sont des matrices identité, et si $\mathbf{C}$ et $\mathbf{D}$ sont des matrices nulles, exprimer le produit $\mathbf{RN}$.

Exercice 1.16: $(\star)$

Soient $\mathbf{A} = \begin{bmatrix} -2 & 3 \\ 6 & 3 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} -1 & 3 & 1 \\ 0 & 4 & 1 \\ 1 & 2 & 2 \end{bmatrix}, \mathbf{C} = \begin{bmatrix} -1 & -2 & 1 & 1 \\ 1 & 4 & 0 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 2 & 0 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} -1 & -2 & 1 & 1 \\ 1 & 4 & 0 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$,

calculer leur déterminant.

Exercice 1.17: ($\star$)
Soient les matrices :

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 4 & 2 \\ -2 & 5 & 13 & 2 \\ -3 & 18 & 42 & -5 \\ 4 & 14 & -21 & -11 \end{bmatrix}, \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 5 & 1 & 0 \\ -4 & 6 & -7 & 1 \end{bmatrix}, \text{ et } \mathbf{U} = \begin{bmatrix} -1 & 1 & 4 & 2 \\ 0 & 3 & 5 & -2 \\ 0 & 0 & 5 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

1. Vérifier que $\mathbf{A} = \mathbf{LU}$.
2. En déduire le déterminant de $\mathbf{A}$.
3. Calculer le déterminant de $\mathbf{B} = \begin{bmatrix} -2 & 1 & -4 & 2 \\ -4 & 5 & -13 & 2 \\ -6 & 18 & -42 & -5 \\ 8 & 14 & 21 & -11 \end{bmatrix}$.

Exercice 1.18: ($\star\star$)
Soit $\mathbf{A} = \begin{bmatrix} -1 & 3 & 1 \\ 2 & 4 & 1 \\ 2 & 7 & 1 \end{bmatrix}$.

1. Calculer $\det(\mathbf{A})$.
2. Est-ce que $\mathbf{A}^\top \mathbf{A}$ est inversible ? Si oui, quel est le déterminant de $(\mathbf{A}^\top \mathbf{A})^{-1}$?

Exercice 1.19: ($\star\star\star$)
Soit $\mathbf{A}_n = \begin{bmatrix} 4 & 4 & 0 & \dots & 0 \\ 1 & 4 & 4 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 4 \\ 0 & \dots & 0 & 1 & 4 \end{bmatrix} \in \mathbb{R}^{n \times n}$.

1. Montrer que $\det(A_{n+2}) = 4(\det(A_{n+1}) - \det(A_n))$ pour tout $n \in \mathbb{N}$.
2. En déduire $\det(A_n)$ pour tout $n \in \mathbb{N}$.

Exercice 1.20: ($\star\star\star$)
Soient $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^{n \times n}$, et $\mathbf{D} \in \mathbb{R}^{m \times m}$.

1. Soit la matrice diagonale par blocs définie par $\mathbf{M}_1 = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}$.

   Montrer que $\det(\mathbf{M}_1) = \det(\mathbf{A}) \det(\mathbf{D})$.
2. Soit la matrice triangulaire par blocs définie par $\mathbf{M}_2 = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}$.

   En supposant que $\mathbf{A}$ est inversible, montrer que $\det(\mathbf{M}_2) = \det(\mathbf{A}) \det(\mathbf{D})$.

Exercice 1.21: $(\star)$
Soit $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$.

1. La matrice $\mathbf{A}$ est-elle inversible ?
2. Calculer $\mathbf{A}\mathbf{A}^\top$. La matrice est-elle inversible ?
3. Calculer $\mathbf{A}^\top\mathbf{A}$. La matrice est-elle inversible ?
4. Calculer $(\mathbf{A}^\top\mathbf{A})^\top$.

Exercice 1.22: $(\star)$
Justifier d'au moins trois façons différentes que la matrice $\mathbf{I}_2$ est inversible.

Exercice 1.23: $(\star)$
Justifier d'au moins trois façons différentes que la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

n'est pas inversible.

Exercice 1.24: $(\star\star)$
LIEN AVEC exercice 8.22 ? ? On considère la matrice $\mathbf{A}_k = \begin{bmatrix} 2 & -1 & k \\ 0 & k & 2 \\ 1 & 4 & 5 \end{bmatrix}$. Pour quelle(s) valeur(s) de $k \in \mathbb{R}$ la matrice $\mathbf{A}_k$ n'est-elle pas inversible ?

Exercice 1.25: une histoire d'inverse $(\star\star)$
Soient $\mathbf{F} \in \mathbb{R}^{n\times n}$ une matrice inversible et $\mathbf{u} \in \mathbb{R}^n$, un vecteur colonne. Montrer que

$$(\mathbf{F}^{-1} + \mathbf{u}\mathbf{u}^\top)^{-1} = \mathbf{F} - \frac{\mathbf{F}\mathbf{u}\mathbf{u}^\top\mathbf{F}}{1 + \mathbf{u}^\top\mathbf{F}\mathbf{u}} \tag{1.10}$$

Exercice 1.26: matrices nilpotentes et inversibilité $(\star\star)$
Soit $\mathbf{A}$ une matrice telle qu'il existe n tel que $\mathbf{A}^n = 0$ avec $\mathbf{A}$ non nulle.

1. Donner un exemple en dimension 2 pour $n = 2$.
2. Montrer que A n'est pas inversible.
3. Calculer et simplifier au maximum l'expression suivante
$$\sum_{i=0}^{n-1} (\mathbf{A}^i - \mathbf{A}^{i+1}).$$
4. Montrer que $(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{n-1})$ est inversible.

Exercice 1.27: $(\star\star)$
Soient les matrices définies par blocs :

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D} & \mathbf{I} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{D} & \mathbf{I} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbf{I} & \mathbf{D} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}$$

avec $\mathbf{I}, \mathbf{A}, \mathbf{D} \in \mathbb{R}^{n\times n}$.

1. $\mathbf{A}$ est-elle triangulaire ? S'il existe, donner son déterminant.
2. Calculer le produit $\mathbf{AB}$.
3. Donner l'inverse de $\mathbf{A}$.
4. Calculer $\mathbf{AC}$.

Exercice 1.28: $(\star\star)$
Exercice à la bonne place ?

1. Soit $\mathbf{A} = \begin{bmatrix} 1 & 12 & -6 \\ 0 & 1 & 0 \\ 0 & 4 & -1 \end{bmatrix}$. $\mathbf{A}$ est-elle inversible ?
2. Montrer que :
$$(\mathbf{A} - \mathbf{I}_3)(\mathbf{A} + \mathbf{I}_3) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
3. De manière générale, soit $\mathbf{B} \in \mathbb{R}^{n\times n}$ une matrice vérifiant :
$$(\mathbf{B} + \mathbf{I}_n)(\mathbf{B} - \mathbf{I}_n) = \mathbf{0}_n$$
Montrer que $\mathbf{B}$ est inversible et que son inverse est $\mathbf{B}^{-1} = \mathbf{B}$.

Exercice 1.29: $(\star\star)$
Montrer qu'une matrice antisymétrique a seulement des zéros sur sa diagonale.

Exercice 1.30: $(\star\star)$
Soient $\mathbf{A}$ et $\mathbf{B}$ deux matrices et $\mathbf{x}$ un vecteur tels que $\mathbf{Ax} = \mathbf{Bx}$. Peut-on en conclure que $\mathbf{A} = \mathbf{B}$?

Exercice 1.31: $(\star\star)$
Soient $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ et $\mathbf{B} = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ deux matrices de $\mathbb{R}^3$

1. Calculer la norme de Frobenius de ces matrices.
2. Vérifier que $|\langle \mathbf{A}, \mathbf{B}\rangle| \leq \|\mathbf{A}\|_{\mathrm{F}} \|\mathbf{B}\|_{\mathrm{F}}$.

3. À quelle inégalité liée au produit scalaire réel vu dans la section 1.1.4 l'inégalité de la question précédente est-elle similaire ?
4. Soient $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$ deux matrices symétriques, montrer que $\operatorname{tr}(\mathbf{AB})^2 \leq \operatorname{tr}(\mathbf{A}^2) \operatorname{tr}(\mathbf{B}^2)$. (Indication : s'aider de la démonstration utilisée pour démontrer l'inégalité cherchée à la question précédente – LIEN.)

Chapitre

5

Espaces vectoriels

Ce chapitre est consacré à l'étude des *espaces vectoriels*, qui constituent des structures algébriques fondamentales en mathématique. Un espace vectoriel est un ensemble dont les éléments, appelés *vecteurs*, peuvent être ajoutés entre eux et multipliés par des *scalaires*, qui sont réels ou complexes dans ce document, selon des règles précises. Cette structure permet de généraliser et d'abstraire les notions de vecteurs et de matrices introduites au chapitre 1 et de formaliser le fait que les vecteurs des ensembles $\mathbb{R}^n$ et $\mathbb{C}^n$, ainsi que les matrices des ensembles $\mathbb{R}^{m\times n}$ et $\mathbb{C}^{m\times n}$, sont des espaces vectoriels. Elle fournit un cadre unifié pour définir et étudier des concepts essentiels tels que la dimension, les bases, ou encore les sous-espaces vectoriels. Les espaces vectoriels se rencontrent dans une grande variété de contextes mathématiques, allant des polynômes aux fonctions, en passant par les suites, les matrices ou encore les solutions de certains types d'équations différentielles.

5.1 Définitions

On commence par introduire la notion d'espace vectoriel, accompagnée de quelques exemples illustratifs, avant d'explorer celle de sous-espace vectoriel.

5.1.1 Définition d'un espace vectoriel

La définition suivante formalise la notion d'espace vectoriel.

Définition 5.1.1 (*espace vectoriel*)

Un *espace vectoriel* V est un ensemble d'éléments, appelés *vecteurs*, muni des opérations :

- d'*addition vectorielle*, notée "+", qui permet d'additionner deux vecteurs de V afin d'obtenir un autre vecteur de V ;
- de *multiplication par un scalaire*, qui permet de multiplier tout vecteur de V par un scalaire afin d'obtenir un autre vecteur de V ;

qui respectent les huit propriétés suivantes :

1. $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$ pour tous vecteurs $\mathbf{x}$ et $\mathbf{y}$ de V ;
2. $\mathbf{x}+(\mathbf{y}+\mathbf{z})=(\mathbf{x}+\mathbf{y})+\mathbf{z}$ pour tous vecteurs $\mathbf{x}$, $\mathbf{y}$ et $\mathbf{z}$ de V ;
3. il existe un vecteur de V, noté $\mathbf{0}$, appelé le *vecteur nul* ou l'*élément nul de l'addition*, tel que $\mathbf{x}+\mathbf{0}=\mathbf{x}$ pour tout vecteur $\mathbf{x}$ de V ;
4. pour tout vecteur $\mathbf{x}$ de V, il existe un vecteur de V, appelé *l'opposé* de $\mathbf{x}$ et noté $-\mathbf{x}$, tel que $\mathbf{x}+(-\mathbf{x})=\mathbf{0}$ (on pourra écrire $\mathbf{x}-\mathbf{x}=\mathbf{0}$) ;
5. $1\mathbf{x}=\mathbf{x}$ pour tout vecteur $\mathbf{x}$ de V ;
6. $(\alpha\beta)\mathbf{x}=\alpha(\beta\mathbf{x})$ pour tous scalaires α et β et tout vecteur $\mathbf{x}$ de V ;
7. $\alpha(\mathbf{x}+\mathbf{y})=\alpha\mathbf{x}+\alpha\mathbf{y}$ pour tout scalaire α et tous vecteurs $\mathbf{x}$ et $\mathbf{y}$ de V ;
8. $(\alpha+\beta)\mathbf{x}=\alpha\mathbf{x}+\beta\mathbf{x}$ pour tous scalaires α et β et tout vecteur $\mathbf{x}$ de V.

On parlera d'*espace vectoriel réel* ou d'*espace vectoriel complexe* selon que le scalaire considéré dans l'opération de multiplication par un scalaire est un nombre réel ou un nombre complexe.

Parfois, un ensemble considéré peut être vu comme un espace vectoriel réel ou complexe selon la façon de le modéliser. C'est par exemple le cas de $\mathbb{C}$ qui peut être vu comme un espace vectoriel complexe ou un espace vectoriel réel. L'étude de cas de la section 5.4 traite ce point plus en détail.

Les opérations d'addition vectorielle et de multiplication par un scalaire permettent d'élargir la définition de combinaison linéaire introduite à la définition 1.1.9.

Définition 5.1.2 (*combinaison linéaire*)

Soient $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ p vecteurs d'un espace vectoriel V, et $\alpha_1, \alpha_2, \ldots, \alpha_p$ p scalaires (réels si V est considéré comme un espace vectoriel réel ou complexes si V est considéré comme un espace vectoriel complexe) :

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_p\mathbf{v}_p = \sum_{k=1}^{p} \alpha_k \mathbf{v}_k$$

est une combinaison linéaire des vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$.

Une combinaison linéaire est donc une somme finie pondérée de vecteurs, mais ces vecteurs peuvent appartenir à des espaces vectoriels que $\mathbb{R}^n$ ou $\mathbb{C}^n$; ils sont, de façon générale, les éléments d'espaces vectoriels.

Le résultat suivant découle naturellement de la notion de combinaison linéaire.

Proposition 5.1.1 (*combinaison linéaire*)

Toute combinaison linéaire de vecteurs d'un espace vectoriel V est également un vecteur de V.

Preuve. On considère la combinaison linéaire $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_p\mathbf{x}_p$ des p vecteurs $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$ de V avec $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ un ensemble de p scalaires. Selon la propriété de la multiplication par un scalaire, chaque élément $\alpha_i\mathbf{x}_i$, pour $i \in [\![1; p]\!]$, est dans V. Ainsi, la combinaison linéaire est une somme de vecteurs de V, ce qui, selon la propriété de l'addition vectorielle, est également dans V. Donc $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \ldots + \alpha_p\mathbf{x}_p \in V$. $\square$

Les huit propriétés de la définition 5.1.1 permettent de formaliser rigoureusement les comportements que l'on associe intuitivement aux vecteurs du plan. L'addition peut ainsi être comprise comme la combinaison de deux vecteurs, et la multiplication par un scalaire comme une dilatation (ou contraction) du vecteur. Ces opérations, familières dans un cadre géométrique, sont à la base des axiomes qui définissent un espace vectoriel. La structure abstraite ainsi construite prolonge cette intuition et en offre une généralisation puissante, applicable bien au-delà du plan ou de l'espace à trois dimensions.

5.1.2 Exemples d'espaces vectoriels

Parmi les exemples d'espaces vectoriels donnés ci-dessous, les exemples I et II formalisent le fait que l'ensemble des vecteurs à composantes réelles et l'ensemble des matrices réelles forment des espaces vectoriels réels. Les mêmes principes s'appliquent pour $\mathbb{C}^m$ et $\mathbb{C}^{m\times n}$, qui forment des espaces vectoriels complexes.

Exemple 5.1.1 *(ensembles qui sont des espaces vectoriels)*

I $\mathbb{R}^n$, l'ensemble des vecteurs colonne à n composantes réelles, est un espace vectoriel réel.

II $\mathbb{R}^{m\times n}$, l'ensemble des matrices réelles de taille $m \times n$, est un espace vectoriel réel [a].

III $\mathcal{S}$, l'ensemble des suites numériques à valeurs réelles, est un espace vectoriel réel.

IV L'ensemble des fonctions réelles ou complexes est un espace vectoriel.

V L'ensemble des fonctions continues sur un intervalle est un espace vectoriel.

VI L'ensemble des fonctions bornées sur $\mathbb{R}$ est un espace vectoriel.

VII $\mathcal{P}_2(\mathbb{R})$, l'ensemble des polynômes réels de degré inférieur ou égal à deux, est un espace vectoriel.

a. Il est donc correct de dire que les matrices sont des vecteurs.

Preuve des points I, III, et IV.

I On note chaque élément de $\mathbb{R}^n$ comme $\mathbf{x} = (x_1, x_2, \dots, x_n)$, avec $x_i \in \mathbb{R}$ pour $i \in [\![1;n]\!]$. Pour deux éléments $\mathbf{x} = (x_1, x_2, \dots, x_n)$ et $\mathbf{y} = (y_1, y_2, \dots, y_n)$ de $\mathbb{R}^n$ et pour un scalaire $\alpha \in \mathbb{R}$, on considère les deux opérations définies comme suit :

- ▷ addition vectorielle : $\mathbf{x}+\mathbf{y} = (x_1+y_1, x_2+y_2, \dots, x_n+y_n)$;
- ▷ multiplication par un scalaire : $\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \dots, \alpha x_n)$.

On doit vérifier la validité des huit propriétés de la définition 5.1.1 :

1 $\mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$ car l'addition des réels est commutative ;
2 $(\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$ car l'addition des réels est associative ;
3 le vecteur nul $\mathbf{0} = (0, 0, \dots, 0) \in \mathbb{R}^n$ satisfait $\mathbf{x}+\mathbf{0} = \mathbf{x}$;
4 pour tout $\mathbf{x} = (x_1, x_2, \dots, x_n)$, le vecteur $-\mathbf{x} = (-x_1, -x_2, \dots, -x_n)$ vérifie $\mathbf{x}+(-\mathbf{x}) = \mathbf{0}$;
5 $1\mathbf{x} = \mathbf{x}$ car $1x_i = x_i$ pour tout $i \in [\![1;n]\!]$;
6 $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$ car $(\alpha\beta)x_i = \alpha(\beta x_i)$ pour tout $i \in [\![1;n]\!]$;
7 $\alpha(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x}+\alpha\mathbf{y}$ car $\alpha(x_i+y_i) = \alpha x_i + \alpha y_i$ pour tout $i \in [\![1;n]\!]$;
8 $(\alpha+\beta)\mathbf{x} = \alpha\mathbf{x}+\beta\mathbf{x}$ car $(\alpha+\beta)x_i = \alpha x_i + \beta x_i$ pour tout $i \in [\![1;n]\!]$.

On a donc montré que $\mathbb{R}^n$ est un espace vectoriel réel.

III Soit $\mathcal{S} = \{(u_n)_{n\in\mathbb{N}} : u_n \in \mathbb{R}\}$ l'ensemble des suites numériques à valeurs réelles. Pour deux éléments $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ de $\mathcal{S}$ et pour un scalaire $\alpha \in \mathbb{R}$, on considère les deux opérations définies comme suit :

- ▷ addition : $(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (u_n+v_n)_{n\in\mathbb{N}}$;

▷ multiplication par un scalaire : $\alpha(u_n)_{n\in\mathbb{N}} = (\alpha u_n)_{n\in\mathbb{N}}$.

On doit vérifier la validité des huit propriétés de la définition 5.1.1 :

1 $(u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}} = (v_n)_{n\in\mathbb{N}} + (u_n)_{n\in\mathbb{N}}$ car $u_n + v_n = v_n + u_n$ pour tout $n \in \mathbb{N}$;

2 $((u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}}) + (w_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}} + ((v_n)_{n\in\mathbb{N}} + (w_n)_{n\in\mathbb{N}})$ car $(u_n + v_n) + w_n = u_n + (v_n + w_n)$ pour tout $n \in \mathbb{N}$;

3 la suite nulle $(0)_{n\in\mathbb{N}}$ est l'élément neutre ;

4 pour tout $(u_n)_{n\in\mathbb{N}} \in \mathcal{S}$, la suite $(-u_n)_{n\in\mathbb{N}}$ vérifie $(u_n)_{n\in\mathbb{N}} + (-u_n)_{n\in\mathbb{N}} = (0)_{n\in\mathbb{N}}$;

5 $1(u_n)_{n\in\mathbb{N}} = (u_n)_{n\in\mathbb{N}}$ car $1u_n = u_n$ pour tout $n \in \mathbb{N}$;

6 $(\alpha\beta)(u_n)_{n\in\mathbb{N}} = \alpha(\beta(u_n)_{n\in\mathbb{N}})$ car $(\alpha\beta)u_n = \alpha(\beta u_n)$pour tout $n \in \mathbb{N}$;

7 $\alpha((u_n)_{n\in\mathbb{N}} + (v_n)_{n\in\mathbb{N}}) = \alpha(u_n)_{n\in\mathbb{N}} + \alpha(v_n)_{n\in\mathbb{N}}$ car $\alpha(u_n + v_n) = \alpha u_n + \alpha v_n$ pour tout $n \in \mathbb{N}$;

8 $(\alpha + \beta)(u_n)_{n\in\mathbb{N}} = \alpha(u_n)_{n\in\mathbb{N}} + \beta(u_n)_{n\in\mathbb{N}}$ car $(\alpha + \beta)u_n = \alpha u_n + \beta u_n$ pour tout $n \in \mathbb{N}$.

On a donc montré que $\mathcal{S}$ est un espace vectoriel réel.

IV Pour deux fonctions réelles ou complexes f et g et pour un scalaire α, on considère les deux opérations suivantes :

▷ addition : $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$;

▷ multiplication par un scalaire : $(\alpha f)(\mathbf{x}) = \alpha f(\mathbf{x})$.

Les huit propriétés de la définition 5.1.1 sont faciles à démontrer, en considérant le vecteur nul comme la fonction nulle $f = \mathbf{0}$, $\mathbf{0}$ étant dans $\mathbb{R}^m \subset \mathbb{C}^m$, et l'opposée d'une fonction f étant $-f$ (c'est-à-dire $(-f)(\mathbf{x}) = -f(\mathbf{x})$).

□

Les preuves des points II, V, VI et VII sont le sujet des exercices 5.3, 5.6, 5.7, et 5.8. De nombreux ensembles ne satisfont pas les propriétés de la définition 5.1.1. À titre d'illustration, on donne ci-dessous deux exemples d'ensembles qui ne sont pas des espaces vectoriels. La preuve du point II est le sujet de l'exercice 5.12.

Exemple 5.1.2 (*ensembles qui ne sont pas des espaces vectoriels*)

I L'ensemble des matrices inversibles n'est pas un espace vectoriel.

II L'ensemble des vecteurs du premier quadrant de $\mathbb{R}^2$ n'est pas un espace vectoriel.

Preuve du point I. La matrice nulle $\mathbf{O}$ est l'élément nul de l'addition entre matrices, mais elle n'est pas inversible, ce qui contredit la troisième propriété de la définition 5.1.1 qui exige que cet élément nul soit dans l'ensemble. □

Bien que, tel qu'on l'a vu, il soit possible de montrer qu'un ensemble forme un espace vectoriel en vérifiant les propriétés de la définition 5.1.1, ceci peut être fastidieux. En pratique, il peut s'avérer très utile d'utiliser la notion de *sous-espace vectoriel*, ce qui constitue l'objet de la prochaine section.

5.1.3 Sous-espaces vectoriels

De nombreuses structures intéressantes en algèbre se révèlent être des sous-espaces vectoriels d'espaces vectoriels connus (voir section 6.1). Tel qu'illustré dans cette section, il est possible de démontrer que certains de ces ensembles sont des espaces vectoriels, sans avoir à vérifier les huit propriétés de la définition 5.1.1.

Définition 5.1.3 (*sous-espace vectoriel*)

Soit V un espace vectoriel, et soit $U \subseteq V$ un sous-ensemble de V. Si U, muni des mêmes opérations d'addition vectorielle et de multiplication par un scalaire que V, est lui-même un espace vectoriel, alors U est appelé un *sous-espace vectoriel* de V.

Le théorème 5.1.1 offre une caractérisation plus simple d'un sous-espace vectoriel, évitant le recours à la définition 5.1.1. Il permet de démontrer qu'un ensemble $U \subseteq V$ est un sous-espace vectoriel à l'aide de propriétés assez évidentes à vérifier.

Théorème 5.1.1 (*caractérisation des sous-espaces vectoriels*)

Soit V un espace vectoriel, et soit U un sous-ensemble de V muni des mêmes opérations d'addition vectorielle et de multiplication par un scalaire que V. Alors U est un sous-espace vectoriel de V si et seulement si les trois conditions suivantes sont satisfaites :

1. U contient le vecteur nul : $\mathbf{0} \in U$;
2. fermeture de l'addition : $\mathbf{x} + \mathbf{y} \in U$ pour tous vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans U ;
3. fermeture de la multiplication par un scalaire : $\alpha\mathbf{x} \in U$ pour tout scalaire α et pour tout $\mathbf{x} \in U$.

Preuve. Soit V un espace vectoriel, et soit U un sous-ensemble de V muni des mêmes opérations d'addition vectorielle et de multiplication par un scalaire que V.

Si U est un espace vectoriel, alors il résulte de la définition 5.1.1 que $\mathbf{0} \in U$, que $\mathbf{x}+\mathbf{y} \in U$ pour tous vecteurs $\mathbf{x}$ et $\mathbf{y}$ dans U, et que $\alpha\mathbf{x} \in U$ pour tout scalaire α et pour tout $\mathbf{x} \in U$. Les trois conditions du théorème 5.1.1 sont donc satisfaites.

Supposons maintenant que les trois conditions du théorème 5.1.1 sont satisfaites, et montrons que U est alors nécessairement un espace vectoriel. Il suffit de prouver que les huit propriétés de la définition 5.1.1 sont vérifiées :

▷ les propriétés 1, 2, 5, 6, 7 et 8 sont vérifiées dans V et donc aussi dans U.

- ▷ par la propriété 1 du théorème 5.1.1, on a $\mathbf{0} \in U$. De plus, pour tout $\mathbf{x} \in U$, on a $\mathbf{x} + \mathbf{0} = \mathbf{x}$ dans V, et donc aussi dans U.
- ▷ pour tout $\mathbf{x} \in U$, son opposé dans V est donné par $-\mathbf{x} = (-1)\mathbf{x}$, et $-\mathbf{x} \in U$ par la propriété 3 du théorème 5.1.1. De plus, $\mathbf{x} - \mathbf{x} = \mathbf{0}$ dans V, et donc aussi dans U.

□

L'exemple suivant illustre l'application du théorème 5.1.1. Plusieurs exercices le font aussi, comme les exercices 5.10 et 5.9.

Exemple 5.1.3

L'ensemble $U = \{(x, y) \in \mathbb{R}^2 : x + y = 0\} \subset V = \mathbb{R}^2$ est un sous-espace vectoriel de $\mathbb{R}^2$. En effet,

1. L'origine $\mathbf{0} = (0, 0) \in U$ car $0 + 0 = 0$.
2. Si $\mathbf{x} = (x_1, x_2) \in U$ et $\mathbf{y} = (y_1, y_2) \in U$, alors $x_1 + x_2 = 0$ et $y_1 + y_2 = 0$. L'addition de $\mathbf{x}$ et $\mathbf{y}$ donne $\mathbf{x} + \mathbf{y} = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ dont la somme des composantes est $x_1 + y_1 + x_2 + y_2 = 0$. Ainsi $\mathbf{x} + \mathbf{y} \in U$.
3. Si $\mathbf{x} = (x_1, x_2) \in U$ (et donc $x_1 + x_2 = 0$) et si $\alpha \in \mathbb{R}$, alors $\alpha\mathbf{x} = (\alpha x_1, \alpha x_2)$ dont la somme des composantes est $\alpha x_1 + \alpha x_2 = \alpha(x_1 + x_2) = 0$. Donc $\alpha\mathbf{x} \in U$.

Toutes les conditions du théorème 5.1.1 sont ainsi satisfaites et l'ensemble des vecteurs de $\mathbb{R}^2$ dont la somme des composantes est nulle forme bien un sous-espace vectoriel de $\mathbb{R}^2$. Graphiquement, ce sous-espace est la droite d'équation $y = -x$.

Remarque

Quelques observations à propos du théorème 5.1.1 :

- ▷ Tout espace vectoriel V admet toujours deux sous-espaces vectoriels particuliers, soit $\{\mathbf{0}\}$ (c'est-à-dire l'ensemble ne contenant que le vecteur nul) et V lui-même. Ces sous-espaces vectoriels sont dits *triviaux*.
- ▷ Les conditions 2 et 3 du théorème 5.1.1 peuvent être remplacées par la seule condition que pour toute paire $(\mathbf{x}, \mathbf{y})$ de U, et tous scalaires α et β, on a

  $$\alpha\mathbf{x} + \beta\mathbf{y} \in U$$

  qui est plus directe à vérifier.
- ▷ Les conditions 2 et 3 du théorème 5.1.1 permettent également de réaliser qu'un sous-espace vectoriel, tout comme un espace vectoriel, est un ensemble qui contient toutes les combinaisons linéaires possibles de ses propres éléments. Cette remarque permet d'exprimer le corollaire 5.1.1 sur l'opérateur $\mathrm{Vect}(\cdot)$.

- ▷ Le seul espace vectoriel qui ne contient pas une infinité de vecteurs est $\{\mathbf{0}\}$ [a].
- ▷ Tout sous-espace vectoriel est un espace vectoriel, et tout espace vectoriel est un sous-espace vectoriel (au moins de lui-même). L'exemple 5.1.4 illustre cela.
- ▷ Lorsqu'on mentionne un sous-espace vectoriel, il faut aussi mentionner le "sur-espace", c'est-à-dire l'espace vectoriel dans lequel est inclus le sous-espace vectoriel étudié. Dans l'exemple 5.1.3, le sur-espace est $\mathbb{R}^2$.
- ▷ Si on veut prouver qu'un ensemble est (ou n'est pas) un espace vectoriel, la vérification prioritaire à faire est l'appartenance du vecteur nul à cet ensemble.

a. Pour les espaces vectoriels $\mathbb{R}$ et $\mathbb{C}$.

Exemple 5.1.4

- ▷ $\mathbb{R}^3$, l'ensemble des vecteurs colonne de trois composantes, est un espace vectoriel, d'où son appellation familière d'*espace à trois dimensions*, ou d'*espace 3D*. De même $\mathbb{R}^2$ est un espace vectoriel (l'*espace 2D*). Cependant, $\mathbb{R}^2$ n'est pas un sous-espace vectoriel de $\mathbb{R}^3$, car $\mathbb{R}^2$ n'est pas un sous-ensemble de $\mathbb{R}^3$. Ils sont en fait fondamentalement différents car les vecteurs de $\mathbb{R}^3$ ont trois composantes, et ceux de $\mathbb{R}^2$ ont deux composantes, et ces deux espaces ne partagent pas les mêmes additions vectorielles ni les mêmes multiplications par un scalaire. Pour résumer, $\mathbb{R}^2$ n'est pas un sous-espace vectoriel de $\mathbb{R}^3$, mais est un sous-espace vectoriel de lui-même.
- ▷ Tout plan de $\mathbb{R}^3$ qui contient $\mathbf{0}$ est un sous-espace vectoriel de $\mathbb{R}^3$, et donc un espace vectoriel. Toute droite (passant par $\mathbf{0}$) d'un plan de $\mathbb{R}^3$ est un sous-espace vectoriel de ce plan, et aussi un sous-espace vectoriel de $\mathbb{R}^3$. Toute droite de $\mathbb{R}^3$ passant par l'*origine* $\mathbf{0}$ est donc aussi un espace vectoriel.
- ▷ Ainsi, dans $\mathbb{R}^3$, les espaces vectoriels possibles sont :
  - $\mathbb{R}^3$ lui-même (un des deux sous-espaces vectoriels triviaux de $\mathbb{R}^3$) ;
  - tous les plans de $\mathbb{R}^3$ qui passent par l'origine $\mathbf{0}$;
  - toutes les droites de $\mathbb{R}^3$ qui passent par l'origine $\mathbf{0}$;
  - $\{\mathbf{0}\}$, l'autre sous-espace vectoriel trivial de $\mathbb{R}^3$.

Pour rappel, l'opérateur $\mathrm{Vect}(\cdot)$ (voir définition 1.1.10) prend une famille de vecteurs en entrée, et retourne l'ensemble de toutes les combinaisons linéaires de cette famille. Le résultat suivant formalise la remarque qu'un sous-espace vectoriel est caractérisé par un ensemble de combinaisons linéaires.

Corollaire 5.1.1 (Vect($\cdot$) *est un sous-espace vectoriel*)

L'opérateur Vect($\cdot$) retourne un sous-espace vectoriel.

Preuve. Étant donnée une famille de vecteurs, l'opérateur Vect($\cdot$) retourne l'ensemble de toutes les combinaisons linéaires de cette famille. Par conséquent, Vect($\cdot$) contient toutes les combinaisons linéaires possibles de ses propres éléments, ce qui démontre, par la remarque de la section 5.1.3, que Vect($\cdot$) est un espace vectoriel. □

Le corollaire 5.1.1 fournit un outil puissant pour prouver qu'un ensemble est un sous-espace vectoriel, comme l'illustre l'exemple 5.1.5 (voir aussi l'exercice 5.5).

Exemple 5.1.5

Soit l'ensemble $U = \{(x, y) \in \mathbb{R}^2 : x + y = 0\} \subset \mathbb{R}^2$ déjà étudié à l'exemple 5.1.3. On a vu que U est la droite de $\mathbb{R}^2$ d'équation $y = -x$, dont un vecteur directeur est $(1, -1)$. Donc $U = \text{Vect}((1, -1))$, ce qui confirme que U est bien un sous-espace vectoriel de $\mathbb{R}^2$.

La caractérisation des sous-espaces vectoriels peut être utilisée pour démontrer la proposition 5.1.2 qui suit.

Proposition 5.1.2 (*les polynômes forment un espace vectoriel*)

Les ensembles de polynômes $\mathcal{P}(\mathbb{R})$ et $\mathcal{P}(\mathbb{C})$ sont des espaces vectoriels.

Preuve. On a vu au point IV de l'exemple 5.1.1 que les fonctions réelles et complexes forment des espaces vectoriels. $\mathcal{P}(\mathbb{R})$ est un sous-ensemble des fonctions réelles d'une variable ($n = 1$) qui ne renvoient qu'un réel ($m = 1$) et $\mathcal{P}(\mathbb{C})$ est un sous-ensemble des fonctions complexes d'une variable ($n = 1$) qui ne renvoient qu'un complexe ($m = 1$) (voir la définition 4.2.1). Ils contiennent tous deux un élément nul (le polynôme nul) et toute combinaison linéaire de polynômes est un polynôme. □

Cette section consacrée aux sous-espaces vectoriels s'achève par l'étude des intersections et des unions de sous-espaces vectoriels.

Proposition 5.1.3 (*intersection de sous-espaces vectoriels*)

Si U_1 et U_2 sont deux sous-espaces vectoriels d'un espace vectoriel V, alors leur intersection $U_1 \cap U_2$ est également un sous-espace vectoriel de V.

Preuve. Pour montrer que $U_1 \cap U_2$ est un sous-espace vectoriel de V, il suffit de vérifier que les trois conditions du théorème 5.1.1 sont satisfaites :

1 Le vecteur nul $\mathbf{0}$ appartient à U_1 et à U_2, et donc à leur intersection : $\mathbf{0} \in U_1 \cap U_2$;

2 si $\mathbf{x}$ et $\mathbf{y}$ sont dans $U_1 \cap U_2$, alors $\mathbf{x} + \mathbf{y} \in U_1$ (car $\mathbf{x}$ et $\mathbf{y}$ sont dans U_1, et $\mathbf{x} + \mathbf{y}$ également, par la propriété de l'addition vectorielle de la définition 5.1.1) et $\mathbf{x}+\mathbf{y} \in U_2$ (argument similaire). Donc $\mathbf{x} + \mathbf{y} \in U_1 \cap U_2$;

3 pour tout scalaire α et tout $\mathbf{x} \in U_1 \cap U_2$, on a $\alpha\mathbf{x} \in U_1$ et $\alpha\mathbf{x} \in U_2$ par les propriétés de la multiplication par un scalaire de la définition 5.1.1. Donc $\alpha\mathbf{x} \in U_1 \cap U_2$.

Les trois conditions étant satisfaites, $U_1 \cap U_2$ est bien un sous-espace vectoriel de V.

□

Remarque

L'union de deux sous-espaces vectoriels n'est en général pas un sous-espace vectoriel. Par exemple, si U_1 et U_2 sont deux droites distinctes de $\mathbb{R}^2$ passant par l'origine, U_1 et U_2 sont des sous-espaces vectoriels de $\mathbb{R}^2$ mais pas leur union $U_1 \cup U_2$. En effet, il est possible que la somme de deux vecteurs de l'union ne soit pas dans l'union, comme l'illustre l'exemple 5.1.6. La propriété de fermeture de addition du théorème 5.1.1 n'est donc pas vérifiée.

Exemple 5.1.6

Soient $U_1 = \{(x, y) \in \mathbb{R}^2 : x + y = 0\}$ et $U_2 = \{(x, y) \in \mathbb{R}^2 : x - y = 0\}$. Ces deux ensembles sont des sous-espaces vectoriels de $\mathbb{R}^2$ correspondant aux droites $y = -x$ et $y = x$. Leur intersection est réduite au vecteur nul : $U_1 \cap U_2 = \{\mathbf{0}\}$. L'intersection de ces deux sous-espaces vectoriels est donc également un sous-espace vectoriel, en accord avec la proposition 5.1.3. En revanche, leur union $U_1 \cup U_2$ ne vérifie pas la propriété de fermeture de l'addition du théorème 5.1.1. Par exemple, $(-2, 2) \in U_1$ et $(1, 1) \in U_2$, mais $(-2, 2) + (1, 1) = (-1, 3) \notin U_1 \cup U_2$. Ainsi, $U_1 \cup U_2$ n'est pas un sous-espace vectoriel de $\mathbb{R}^2$. La figure 5.1 illustre ce propos.

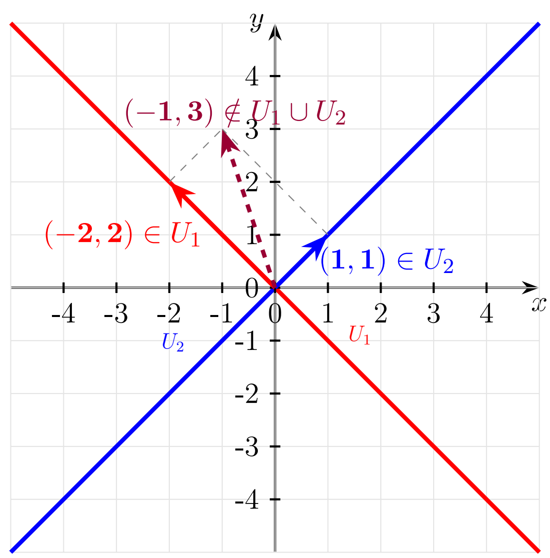

FIGURE 5.1 – Deux sous-espaces vectoriels (définis dans l'exemple 5.1.6) dont l'union n'est pas un sous-espace vectoriel.

5.2 Bases

Une base d'un espace vectoriel est une famille aussi petite que possible d'éléments qui permet de définir sans équivoque tout vecteur de cet espace. Une telle base permet, entre autres, d'effectuer des opérations liant des vecteurs de l'espace. Afin d'exprimer rigoureusement la notion de base d'un espace vectoriel, il est nécessaire tout d'abord d'introduire les notions d'indépendance linéaire (section 5.2.1) et d'ensembles générateurs (section 5.2.2).

5.2.1 Indépendance linéaire

La notion de famille "aussi petite que possible" prend ses racines dans le concept d'indépendance linéaire. Cette notion se décline en plusieurs définitions vues ci-dessous, qui donnent une grande flexibilité selon le contexte d'utilisation.

Définition 5.2.1 (*indépendance linéaire*)

Les p vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ sont *linéairement indépendants* si la combinaison linéaire $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_p$ n'est égale au vecteur nul que lorsque les poids $\alpha_1, \alpha_2, \ldots, \alpha_p$ sont tous égaux à zéro (c'est la *combinaison triviale*).

Définition 5.2.2 (*dépendance linéaire*)

Les p vecteurs $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ sont *linéairement dépendants* s'ils ne sont pas linéairement indépendants. En d'autres termes, les p vecteurs sont linéairement dépendants s'il est possible de les combiner pour obtenir le vecteur nul avec au moins un poids non nul (on parle alors d'une *combinaison non-triviale*). Dans le cas $p = 2$, si les deux vecteurs $\mathbf{v}_1$ et $\mathbf{v}_2$ sont linéairement dépendants, on dira qu'ils sont *colinéaires* [a]. L'un est alors le multiple de l'autre : $\mathbf{v}_1 = \alpha \mathbf{v}_2$ avec α un scalaire.

a. ceci est valide même quand un des vecteurs, ou les deux, sont nuls.

Pour résumer, des vecteurs sont linéairement indépendants si seule leur combinaison triviale donne le vecteur nul, alors qu'ils sont linéairement dépendants s'il existe une combinaison non-triviale qui donne le vecteur nul.

Si les p vecteurs sont placés dans un ensemble $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$, alors on pourra directement qualifier E d'*ensemble indépendant* ou *d'ensemble dépendant*. Si les vecteurs sont placés dans la famille $F = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$, on emploiera pour F les termes de *famille libre* et de *famille liée*. Ceci est formalisé dans les quatre définitions suivantes.

Définition 5.2.3 (*ensemble indépendant*)
Un ensemble de vecteurs est indépendant si les vecteurs qui le composent sont linéairement indépendants.

Définition 5.2.4 (*ensemble dépendant*)
Un ensemble de vecteurs est dépendant s'il n'est pas indépendant.

Définition 5.2.5 (*famille libre*)
Une famille de vecteurs est libre si les vecteurs qui la composent sont linéairement indépendants.

Définition 5.2.6 (*famille liée*)
Une famille de vecteurs est liée si elle n'est pas libre.

Ces concepts sont illustrés à l'aide des deux exemples suivants.

Exemple 5.2.1
Considérons les trois vecteurs $\mathbf{e}_1 = (1,0,0)$, $\mathbf{e}_2 = (0,1,0)$ et $\mathbf{e}_3 = (0,0,1)$ dans $\mathbb{R}^3$. Étudier l'indépendance linéaire de ces vecteurs revient à examiner les solutions de l'équation $\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \alpha_3\mathbf{e}_3 = \mathbf{0}$ dont le développement donne

$$\alpha_1\mathbf{e}_1+\alpha_2\mathbf{e}_2+\alpha_3\mathbf{e}_3 = \alpha_1\begin{bmatrix}1\\0\\0\end{bmatrix}+\alpha_2\begin{bmatrix}0\\1\\0\end{bmatrix}+\alpha_3\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}\alpha_1\\0\\0\end{bmatrix}+\begin{bmatrix}0\\\alpha_2\\0\end{bmatrix}+\begin{bmatrix}0\\0\\\alpha_3\end{bmatrix} = \begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

On a donc $\alpha_1 = \alpha_2 = \alpha_3 = 0$ comme solution unique. L'indépendance linéaire de ces trois vecteurs est ainsi établie.

Exemple 5.2.2
Soient les vecteurs $\mathbf{v}_1 = (1,1)$, $\mathbf{v}_2 = (0,2)$ et $\mathbf{v}_3 = (1,3)$. Comme $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$, et donc $\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 = \mathbf{0}$, on en déduit que $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ sont linéairement dépendants et que $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ est une famille liée.

Il convient de noter, pour conclure cette section, que les ensembles indépendants et les familles libres ne peuvent pas contenir le vecteur nul, ce qui est formalisé par la proposition 5.2.1 suivante.

Proposition 5.2.1 *(le vecteur* $\mathbf{0}$ *rend tout ensemble dépendant et toute famille liée)*
Un ensemble de vecteurs qui contient le vecteur nul est dépendant. De même, une famille de vecteurs qui contient le vecteur nul est liée.

Preuve. Soit E l'ensemble de vecteurs $\{\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ contenant le vecteur nul, et soit $F = (\mathbf{0}, \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$. On peut facilement combiner ces vecteurs pour obtenir $\mathbf{0}$ en utilisant un poids α_0 non nul : $\alpha_0\mathbf{0} + \sum_{k=1}^{p} 0\mathbf{v}_k = \alpha_0\mathbf{0} = \mathbf{0}$. □

5.2.2 Ensembles générateurs

La deuxième notion nécessaire à la définition d'une base est celle d'ensemble générateur, qui s'exprime aussi pour les familles. Elle permet de comprendre comment un ensemble fini de vecteurs permet d'exprimer tous les vecteurs possibles d'un espace vectoriel.

Définition 5.2.7 *(famille génératrice)*
Une famille de vecteurs $F = (\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p)$ est une *famille génératrice* d'un espace vectoriel V si chaque vecteur $\mathbf{x} \in V$ peut s'exprimer comme une combinaison linéaire de vecteurs de F. On dit alors que F *génère* V, et $V = \mathrm{Vect}(F)$.

Définition 5.2.8 *(ensemble générateur)*
Un ensemble de vecteurs $E = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$ *génère* un espace vectoriel V si chaque vecteur $\mathbf{x} \in V$ peut s'exprimer comme une combinaison linéaire de vecteurs de E.

On note à nouveau l'usage de l'opérateur $\mathrm{Vect}(\cdot)$ de la définition définition 1.1.10 qui prend une famille de vecteurs en entrée et qui renvoie un ensemble qui est toujours un espace vectoriel.

On pourra dire que "des vecteurs génèrent un espace", ce qui signifie que l'ensemble (ou la famille) constitué de ces vecteurs génère l'espace en question.

Exemple 5.2.3
La famille $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \subset \mathbb{R}^2$ avec $\mathbf{v}_1 = \mathbf{e}_1 = (1, 0)$, $\mathbf{v}_2 = \mathbf{e}_2 = (0, 1)$ et $\mathbf{v}_3 = (1, 1)$ génère $\mathbb{R}^2$, ce qui s'écrit par l'équation $\mathrm{Vect}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathbb{R}^2$. En effet, pour tout $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$, on a $\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + 0\mathbf{v}_3$.

La proposition 5.2.2 qui suit énonce le résultat évident que toute famille génératrice est contenue dans l'espace qu'elle génère.

Proposition 5.2.2 (*les vecteurs d'une famille appartiennent à l'espace qu'elle génère*)
Si la famille F génère l'espace vectoriel V, alors $F \subseteq V$.

Preuve. Soit F une famille de vecteurs, soit $V = \text{Vect}(F)$, et soit $\mathbf{v}$ un vecteur de F. Étant donné que l'espace vectoriel V est l'ensemble de toutes les combinaisons linéaires des vecteurs de F, y compris la combinaison $1\mathbf{v}$, on déduit que $\mathbf{v} \in V$. Comme ceci est vrai pour tout $\mathbf{v} \in F$, il s'ensuit que $F \subseteq V$. □

5.2.3 Bases

Les notions d'indépendance linéaire et d'ensembles générateurs ayant été introduites, il est maintenant possible de définir ce qu'est une base d'un espace vectoriel.

Définition 5.2.9 (*base d'un espace vectoriel*)
Une famille $\mathcal{B}$ *forme* une *base* de l'espace vectoriel V si les deux conditions suivantes sont satisfaites :

1. les vecteurs de $\mathcal{B}$ sont linéairement indépendants ;
2. les vecteurs de $\mathcal{B}$ génèrent V.

Une base $\mathcal{B}$ d'un espace vectoriel est donc une famille génératrice libre. L'exemple 5.2.4 illustre le fait que les bases ne sont pas uniques. Une autre façon de faire ce constat est de simplement remarquer que si $(\mathbf{v})$ est une base d'un espace vectoriel V, alors $(\alpha\mathbf{v})$ aussi, pour tout $\alpha \neq 0$. Cela signifie que si un espace vectoriel admet une base non vide (voir proposition 5.2.3 pour l'autre cas), alors cet espace possède une infinité de bases.

Exemple 5.2.4
La famille $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ de l'exemple 5.2.3 n'est pas une base de $\mathbb{R}^2$ car elle n'est pas libre (car $\mathbf{v}_3 = \mathbf{v}_1 + \mathbf{v}_2$). On a constaté que chaque vecteur $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ peut s'écrire comme $\mathbf{x} = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$, donc $\text{Vect}(\mathbf{v}_1, \mathbf{v}_2) = \mathbb{R}^2$. De plus, comme $\mathbf{v}_1 = \mathbf{e}_1$ et $\mathbf{v}_2 = \mathbf{e}_2$ sont indépendants, $(\mathbf{v}_1, \mathbf{v}_2) = (\mathbf{e}_1, \mathbf{e}_2) = \mathcal{E}$ est une base de $\mathbb{R}^2$, appelée base canonique. La famille $(\mathbf{v}_1, \mathbf{v}_3)$ est une autre famille génératrice de $\mathbb{R}^2$. En effet, chaque vecteur $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ peut s'écrire comme

$$\mathbf{x} = (x_1 - x_2)\mathbf{v}_1 + x_2\mathbf{v}_3 = (x_1 - x_2)\begin{bmatrix}1\\0\end{bmatrix} + x_2\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}x_1 - x_2 + x_2\\0 + x_2\end{bmatrix} = \begin{bmatrix}x_1\\x_2\end{bmatrix}.$$

De plus, $(\mathbf{v}_1, \mathbf{v}_3)$ est une famille libre car les vecteurs $\mathbf{v}_1$ et $\mathbf{v}_3$ ne sont pas colinéaires. On peut donc conclure que $(\mathbf{v}_1, \mathbf{v}_3)$ est une autre base de $\mathbb{R}^2$.

Proposition 5.2.3 (*base de* $\{\mathbf{0}\}$)

L'espace vectoriel $\{\mathbf{0}\}$ possède une unique base qui est $\emptyset$ (c'est-à-dire l'ensemble vide).

Preuve. Soit $\mathcal{B}$ une base de $\{\mathbf{0}\}$. $\mathcal{B}$ est donc une famille génératrice de $\{\mathbf{0}\}$, et selon la proposition 5.2.2, $\mathcal{B} \subseteq \{\mathbf{0}\}$: les vecteurs de $\mathcal{B}$ ne peuvent être que le vecteur nul. Or, $\mathcal{B}$ est aussi une famille libre, et selon la proposition 5.2.1, aucune famille libre ne peut contenir le vecteur nul. La seule option possible est donc d'avoir $\mathcal{B} = \emptyset$. □

Le théorème 5.2.1 qui suit est fondamental car il montre, en particulier, que chaque vecteur d'un espace vectoriel peut s'écrire de façon unique dans une base donnée de cet espace.

Théorème 5.2.1

Soit $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p)$, une famille de vecteurs de l'espace vectoriel V, et soit $\mathbf{x} \in V$ un vecteur quelconque de V. La famille $\mathcal{B}$ est une base de V si et seulement s'il existe une unique famille de scalaires $[\mathbf{x}]_{\mathcal{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_p)$ telle que

$$\mathbf{x} = \sum_{i=1}^{p} \alpha_i \mathbf{b}_i \ .$$

La notation $[\mathbf{x}]_{\mathcal{B}}$ désigne les *coordonnées* uniques de $\mathbf{x}$ dans la base $\mathcal{B}$.

Preuve. On montre tout d'abord que si $\mathcal{B}$ est une base, alors $[\mathbf{x}]_{\mathcal{B}}$ est unique, quel que soit $\mathbf{x} \in V$. On raisonne par contradiction. Supposons qu'il existe deux familles de scalaires $(\alpha_1, \alpha_2, \ldots, \alpha_p)$ et $(\beta_1, \beta_2, \ldots, \beta_p)$ distinctes telles que

$$\mathbf{x} = \sum_{i=1}^{p} \alpha_i \mathbf{b}_i \ , \tag{5.1a}$$

$$\mathbf{x} = \sum_{i=1}^{p} \beta_i \mathbf{b}_i \ . \tag{5.1b}$$

En soustrayant l'équation (5.1a) à (5.1b), on obtient

$$\mathbf{x} - \mathbf{x} = \mathbf{0} = \sum_{i=1}^{p} (\beta_i - \alpha_i) \mathbf{b}_i$$

avec au moins un coefficient $(\beta_i - \alpha_i) \neq 0$ pour $i \in [\![1; p]\!]$ car les deux familles de scalaires sont distinctes. On a donc une combinaison linéaire non triviale des éléments de $\mathcal{B}$ qui donne le vecteur nul. La famille $\mathcal{B}$ est donc liée, ce qui contredit le fait que $\mathcal{B}$ est une base.

Réciproquement, supposons que pour tout vecteur $\mathbf{x} \in V$, il existe une unique famille de scalaires $(\alpha_1, \alpha_2, \ldots, \alpha_p) \subseteq V$ telle que

$$\mathbf{x} = \sum_{i=1}^{p} \alpha_i \mathbf{b}_i \ . \tag{5.2}$$

Pour montrer que $\mathcal{B}$ est une base, il est d'abord aisé de remarquer que $\mathcal{B}$ est une famille génératrice de V car tout $\mathbf{x} \in V$ peut s'exprimer comme combinaison linéaire de vecteurs de $\mathcal{B}$. Il reste à montrer que $\mathcal{B}$ est libre. Pour cela, exploitons le fait que (5.2) est valide pour tout $\mathbf{x} \in V$, incluant $\mathbf{0}$. La combinaison triviale est donc l'unique combinaison linéaire des vecteurs de $\mathcal{B}$ donnant le vecteur nul, ce qui prouve que $\mathcal{B}$ est une famille libre en plus d'être génératrice de V. C'est donc une base de V. □

Remarque

Pour que le résultat énoncé dans le théorème précédent soit valide, il est important de conserver l'ordre des vecteurs $\mathbf{b}_1$, $\mathbf{b}_2$, …, $\mathbf{b}_p$ dans $\mathcal{B}$, ceci afin de s'assurer que chaque α_i, pour $i \in [\![1;p]\!]$, corresponde au $\mathbf{b}_i$ de même indice dans la combinaison linéaire qui donne $\mathbf{x}$. Ceci justifie le fait que les bases sont des familles (avec ordre) et non des ensembles (sans ordre).

Il est à noter que la section 5.2.5 porte sur les coordonnées de vecteurs dans plusieurs bases, et indique comment exprimer un même vecteur dans des bases différentes.

Proposition 5.2.4 *(base canonique de $\mathbb{R}^n$)*

La famille $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$ (notée aussi $\mathcal{E}_n$ en cas d'ambiguïté), où $\mathbf{e}_j$ est la j-ième colonne de $\mathbf{I}$, pour $j \in [\![1;n]\!]$, est une base de $\mathbb{R}^n$, appelée la *base canonique* de $\mathbb{R}^n$.

Preuve. Soit $\mathbf{x}$ un vecteur quelconque de $\mathbb{R}^n$. Selon le théorème 5.2.1, $\mathcal{E}$ est une base de $\mathbb{R}^n$ s'il n'existe qu'une seule combinaison linéaire des vecteurs de $\mathcal{E}$ qui donne $\mathbf{x}$. Une telle combinaison s'écrit $x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n = \mathbf{x}$, avec $(x_1, x_2, \ldots, x_n)$ une famille de scalaires, et elle correspond au *système d'équations linéaires*

$$\mathbf{I}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{x}$$

qui fait apparaître les coordonnées de $\mathbf{x}$ comme étant x_1, x_2, …, x_n. L'écriture unique de $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ est donc une preuve que $\mathcal{E}$ est une base de $\mathbb{R}^n$. □

Par défaut, la base canonique est celle utilisée pour les coordonnées d'un vecteur de $\mathbb{R}^n$:

$$\mathbf{x} = [\mathbf{x}]_{\mathcal{E}} = (x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n \in \mathbb{R}^n \ .$$

Remarque (*base canonique de* $\mathbb{C}^n$)

La proposition 5.2.4 est tout aussi valable pour $\mathbb{C}^n$ en tant qu'espace vectoriel complexe : la base canonique $\mathcal{E}_n$ est une base de $\mathbb{C}^n$. La preuve est identique.

L'exemple 5.2.5 suivant illustre cela.

Exemple 5.2.5

$\mathbf{z} = (1, \mathrm{i}+2, -1-\mathrm{i}) \in \mathbb{C}^3$ est la combinaison unique suivante des vecteurs de $\mathcal{E}_3$:

$$\begin{aligned} \mathbf{z} = [\mathbf{z}]_{\mathcal{E}_3} = 1\mathbf{e}_1 + (\mathrm{i}+2)\mathbf{e}_2 + (-1-\mathrm{i})\mathbf{e}_3 &= 1\begin{bmatrix}1\\0\\0\end{bmatrix} + (\mathrm{i}+2)\begin{bmatrix}0\\1\\0\end{bmatrix} + (-1-\mathrm{i})\begin{bmatrix}0\\0\\1\end{bmatrix} \\ &= \begin{bmatrix}1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\\mathrm{i}+2\\0\end{bmatrix} + \begin{bmatrix}0\\0\\-1-\mathrm{i}\end{bmatrix} = \begin{bmatrix}1\\\mathrm{i}+2\\-1-\mathrm{i}\end{bmatrix} = (1, \mathrm{i}+2, -1-\mathrm{i}) \ . \end{aligned}$$

On termine cette section en définissant la base canonique des polynômes. Cette définition est utilisée pour l'exercice 5.11.

Définition 5.2.10 (*base canonique des polynômes*)

La famille $(1, x, x^2, \ldots, x^n)$, notée $\mathcal{E}$ ou $\mathcal{E}_n$ en cas d'ambiguïté, est appelée **la base canonique** de $\mathcal{P}_n(\mathbb{R})$ ou $\mathcal{P}_n(\mathbb{C})$.

Les *bases orthogonales*, particulièrement pratiques, seront introduites à la section 5.3.2.

5.2.4 Dimension d'un espace vectoriel

La *dimension d'un espace vectoriel* est une notion fondamentale de l'algèbre. Ce concept est simple et intuitif, mais il faut prendre garde au fait que le terme “dimension” appartient au langage courant dans bien des contextes différents. Il faut donc l'utiliser adéquatement. Par exemple, la notion de dimension est toujours associée à un espace vectoriel. Il ne faut donc pas utiliser ce terme pour parler de la *taille* d'un vecteur ou d'une matrice.

La définition de la dimension est basée sur le résultat important du lemme 5.2.1 qui indique que les familles libres d'un espace vectoriel ne contiennent jamais plus de vecteurs que les familles génératrices de cet espace. Ce lemme est une version simplifiée du théorème de l'échange, aussi appelé lemme de Steinitz. La conséquence directe de ce résultat, exposée au corollaire 5.2.1, est que toutes les bases d'un espace vectoriel possèdent exactement le même nombre de vecteurs, et ce nombre est la dimension de l'espace.

Lemme 5.2.1 (*lemme de l'échange*)

Soit $U \subseteq V$ une famille libre d'un espace vectoriel V, comportant m vecteurs, et soit $W \subseteq V$ une famille génératrice de n vecteurs de V. Alors

$$m \leq n \ .$$

La preuve du lemme 5.2.1 ci-dessous se fait par récurrence. Elle est d'un niveau avancé, et exige que la section B.6 soit maîtrisée.

Preuve du lemme de l'échange. Soit $U = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$ une famille libre de V et soit $W = (\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n)$ une famille génératrice de V. On établit par récurrence sur m que $m \leq n$ et que, après ré-ordonnancement éventuel des $\mathbf{w}_j$, pour $j \in [\![1; n]\!]$, la famille $V_m = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m, \mathbf{w}_{m+1}, \mathbf{w}_{m+2}, \ldots, \mathbf{w}_n)$ génère V.

- ▷ **Initialisation.** Si $m = 0$, l'énoncé est immédiat car $n \geq 0$, $V_0 = W$, et $\mathrm{Vect}(W) = V$.
- ▷ **Hypothèse de récurrence.** On suppose la propriété vraie pour $m - 1 \geq 0$, ce qui implique $m - 1 \leq n$. Aussi, on réordonne les $\mathbf{w}_j$, pour $j \in [\![1; n]\!]$, de sorte que

$$\mathrm{Vect}(V_{m-1}) = \mathrm{Vect}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1}, \mathbf{w}_m, \mathbf{w}_{m+1}, \ldots, \mathbf{w}_n) = V \ .$$

- ▷ **On montre que la propriété est vraie pour m.** On note tout d'abord que $m \leq n$. En effet, si tel n'était pas le cas, on aurait $m - 1 = n$ et $V_{m-1} = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1})$, ce qui impliquerait que $\mathbf{u}_m \in V = \mathrm{Vect}(V_{m-1})$, et $\mathbf{u}_m$ serait donc une combinaison linéaire des vecteurs $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1}$, ce qui est impossible car U est une famille libre.

  Étant donné que $\mathbf{u}_m$ appartient à $\mathrm{Vect}(V_{m-1})$, il existe des scalaires $\alpha_1, \alpha_2, \ldots, \alpha_n$ tels que

$$\mathbf{u}_m = \sum_{i=1}^{m-1} \alpha_i \mathbf{u}_i + \sum_{j=m}^{n} \alpha_j \mathbf{w}_j \ .$$

  Au moins un des poids α_j, pour $j \geq m$, est non nul, sinon cette égalité contredirait l'indépendance linéaire des $\mathbf{u}_1$, $\mathbf{u}_2$, ..., $\mathbf{u}_m$. En réordonnant si nécessaire les $\mathbf{w}_j$, pour $j \in [\![m; n]\!]$, on peut supposer $\alpha_m \neq 0$ et on obtient

$$\mathbf{w}_m = \frac{1}{\alpha_m}\Big(\mathbf{u}_m - \sum_{i=1}^{m-1} \alpha_i \mathbf{u}_i - \sum_{j=m+1}^{n} \alpha_j \mathbf{w}_j\Big) \ .$$

  Ainsi, $\mathbf{w}_m \in \mathrm{Vect}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m, \mathbf{w}_{m+1}, \mathbf{w}_{m+2}, \ldots, \mathbf{w}_n) = \mathrm{Vect}(V_m)$. De plus, les autres vecteurs $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{m-1}, \mathbf{w}_{m+1}, \mathbf{w}_{m+2}, \ldots, \mathbf{w}_n$ de la famille V_{m-1} sont dans

$V_m \subseteq \mathrm{Vect}(V_m)$. Par conséquent, $\mathrm{Vect}(V_{m-1}) \subseteq \mathrm{Vect}(V_m)$. Comme V_{m-1} génère V (par hypothèse de récurrence), on déduit que

$$V = \mathrm{Vect}(V_{m-1}) \subseteq \mathrm{Vect}(V_m) \subseteq V$$

et donc $\mathrm{Vect}(V_m) = V$. La propriété est ainsi démontrée.

$\square$

Il est désormais simple d'énoncer le corollaire 5.2.1 selon lequel toutes les bases d'un espace vectoriel possèdent le même nombre de vecteurs.

Corollaire 5.2.1

Si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases d'un même espace vectoriel V, alors

$$|\mathcal{B}| = |\mathcal{C}| \ .$$

Preuve. Étant donné que $\mathcal{B}$ est une famille libre et que $\mathcal{C}$ est une famille génératrice de V, on déduit du lemme 5.2.1 que $|\mathcal{B}| \leq |\mathcal{C}|$. De même, comme $\mathcal{C}$ est libre et $\mathcal{B}$ génère V, on déduit que $|\mathcal{C}| \leq |\mathcal{B}|$. Ainsi $|\mathcal{B}| = |\mathcal{C}|$. $\square$

La définition de la dimension d'un espace vectoriel peut maintenant être donnée.

Définition 5.2.11 (*dimension d'un espace vectoriel*)

La dimension d'un espace vectoriel V est notée $\dim(V)$ et correspond au nombre de vecteurs de n'importe quelle base de V.

Illustrons cette définition sur le simple exemple suivant.

Exemple 5.2.6 (*dimension d'un espace vectoriel*)

Soit $V = \mathrm{Vect}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ un espace vectoriel où $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (3, 2, 1)$ et $\mathbf{v}_3 = (2, 0, 0)$. Pour déterminer la dimension de V, il suffit de donner une base de cet espace. La famille $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ n'est pas libre puisque $\mathbf{v}_3 = \mathbf{v}_2 - \mathbf{v}_1$. Donc $V = \mathrm{Vect}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \mathrm{Vect}(\mathbf{v}_1, \mathbf{v}_2)$, et comme $(\mathbf{v}_1, \mathbf{v}_2)$ est une famille libre (car $\mathbf{v}_1$ et $\mathbf{v}_2$ ne sont pas colinéaires), on déduit que $(\mathbf{v}_1, \mathbf{v}_2)$ est une base de V. La dimension de V est donc $\dim(V) = 2$ (V est un plan de $\mathbb{R}^3$).

Remarque (*dimension de* $\{\mathbf{0}\}$)

On a vu à la proposition 5.2.3 que la seule base de l'espace vectoriel $\{\mathbf{0}\}$ est l'ensemble vide dont la cardinalité est zéro. Ainsi $\{\mathbf{0}\}$ est de dimension 0 : $\dim(\{\mathbf{0}\}) = |\emptyset| = 0$.

Le concept de dimension permet, entre autres, de définir formellement certaines notions intuitives. Par exemple :

- ▷ $\dim(\mathbb{R}^n) = \dim(\mathbb{C}^n) = n$ car $|\mathcal{E}_n| = n$ (voir proposition 5.2.4) ;
- ▷ l'espace 2D est $\mathbb{R}^2$ et l'espace 3D est $\mathbb{R}^3$ ("D" est bien sûr pour "dimension") ;
- ▷ tout sous-espace vectoriel de $\mathbb{R}^n$ de dimension 1 est une *droite* de $\mathbb{R}^n$ (passant par l'origine) ;
- ▷ tout sous-espace vectoriel de $\mathbb{R}^n$ de dimension 2 est un *plan* de $\mathbb{R}^n$ (passant par l'origine) ;
- ▷ tout sous-espace vectoriel de $\mathbb{R}^n$ de dimension $n-1$ est un *hyperplan* de $\mathbb{R}^n$;
- ▷ les hyperplans de $\mathbb{R}^3$ sont donc des plans.

À des fins de simplicité, dans cet ouvrage, les droites, plans, et hyperplans, sont toujours des espaces vectoriels, et contiennent donc toujours le vecteur nul.

Cette section se termine par une série de résultats intuitifs découlant directement du concept de dimension. Tout d'abord, la proposition 5.2.5 prouve qu'un espace vectoriel de dimension n ne peut pas contenir plus de n vecteurs linéairement indépendants.

Proposition 5.2.5

Si V est un espace vectoriel de dimension n, alors toute famille libre de V contient au plus n vecteurs.

Preuve. Soit $F \subseteq V$ une famille libre de V. Selon le lemme 5.2.1, toute famille libre de V possède moins de vecteurs que toute famille génératrice de V. Comme $\dim(V) = n$, on peut trouver des familles génératrices de n vecteurs. Ainsi $|F| \leq \dim(V) = n$. □

Remarque

Il découle de la proposition 5.2.5 que l'espace 3D (c'est-à-dire $\mathbb{R}^3$) ne peut pas avoir plus de trois vecteurs linéairement indépendants. De même, dans l'espace 2D, qui est $\mathbb{R}^2$, on ne peut avoir plus de deux vecteurs linéairement indépendants.

De façon générale, la proposition 5.2.5 est très utile pour, parfois, prouver simplement qu'une famille de vecteurs est liée.

La proposition 5.2.6 stipule qu'il est impossible de générer un espace vectoriel de dimension n avec moins de n vecteurs.

Proposition 5.2.6

Si V est un espace vectoriel de dimension n, alors toute famille génératrice de V doit contenir au moins n vecteurs.

Preuve. Soit $F \subseteq V$ une famille génératrice de V. Selon le lemme 5.2.1, toute famille libre de V possède moins de vecteurs que toute famille génératrice de V. Comme $\dim(V) = n$, on peut trouver des familles libres de n vecteurs. Ainsi $|F| \geq \dim(V) = n$. □

Le résultat suivant est un corollaire direct du lemme 5.2.1. On ne le mentionne qu'ici car il est très utile pour prouver la proposition 5.2.7.

Corollaire 5.2.2 *(de la base incomplète)*

Si $U = (\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$ est une famille libre d'un espace vectoriel V dont une base est $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$, alors il existe une sous-famille de vecteurs $B \subseteq \mathcal{B}$ telle que $|B| = |\mathcal{B}| - |U|$ et $B \cup U$ est une base de V (c'est à dire qu'il est possible de compléter la famille U avec la famille B pour obtenir une base de V).

Preuve. Poser $B = V_m \setminus U = (\mathbf{b}_{m+1}, \mathbf{b}_{m+2}, \ldots, \mathbf{b}_n)$ dans la preuve du lemme 5.2.1. Il faut noter que dans la preuve, $W = \mathcal{B}$ est bien génératrice car c'est une base de V. □

La proposition qui suit permet de prouver qu'une famille libre est une base, sans avoir à montrer qu'elle est génératrice. Elle est utilisée, par exemple, pour répondre à l'exercice 5.13.

Proposition 5.2.7

Une famille libre de p vecteurs d'un espace vectoriel de dimension p est nécessairement une base de cet espace.

Preuve. Puisqu'une famille libre de p vecteurs est déjà de la même taille que la dimension de l'espace, elle est nécessairement une base. Si elle ne l'était pas, elle pourrait être complétée en une base, mais cela impliquerait d'ajouter des vecteurs, selon le corollaire 5.2.2, ce qui n'est pas possible car on a déjà atteint le nombre maximal de vecteurs linéairement indépendants pour cet espace. □

La proposition 5.2.8 suivante formalise le fait que si un sous-espace vectoriel est de même dimension que son sur-espace, alors il est nécessairement égal à son sur-espace.

Proposition 5.2.8

Si W est un sous-espace vectoriel d'un espace vectoriel V et si $\dim(W) = \dim(V)$, alors $W = V$.

Preuve. Soit V un espace vectoriel et soit $W \subseteq V$ un sous-espace vectoriel de V tel que $\dim(W) = \dim(V)$. Soit $\mathcal{B}$ une base quelconque de W. On a donc $W = \mathrm{Vect}(\mathcal{B})$. Comme tous les vecteurs de $\mathcal{B}$ sont aussi dans V, $\mathcal{B}$ est une famille libre de V. Et selon la proposition 5.2.7, c'est aussi une base de V, et $V = \mathrm{Vect}(\mathcal{B})$. Donc $W = V$. □

La proposition 5.2.8 permet, par exemple, de déduire que si un sous-espace vectoriel de $\mathbb{R}^3$ est de dimension trois, ce sous-espace est $\mathbb{R}^3$ lui-même.

5.2.5 Changements de base

Il a été établi au théorème 5.2.1 que pour n'importe quel espace vectoriel V et n'importe quelle base $\mathcal{B} \subseteq V$, il est possible d'exprimer chaque vecteur $\mathbf{x} \in V$ de façon unique à l'aide du vecteur $[\mathbf{x}]_{\mathcal{B}}$ dont les composantes sont les coordonnées de $\mathbf{x}$ dans la base $\mathcal{B}$. Certaines bases sont assurément plus pratiques que d'autres, et il convient donc de savoir exprimer des vecteurs dans plusieurs bases différentes. Ainsi, si on considère une autre base $\mathcal{C}$ de V, les coordonnées de $\mathbf{x}$ dans $\mathcal{C}$ seront différentes des coordonnées de $\mathbf{x}$ dans $\mathcal{B}$: $[\mathbf{x}]_{\mathcal{B}} \neq [\mathbf{x}]_{\mathcal{C}}$. Cette section introduit le concept de *changement de base*, pour des espaces vectoriels dans $\mathbb{R}^n$ ou dans $\mathbb{C}^n$. Ces changements de base permettent d'obtenir $[\mathbf{x}]_{\mathcal{B}}$ à partir de $[\mathbf{x}]_{\mathcal{C}}$, et inversement. Pour simplifier la présentation, le reste de cette section sur les changements de base se concentre sur $\mathbb{R}^n$, mais tous les résultats peuvent aussi s'appliquer à $\mathbb{C}^n$, ainsi qu'illustré dans l'exemple 5.2.9.

Étant donné un vecteur de $\mathbb{R}^2$ dont on connaît les composantes dans la base canonique $\mathcal{E}$, l'exemple suivant illustre comment exprimer ce vecteur dans une autre base $\mathcal{B}$ de $\mathbb{R}^2$.

Exemple 5.2.7 (*exprimer* $\mathbf{x}$ *dans une autre base que la base canonique*)
Soit $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2) = ((1,1),(0,2))$. Étant donné que $\mathbf{b}_1$ et $\mathbf{b}_2$ ne sont pas colinéaires, $\mathcal{B}$ est une famille libre de dimension 2, ce qui prouve que $\mathcal{B}$ est une base de $\mathbb{R}^2$.
Soit $\mathbf{x} = (5,6) \in \mathbb{R}^2$ dont les composantes sont les coordonnées de $\mathbf{x}$ dans la base canonique : $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}}$. On veut déterminer $[\mathbf{x}]_{\mathcal{B}}$, c'est-à-dire les coordonnées de $\mathbf{x}$ dans $\mathcal{B}$, et pour cela, on doit trouver les scalaires α_1 et α_2 tels que

$$\mathbf{x} = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2$$

ce qui s'écrit sous forme matricielle

$$[\mathbf{b}_1 \mathbf{b}_2] \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \mathbf{x} \ .$$

C'est un système d'équations linéaires, dont la résolution est le sujet du chapitre 8, mais celui-ci peut déjà être résolu grâce au théorème 1.4.2, car la matrice $\mathbf{B} = [\mathbf{b}_1 \ \ \mathbf{b}_2] = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$ est inversible. Son inverse est facile à calculer (voir la proposition 8.5.1), ce qui donne

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \mathbf{B}^{-1}\mathbf{x} = \frac{1}{2} \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 5 \\ \frac{1}{2} \end{bmatrix}$$

et donc

$$\mathbf{x} = 5\mathbf{b}_1 + \frac{1}{2}\mathbf{b}_2 \ ,$$

c'est-à-dire $[\mathbf{x}]_{\mathcal{B}} = \left(5, \frac{1}{2}\right)$.

La proposition qui suit indique comment généraliser l'exemple ci-dessus afin d'exprimer les vecteurs d'un sous-espace vectoriel de $\mathbb{R}^n$ dans n'importe quelle base de ce sous-espace.

Proposition 5.2.9 *(liens entre* $\mathbf{x}$ *et* $[\mathbf{x}]_{\mathcal{B}}$*)*

Soit $\mathbf{x} \in V$ avec V un sous-espace vectoriel de dimension p de $\mathbb{R}^n$, dont une base est $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p)$. Ce vecteur $\mathbf{x}$ peut s'exprimer comme

$$\mathbf{x} = \mathbf{B}[\mathbf{x}]_{\mathcal{B}}$$

avec $\mathbf{B} = [\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_p] \in \mathbb{R}^{n\times p}$ et $[\mathbf{x}]_{\mathcal{B}} \in \mathbb{R}^p$, les coordonnées de $\mathbf{x}$ dans $\mathcal{B}$.
Si de plus $p = n$, alors $\mathcal{B}$ est une base de $\mathbb{R}^n$ et

$$[\mathbf{x}]_{\mathcal{B}} = \mathbf{B}^{-1}[\mathbf{x}]_{\mathcal{E}} = \mathbf{B}^{-1}\mathbf{x} \in \mathbb{R}^n\ .$$

Preuve. Dans la base $\mathcal{B}$, les coordonnées de $\mathbf{x} \in V$ sont $[\mathbf{x}]_{\mathcal{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_p)$ et on a $\mathbf{x} = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \ldots + \alpha_p\mathbf{b}_p$, ce qui correspond au produit d'une matrice par un vecteur vu à la section 1.2.5, soit $\mathbf{x} = \mathbf{B}[\mathbf{x}]_{\mathcal{B}}$. Si $p = n$, alors la proposition 5.2.7 nous assure que $\mathcal{B}$ est une base de $\mathbb{R}^n$, de plus $\mathbf{B}$ est carrée et ses colonnes sont linéairement indépendantes (car $\mathcal{B}$ est une base). $\mathbf{B}$ est donc inversible selon le cas 6 du théorème 1.4.1. Il découle du théorème 1.4.2 qu'il existe une solution unique au système $\mathbf{B}[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}$ donnée par $[\mathbf{x}]_{\mathcal{B}} = \mathbf{B}^{-1}\mathbf{x}$. De plus, en considérant la base canonique $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$, on a $\mathbf{x} = \mathbf{I}[\mathbf{x}]_{\mathcal{E}} = [\mathbf{x}]_{\mathcal{E}}$. $\square$

Considérons maintenant une base $\mathcal{C}$, et supposons qu'on s'intéresse à obtenir les coordonnées d'un vecteur dans cette base, en connaissant ses coordonnées dans une autre base $\mathcal{B}$. Ce processus est illustré sur l'exemple 5.2.8, qui est la suite de l'exemple 5.2.7. Tout ceci sera par la suite formalisé à l'aide du lemme 5.2.2 et du théorème 5.2.2.

Exemple 5.2.8 *(exprimer* $\mathbf{x}$ *dans deux bases autres que la base canonique)*

Soient $\mathcal{B} = ((1,1),(0,2))$ et $\mathbf{x} = (5,6)$ la base et le vecteur de $\mathbb{R}^2$ de l'exemple 5.2.7. Soit $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2) = ((-1,3),(1,12))$ une autre base de $\mathbb{R}^2$. On veut exprimer $\mathbf{x}$ dans la base $\mathcal{C}$ connaissant ses coordonnées dans la base $\mathcal{B}$. Une option consiste à utiliser la proposition 5.2.9 pour avoir

$$[\mathbf{x}]_{\mathcal{C}} = \mathbf{C}^{-1}\mathbf{x} \tag{5.3}$$

avec $\mathbf{C} = [\mathbf{c}_1\ \mathbf{c}_2]$, ce qui donne

$$[\mathbf{x}]_\mathcal{C} = \frac{1}{15}\begin{bmatrix} -12 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \frac{1}{15}\begin{bmatrix} -54 \\ 21 \end{bmatrix} .$$

Pour exprimer $[\mathbf{x}]_\mathcal{C}$ en fonction de $[\mathbf{x}]_\mathcal{B}$, il suffit d'injecter $\mathbf{x} = \mathbf{B}[\mathbf{x}]_\mathcal{B}$ dans (5.3) et on obtient

$$[\mathbf{x}]_\mathcal{C} = \mathbf{C}^{-1}\mathbf{B}[\mathbf{x}]_\mathcal{B}$$

c'est-à-dire

$$[\mathbf{x}]_\mathcal{C} = \frac{1}{15}\begin{bmatrix} -12 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ \frac{1}{2} \end{bmatrix} = \frac{1}{15}\begin{bmatrix} -11 & 2 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ \frac{1}{2} \end{bmatrix} = \frac{1}{15}\begin{bmatrix} -54 \\ 21 \end{bmatrix} .$$

La matrice qui permet de "passer" de $[\mathbf{x}]_\mathcal{B}$ à $[\mathbf{x}]_\mathcal{C}$ est notée $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$, et on a donc

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{C}^{-1}\mathbf{B} = \tfrac{1}{15}\begin{bmatrix} -11 & 2 \\ 4 & 2 \end{bmatrix} .$$

Avant de généraliser l'exemple 5.2.8 au théorème 5.2.2, il convient de réaliser que les bases en jeu dans cet exemple sont des bases de $\mathbb{R}^2$, ce qui signifie que les matrices étudiées sont carrées. Dans le cas général, on s'intéresse à des bases de sous-espaces vectoriels de $\mathbb{R}^n$ de dimension $p \leq n$, comme par exemple à l'exercice 5.13.

Le lemme 5.2.2 suivant est utile pour prouver le théorème 5.2.2. Il établit, d'une certaine manière, une propriété de "distributivité" des coordonnées d'une combinaison linéaire dans une certaine base.

Lemme 5.2.2 (*distributivité de* $[\cdot]_\mathcal{B}$)

Soient p vecteurs $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_p$ d'un espace vectoriel V dont $\mathcal{B}$ est une base, et soient p scalaires α_1, α_2, …, α_p. On a

$$[\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_p\mathbf{v}_p]_\mathcal{B} = \alpha_1[\mathbf{v}_1]_\mathcal{B} + \alpha_2[\mathbf{v}_2]_\mathcal{B} + \ldots + \alpha_p[\mathbf{v}_p]_\mathcal{B} .$$

La preuve du lemme 5.2.2 est technique et demande d'absolument maîtriser l'usage de l'opérateur de somme $\sum$, ainsi que les double sommes $\sum\sum$.

Preuve. On exprime tout d'abord $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_p$ dans la base $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ de V, où $n = \dim(V)$:

$$\mathbf{v}_1 = \gamma_1^1\mathbf{b}_1 + \gamma_2^1\mathbf{b}_2 + \ldots + \gamma_n^1\mathbf{b}_n = \sum_{i=1}^n \gamma_i^1\mathbf{b}_i$$

$$\mathbf{v}_2 = \gamma_1^2\mathbf{b}_1 + \gamma_2^2\mathbf{b}_2 + \ldots + \gamma_n^2\mathbf{b}_n = \sum_{i=1}^{n} \gamma_i^2\mathbf{b}_i$$

$$\vdots$$

$$\mathbf{v}_p = \gamma_1^p\mathbf{b}_1 + \gamma_2^p\mathbf{b}_2 + \ldots + \gamma_n^p\mathbf{b}_n = \sum_{i=1}^{n} \gamma_i^p\mathbf{b}_i$$

où γ_i^j est un scalaire pour tout $i \in [\![1;n]\!]$ et pour tout $j \in [\![1;p]\!]$. Considérons maintenant le vecteur $\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_p\mathbf{v}_p$. Il peut s'écrire comme suit :

$$\begin{aligned}
\mathbf{x} = \alpha_1 \left(\sum_{i=1}^{n} \gamma_i^1\mathbf{b}_i\right) + \alpha_2 \left(\sum_{i=1}^{n} \gamma_i^2\mathbf{b}_i\right) + \ldots + \alpha_p \left(\sum_{i=1}^{n} \gamma_i^p\mathbf{b}_i\right) \\
= \sum_{j=1}^{p} \alpha_j \left(\sum_{i=1}^{n} \gamma_i^j\mathbf{b}_i\right) \\
= \sum_{j=1}^{p} \sum_{i=1}^{n} \alpha_j\gamma_i^j\mathbf{b}_i \\
= \sum_{i=1}^{n} \sum_{j=1}^{p} \alpha_j\gamma_i^j\mathbf{b}_i \\
= \sum_{i=1}^{n} \left(\sum_{j=1}^{p} \alpha_j\gamma_i^j\right)\mathbf{b}_i \ .
\end{aligned}$$

On a ainsi réussi à exprimer $\mathbf{x}$ comme combinaison linéaire des vecteurs de $\mathcal{B}$:

$$[\mathbf{x}]_{\mathcal{B}} = \left(\sum_{j=1}^{p} \alpha_j\gamma_1^j, \sum_{j=1}^{p} \alpha_j\gamma_2^j, \ldots, \sum_{j=1}^{p} \alpha_j\gamma_n^j\right) = \begin{bmatrix} \sum_{j=1}^{p} \alpha_j\gamma_1^j \\ \sum_{j=1}^{p} \alpha_j\gamma_2^j \\ \vdots \\ \sum_{j=1}^{p} \alpha_j\gamma_n^j \end{bmatrix} .$$

Il ne reste plus qu'à développer $\alpha_1[\mathbf{v}_1]_{\mathcal{B}} + \alpha_2[\mathbf{v}_2]_{\mathcal{B}} + \ldots + \alpha_p[\mathbf{v}_p]_{\mathcal{B}}$ et à vérifier qu'on obtient

$[\mathbf{x}]_{\mathcal{B}}$. C'est bien le cas, car

$$\alpha_1 \begin{bmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \vdots \\ \gamma_n^1 \end{bmatrix} + \alpha_2 \begin{bmatrix} \gamma_1^2 \\ \gamma_2^2 \\ \vdots \\ \gamma_n^2 \end{bmatrix} + \cdots + \alpha_p \begin{bmatrix} \gamma_1^p \\ \gamma_2^p \\ \vdots \\ \gamma_n^p \end{bmatrix} = \begin{bmatrix} \alpha_1\gamma_1^1 + \alpha_2\gamma_1^2 + \ldots + \alpha_p\gamma_1^p \\ \alpha_1\gamma_2^1 + \alpha_2\gamma_2^2 + \ldots + \alpha_p\gamma_2^p \\ \vdots \\ \alpha_1\gamma_n^1 + \alpha_2\gamma_n^2 + \ldots + \alpha_p\gamma_n^p \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{p} \alpha_j\gamma_1^j \\ \sum_{j=1}^{p} \alpha_j\gamma_2^j \\ \vdots \\ \sum_{j=1}^{p} \alpha_j\gamma_n^j \end{bmatrix} = [\mathbf{x}]_{\mathcal{B}} \ .$$

$\square$

Il est désormais aisé d'introduire formellement la matrice de changement de base.

Théorème 5.2.2 (*matrice de changement de base*)

Soit V un sous-espace vectoriel de dimension p de $\mathbb{R}^n$, soit $\mathbf{x} \in V$ et soient $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_p)$ et $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_p)$ deux bases de V. Les coordonnées de $\mathbf{x}$ dans $\mathcal{C}$ sont données par

$$[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} \in \mathbb{R}^p$$

avec

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \big[[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}} \cdots [\mathbf{b}_p]_{\mathcal{C}}\big] \in \mathbb{R}^{p\times p}$$

qui est appelée la matrice de changement de base de $\mathcal{B}$ à $\mathcal{C}$.
Si de plus $p = n$, alors $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{C}^{-1}\mathbf{B}$, avec $\mathbf{B} = [\mathbf{b}_1\ \mathbf{b}_2\ \cdots\ \mathbf{b}_n]$ et $\mathbf{C} = [\mathbf{c}_1\ \mathbf{c}_2\ \cdots\ \mathbf{c}_n]$ dans $\mathbb{R}^{n\times n}$.

Preuve. Soient $\alpha_1, \alpha_2, \ldots, \alpha_p$ les coordonnées de $\mathbf{x}$ dans la base $\mathcal{B}$: $[\mathbf{x}]_{\mathcal{B}} = (\alpha_1, \alpha_2, \ldots, \alpha_p)$. La proposition 5.2.9 donne $\mathbf{x} = \mathbf{B}[\mathbf{x}]_{\mathcal{B}}$. On a donc

$$[\mathbf{x}]_{\mathcal{C}} = \big[\mathbf{B}[\mathbf{x}]_{\mathcal{B}}\big]_{\mathcal{C}} = \big[\mathbf{B}(\alpha_1, \alpha_2, \ldots, \alpha_p)\big]_{\mathcal{C}} = \big[\alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \ldots + \alpha_p\mathbf{b}_p\big]_{\mathcal{C}}$$

ce qui, d'après le lemme 5.2.2, donne

$$[\mathbf{x}]_{\mathcal{C}} = \alpha_1[\mathbf{b}_1]_{\mathcal{C}} + \alpha_2[\mathbf{b}_2]_{\mathcal{C}} + \ldots + \alpha_p[\mathbf{b}_p]_{\mathcal{C}} = \big[[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}} \cdots [\mathbf{b}_p]_{\mathcal{C}}\big] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} = \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} \ .$$

Si $p = n$, alors $[\mathbf{x}]_{\mathcal{C}} = \mathbf{C}^{-1}\mathbf{x}$ (proposition 5.2.9) et comme $\mathbf{x} = \mathbf{B}[\mathbf{x}]_{\mathcal{B}}$ (par la même proposition 5.2.9), on a $[\mathbf{x}]_{\mathcal{C}} = \mathbf{C}^{-1}\mathbf{B}[\mathbf{x}]_{\mathcal{B}}$, et donc $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{C}^{-1}\mathbf{B}$. $\square$

On peut constater que le théorème 5.2.2 généralise la proposition 5.2.9, lorsque $\mathcal{C} = \mathcal{E}$. En effet, on a alors $[\mathbf{x}]_{\mathcal{C}} = [\mathbf{x}]_{\mathcal{E}} = \mathbf{x}$ et $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{P}_{\mathcal{E}\leftarrow\mathcal{B}} = \big[[\mathbf{b}_1]_{\mathcal{E}}\ [\mathbf{b}_2]_{\mathcal{E}} \cdots [\mathbf{b}_p]_{\mathcal{E}}\big] = [\mathbf{b}_1\ \mathbf{b}_2 \cdots \mathbf{b}_p] =$

$\mathbf{B}$, et donc $\mathbf{x} = \mathbf{B}[\mathbf{x}]_{\mathcal{B}}$. On montrera à la proposition 5.2.10 que la matrice de changement de base est inversible et que son inverse est aussi une matrice de changement de base.

Remarque (*calcul de la matrice de changement de base*)

Pour calculer $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \big[[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}} \cdots [\mathbf{b}_p]_{\mathcal{C}}\big]$, il faut exprimer les vecteurs de $\mathcal{B}$ dans la base $\mathcal{C}$. On commence par $[\mathbf{b}_1]_{\mathcal{C}}$ qui est la solution du système d'équations linéaires $\mathbf{C}[\mathbf{b}_1]_{\mathcal{C}} = \mathbf{b}_1$. On doit ensuite résoudre $\mathbf{C}[\mathbf{b}_2]_{\mathcal{C}} = \mathbf{b}_2$ pour trouver $[\mathbf{b}_2]_{\mathcal{C}}$, et ainsi de suite. Il faut donc résoudre p systèmes. Dans cette situation, on peut résoudre tous ces systèmes en parallèle, en considérant la *matrice augmentée*

$$\big[\mathbf{C} \mid \mathbf{B}\big] = \big[\mathbf{c}_1\ \mathbf{c}_2\ \cdots\ \mathbf{c}_p \mid \mathbf{b}_1\ \mathbf{b}_2\ \cdots \mathbf{b}_p\big] \sim \big[\mathbf{I} \mid \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}\big]\ .$$

Les matrices augmentées sont formellement définies à la définition 8.1.6. Elles représentent des systèmes dont la résolution est expliquée à la section 8.2. Un exemple est donné à l'exercice 5.14.

L'exemple 5.2.9 suivant illustre le concept de changement de base pour des bases de l'espace vectoriel complexe $\mathbb{C}^3$.

Exemple 5.2.9 (*changement de base dans un espace complexe*)

Soient $\mathcal{B}$ et $\mathcal{C}$ deux bases de $\mathbb{C}^3$ avec

$$\mathcal{B}{=}(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3){=}\left(\begin{bmatrix}\mathrm{i}\\-\mathrm{i}\\\mathrm{i}\end{bmatrix},\begin{bmatrix}0\\1\\3\end{bmatrix},\begin{bmatrix}0\\0\\-2i\end{bmatrix}\right)\ \text{et}\ \mathcal{C}{=}(\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3){=}\left(\begin{bmatrix}\mathrm{i}\\0\\0\end{bmatrix},\begin{bmatrix}0\\-2\\0\end{bmatrix},\begin{bmatrix}0\\0\\3\end{bmatrix}\right).$$

Pour trouver la matrice de changement de base $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$, il faut exprimer les vecteurs de la base $\mathcal{B}$ selon les vecteurs de la base $\mathcal{C}$.
Pour cela, il faut résoudre le système $[\mathbf{C} \mid \mathbf{B}]$ avec $\mathbf{B} = (\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3)$ et $\mathbf{C} = (\mathbf{c}_1\ \mathbf{c}_2\ \mathbf{c}_3)$, qui est évident à résoudre car $\mathbf{C}$ est diagonale. On obtient alors

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = [[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}}\ [\mathbf{b}_3]_{\mathcal{C}}] = \begin{bmatrix}1 & 0 & 0\\ \mathrm{i}/2 & -1/2 & 0\\ \mathrm{i}/3 & 1 & -2\mathrm{i}/3\end{bmatrix}\ .$$

Pour avoir $\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}$, il suffit d'exprimer les vecteurs de $\mathcal{C}$ selon la base $\mathcal{B}$ ou de prendre l'inverse de $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$. La première option est la plus simple car $\mathbf{B}$ est triangulaire, mais dans les deux cas, on obtient

$$\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}} = [[\mathbf{c}_1]_{\mathcal{B}}\ [\mathbf{c}_2]_{\mathcal{B}}\ [\mathbf{c}_3]_{\mathcal{B}}] = \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}^{-1} = \begin{bmatrix}1 & 0 & 0\\ \mathrm{i} & -2 & 0\\ 2 & 3\mathrm{i} & 3\mathrm{i}/2\end{bmatrix}\ .$$

Les résultats qui suivent sont des extensions pratiques du théorème 5.2.2. La proposition 5.2.10 caractérise l'inverse des matrices de changement de base, et la proposition 5.2.11 permet de considérer les changements de base avec trois bases.

Proposition 5.2.10 (*inverse d'une matrice de changement de base*)
Si $\mathcal{B}$ et $\mathcal{C}$ dont deux bases d'un même espace vectoriel V de dimension p. Alors la matrice $p \times p$ de changement de base de $\mathcal{B}$ à $\mathcal{C}$ est inversible et son inverse est la matrice de changement de base de $\mathcal{C}$ à $\mathcal{B}$:

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}^{-1} = \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}} \ .$$

Preuve. Soit $\mathbf{x} \in V$. Tel qu'établi dans le théorème 5.2.2, on a $[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$ ainsi que $[\mathbf{x}]_{\mathcal{B}} = \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}[\mathbf{x}]_{\mathcal{C}}$. On a donc les deux égalités suivantes :

$$[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}[\mathbf{x}]_{\mathcal{C}} \tag{5.4}$$

$$[\mathbf{x}]_{\mathcal{B}} = \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} \ . \tag{5.5}$$

Ces deux égalités (5.4) et (5.5) sont vraies pour tout $\mathbf{x} \in V$ et en particulier pour $\mathbf{b}_i$ et $\mathbf{c}_i$ qui sont tels que $[\mathbf{c}_i]_{\mathcal{C}} = [\mathbf{b}_i]_{\mathcal{B}} = \mathbf{e}_i$ pour tout $i \in [\![1;p]\!]$. Ainsi, $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}} = \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{I}_p$, ce qui, selon le cas 1 du théorème 1.4.1, montre que $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$ est inversible et que son inverse est $\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}$. □

Proposition 5.2.11 (*composition de matrices de changement de base*)
Si $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$ sont trois bases d'un espace vectoriel V, alors $\mathbf{P}_{\mathcal{D}\leftarrow\mathcal{B}} = \mathbf{P}_{\mathcal{D}\leftarrow\mathcal{C}}\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$.

Preuve. Soit $\mathbf{x}$ un vecteur quelconque de V, Il résulte du théorème 5.2.2 qu'on peut obtenir $[\mathbf{x}]_{\mathcal{D}}$ à l'aide de $[\mathbf{x}]_{\mathcal{B}}$: $[\mathbf{x}]_{\mathcal{D}} = \mathbf{P}_{\mathcal{D}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$. Aussi, $[\mathbf{x}]_{\mathcal{D}} = \mathbf{P}_{\mathcal{D}\leftarrow\mathcal{C}}[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{D}\leftarrow\mathcal{C}}\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$. On a donc $\mathbf{P}_{\mathcal{D}\leftarrow\mathcal{B}} = \mathbf{P}_{\mathcal{D}\leftarrow\mathcal{C}}\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}$. □

Un cas particulier intéressant de la proposition 5.2.11 est d'utiliser la base euclidienne de la sorte :

$$\begin{aligned}\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} &= \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{E}}\mathbf{P}_{\mathcal{E}\leftarrow\mathcal{B}} \\ &= \mathbf{P}_{\mathcal{E}\leftarrow\mathcal{C}}^{-1}\mathbf{P}_{\mathcal{E}\leftarrow\mathcal{B}}\end{aligned}$$

et si $p = n$, on retombe sur

$$\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \mathbf{C}^{-1}\mathbf{B} \ .$$

Résumé des formules de changements de base

Pour conclure cette section, on donne ci-dessous un résumé des différentes formules de changement de base, pour le cas $p = n$.

▷ $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$ $\mathcal{E} = (\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n)$

▷ $\mathbf{B} = \mathbf{P}_{\mathcal{E}\leftarrow\mathcal{B}} = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n]$

$\mathbf{C} = \mathbf{P}_{\mathcal{E}\leftarrow\mathcal{C}} = [\mathbf{c}_1\ \mathbf{c}_2\ \ldots\ \mathbf{c}_n]$

▷ $\mathbf{x} = \mathbf{B}[\mathbf{x}]_{\mathcal{B}} = \mathbf{C}[\mathbf{x}]_{\mathcal{C}}$

▷ $[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$

$[\mathbf{x}]_{\mathcal{B}} = \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}[\mathbf{x}]_{\mathcal{C}} = \mathbf{P}^{-1}_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{C}}$

▷ $\mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}} = \big[[\mathbf{b}_1]_{\mathcal{C}}\ [\mathbf{b}_2]_{\mathcal{C}}\ \ldots\ [\mathbf{b}_n]_{\mathcal{C}}\big] = \mathbf{C}^{-1}\mathbf{B}$

$\mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}} = \big[[\mathbf{c}_1]_{\mathcal{B}}\ [\mathbf{c}_2]_{\mathcal{B}}\ \ldots\ [\mathbf{c}_n]_{\mathcal{B}}\big] = \mathbf{B}^{-1}\mathbf{C} = \mathbf{P}^{-1}_{\mathcal{C}\leftarrow\mathcal{B}}$

▷ $\Big[\mathbf{C}\ \mathbf{B}\Big] = \Big[\mathbf{c}_1\ \mathbf{c}_2\ \ldots\ \mathbf{c}_n\ \mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n\Big] \sim \Big[\mathbf{I}\ \mathbf{P}_{\mathcal{C}\leftarrow\mathcal{B}}\Big]$

$\Big[\mathbf{B}\ \mathbf{C}\Big] = \Big[\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n\ \mathbf{c}_1\ \mathbf{c}_2\ \ldots\ \mathbf{c}_n\Big] \sim \Big[\mathbf{I}\ \mathbf{P}_{\mathcal{B}\leftarrow\mathcal{C}}\Big]$

Annexe

A

Solutions des exercices

Série d'exercices 0: Notations

Solution 0.1: (énoncé)
Les expressions valides sont :

1 $\mathbf{x} \in A$;
2 $\mathbf{x} \in F$;
5 $\{\mathbf{x}, \mathbf{y}\} = A$;
7 $\{\mathbf{x}, \mathbf{y}\} \subseteq A$;
8 $\mathbf{x} \in \mathbb{R}^2$;
13 $\mathbf{x} = (1, 2)$;
18 $\mathbf{x} = [1\ 2]^\top$;
19 $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

Solution 0.2: (énoncé)

1. ▷ $(5, 2, 5)$ est une famille et cette écriture est déjà correcte. Attention, ce n'est pas la même famille que $(2, 5, 5)$ ni $(2, 5)$ car avec les familles, l'ordre des éléments et les répétitions sont pris en compte ;
   ▷ $\{\mathbf{e}_2, \mathbf{e}_1, \mathbf{e}_3\}$ est un ensemble. Cette écriture est correcte mais il est plus intuitif d'écrire $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ même si l'ordre des vecteurs n'est pas pris en compte ;
   ▷ $\{p, p+1, \dots, n\} = [\![p; n]\!]$ est un ensemble.
2. $\{-3, -2, \dots, 4\} = [\![-3; 4]\!]$.
3. $\mathcal{E}$ est une famille.

Solution 0.3: (énoncé)

1. A est un ensemble et B une famille.
2. On a
   (a) $A \cup B = \{1, 3, 4, 5\}$;
   (b) $A \cap B = \{1, 3\}$;
   (c) $B \setminus A = \{4\}$;
   (d) $A \setminus B = \{5\}$.

   Ces opérations mélangent des ensembles et des familles, et par convention, on considère que les résultats sont des ensembles.

Solution 0.4: (énoncé)
Comme l'opérateur $\mathrm{Vect}(\cdot)$ prend une famille en entrée et retourne un ensemble, seules les formulations suivantes sont valides :

2 $\mathrm{Vect}(F) = A$ (famille en entrée, ensemble en sortie) ;

6 $\mathrm{Vect}(\mathbf{x}, \mathbf{y}) \subset A$ (un ensemble sous-ensemble d'un autre) ;

7 $\mathrm{Vect}(\mathbf{x}, \mathbf{y}) = A$ (deux ensembles égaux) ;

9 $\mathrm{Vect}((\mathbf{x}, \mathbf{y})) = A$ (idem).

Solution 0.5: (énoncé)

1. $2 \in B$;
2. $\{2\} \subset A$;
3. Aucun lien ;
4. $\{1, 6\} \subset A$;
5. Aucun lien ;
6. $B \subset A$.

Solution 0.6: (énoncé)

1. $2 \in G$;
2. $(2) \subset F$;
3. Aucun lien ;
4. $G \subset F$;
5. Aucun lien (l'ordre diffère).
6. $F \subset (1, 2, 2, 4, 6)$.

Solution 0.7: (énoncé)

1. $\mathbf{u}_3 \in G$;
2. $-1 \in \mathbf{u}_2$;
3. $(\mathbf{u}_2) \subset F$;
4. Aucun lien ;
5. $G \subset F$;
6. Aucun lien (l'ordre diffère) ;
7. $F \subset (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_2, \mathbf{u}_3)$.

Série d'exercices 1: Matrices et vecteurs

Solution 1.1: (énoncé)

1. **Faux.** Le produit $\mathbf{AB}$ n'est pas défini ;
2. **Vrai.** $\mathrm{tr}(\mathbf{B}^{-1}\mathbf{AB}) = \mathrm{tr}(\mathbf{ABB}^{-1}) = \mathrm{tr}(\mathbf{AI}) = \mathrm{tr}(\mathbf{A})$;

3. **Faux.** Un contre-exemple simple est la matrice $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ qui est triangulaire mais pas inversible ;
4. **Vrai.** Avec $\mathbf{C} = (\mathbf{AB})^{-1}$, $\mathbf{BC}$ est l'inverse de $\mathbf{A}$ et $\mathbf{CA}$ celle de $\mathbf{B}$. On peut aussi le montrer avec le déterminant ;
5. **Faux.** Un contre-exemple est $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$ et $\mathbf{b} = (1, 1, 1)$; Le système $\mathbf{Ax} = \mathbf{b}$ possède l'unique solution $\mathbf{x} = (1, 1)$ et pourtant $\mathbf{A}$ n'est pas inversible (elle n'est pas carrée) ;
6. **Faux.** Pour construire un contre-exemple simple, on peut voir qu'un seul nombre définit un bloc de taille 1×1. Si $a \neq 0$, la matrice $[a]$ est inversible : son inverse est donnée par $[a]^{-1} = [a^{-1}] = [1/a]$. On peut ainsi considérer la matrice $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ dont les quatre blocs $[1]$ sont inversibles, mais qui n'est pas inversible. Remarquer que l'énoncé est également faux en supposant la matrice entière inversible : considérer $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ qui est inversible mais son inverse n'est pas obtenue en inversant les blocs ;
7. **Vrai.** Les opérations élémentaires sur les lignes d'une matrice $\mathbf{A} \in \mathbb{R}^{n \times n}$ affectent son déterminant de manière bien définie :
   - ▷ Multiplier une ligne par un scalaire non nul $\alpha \in \mathbb{R}^*$ multiplie le déterminant par α ;
   - ▷ Permuter deux lignes multiplie le déterminant par -1 ;
   - ▷ Ajouter un multiple d'une ligne à une autre laisse le déterminant inchangé.

   Appliquer successivement les opérations suivantes à la matrice $\mathbf{A}$:

   (a) La deuxième ligne est multipliée par α : le déterminant devient $\alpha \det(\mathbf{A})$;
   (b) La première et la deuxième ligne sont permutées : le déterminant devient $-\alpha \det(\mathbf{A})$;
   (c) Trois fois la troisième ligne sont ajoutées à la première : le déterminant reste $-\alpha \det(\mathbf{A})$.

   La matrice $\mathbf{A}'$ issue de ces trois transformations vérifie donc $\det(\mathbf{A}') = -\alpha \det(\mathbf{A})$. L'égalité mentionnée dans l'énoncé est correcte, ce qui rend l'affirmation vraie.

Solution 1.2: (énoncé)

1. $\mathbf{Ab} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \\ 1 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} + 1 \begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \\ 6 \end{bmatrix} = \mathbf{c}$;

2. D'après ce qu'on vient d'écrire, on peut déduire que $\mathbf{c}$ est égal à la somme des colonnes de $\mathbf{A}$: $\mathbf{c} = \begin{bmatrix}1\\2\\1\end{bmatrix} + \begin{bmatrix}2\\3\\4\end{bmatrix} + \begin{bmatrix}3\\3\\1\end{bmatrix}$;
3. D'après l'énoncé, $\mathbf{b}$ est une combinaison linéaire des colonnes de $\mathbf{A}$, qui s'exprime comme $4\mathbf{b} = \mathbf{A}(-1, 1, 1)$. On a donc $\mathbf{b} = \frac{1}{4}\mathbf{A}(-1, 1, 1)$ et $\mathbf{x} = \frac{1}{4}(-1, 1, 1)$.

Solution 1.3: (énoncé)

$$\mathbf{Ax} = x_1 \begin{bmatrix}-1\\0\\1\\7\end{bmatrix} + x_2 \begin{bmatrix}1\\1\\-2\\3\end{bmatrix} + x_3 \begin{bmatrix}2\\-3\\1\\-1\end{bmatrix}$$

Solution 1.4: (énoncé)

$$\mathbf{A} = \begin{bmatrix}2 & 1 & -1 & 6\\1 & 1 & 0 & -4\end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix}$$

Solution 1.5: (énoncé)

1. $\|\mathbf{a}\| = 2$, $\|\mathbf{b}\| = 3\sqrt{2}$ et $\|\mathbf{c}\| = 2$
2. $\langle \mathbf{a}, \mathbf{b}\rangle = 2{*}3{+}0{*}3 = 2{*}3\sqrt{2}{*}\cos(\frac{\pi}{4}) = 6$ $\langle \mathbf{a}, \mathbf{c}\rangle = -\sqrt{3}{*}2{+}1{*}0 = 2{*}2{*}\cos(\frac{5\pi}{6}) = -2\sqrt{3}$
3. $\langle \mathbf{b}, \mathbf{c}\rangle = \frac{3\sqrt{3}+3}{2}\langle \mathbf{a}, \mathbf{c}\rangle + 3\langle \mathbf{c}, \mathbf{c}\rangle = -(3\sqrt{3}+3)\sqrt{3} + 3 * 2 = -3 + 3\sqrt{3}$.

Solution 1.6: (énoncé)

1. $\|\mathbf{x}+\mathbf{y}\|^2 = \langle \mathbf{x}+\mathbf{y}, \mathbf{x}+\mathbf{y}\rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 + 2\langle \mathbf{x}, \mathbf{y}\rangle$ $\|\mathbf{x}-\mathbf{y}\|^2 = \langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle$
2. On reprend la première expression mais ici, on sait $\langle \mathbf{x}, \mathbf{y}\rangle = 0$.

Solution 1.7: (énoncé)

1. $\mathbf{x}+\mathbf{y} = (1+\alpha)\mathbf{x}$, donc $\|\mathbf{x}+\mathbf{y}\| = |1+\alpha|\,\|\mathbf{x}\| = (1+\alpha)\|\mathbf{x}\|$ (puisque $1+\alpha > 0$). De plus, $\|\mathbf{y}\| = |\alpha|\,\|\mathbf{x}\| = \alpha\|\mathbf{x}\|$. Ainsi, $\|\mathbf{x}+\mathbf{y}\| = (1{+}\alpha)\|\mathbf{x}\| = \|\mathbf{x}\|{+}\alpha\|\mathbf{x}\| = \|\mathbf{x}\|{+}\|\mathbf{y}\|$.
2. Si $\mathbf{y} = \alpha\mathbf{x}$ avec $-1 < \alpha < 0$. Ici, $\|\mathbf{y}\| = |\alpha|\,\|\mathbf{x}\| = -\alpha\|\mathbf{x}\|$ et $1+\alpha > 0$. $\|\mathbf{x}+\mathbf{y}\| = (1+\alpha)\|\mathbf{x}\| = \|\mathbf{x}\| + \alpha\|\mathbf{x}\| = \|\mathbf{x}\| - \|\mathbf{y}\|$.

3. On pose $\alpha = \dfrac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|^2}$. Alors, $\langle \mathbf{y} - \alpha\mathbf{x}, \mathbf{y} - \alpha\mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{y} - \alpha\mathbf{x} \rangle - \alpha \langle \mathbf{x}, \mathbf{y} - \alpha\mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{y} \rangle - \alpha \langle \mathbf{y}, \mathbf{x} \rangle - \alpha \langle \mathbf{x}, \mathbf{y} \rangle + \alpha^2 \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{y}\|^2 - \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\|\mathbf{x}\|^2} = 0$ par définition de α et hypothèse de l'énoncé. De fait, le seul vecteur de norme 0 est le vecteur nul donc : $\mathbf{y} = \alpha\mathbf{x}$.
4. On développe l'équation au carré des deux côté pour retrouver l'égalité.
5. On a égalité si et seulement si il existe $\alpha \geq 0$ tel que $\mathbf{y} = \alpha\mathbf{x}$. (Les cas triviaux $\mathbf{x} = 0$ ou $\mathbf{y} = 0$ sont inclus.)

Solution 1.8: (énoncé)

On commence par remarquer $\mathbf{u}$ et $\mathbf{v}$ ne sont pas colinéaires, donc $\mathrm{Vect}(\mathbf{u}, \mathbf{v})$ est bien un plan de $\mathbb{R}^3$. Puis, si $\mathbf{x} = (x, y, z) \in \mathrm{Vect}(\mathbf{u}, \mathbf{v})$, alors

$$\mathbf{x} = \alpha\mathbf{u} + \beta\mathbf{v} = \alpha \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha \\ 2\alpha + \beta \\ 3\alpha + \beta \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

avec α et β deux réels.

- ▷ Méthode 1 : $z = 3\alpha + \beta = \alpha + 2\alpha + \beta = x + y$ est l'équation du plan.
- ▷ Méthode 2 : $\mathbf{u} \wedge \mathbf{v} = (-1, -1, 1)$ est orthogonal au plan, donc son équation est donnée par $-x - y + z = 0$, soit $z = x + y$.

Solution 1.9: (énoncé)

1. $\mathbf{Ax}$:
   $\mathbf{A} \in \mathbb{R}^{3\times 2}$, $\mathbf{x} \in \mathbb{R}^2$, donc le produit est bien défini (multiplication d'une matrice 3×2 par un vecteur colonne de taille 2).
   Résultat : vecteur colonne dans $\mathbb{R}^3$.
2. $\mathbf{BA}$:
   $\mathbf{B} \in \mathbb{R}^{2\times 4}$ et $\mathbf{A} \in \mathbb{R}^{3\times 2}$.
   Ici, les dimensions internes ne coïncident pas : $4 \neq 3$.
   $\mathbf{BA}$ n'est pas définie, car le nombre de colonnes de $\mathbf{B}$ ne correspond pas au nombre de lignes de $\mathbf{A}$.
3. $\mathbf{A} + \mathbf{y}$:
   $\mathbf{A} \in \mathbb{R}^{3\times 2}$, $\mathbf{y} \in \mathbb{R}^3$.
   Irréalisable : $\mathbf{A}$ est une matrice 3×2, $\mathbf{y}$ un vecteur colonne de taille 3.
   Il n'est pas possible de les additionner.
   $\mathbf{A} + \mathbf{y}$ n'est pas définie.
4. $\mathbf{A}^\top\mathbf{y}$:
   $\mathbf{A} \in \mathbb{R}^{3\times 2} \Rightarrow \mathbf{A}^\top \in \mathbb{R}^{2\times 3}$, $\mathbf{y} \in \mathbb{R}^3$.

Le produit $\mathbf{A}^\top\mathbf{y}$ est bien défini : 2×3 par 3×1.
Résultat : vecteur colonne de taille 2, donc dans $\mathbb{R}^2$.

5. $\mathbf{x}^\top\mathbf{A}$:
   $\mathbf{x}\in\mathbb{R}^2 \Rightarrow \mathbf{x}^\top \in \mathbb{R}^{1\times 2}$, $\mathbf{A}\in\mathbb{R}^{3\times 2}$.
   Ici, les tailles ne sont pas compatibles pour un produit direct (on a 1×2 et 3×2), il faudrait que le nombre de colonnes de $\mathbf{x}^\top$ (2) corresponde au nombre de lignes de $\mathbf{A}$ (3).
   Ainsi $\mathbf{x}^\top\mathbf{A}$ **n'est pas définie**.

Solution 1.10: (énoncé)

1. $\mathbf{AB} = \begin{bmatrix} 25 & 9 \\ 4 & 12 \end{bmatrix}$
2. $\mathbf{BA} = \begin{bmatrix} 10 & -6 & -2 \\ 9 & 3 & 8 \\ 6 & 18 & 24 \end{bmatrix}$

Solution 1.11: (énoncé)

1. La matrice $\mathbf{A}$ est de taille 3×3 et $\mathbf{B}^\top$ est de taille 2×3. Ces matrices n'ont pas les mêmes tailles. L'opération $\mathbf{A}+\mathbf{B}^\top$ n'est donc pas définie.
2. Le produit $\mathbf{CB}$ est défini puisque $\mathbf{C}$ est 2×3 et $\mathbf{B}$ est 3×2. Le résultat est une matrice 2×2. Ainsi, $(\mathbf{CB})^\top$ est de taille 2×2. La matrice $\mathbf{A}$ est de taille 3×3, donc l'addition $(\mathbf{CB})^\top+\mathbf{A}$ n'est pas définie car les tailles ne sont pas compatibles.
3. Le produit $\mathbf{Au}$ est défini : $\mathbf{A}$ est 3×3 et $\mathbf{u}$ est 3×1. Le résultat est un vecteur colonne 3×1. Le produit $\mathbf{Bv}$ est aussi défini : $\mathbf{B}$ est 3×2 et $\mathbf{v}$ est 2×1, donc le produit donne un vecteur colonne 3×1. L'addition $\mathbf{Au}+\mathbf{Bv}$ est donc bien définie et donne un vecteur 3×1 :

$$\mathbf{Au} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ -1 & 1 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1\cdot 2 + 2\cdot(-1) + (-1)\cdot 3 \\ 0\cdot 2 + 3\cdot(-1) + 2\cdot 3 \\ -1\cdot 2 + 1\cdot(-1) + 4\cdot 3 \end{bmatrix} = \begin{bmatrix} 2-2-3 \\ -3+6 \\ -2-1+12 \end{bmatrix} = \begin{bmatrix} -3 \\ 3 \\ 9 \end{bmatrix}$$

$$\mathbf{Bv} = \begin{bmatrix} 2 & 0 \\ -1 & 3 \\ 1 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2\cdot 1 + 0\cdot 4 \\ -1\cdot 1 + 3\cdot 4 \\ 1\cdot 1 + (-2)\cdot 4 \end{bmatrix} = \begin{bmatrix} 2 \\ -1+12 \\ 1-8 \end{bmatrix} = \begin{bmatrix} 2 \\ 11 \\ -7 \end{bmatrix}$$

$$\mathbf{Au}+\mathbf{Bv} = \begin{bmatrix} -3 \\ 3 \\ 9 \end{bmatrix} + \begin{bmatrix} 2 \\ 11 \\ -7 \end{bmatrix} = \begin{bmatrix} -1 \\ 14 \\ 2 \end{bmatrix}$$

4. Le produit $\mathbf{Au}$ donne un vecteur 3×1 comme vu ci-dessus. La matrice $\mathbf{C}$ est de taille 2×3, donc $\mathbf{C}(\mathbf{Au})$ est défini et donne un vecteur 2×1. Ensuite, $2\mathbf{v}$ est également un vecteur 2×1. L'opération $\mathbf{C}(\mathbf{Au}) - 2\mathbf{v}$ est donc bien définie :

$$\mathbf{C}(\mathbf{Au}) = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ 3 \\ 9 \end{bmatrix} = \begin{bmatrix} 1\cdot(-3) + (-1)\cdot 3 + 2\cdot 9 \\ 3\cdot(-3) + 0\cdot 3 + 1\cdot 9 \end{bmatrix} = \begin{bmatrix} -3-3+18 \\ -9+9 \end{bmatrix} = \begin{bmatrix} 12 \\ 0 \end{bmatrix}$$

$$2\mathbf{v} = 2\cdot\begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 8 \end{bmatrix} \quad \Rightarrow \quad \mathbf{C}(\mathbf{Au}) - 2\mathbf{v} = \begin{bmatrix} 12 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 8 \end{bmatrix} = \begin{bmatrix} 10 \\ -8 \end{bmatrix}$$

5. La somme $\mathbf{A} + \mathbf{A}^\top$ est définie car $\mathbf{A}$ est carrée (3×3). Le résultat est une matrice symétrique 3×3. Le produit $(\mathbf{A} + \mathbf{A}^\top)\mathbf{u}$ est donc bien défini et donne un vecteur 3×1 :

$$\mathbf{A} + \mathbf{A}^\top = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 2 \\ -1 & 1 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 1 \\ -1 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 6 & 3 \\ -2 & 3 & 8 \end{bmatrix}$$

$$(\mathbf{A}+\mathbf{A}^\top)\mathbf{u} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 6 & 3 \\ -2 & 3 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2\cdot 2 + 2\cdot(-1) + (-2)\cdot 3 \\ 2\cdot 2 + 6\cdot(-1) + 3\cdot 3 \\ -2\cdot 2 + 3\cdot(-1) + 8\cdot 3 \end{bmatrix} = \begin{bmatrix} 4-2-6 \\ 4-6+9 \\ -4-3+24 \end{bmatrix} = \begin{bmatrix} -4 \\ 7 \\ 17 \end{bmatrix}$$

Solution 1.12: (énoncé)

1. $\mathbf{AB} = \begin{bmatrix} -\cos(a)\cos(b) - \sin(a)\sin(b) & -\cos(a)\sin(b) - \sin(a)\cos(b) \\ -\sin(a)\cos(b) + \cos(a)\sin(b) & -\sin(a)\sin(b) + \cos(a)\cos(b) \end{bmatrix} = \begin{bmatrix} -\cos(a-b) & -\sin(a+b) \\ \sin(b-a) & \cos(a+b) \end{bmatrix}$
2. $\mathbf{A}^2 = \begin{bmatrix} \cos(2a) & -\sin(2a) \\ \sin(2a) & \cos(2a) \end{bmatrix}$
3. Par récurrence : $\mathbf{A}^n = \begin{bmatrix} \cos(na) & -\sin(na) \\ \sin(na) & \cos(na) \end{bmatrix}$

Solution 1.13: (énoncé)

On cherche une matrice de cette forme pour que le produit existe : $\mathbf{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

On veut : $\begin{bmatrix} a+c & b+d \\ a+c & b+d \end{bmatrix} = \begin{bmatrix} a+b & a+b \\ c+d & c+d \end{bmatrix}$. On veut donc : $a+b = a+c$, $a+b = b+d$, $c+d = a+c$, $b+d = c+d$. Ce sont donc des matrices de la forme : $\begin{bmatrix} a & b \\ b & a \end{bmatrix}$.

Solution 1.14: (énoncé)
$a_{i,i+1} = 1$ pour $i = 1, \dots, n-1$ et $a_{i,j} = 0$ ailleurs.
Pour $k = 2$, Donc pour $\mathbf{A}^2$ si $n > 2$, les seuls coefficients $a_{i,j}$ non nuls sont ceux tels que $i - j = 2$, et si $n = 2$, $\mathbf{A}^2 = \mathbf{O}$.
Pour $k = 3$, Donc pour $\mathbf{A}^3$ si $n > 3$, les seuls coefficients $a_{i,j}$ non nuls sont ceux tels que $i - j = 3$, et si $n = 3$, $\mathbf{A}^3 = \mathbf{O}$.
etc.

Solution 1.15: (énoncé)

1. $\mathbf{F} \in \mathbb{R}^{r\times(n-r)}$, $\mathbf{C} \in \mathbb{R}^{(m-r)\times r}$, $\mathbf{D} \in \mathbb{R}^{(m-r)\times(n-r)}$, et $\mathbf{B} \in \mathbb{R}^{(n-r)\times(n-r)}$.
2. $\mathbf{RN} = \begin{bmatrix} -\mathbf{AF} + \mathbf{FB} \\ -\mathbf{CF} + \mathbf{DB} \end{bmatrix} \in \mathbb{R}^{m\times(n-r)}$.
3. On donne d'abord $\mathbf{A} = \mathbf{I}_r$, $\mathbf{B} = \mathbf{I}_{n-r}$, $\mathbf{C} = \mathbf{O}_{m-r,r}$ et $\mathbf{D} = \mathbf{O}_{m-r,n-r}$, puis

$$\mathbf{RN} = \begin{bmatrix} -\mathbf{I}_r\mathbf{F} + \mathbf{F}\mathbf{I}_{n-r} \\ -\mathbf{O}_{m-r,r}\mathbf{F} + \mathbf{O}_{m-r,n-r}\mathbf{I}_{n-r} \end{bmatrix} = \begin{bmatrix} -\mathbf{O}_{r,n-r} \\ \mathbf{O}_{m-r,n-r} \end{bmatrix} = \mathbf{O}_{m,n-r} \ .$$

Solution 1.16: (énoncé)

1. $\det(\mathbf{A}) = -2\cdot 3 - 6\cdot 3 = -24$
2. $\det(\mathbf{B}) = -1\cdot(4\cdot 3 - 5\cdot 1) = -7$ en additionnant les lignes 1 et 3 et en développant par rapport à la première colonne.
3. $\det(\mathbf{C}) = 0$. La deuxième et la troisième lignes de $\mathbf{C}$ sont égales donc la matrice n'est pas inversible donc le déterminant est nul.
4. $\det(\mathbf{C}) = 1\det(\mathbf{D}) = (-1)^{1+2}4(1\cdot 1 - 1\cdot 1) + (-1)^{2+2}2(1\cdot 1 - 2\cdot 1) = -2$ en développant par rapport à la troisième colonne avec $\mathbf{D} = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, en développant ensuite par la deuxième colonne de $\mathbf{D}$.

Solution 1.17: (énoncé)

1. C'est la décomposition $\mathbf{LU}$ de $\mathbf{A}$ vue au chapitre chapitre 6
2. Comme $\mathbf{L}$ et $\mathbf{U}$ sont triangulaires leur déterminants sont simples à calculer $\det(\mathbf{A}) = \det(\mathbf{L})\det(\mathbf{U}) = 1\cdot(-1)\cdot 3\cdot 5\cdot 2 = -30$
3. La première colonne de $\mathbf{A}$ est multipliée par 2 et la troisième par -1 pour obtenir $\mathbf{B}$ donc $\det(\mathbf{B}) = 2\cdot(-1)\cdot\det(\mathbf{A}) = 60$

Solution 1.18: (énoncé)

1. On applique le développement par les cofacteurs sur la première ligne de $\mathbf{A} = \begin{bmatrix} -1 & 3 & 1 \\ 2 & 4 & 1 \\ 2 & 7 & 1 \end{bmatrix}$:

$$\begin{aligned} \det(\mathbf{A}) &= -1 \cdot (4 \cdot 1 - 1 \cdot 7) - 3 \cdot (2 \cdot 1 - 1 \cdot 2) + 1 \cdot (2 \cdot 7 - 4 \cdot 2) \\ &= -1(-3) - 3(0) + 1(14 - 8) = 3 + 6 = 9 \end{aligned}$$

2. A est inversible car son déterminant est non nul, de plus $A^\top$ l'est aussi car son déterminant est égal à celui de A, enfin $\mathbf{A}^\top\mathbf{A}$ est inversible comme produit de matrices inversibles. On a :

$$\det((\mathbf{A}^\top\mathbf{A})^{-1}) = \frac{1}{\det(\mathbf{A}^\top\mathbf{A})} = \frac{1}{\det(\mathbf{A})^2} = \frac{1}{81}$$

Solution 1.19: (énoncé)

1. On développe $\mathbf{A}_{n+2}$ d'abord par la première ligne, cela donne le facteur $4\det(\mathbf{A}_{n+1})$ puis pour obtenir le deuxième il faut développer la matrice résultante par la première colonne.
2. Par récurrence : $\det(\mathbf{A}_n) = (n + 1)2^n$

Solution 1.20: (énoncé)
Attention ! Il est impossible d'appliquer la formule du déterminant d'une matrice 2×2 sur des matrices construites par blocs. Généralement, on aura :

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \neq \det(\mathbf{A})\det(\mathbf{D}) - \det(\mathbf{C})\det(\mathbf{B}).$$

1. L'objectif est de se ramener à une forme sur laquelle on peut s'imaginer appliquer un développement selon une ligne/colonne. Observer d'abord que

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix}.$$

Traiter d'abord la première matrice : un développement récursif selon la dernière ligne donne

$$\begin{aligned}
\begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_m \end{bmatrix} &= 1 \cdot \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{m-1} \end{bmatrix} \\
&= 1 \cdot 1 \cdot \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{m-2} \end{bmatrix} \\
&= \ldots \\
&= \underbrace{1 \cdot 1 \cdot \ldots \cdot 1}_{m \text{ fois}} \cdot \det(\mathbf{A}) \\
&= \det(\mathbf{A}).
\end{aligned}$$

La même logique (selon la première ligne cette fois) donne :

$$\begin{aligned}
\begin{bmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix} &= 1 \cdot \begin{bmatrix} \mathbf{I}_{n-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix} \\
&= 1 \cdot 1 \cdot \begin{bmatrix} \mathbf{I}_{n-2} & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix} \\
&= \ldots \\
&= \underbrace{1 \cdot 1 \cdot \ldots \cdot 1}_{n \text{ fois}} \cdot \det(\mathbf{C}) \\
&= \det(\mathbf{C}).
\end{aligned}$$

Finalement,

$$\det(\mathbf{M}_1) = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_m \end{bmatrix} \cdot \begin{bmatrix} \mathbf{I}_n & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix} = \det(\mathbf{A}) \cdot \det(\mathbf{C}).$$

2. En observant que

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{I}_n & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{O} & \mathbf{I}_m \end{bmatrix},$$

on peut écrire :

$$\det(\mathbf{M}) = \begin{bmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{O} & \mathbf{C} \end{bmatrix} \cdot \underbrace{\begin{bmatrix} \mathbf{I}_n & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I}_m \end{bmatrix}}_{=1},$$

et ceci vaut bien $\det(\mathbf{A}) \cdot \det(\mathbf{C})$ par la question précédente.

Solution 1.21: (énoncé)

1. La matrice $\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$ n'est pas carrée, donc elle n'est pas inversible.

2.
$$\mathbf{A}\mathbf{A}^\top = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 10 \end{bmatrix}$$

Cette matrice est inversible car son déterminant est non nul :
$$\det(\mathbf{A}\mathbf{A}^\top) = 5 \cdot 10 - 3 \cdot 3 = 50 - 9 = 41 \neq 0$$

3. $\mathbf{A}^\top\mathbf{A} = \begin{bmatrix} 10 & 2 & -3 \\ 2 & 4 & 0 \\ -3 & 0 & 1 \end{bmatrix}$ n'est pas inversible car son déterminant est nul.

4. On a
$$(\mathbf{A}^\top\mathbf{A})^\top = \mathbf{A}^\top(\mathbf{A})^\top = \mathbf{A}^\top\mathbf{A}$$

Solution 1.22: (énoncé)

Note : quand les notes seront écrites sur la caractérisation des matrices inversibles, il serait bon de revoir cette correction pour mentionner les numéros des propriétés donnés dans les notes de cours. Il faut d'abord noter que $\mathbf{I}_2$ est carrée, première étape obligatoire pour qu'elle puisse être inversible.

1. $\mathbf{I}_2$ peut être réduite en elle-même par une élimination de Gauss, qui n'effectue aucune étape. Elle est donc inversible. : ce n'est pas dans la caractérisation
2. Chacun des coefficients non nuls de $\mathbf{I}_2$ est un pivot. Il y en a évidemment 2, soit un pivot par colonne. $\mathbf{I}_2$ est donc inversible. (point 7)
3. En cherchant à résoudre $\mathbf{I}_2\mathbf{x} = \mathbf{0}$, on peut :
   (a) Remarquer que $\mathbf{I}_2\mathbf{x} = \mathbf{x}$, d'où $\mathbf{I}_2\mathbf{x} = 0 \Leftrightarrow \mathbf{x} = 0$ et donc $\text{Ker}(\mathbf{I}_2) = \{\mathbf{0}\}$
   (b) Observer que la matrice augmentée du système est déjà sous forme échelonnée :
   $$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$
   et donc, la seule solution à $\mathbf{I}_2\mathbf{x} = \mathbf{0}$ est $\mathbf{x} = \mathbf{0}$, d'où $\text{Ker}(\mathbf{A}) = \{\mathbf{0}\}$.

   Donc $\mathbf{I}_2$ est inversible. (point 9)
4. Les colonnes de $\mathbf{I}_2$ sont les vecteurs $(1, 0)$ et $(0, 1)$, qui sont évidemment linéairement indépendantes. (point 6)
5. Si $\mathbf{b} \in \mathbb{R}^2$, l'équation $\mathbf{I}_2\mathbf{x} = \mathbf{b}$ se réduit en $\mathbf{x} = \mathbf{b}$, elle admet donc la solution $\mathbf{x} = \mathbf{b}$. Puisque cette équation admet au moins une solution pour tout $\mathbf{b} \in \mathbb{R}^2$ (et que $\mathbf{I}_2$ est carrée), $\mathbf{I}_2$ est inversible. (point 12)
6. Il est mentionné plus haut que les colonnes de $\mathbf{I}_2$ sont linéairement indépendantes. Elles constituent donc une famille de deux vecteurs linéairement indépendants dans $\mathbb{R}^2$. Elles engendrent donc $\mathbb{R}^2$, ainsi $\mathbf{I}_2$ est inversible. (point 10)

7. $\det(\mathbf{I}_2) = 1$ car c'est une matrice diagonale avec des 1 sur la diagonale, donc $\mathbf{I}_2$ est inversible. (point 4)
8. En posant $\mathbf{C} = \mathbf{I}_2$, on constate que $\mathbf{I}_2\mathbf{C} = \mathbf{C}\mathbf{I}_2 = \mathbf{I}_2$, donc $\mathbf{I}_2$ est inversible, et elle est sa propre inverse. (point 1)

Solution 1.23: (énoncé)
Note : quand les notes seront écrites sur la caractérisation des matrices inversibles, il serait bon de revoir cette correction pour mentionner les numéros des propriétés donnés dans les notes de cours. Il faut d'abord noter que $\mathbf{A}$ est carrée, ce qui oblige à justifier qu'elle n'est pas inversible d'une autre façon.

1. En appliquant l'élimination de Gauss sur $\mathbf{A}$, on obtient la forme échelonnée
$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$
$\mathbf{A}$ ne peut donc pas être réduite à $\mathbf{I}_2$, elle n'est pas inversible. idem : pas vu dans le caractérisation
2. La forme échelonnée ci-dessus montre bien que $\mathbf{A}$ n'admet qu'une seule position de pivot. Puisqu'elle a deux colonnes, elle n'est pas inversible. (point 7)
3. En cherchant à résoudre $\mathbf{Ax} = \mathbf{0}$, on obtient la matrice augmentée suivante sous forme échelonnée :
$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$
Le système est donc réalisable, et admet une infinité de solutions, dont par exemple $(1, -1)$. Puisque $\operatorname{Ker}(\mathbf{A}) \neq \{\mathbf{0}\}$, $\mathbf{A}$ n'est pas inversible. (point 9)
4. Les deux colonnes de $\mathbf{A}$ sont identiques, donc elles sont évidemment linéairement dépendantes. $\mathbf{A}$ n'est donc pas inversible. (point 6)
5. Avec $\mathbf{b} = (1, 0)$, le système $\mathbf{Ax} = \mathbf{b}$ se réduit en la matrice augmentée suivante :
$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}.$$
Ce système est donc irréalisable. Puisqu'il existe un $\mathbf{b} \in \mathbb{R}^2$ tel que $\mathbf{Ax} = \mathbf{b}$ n'admette aucune solution, $\mathbf{A}$ n'est pas inversible. (point 12)
6. Les colonnes de $\mathbf{A}$ sont linéairement dépendantes. Deux vecteurs linéairement dépendants ne peuvent pas engendrer $\mathbb{R}^2$, donc les colonnes de $\mathbf{A}$ n'engendrent pas $\mathbb{R}^2$. $\mathbf{A}$ n'est donc pas inversible. (point 10)
7. Supposons qu'il existe une matrice
$$\mathbf{C} = \begin{bmatrix} c_1 & c_3 \\ c_2 & c_4 \end{bmatrix}$$

telle que $\mathbf{AC} = \mathbf{I}_2$. En appliquant la règle ligne-colonne, cette égalité se réécrit sous la forme :

$$\begin{cases} c_1 + c_3 = 1 \\ c_1 + c_3 = 0 \\ c_2 + c_4 = 1 \\ c_2 + c_4 = 0 \end{cases},$$

qui est évidemment irréalisable. Puisqu'une telle matrice $\mathbf{C}$ n'existe pas, $\mathbf{A}$ n'est pas inversible. (point 2)

8. $\det(\mathbf{A}) = 1 \times 1 - 1 \times 1 = 0$ donc $\mathbf{A}$ n'est pas inversible. (point 4)

Solution 1.24: (énoncé)

1. On calcule le déterminant de $\mathbf{A}$ en développant selon la deuxième ligne :

$$\det(\mathbf{A}) = k \det\begin{bmatrix} 2 & k \\ 1 & 5 \end{bmatrix} - 2 \det\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix} = k(10-k) - 2(8+1) = -k^2 + 10k - 18.$$

Les racines de ce polynôme sont $k = 5 - \sqrt{7}$ et $k = 5 + \sqrt{7}$.
Conclusion : $\mathbf{A}$ n'est pas inversible si et seulement si $k \in \{5 - \sqrt{7}, 5 + \sqrt{7}\}$.

Solution 1.25: (énoncé)
On prend le terme de droite et on le multiplie à droite par $(\mathbf{F}^{-1} + \mathbf{uu}^\top)$. On obtient :

$$\begin{aligned}
\left(\mathbf{F} - \frac{\mathbf{Fuu}^\top\mathbf{F}}{1 + \mathbf{u}^\top\mathbf{Fu}}\right)(\mathbf{F}^{-1} + \mathbf{uu}^\top) &= \mathbf{FF}^{-1} + \mathbf{Fuu}^\top - \frac{\mathbf{Fuu}^\top\mathbf{F}}{1 + \mathbf{u}^\top\mathbf{Fu}}\mathbf{F}^{-1} - \frac{\mathbf{Fuu}^\top\mathbf{F}}{1 + \mathbf{u}^\top\mathbf{Fu}}\mathbf{uu}^\top, \\
&= \mathbf{I} + \mathbf{Fuu}^\top - \frac{\mathbf{Fuu}^\top}{1 + \mathbf{u}^\top\mathbf{Fu}} - \frac{\mathbf{Fuu}^\top\mathbf{F}}{1 + \mathbf{u}^\top\mathbf{Fu}}\mathbf{uu}^\top, \\
&= \mathbf{I} + \mathbf{Fuu}^\top - \mathbf{Fu}\frac{1}{1 + \mathbf{u}^\top\mathbf{Fu}}\mathbf{u}^\top - \mathbf{Fu}\frac{\mathbf{u}^\top\mathbf{Fu}}{1 + \mathbf{u}^\top\mathbf{Fu}}\mathbf{u}^\top, \\
&= \mathbf{I}.
\end{aligned}$$

On a trouvé une matrice telle que le produit des deux est la matrice identité donc elles sont bien inverses l'un de l'autre.

Solution 1.26: (énoncé)
Soit $\mathbf{A}$ une matrice carrée telle qu'il existe n tel que $\mathbf{A}^n = 0$ avec $\mathbf{A}$ non nulle.

1. $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

2. On a $\det(\mathbf{A}^n) = 0$, c'est-à-dire : $\det(\mathbf{A})^n = 0$ donc $\det(\mathbf{A}) = 0$. Alors $\mathbf{A}$ n'est pas inversible.
3. $\sum_{i=0}^{n-1}(\mathbf{A}^i - \mathbf{A}^{i+1}) = \sum_{i=0}^{n-1}(\mathbf{A}^i) - \sum_{i=0}^{n-1}(\mathbf{A}^{i+1}) = \sum_{i=0}^{n-1}(\mathbf{A}^i) - \sum_{i=1}^{n}(\mathbf{A}^i) = \mathbf{I} - \mathbf{A}^n$
4. $(\mathbf{I} - \mathbf{A})(\mathbf{I} + \mathbf{A} + \mathbf{A}^2 + ... + \mathbf{A}^{n-1}) = \mathbf{I}$.

Solution 1.27: (énoncé)

1. La matrice $\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D} & \mathbf{I} \end{bmatrix}$ est triangulaire inférieure. Tous les blocs diagonaux sont égaux à $\mathbf{I}$, donc $\det(\mathbf{A}) = \det(\mathbf{I})\det(\mathbf{I}) = 1$.
2. $$\mathbf{AB} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ -\mathbf{D} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D} - \mathbf{D} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}$$
3. Puisque $\mathbf{AB} = \mathbf{I}$, on a $\mathbf{A}^{-1} = \mathbf{B}$.
4. $$\mathbf{AC} = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{D} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{D} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{D} \\ \mathbf{D} & \mathbf{D}^2 + \mathbf{I} \end{bmatrix}$$

Solution 1.28: (énoncé)

1. Son déterminant est égal à 1 donc elle est inversible.
2. Montrons que $(\mathbf{A} - \mathbf{I}_3)(\mathbf{A} + \mathbf{I}_3) = \mathbf{0}_3$.
$$(\mathbf{A} - \mathbf{I}_3)(\mathbf{A} + \mathbf{I}_3) = \begin{bmatrix} 0 & 12 & -6 \\ 0 & 0 & 0 \\ 0 & 4 & -2 \end{bmatrix} \begin{bmatrix} 2 & 12 & -6 \\ 0 & 2 & 0 \\ 0 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
3. Soit $\mathbf{B} \in \mathbb{R}^{n\times n}$ telle que $(\mathbf{B} + \mathbf{I}_n)(\mathbf{B} - \mathbf{I}_n) = \mathbf{O}_n$.
$$(\mathbf{B} + \mathbf{I}_n)(\mathbf{B} - \mathbf{I}_n) = \mathbf{B}^2 - \mathbf{I}_n = \mathbf{O}_n \quad \Rightarrow \quad \mathbf{B}^2 = \mathbf{I}_n$$
Donc $\mathbf{B}^{-1} = \mathbf{B}$, c'est-à-dire que $\mathbf{B}$ est inversible et que son inverse est elle-même.

Solution 1.29: (énoncé)
Soit $\mathbf{A} \in \mathbb{R}^{n\times n}$. Lors de la transposition, les termes de la diagonale restent les mêmes. Pour avoir $\mathbf{A}^\top = -\mathbf{A}$, sur la diagonale, on doit avoir $\mathbf{A}(i,i) = -\mathbf{A}(i,i)$ pour tout $i \in [\![1;n]\!]$ en regardant les deux côtés de l'équation. Ce qui veut dire que $2\mathbf{A}(i,i) = 0$ pour tout $i \in [\![1;n]\!]$. en ajoutant $\mathbf{A}(i,i)$ de chaque côté de l'équation. Et donc, $\mathbf{A}(i,i) = 0$ pour tout $i \in [\![1;n]\!]$.

Solution 1.30: (énoncé)

Non, ce n'est vrai que si $\mathbf{A}$ et $\mathbf{B}$ sont inversibles. Par exemple, si $\mathbf{A} = \mathbf{O} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$, et $\mathbf{x} = (1, 1)$, alors $\mathbf{Ax} = \mathbf{Bx} = \mathbf{0} = (0, 0)$ et $\mathbf{A}$ et $\mathbf{B}$ sont différentes (et non inversibles).

Solution 1.31: (énoncé)

1. $\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\operatorname{tr}(\mathbf{A}^\top \mathbf{A})} = \sqrt{18} = 3\sqrt{2}$ et $\|\mathbf{B}\|_{\mathrm{F}} = \sqrt{\operatorname{tr}(\mathbf{B}^\top \mathbf{B})} = \sqrt{10}$
2. $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr}(\mathbf{A}^\top \mathbf{B}) = 9$ et on a bien $9 \leq 3\sqrt{2}\sqrt{10} = 12\sqrt{5}$
3. Cauchy-Schwarz.
4. Même démonstration que le théorème 1.1.1 mais avec la norme matricielle (définition 1.2.18) et en utilisant le fait que les matrices sont symétriques.

Série d'exercices 2: Élimination

Solution 2.1: (énoncé)

TODO.

Solution 2.2: (énoncé)

TODO.

Solution 2.3: (énoncé)

TODO.

Solution 2.4: (énoncé)

TODO.

Solution 2.5: (énoncé)

TODO.

Solution 2.6: (énoncé)

Pour chacun des exemples suivant, on s'autorise à faire plusieurs opérations d'élimination à la fois. Les pivots sont encadrés.

Pour $\mathbf{M}_1$: