Algèbre Linéaire
et
Analyse de Données

Licence 2 - MIASHS

Guillaume Metzler

Université Lumière Lyon 2
Laboratoire ERIC, UR 3083, Lyon

guillaume.metzler@univ-lyon2.fr

Printemps 2022

Introduction

Motivations

- Apprendre des techniques permettant de dégager les informations présentes dans un jeu de données $\rightarrow$ en dégager la substantifique moelle !
- Partir de la représentation la plus classique d'un jeu de données, *i.e.* sous forme d'un tableau, et en apprendre une représentation synthétique sous forme de graphiques.
- Nous verrons aussi quelles techniques adopter en fonction de la nature de nos données.
- On va donc voir comment synthétiser l'information à des fins d'**interprétations** mais aussi de **visualisation**.

Motivations

- On va s'éloigner de la simple statistique descriptive pour se plonger dans l'étude de variables dites **multidimensionnelles**.
- Pourquoi multidimensionelles ? Parce qu'un objet ou un individu pourra être représenté par une multitude d'informations ou **caractéristiques**.
- Toutes ces informations sont structurées dans des **tableaux**/**matrices** dont les éléments peuvent à la fois être des **chiffres** ou du **texte**.

Motivations

Les applications sont nombreuses dans le domaine des sciences sociales :

- **Marketing :** déterminer des profils clients dans un panel - identifier les cibles prioritaires et les meilleurs arguments dans des campagnes publicitaires
- **Sociologie :** identifier les profils d'individus les plus sensibles aux fake news dans la population - détecter des communauté d'individus dans les réseaux sociaux
- **Economie** : identification des facteurs clefs pour la prise de décision dans un but commercial
- ...

On retrouve également des applications dans le domaine médical, biologique, de détection de fraudes, ...

Comment faire cela ?

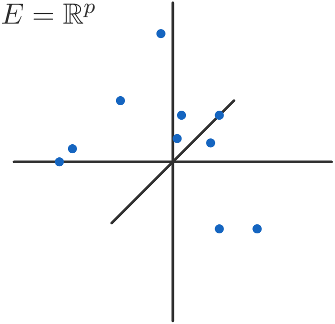
Nos données sont représentées sous une forme structurée, *i.e.* des tableaux ou des matrices de taille $n \times p$.

$$X = \begin{array}{c} \\ \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_n \end{array} \begin{array}{c} \begin{array}{ccccc} \mathbf{v}_1 & \cdots & \mathbf{v}_k & \cdots & \mathbf{v}_p \end{array} \\ \begin{pmatrix} x_{11} & \cdots & x_{1k} & \cdots & x_{1p} \\ \vdots & & \vdots & & \vdots \\ x_{i1} & \cdots & x_{ik} & \cdots & x_{ip} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & \cdots & x_{nk} & \cdots & x_{np} \end{pmatrix} \end{array},$$

Chaque ligne i représentera un **individu** qui sera décrit par un ensemble de p **caractéristiques** ou **attributs** $\mathbf{v}_j$ que l'on appellera également **variables** (ou **features**).
Individu $\rightarrow$ un vecteur de taille p, *i.e.* $\mathbf{x}_i \in \mathbb{R}^p$.
Variables $\rightarrow$ un vecteur de taille n, *i.e.* $\mathbf{v}_j \in \mathbb{R}^n$.

Deux visualisation possibles

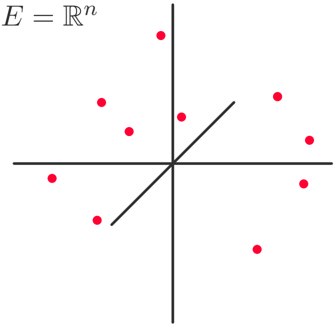
La tableau précédent suggère deux visualisation possibles de notre jeu de données



Une première représentation dans l'**espace des variables**, ainsi chaque individu est représenté par un vecteur de $\mathbb{R}^p$.

Deux visualisation possibles

On peut également faire le choix de représenter les **variables** dans l'espace des **individus**.



Ainsi chaque **variable** est représenté par un vecteur de $\mathbb{R}^n$.

Motivations

Nous verrons que les deux représentations permettent de dégager des informations différentes dans notre jeu de données : on pourra, par exemple, détecter d'éventuelles **corrélations** entre des variables ou des **groupes d'individus**.

D'un point de vue technique nous verrons également que ces deux représentations là sont très liées et qu'en étudiant simplement l'une de ces représentations nous sommes capables d'en déduire des informations sur l'autre représentation.

Objectifs

- Il est difficile d'interpréter et de synthétiser des informations se trouvant dans des espaces à **grande dimension**. Notre vision humaine nous limite à la visualisation en 3 dimension.
- Il va donc falloir développer des techniques qui permettant la synthèse et la visualisation de ces résultats dans des espaces de dimension faible (2 ou 3 en général) tout en préservant l'information qui était présente dans la représentation initiale.
- La quantitée d'informations sera représentée par les notions de **variance** (quand on étudiera les variables) ou encore par la notion de **distances entre les individus** (si on étudie les individus et non les variables).

Un exemple pour motiver ce cours I

Nous demandons aux étudiants qui suivent ce cours s'ils sont satisfaits, à travers 5 critères sur lesquels ils doivent se positionner sur une échelle de 1 à 5, 1 correspondant à "très insatisfait" et 5 à "très satisfait". Voici ces 5 critères :

1. Clarté du cours écrit
2. Fluidité de la lecture du cours écrit
3. Facilité à comprendre les exemples du cours
4. Clarté des vidéos
5. Satisfaction vis à vis de l'enseignant dans la vidéo

Un exemple pour motiver ce cours II

	Q1	Q2	Q3	Q4	Q5
Individu 1	3	3	3	5	5
Individu 2	2	3	1	5	4
Individu 3	2	3	3	4	5
Individu 4	1	1	1	1	1
Individu 5	5	5	4	3	3
Individu 6	4	5	5	2	3
Individu 7	5	5	5	3	3
Individu 8	1	1	1	1	1

TABLE – Résultats du questionnaire sur un ensemble de 8 individus.

Un exemple pour motiver ce cours III

Les outils statistiques de statistiques descriptives nous permettent simplement de décrire un individu *moyen*.

Il laisse à penser que l'échantillon est globalement moyennement satisfait vis à vis de l'ensemble des critères étudiés.

	Q1	Q2	Q3	Q4	Q5
Moyenne	2.875	3.25	3	3	3.125

TABLE – Réponse moyenne sur les différents critères vis à vis des réponses fournies au questionnaire et présentées dans la Table 1.

En réalité, différents profils peuvent être extraits de cette table...

Un exemple pour motiver ce cours IV

	Q1	Q2	Q3	Q4	Q5
Individu 1	3	3	3	5	5
Individu 2	2	3	1	5	4
Individu 3	2	3	3	4	5
Individu 4	1	1	1	1	1
Individu 5	5	5	4	4	3
Individu 6	4	5	5	2	3
Individu 7	5	5	5	3	3
Individu 8	1	1	1	1	1

TABLE – Création de groupes en fonction des réponses des individus. On identifie trois profils différents en fonction de la nature des réponses, identifiés en marron, gris et blanc.

Un exemple pour motiver ce cours V

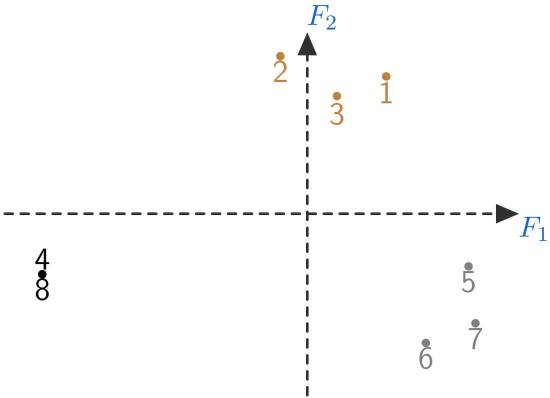


FIGURE – Représentation des individus, *i.e.* des réponses fournies au questionnaire, dans un espace à deux dimensions obtenu par l'ACP.

Quels outils ?

Pour cela on aura besoin de techniques basées sur l'**Algèbre Linéaire** et nous allons, entre autre, mobiliser ce que vous avez pu voir sur les matrices :

- Propriétés des matrices
- Représentation des vecteurs dans une base
- Réduction des endomorphismes

Mais aussi d'outils géométriques pour mieux appréhender et utiliser les notions de distances :

- Espace euclidien
- Projections
- Produit scalaire - norme

Quelles techniques ?

Les techniques d'**analyse de données** que nous allons étudier vont mobiliser les outils d'algèbre linéaire qui seront présentés.

D'un point de vue technique ou d'un point de vue mathématique, il n'y aura rien de nouveau, il faudra simplement comprendre et interpréter les résultats obtenus avec nos outils mathématiques.

Enfin, nous verrons bien sûr quels outils employés et dans quelles circonstances ainsi que l'interprétation des résultats en fonction du type de données.

Présentations des enseignants

- **Chargé de CM :** Guillaume METZLER - MCF Informatique - Enseignant à l'ICOM et rattaché au Laboratoire ERIC (Batîment K - Bureau K73)
- **Intervenants en TD :**
  - Mickaël LALLOUCHE - PRAG en Mathématiques - Spécialité en Géométrie Algébrique - UFR ASSP
  - Martial AMOVIN - Doctorant en $3^{\text{ème}}$ année en Mathématiques Appliquées - Laboratoire ERIC (Bâtiment K - Bureau K69)
  - Moi-même

Déroulement du cours

Programme :

- 12 séances de CM portant sur l'Algèbre Linéaire et l'Analyse de Données.
  Une séance est également réservée pour effectuer un partiel de 1h45 sur la première partie du cours, consacrée à l'Algèbre Linéaire

- 6 séances de TD sur la partie Algèbre Linéaire puis 5 séances de TIC/TP sur la partie Analyses de Données (sur machine avec R).
  La dernière séances de TIC/TP sera évaluée à nouveau.

Au total, vous aurez donc deux évaluations de 1h45.

Summary

Cours et Supports

Le cours se compose :

- des planches pour la présentation des cours magistraux
- d'un polycopié qui reprend de façon détaillée le contenu des slides
- d'un petit poly avec des exercices

L'ensemble de ces supports sont disponibles à l'adresse suivante (les slides seront régulièrement mis à jour) :

https://guillaumemetzler.github.io/aladd_lyon2.html

Attention : présence aux TD obligatoires ! Aucune correction ne sera mise en ligne.

Algèbre Linéaire

Contenu

Une première partie plutôt *mathématiques* mais qui n'ira pas dans les détails des démonstrations (elles sont disponibles dans le poly).

- Présentation de la notion d'espace vectoriels et des applications linéaires - fondamentaux en algèbre linéaire
- Algèbre linéaire en dimension finie - bases - représentation des vecteurs et applications linéaires dans une base
- Rappels sur le calcul matriciel et la résolution de systèmes liénaires (pivot de gauss)
- Diagonalisation (ou réduction) des endomorphismes diagonalisables - valeurs propres et vecteurs propres d'une matrice
- Etude des formes quadratiques - géométrie eucldienne - produit scalaire

A voir comme une présentation des outils.

Algèbre Linéaire

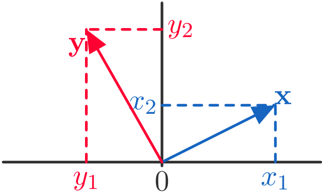
Espaces vectoriels et applications linéaires

Introduction

En algèbre linéaire nous manipulons essentiellement des **vecteurs**, en général notés $\mathbf{x}$, qui vivent dans un certain espace que l'on nomme **espace vectoriel**.

$E = \mathbb{R}$

$E = \mathbb{R}^2$ ou $\mathbb{C}$



Dans de tels espaces, les objets, comme $\mathbf{x}$, sont décrits par des *coordonnées*, *i.e.* $\mathbf{x} = (x_1, x_2, \cdots, x_n)$.

Introduction

Ces objets se retrouvent dans différents contextes

- dans différentes branches des mathématiques : *algèbre linéaire* - *analyse avec l'études des fonctions* - **probabilités** - *statistiques* - *optimisation*
- en physique quantique ou probabiliste par exemple : lorsque l'on étudie certains opérateurs qui peuvent décrire le mouvement d'une particule.
- en machine learning : pour la construction de modèles de prévisions ou qui peuvent servir à effectuer des tâches de façon automatique
- ...

Mais qu'est-ce qu'un espace vectoriel d'un point de vue mathématique ?

Espaces Vectoriels

Définition 2.1: Espace vectoriel

Soit E un ensemble muni d'une loi de composition interne notée "$+$" et d'une loi externe notée "$\cdot$" définie sur $\mathbb{K} \times E$ par

$$\begin{array}{rcl} \mathbb{K} \times E & \to & E \\ (\alpha, \mathbf{x}) & \to & \alpha \cdot \mathbf{x}. \end{array}$$

On dit que E est un $\mathbb{K}$-espace vectoriel (*e.v.*) si les propriétés suivantes sont vérifiées :

i) $(E, +)$ est un groupe abélien (*i.e.* commutatif)

ii) $\forall \mathbf{x} \in E,\ 1 \cdot \mathbf{x} = \mathbf{x}.$

iii) $\forall (\alpha, \beta) \in \mathbb{K}^2,\ \forall \mathbf{x} \in E,\ (\alpha + \beta) \cdot \mathbf{x} = \alpha \cdot \mathbf{x} + \beta \cdot \mathbf{x}.$

iv) $\forall \alpha \in \mathbb{K},\ \forall \mathbf{x}, \mathbf{x}' \in E,\ \alpha \cdot (\mathbf{x} + \mathbf{x}') = \alpha \cdot \mathbf{x} + \alpha \cdot \mathbf{x}'.$

v) $\forall (\alpha, \beta) \in \mathbb{K}^2,\ \forall \mathbf{x} \in E,\ \alpha \cdot (\beta \cdot \mathbf{x})) = (\alpha\beta) \cdot \mathbf{x}.$

Espaces Vectoriels

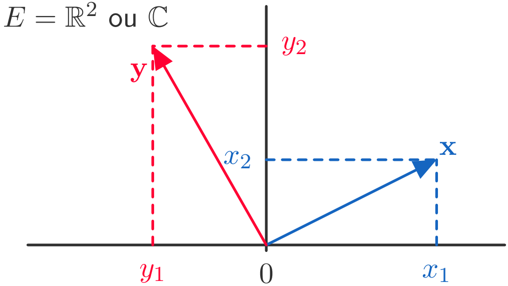
Un groupe abélien est défini par les points suivants :

1. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E,\ (\mathbf{x}+\mathbf{y})+\mathbf{z} = \mathbf{x}+(\mathbf{y}+\mathbf{z})$, c'est l'*associativité*.
2. $\forall \mathbf{x}, \mathbf{y} \in E,\ \mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$, c'est la *commutativité*.
3. il existe un élément, noté $\mathbf{0}$ qui vérifie, pour tout $\mathbf{x} \in E : \mathbf{0}+\mathbf{x} = \mathbf{x}$.
4. $\forall \mathbf{x} \in E$, il existe un élément $\mathbf{x}' \in E$ qui vérifie $\mathbf{x}+\mathbf{x}' = \mathbf{0}$. C'est l'existence d'un *inverse*, en général noté $-\mathbf{x}$.

Dans le cadre de ce cours, les lois "$+$" et "$\cdot$" évoquées font références aux lois additive et multiplicative que l'on connait.

Quelques exemples

L'espace $\mathbb{R}^2$ est un espace vectoriel



Les éléments s'écrivent $\mathbf{x} = (x_1, x_2)$ et l'origine de cet espace est le point $\mathbf{0} = (0, 0)$. En particulier, nous aurons aussi

$$\mathbf{x} + \mathbf{y} = (x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2).$$

On fait la somme composante par composante.

Exemples

Ce que l'on vient de voir dans l'espace $\mathbb{R}^2$ se généralise à $\mathbb{R}^n$.

On peut également considérer l'espace vectoriel des matrices carrées de taille 2, *i.e.* $\mathscr{M}_2(nsetR)$:

- $\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} + \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = \begin{pmatrix} x_{11} + y_{11} & x_{12} + y_{12} \\ x_{21} + y_{21} & x_{22} + y_{22} \end{pmatrix}$
- $\lambda \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} \lambda x_{11} & \lambda x_{12} \\ \lambda x_{21} & \lambda x_{22} \end{pmatrix}$

Quel serait, selon vous, l'élément neutre pour la loi "$+$" et l'opposé pour cette même loi ?

On pourrait aussi étudier l'espace vectoriel des fonctions.

Sous-espace vectoriel

Définition 2.2: Sous-espace vectoriel

Soit E un $\mathbb{K}$-espace vectoriel soit F une partie de E (on peut aussi dire, un sous ensemble de E). F est un **sous-espace vectoriel** de E si F est lui même doté d'une structure d'espace vectoriel pour les lois induites par les lois définies par E.

Proposition 2.1

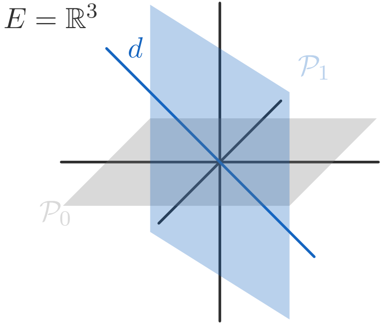
Soit E un $\mathbb{K}$-espace vectoriel et soit F un sous-ensemble de E. F est un **sous-espace** vectoriel de E *si et seulement si* les deux propriétés suivantes sont vraies :

i) $F \neq \emptyset$, *i.e.* F est non vide,

ii) $\forall (\mathbf{x}, \mathbf{x}') \in F^2$, $\forall (\alpha, \beta) \in \mathbb{R}^2$, $\alpha \mathbf{x} + \beta \mathbf{x}' \in F$, *i.e.* F est stable combinaison linéaire.

Exemples

Prenons un exemple concret avec $E = \mathbb{R}^3$, alors toutes les *droites* et tous les *plans* qui passent par l'origine sont des sous-espaces vectoriels de E.



Dans cet exemple, les *plans vectoriels* $\mathcal{P}_0, \mathcal{P}_1$ et la droite vectorielle d sont des sous espaces de $E = \mathbb{R}^3$.

Exemples

Plus formellement :

- une droite vectorielle, *i.e.* un espace de la forme $\lambda \mathbf{x}, \forall \lambda \in \mathbb{R}$ est un sous espace vectoriel de $\mathbb{R}^n$.
- un plan vectoriel (ou hyperplan de $\mathbb{R}^3$ dans ce cas) est un espace de la forme : $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2$, est un sous espace vectoriel.
- plus généralement, les espaces définis comme une combinaisons linéaire d'un ensemble de vecteur de $\mathbb{R}^n$ sont des sous-espaces de $\mathbb{R}^n$, *i.e.*

$$F = \left\{ \mathbf{x} \in E \mid \exists \lambda_i, i \in [\![1, p]\!] \text{ tel que } \mathbf{x} = \sum_{i=1}^{p} \lambda_i \mathbf{x}_i \right\}$$

est un sous-espace vectoriel de $E = \mathbb{R}^n$.

Propriétés des sous-espaces vectoriels

Proposition 2.2: Sous-espace vectoriel et intersection

Soit E un $\mathbb{K}$ espace vectoriel, et soient F et G deux sous-espaces vectoriels de E.
Alors $F \cap G$ est un sous-espace-vectoriel de E.

Preuve : La démonstrations est laissée à titre d'exercice. Il suffit de montrer que $\mathbf{0}$ appartient à cet espace et qu'il est stable par combinaison linéaire.

On pourra aussi montrer que l'**union** de deux sous-espaces vectoriels de E n'est pas un sous-espace vectoriel de E et que le complémentaire d'un sous-espace vectoriel de E n'est pas un sous-espace vectoriel.

Espaces supplémentaires

Définition 2.3: Espaces supplémentaires

Soit E un $\mathbb{K}$-espace vectoriel, et soient F et G deux sous-espaces de E. F et G sont dits **supplémentaires** si les deux propriétés suivantes sont vérifiées :

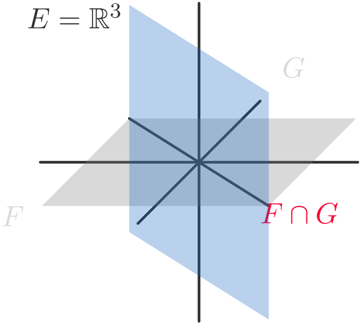
i) $F \cap G = \{0_E\}$,

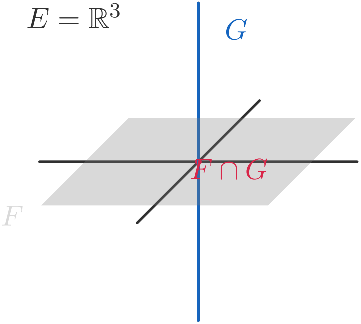
ii) $F + G = E$,

où $F + G$ désigne le sous-espace vectoriel engendré par l'ensemble des vecteurs de F et G.
On notera alors $E = F \oplus G$.

Espaces Supplémentaires

Soit $E = \mathbb{R}^3$ un $\mathbb{R}$-espace vectoriel et soient F et G des sous-espaces de E. Considérons les deux graphes ci-dessous :





Gauche : deux sous-espaces vectoriels de E mais ces derniers ne sont pas supplémentaires, l'intersection (en jaune) n'est pas réduite à $\{0_E\}$.
Droite : deux sous-espaces supplémentaires dans $\mathbb{R}^3$ (une droite et un plan).

Espaces supplémentaires

Proposition 2.3: Somme Supplémentaire

Soit E un $\mathbb{K}$-espace vectoriel, et soient F et G deux sous-espaces de E, alors $E = F \oplus G$ si et seulement si $\forall \mathbf{x} \in E \mid \exists!(\mathbf{x}_1, \mathbf{x}_2) \in F \times G,$ tel que $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$.

Cette notion sera importante lorsque l'on va chercher à décomposer en éléments indépendants l'information présente dans un jeu de données. On la retrouvera plus tard lorsque l'on va aborder la réduction des endomorphismes,*i.e.* des applications linéaires d'un espace dans lui même.

Regardons de suite de quoi il s'agit ! !

Applications linéaires

Définition 2.4: Application linéaire

Soient E et E' deux $\mathbb{K}$-espaces vectoriels, et soit f une application de E dans E'. On dit que f est une application linéaire si :

$$\forall (\alpha, \beta) \in \mathbb{K}^2,\ \forall (\mathbf{x}, \mathbf{x}') \in E^2,\ f(\alpha \mathbf{x} + \beta \mathbf{x}') = \alpha f(\mathbf{x}) + \beta f(\mathbf{x}').$$

On peut résumer la définition d'application linéaire de la façon suivante : *l'image d'une combinaison linéaire par cette application est la combinaison linéaire des images de cette application*.

Cette définition se généralise (par récurrence) pour une combinaison linéaire de n vecteurs.

Exemple

On considère l'application f définie de $\mathbb{R}^3$ dans $\mathbb{R}^2$ par

$$f(x_1, x_2, x_3) = (x_1 + x_2, x_1 + x_3).$$

Cette application est bien linéaire. En effet, pour tout $\alpha, \beta \in \mathbb{R}$, nous avons

$$\begin{aligned}
f(\alpha\mathbf{x} + \beta\mathbf{y}) &= f(\alpha(x_1, x_2, x_3) + \beta(y_1, y_2, y_3)), \\
&\quad \downarrow \text{ on développe} \\
&= (\alpha(x_1 + x_2) + \beta(y_1 + y_2), \alpha(y_1 + y_2) + \beta(y_1 + y_3)), \\
&= (\alpha x_1 + \alpha x_2 + \beta y_1 + \beta y_2, \alpha x_1 + \alpha x_3 + \beta y_1 + \beta y_3), \\
&\quad \downarrow \text{ on décompose} \\
&= (\alpha x_1 + \alpha x_2, \alpha x_1 + \alpha x_3) + (\beta y_1 + \beta y_2, \beta y_1 + \beta y_3), \\
&= \alpha(x_1 + x_2, x_1 + x_3) + \beta(y_1 + y_2, y_1 + y_3), \\
&= \alpha f(\mathbf{x}) + \beta f(\mathbf{y}).
\end{aligned}$$

Noyau et image d'une application linéaire

Définition 2.5: Noyau et Image

Soient E et E' deux $\mathbb{K}$-espaces vectoriels et soit f une application linéaire de E dans E'. On appelle :

- *noyau de f*, noté *Ker(f)*, le sous-espace vectoriel de E défini par :

$$Ker(f) = \{\mathbf{x} \in E \mid f(\mathbf{x}) = 0_E\}.$$

- *image de f*, notée *Im(f)*, le sous espace vectoriel de E' défini par :

$$Im(f) = \{\mathbf{y} \in E' \mid \exists \mathbf{x} \in E \quad \text{tel que} \quad f(\mathbf{x}) = \mathbf{y}\}.$$

Cette définition cache aussi le fait que l'image et le noyau d'une application linéaire sont aussi des sous-espaces vectoriels.

Nature d'une application

Définition 2.6: Nature d'une application

Soient X et Y deux ensembles non vides et f une application de X dans Y. Alors :

- f est dite **injective** si tout élément de Y admet **au plus** un antécédent dans X. Ce que l'on peut aussi formuler :

$$\forall x, x' \in X,\ x \neq x' \implies f(x) \neq f(x')$$

  ou encore

$$\forall x, x' \in X,\ f(x) = f(x') \implies x = x'.$$

- f est dite **surjective** si tout élément Y admet *au moins* un antécédent dans X, *i.e.* :

$$\forall y \in Y,\ \exists x \in X \quad \text{tel que} \quad f(x) = y.$$

- f est dite **bijective** si elle est à la fois *injective et surjective*, ce que l'on peut écrire :

$$\forall y \in Y,\ \exists! x \in X \quad \text{tel que} \quad f(x) = y.$$

Exemples I

Considérons l'application f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(\mathbf{x}) = f(x_1, x_2) = x_1.$$

L'application n'est pas injective, car les vecteurs $(0, 1)$ et $(0, 2)$ ont la même image par l'application f.

On peut en revanche montrer que l'application est surjective. En effet, quelque soit x_1, le vecteur (x_1, t) est bien un antécédent de x_1 par l'application f, pour tout $t \in \mathbb{R}$.

Cette application n'est pas bijective.

Exemples II

Considérons maintenant l'application f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par :

$$f(\mathbf{x}) = f(x_1, x_2) = (x_1, 2x_2).$$

Pour montrer que l'implication est injective, on considère deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ tels que $f(\mathbf{x}) = f(\mathbf{y})$ et montrons que $\mathbf{x} = \mathbf{y}$.

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{y}), \\ &\downarrow \text{ définition de } f \\ (x_1, 2x_2) &= (y_1, 2y_2), \\ (x_1 - y_1, 2(x_2 - y_2)) &= 0. \end{aligned}$$

Cette dernière égalité implique que $x_1 = y_1$ et $x_2 = y_2$, donc $\mathbf{x} = \mathbf{y}$.

Exemples III

Pour la surjectivité, on considère un vecteur $\mathbf{y} \in \mathbb{R}^2$ et on montre que le vecteur $\mathbf{x} = (y_1, \frac{1}{2}y_2)$ vérifie bien $f(\mathbf{x}) = \mathbf{y}$.

L'application linéaire étant surjective et injective, elle est donc bijective.

Point vocabulaire

Soient E et E' des $\mathbb{K}$-espaces vectoriels et soit f une application linéaire de E dans E', *i.e.* $f \in \mathscr{L}(E, E')$: ensemble des applications linéaires de E dans E'. De plus :

- dans le cas où $E = E'$, f est appelée *endomorphisme de E* et notera $\mathscr{L}(E)$ l'ensemble des endomorphismes de E,
- dans le cas où $E' = K$, f est appelée *forme linéaire sur E*.

De plus :

- si f est bijective, alors f est appelée *isomorphisme de E dans E'*,
- enfin, si f est un isomorphisme de E dans lui même (donc un endomorphisme bijectif), alors f est appelée *automorphisme* de E et on note $\mathscr{G}\mathscr{L}(E)$ l'ensemble des automorphismes de E. f est alors inversible et on note f^{-1} son inverse. Il vérifie $f(f^{-1}(\mathbf{x})) = f^{-1}(f(\mathbf{x})) = \mathbf{x}$

Caractérisation des applications linéaires

La vérification des définitions peut parfois se révéler complexe. Heureusement, il existe une caractérisation simple des des applications linéaires injectives.

Proposition 2.4: Caractérisation de l'injectivité

Soient E et E' deux $\mathbb{K}$-espaces vectoriels et soit f une application linéaire de E dans E', alors f est injective *si et seulement si* $Ker(f) = \{0\}$, *i.e.* si son noyau est réduit au vecteur nul.

En dimension finie (prochaine section) on verra comment caractériser simplement les applications en raisonnant sur la **dimension** des espaces.

Pourquoi ces notions ?

L'étude de ces sous-espaces est très important pour étudier les caractéristiques de certaines applications.

- On pourra essayer d'exprimer plus simplement nos applications linéaires.
- On pourra par exemple en déduire plus tard, si des variables sont corrélées et ainsi réduire naturellement la dimension de l'espace de représentation de nos données.

Afin de compléter et terminer cette première section, on va s'intéresser plus précisément à une application linéaire en particulier : **les projections**.

Projecteurs

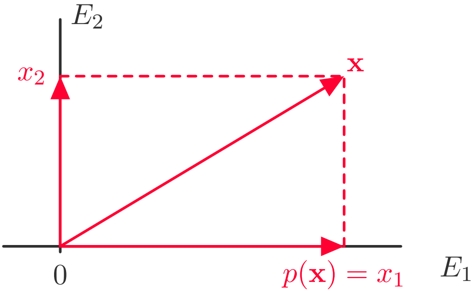
Définition 2.7: Projecteur

Soit E un $\mathbb{K}$-espace vectoriel, et soient E_1 et E_2 des sous-espaces vectoriels supplémentaires de E. On appelle **projecteur sur E_1 parallèlement E_2**, l'application $p : E \to E$ qui, à un vecteur $\mathbf{x} \in E$ se décomposant comme $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ avec $(\mathbf{x}_1, \mathbf{x}_2) \in E_1 \times E_2$, associe le vecteur $\mathbf{x}_1$.

Si la définition paraît abstraite, elle est en fait très simple à comprendre avec un exemple.

Exemple

Soit E un espace vectoriel, disons $E = \mathbb{R}^2$ et considérons le vecteur $\mathbf{x} = (x_1, x_2) = (5, 3) = (5, 0) + (0, 3) \in E$. Soit la projection p qui consiste à conserver uniquement la première composante composante de ce vecteur, alors $p(\mathbf{x}) = (5, 0)$.



Propriétés des projecteurs

On peut résumer cette définition par le schéma suivant

$$\begin{array}{rcl} p: & E = E_1 \oplus E_2 & \to E. \\ & \mathbf{x} = \underbrace{\mathbf{x}_1}_{\in E_1} + \underbrace{\mathbf{x}_2}_{\in E_2} & \mapsto \mathbf{x}_1 \end{array}$$

La proposition suivante montre que les projecteurs ont des images et noyaux qui sont faciles à déterminer à partir de leur définition.

Proposition 2.5: Propriété projecteur

Soit E un $\mathbb{K}$-espace vectoriel, et soient E_1 et E_2 deux sous-espaces supplémentaires de E. Soit p le projecteur sur E_1 parallèlement à E_2. Alors p est un endomorphisme de E dont le noyau et l'image sont :

$$Ker(p) = E_2 \text{ et } Im(p) = E_1.$$

Pour finir

Que pensez-vous de l'application $p \circ p$? Autrement dit, si je prends un vecteurs $\mathbf{x}$ que lui applique un projecteur et que j'applique à nouveau cette projection sur le résultat obtenu par la première projection ?

Jusqu'ici nous n'avons jamais pris en compte la taille de nos espaces, *i.e.* on n'a jamais fixé la taille de nos vecteurs dans tous les résultats qui précèdent.

Que deviennent nos résultats si on étudie des objets avec une taille déterminée ?

Algèbre Linéaire

Espaces vectoriels de dimension finie

Famille libre et famille liée

Définition 2.8: Familles libres, liées

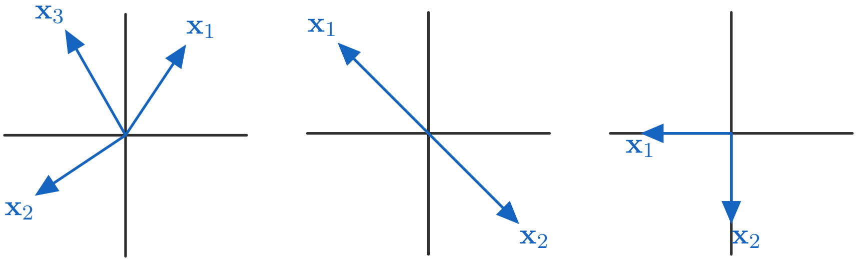
Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ une famille de vecteurs d'un espace vectoriel E. Cette famille est dite :

- **libre** si $\forall (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n,\ \sum_{k=1}^{n} \lambda_k \mathbf{x}_k = 0 \implies \lambda_k = 0\ \forall k \in [\![1, n]\!]$.
- **liée** si elle n'est pas libre, c'est-à-dire s'il existe $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n \setminus \{0\}^n$ tel que :

$$(\lambda_1, \ldots, \lambda_n) \neq (0, \ldots, 0) \quad \text{et} \quad \sum_{k=1}^{n} \lambda_k \mathbf{x}_k = 0_E.$$

Exemples

Soit $E = \mathbb{R}^2$ un espace vectoriel et considérons les graphes ci-dessous avec des familles de vecteurs $(\mathbf{x}_1, \mathbf{x}_2)$ ou $(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$.



A votre avis, dans quel(s) cas les vecteurs forment une famille libre ou liée ?

Exemples

On va simplement regarder s'il existe un lien entre les différents vecteurs.

- Dans le premier cas, nous avons trois vecteurs distincts dans un espace à deux dimensions, on peut donc écrire l'un des vecteurs comme combinaison linéaire des deux autres, la **famille n'est donc pas libre, elle est liée**
- Dans le deuxième cas, les vecteurs sont *colinéaires* et on a la relation $\mathbf{x}_1 = -\mathbf{x}_2$, à nouveau, **la famille est liée**.
- Dans le dernier cas, **la famille est bien libre** car les deux vecteurs sont orthogonaux.

On verra plus tard que le troisième exemple montre le cas d'une famille de vecteurs que l'on appellera **base** de l'espace vectoriel.

Caractérisation des familles liées

Proposition 2.6: Caractérisation famille liée

Soit $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ une famille d'au moins deux vecteurs de E, cette famille est *liée* si et seulement si l'un de ses vecteurs peut s'écrire comme combinaison linéaire des autres vecteurs de cette même famille.

Remarquons, par contraposée, qu'une famille est dite *libre* si aucun de ses vecteurs ne peut s'écrire comme combinaison linéaire de ses autres vecteurs (on parle *d'indépendance linéaire*).

Famille génératrice et base

Définition 2.9: Famille génératrice

Soit $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ une famille de vecteurs de E, cette famille est *génératrice* si $Vect(\mathbf{x}_1, \ldots, \mathbf{x}_n) = E$. C'est-à-dire, si tout élément $\mathbf{x}$ de E peut s'écrire comme une combinaison linéaire des éléments de cette famille :

$$\forall \mathbf{x} \in E,\ \exists (\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n, \quad \text{tel que} \quad \mathbf{x} = \sum_{k=1}^{n} \lambda_k \mathbf{x}_k.$$

Définition 2.10: Base

On appelle base de E toute famille d'éléments de E à la fois libre et génératrice.

Exemples

- Dans $\mathbb{R}^2$, les vecteurs $\mathbf{e}_1 = (1, 0)$ et $\mathbf{e}_2 = (0, 1)$ forment une famille libre et génératrice de $\mathbb{R}^2$.
  En effet :
  $$\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 = \mathbf{0} \iff (\lambda_1, \lambda_2) = \mathbf{0},$$
  i.e. si seulement si $\lambda_1 = \lambda_2 = 0$.

  De plus tout vecteur $\mathbf{x} = (x_1, x_2)$ de $\mathbb{R}^2$ peut s'écrire comme combinaison linéaire des vecteurs $\mathbf{e}_1$ et $\mathbf{e}_2$
  $$\mathbf{x} = (x_1, x_2) = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2.$$

Exemples

- La famille de vecteurs $\mathbf{v}_1 = (-1, 1)$ et $\mathbf{e}_2 = (1, 1)$ forme également une base de $\mathbb{R}^2$. En effet

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 = \mathbf{0} \iff (-\lambda_1 + \lambda_2, \lambda_1 + \lambda_2) = \mathbf{0}.$$

En résolvant le système linéaire :

$$\begin{cases} -\lambda_1 + \lambda_2 & = & 0 \\ \lambda_1 + \lambda_2 & = & 0 \end{cases} \iff \begin{cases} -\lambda_1 + \lambda_2 & = & 0 \\ 2\lambda_2 & = & 0 \end{cases}$$

On en déduit $\lambda_1 = \lambda_2 = 0$. De plus si $\mathbf{x} = (x_1, x_2)$ est un vecteur de $\mathbb{R}^2$, alors il peut s'écrire

$$\mathbf{x} = \frac{x_2 - x_1}{2} \mathbf{v}_1 + \frac{x_2 - x_1}{2} \mathbf{v}_2$$

Exemples

- On pourra en revanche vérifier que la famille de vecteurs

$$\mathbf{v}_1 = (-2, 2) \text{ et } \mathbf{v}_2 = (-1, 1)$$

ne forment pas une famille libre de $\mathbb{R}^2$. Ce n'est donc pas une base de $\mathbb{R}^2$.

- De la même façon, on pourra vérifier que la famille

$$\mathbf{v}_1 = (-1, 0, 1) \text{ et } \mathbf{v}_2 = (1, 0, 0)$$

ne forme pas une famille génératrice de $\mathbb{R}^3$.

Caractérisation des bases

Proposition 2.7: Caractérisation base

Soit $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ une famille de vecteurs de E, cette famille est une base de E si et seulement si

$$\forall \mathbf{x} \in E,\ \exists!(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad \text{tel que} \quad \mathbf{x} = \sum_{k=1}^{n} \lambda_k \mathbf{x}_k.$$

Remarquez bien la différence avec la définition de famille génératrice ! L'écriture de tout élément de $\mathbf{x}$ de E s'exprime de façon **unique** comme élément d'une base de E, alors que cette décomposition n'est pas unique pour une famille génératrice.

Base canonique de $\mathbb{R}^n$

Soit $n \in \mathbb{N}^\star$, l'espace $\mathbb{R}^n$ est l'ensemble des n-uplets $(x_1, \ldots, x_n)$ où pour tout $j \in [\![1, n]\!],\ \mathbf{x}_k \in \mathbb{R}$. On définit alors la famille $(\mathbf{e}_1, \ldots, \mathbf{e}_n) \in (\mathbb{R}^n)^n$, par :

$$\begin{aligned}
\mathbf{e}_1 &= (1, 0, 0, \ldots, 0, 0), \\
\mathbf{e}_2 &= (0, 1, 0, \ldots, 0, 0), \\
&\vdots \\
\mathbf{e}_n &= (0, 0, 0, \ldots, 0, 1).
\end{aligned}$$

Tout vecteur $\mathbf{x} = (x_1, \ldots, x_n)$ se décompose alors de manière unique comme combinaison linéaire des vecteurs de la base canonique :

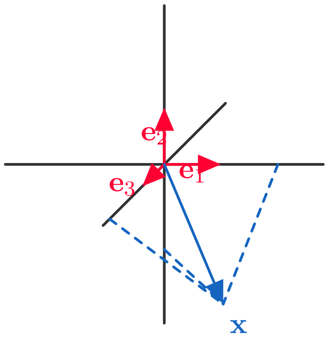
$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n.$$

Les éléments $(x_1, \ldots, x_n)$ sont donc appelées coordonnées du vecteur $\mathbf{x}$ dans la base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$, cette base est appelée *base canonique* de $\mathbb{R}^n$.

Base canonique de $\mathbb{R}^n$

Définir un vecteur dans $\mathbb{R}^n$ revient donc à déterminer exactement ses coordonnées dans la base de E.

Dans l'exemple de gauche, nous avons $\mathbf{x} = 2\mathbf{e}_1 - 1.5\mathbf{e}_2 + 2.5\mathbf{e}_3$ et dans l'exemple de droite, les coordonnées sont plus complexes à déterminer car nous n'utilisons pas la base canonique de $\mathbb{R}^n$.



Base incomplète

Théorème 2.1: Théorème de la base incomplète

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie, alors toute famille libre de E peut être complétée en une base de E.

Prenons par exemple la famille de vecteurs $\mathbf{v}_1 = (1, -1, 0)$ et $\mathbf{v}_2 = (2, 5, 0)$.

On montre que deux vecteurs forment une libre de $\mathbb{R}^3$ mais n'est pas une famille génératrice de $\mathbb{R}^3$.

On peut cependant lui ajouter le vecteur $\mathbf{v}_3 = (0, 0, 1)$ afin que cette famille libre devienne une base de $\mathbb{R}^3$.

Espace vectoriel de dimension finie

Définition 2.11: Dimension finie

Un espace vectoriel vectoriel E est dit de *dimension finie* s'il admet une famille génératrice finie.

Un point très important est que **tout espace vectoriel de dimension fini admet une base (qui est finie)**.

Théorème 2.2: Définition de la dimension

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie, alors toutes les bases de E ont le même nombre fini d'éléments, ce nombre d'éléments est appelé *dimension de l'espace vectoriel* E, il est noté $dim(E)$.

Exemples

L'exemple présenté précédemment permet de montrer que les espaces vectoriels $\mathbb{R}^n,\ n \in \mathbb{N}^\star$ sont des espaces vectoriels de dimension n. Nous avons également exhibé une base pour de tels espaces vectoriels comme la base canonique.

Dans le cas où E est réduit à 0_E, on dira que E est un espace vectoriel de dimension 0 (réduit à un point). Cela revient à considérer que la famille vide $\emptyset$ est la seule base de E.

L'espace des matrices carrées d'ordre 2 est un espace de dimension 4 dont une base est donnée par les matrices suivantes :

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{ et } \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Dimension des sous-espaces

Proposition 2.8: Dimension sous-espace vectoriel

Soit E un espace vectoriel de dimension finie, alors tout sous espace vectoriel F de E est aussi de dimension finie et $dim(F) \leq dim(E)$. De plus, on l'équivalence $E = F \iff dim(E) = dim(F)$

Proposition 2.9: Dimension espaces supplémentaires

Soit un E un espace vectoriel de dimension finie et soient F et G des sous-espaces supplémentaires de E, alors $dim(E) = dim(F) + dim(G)$.

Dimension des sous-espaces

La proposition précédente est un cas particulier du résultat suivant qui est valable quelque soit la nature des sous-espaces F et G d'un espace vectoriel E.

Proposition 2.10: Dimension somme de sous-espaces

Soit E un espace vectoriel de dimension finie et soient F et G des sous-espaces vectoriels de E, alors

$$dim(F+G) = dim(F) + dim(G) - dim(F \cap G).$$

Hyperplans

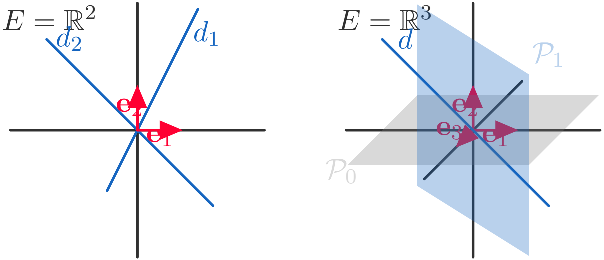
Définition 2.12: Hyperplan

Soit E un espace vectoriel de dimension finie $n \geq 2$. On appelle *hyperplan* de E tout sous-espace vectoriel de E de dimension $n-1$.

En tant que sous-espace vectoriel, les hyperplans doivent donc nécessairement contenir le vecteur nul, *i.e.* le vecteur 0_E.

Exemple

L'exemple le plus simple de sous-espace vectoriel que l'on puisse imaginer est **une droite dans un plan**. En effet, un plan est un espace de dimension 2 et la droite est un espace de dimension 1. Les droites d_1 et d_2 sont des sous espaces vectoriels de $\mathbb{R}^2$.



Même chose avec les plans dans $E = \mathbb{R}^3$, les plans $\mathcal{P}_0$ et $\mathcal{P}_1$ sont des hyperplans de $\mathbb{R}^3$; en revanche la droite d n'est pas un hyperplan de $\mathbb{R}^3$.

Retour sur les familles de vecteurs

On se rappelle qu'un espace vectoriel est engendré par une famille de vecteurs. On peut définir une caractéristique de cette dernière qui permet de faire le lien entre l'espace engendré par la famille de vecteurs et la dimension de l'espace engendré par cette même famille.

Définition 2.13: Rang d'une famille

Soit un espace vectoriel et $\mathscr{F} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ une famille de vecteurs de E. On appelle *rang* de la famille $\mathscr{F}$, noté $rang(\mathscr{F})$ ou encore $rg(\mathscr{F})$, la dimension de l'espace vectoriel engendré par $\mathscr{F}$, *i.e.*

$$rg(\mathbf{x}_1, \ldots, \mathbf{x}_n) = dim\,(Vect(\mathbf{x}_1, \ldots, \mathbf{x}_n)).$$

Rang d'une famille de vecteurs

Cette notion de rang est fondamentale et reviendra lorsque nous reviendrons sur les applications linéaires et leurs représentations avec le théorème du rang mais aussi lorsque nous introduirons les matrices dans la prochaine section.

Proposition 2.11: Propriétés rang d'une famille

Soit E un espace vectoriel et soit $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ une famille de vecteurs de E, alors :

- $rg(\mathbf{x}_1, \dots, \mathbf{x}_n) \leq n$,
- $rg(\mathbf{x}_1, \dots, \mathbf{x}_n) = n$ si et seulement si $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ est libre.

Représentation application linéaire

On considère un espace vectoriel E de dimension finie qui admet une base $(\mathbf{x}_1, \dots, \mathbf{x}_n)$, F un autre espace vectoriel et considérons $f \in \mathscr{L}(E, F)$ une application linéaire. Soit $\mathbf{x}$ un élément de E alors $\mathbf{x}$ peut s'écrire de façon unique $\mathbf{x} = \sum_{k=1}^{n} \lambda_k \mathbf{x}_k$.
Si on applique la fonction f au vecteur $\mathbf{x}$, on a

$$f(\mathbf{x}) = \sum_{k=1}^{n} \lambda_k f(\mathbf{x}_k).$$

L'application linéaire est donc entièrement déterminée par les images $(f(\mathbf{x}_1), \dots, f(\mathbf{x}_n))$ des vecteurs de bases $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ de E.

Morphisme et famille

Via l'étude de cette nouvelle famille de vecteurs nous sommes en mesures d'en déduire des informations sur les propriétés du morphisme, *i.e.* de l'application linéaire f.

Proposition 2.12: Caractérisation morphisme et famille

Soit E un espace vectoriel de dimension finie $n \in \mathbb{N}^\star$ muni d'une base $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$. Soit E' un espace vectoriel de soit f une application linéaire de E dans E', alors :

i) f est injective si et seulement si la famille $(f(\mathbf{e}_1), \ldots f(\mathbf{e}_n))$ est une famille libre de E'.

ii) f est surjective si et seulement si $(f(\mathbf{e}_1), \ldots f(\mathbf{e}_n))$ est une famille génératrice de E'.

iii) f est bijective si et seulement si $(f(\mathbf{e}_1), \ldots f(\mathbf{e}_n))$ est une base de E'

Retour à la notion de rang

Revenons maintenant à la dimension de rang. Nous avions précédemment introduit la notion de rang d'une famille de vecteurs. On va voir qu'il est également possible de définir le rang d'une application linéaire en regardant le rang de l'image des vecteurs d'une base de E par cette application linéaire.

Définition 2.14: Rang application linéaire

Soient E et E' deux $\mathbb{K}$-espaces vectoriels de dimensions finies, et $f \in \mathscr{L}(E, E')$. On appelle *rang de* f, noté $rang(f)$ ou $rg(f)$, la dimension de l'espace $Im(f)$.

Théorème du rang

Nous pouvons même relier le rang d'une application linéaire $f \in \mathscr{L}(E, E')$ à la dimension de l'espace de départ E et la dimension du noyau, c'est ce nous donne le **théorème du rang**.

Théorème 2.3: Théorème du rang

Soient E et E' deux $\mathbb{K}$-espaces vectoriels de dimensions finies et soit $f \in \mathscr{L}(E, E')$, alors $dim\,(Ker(f)) + dim\,(Im(f)) = dim(E)$.

Caractérisation morphismes

Proposition 2.13: Morphismes et dimension

Soient E et E' deux espaces vectoriels de dimensions finies tels que $dim(E) = dim(E')$. Soit $f \in \mathscr{L}(E, E')$, il y a alors équivalence entre les trois propriétés suivantes :

i) f est injective

ii) f est surjective

iii) f est bijective

Preuve : Ce résultat découle immédiatement du théorème du rang.

Ce résultat se révèle cependant très utile lorsque l'on cherche à montrer qu'une application linéaire entre deux espaces de même dimension réalise un isomorphisme. En effet, il suffira simplement de montrer qu'elle est injective **ou** surjective.

Pour finir

On a vu quelques résultats théoriques sur les espaces vectoriels de dimension finie pour finir sur l'étude des applications linéaires dans de tels espaces.

L'étude des familles de vecteurs a été point clef de cette section mais son utilité est restée très abstraite.

Dans la prochaine section, nous allons voir comment les propriétés de ces familles de vecteurs avec des objets plus pratiques : les **matrices**.

Algèbre Linéaire

Matrices et calcul matriciel

Définition

Définition 2.15: Matrice

On appelle *matrice à n lignes et p colonnes* à coefficients dans $\mathbb{K}$, toute application de $[\![1, n]\!] \times [\![1, p]\!]$ dans $\mathbb{K}$.
Une telle matrice, notée A, se note alors

$$A = (a_{ij})_{i,j=1}^{n,p} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}.$$

La matrice A est également appelée matrice de type (n, p) pour dire qu'elle comporte n lignes et p colonnes.

Matrice carrée

Définition 2.16: Matrices carrées

On appelle **matrice carrée** toute matrice de type (n, n). Ce type de matrice est dit matrice d'ordre n.
Une matrice carrée A est dite **diagonale** si tous les éléments si pour tout $i \neq j,\ a_{ij} = 0$.
Une matrice carrée A est dite **triangulaire supérieure** (respectivement **triangulaire inférieure**) si pour tout $i > j$ (respectivement pour tout $i < j$) $a_{ij} = 0$.

Le plus souvent, nous seront amenés à étudier des matrices carrées lorsque nous ferons dans la décomposition en valeurs propres, mais ce n'est pas toujours le cas en réalité comme nous le verrons dans la partie *Analyse de Données*.

Exemple

Les matrices A, B, C et D suivantes sont respectivement des matrices carrées de type $(3, 3)$ quelconque, diagonale, triangulaire supérieure et triangulaire inférieure.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \pi \end{pmatrix},$$

$$C = \begin{pmatrix} \mathsf{e} & 3 & \ln(2) \\ 0 & 1 & 0 \\ 0 & 0 & \pi \end{pmatrix}, \quad D = \begin{pmatrix} \mathsf{e} & 0 & 0 \\ 1 & 2 & 0 \\ 0 & \gamma & 0.12 \end{pmatrix}.$$

La matrice diagonale ne comportant que des 1 est appelée **matrice identitée**, elle est notée I_n.

Structure de l'espace des matrices

La définition suivante montre que l'espace des matrices a une structure bien particulière que l'on a déjà rencontré.

Définition 2.17: Structure de l'espace des matrices

On note $\mathscr{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n, p) à coefficients dans $\mathbb{K}$. L'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$ est quant à lui noté $\mathscr{M}_N(\mathbb{K})$.

On en déduit même que l'ensemble des matrices d'ordre (n, p) muni de la multiplication externe et d'une loi additive interne à une structure d'espace vectoriel.
Cet espace est de dimension $(n \times p)$.

Propriétés du produit matriciel

Proposition 2.14: Propriétés

- Le produit matriciel est associatif, *i.e.* pour toute matrice A, B et C telles que $A \in \mathscr{M}_{m,n}(\mathbb{K})$, $B \in \mathscr{M}_{n,p}(\mathbb{K})$ et $C \in \mathscr{M}_{p,q}(\mathbb{K})$, alors

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

- En général, le produit matriciel n'est pas commutatif, *i.e.* si on considère $A \in \mathscr{M}_n(\mathbb{K})$, $B \in \mathscr{M}_n(\mathbb{K})$, on a souvent

$$AB = BA$$

On prendra garde à la dimension des matrices lors de la multiplication, il faut vérifier que dimensions concordent ! ! !

Transposition

Définition 2.18: Transposition d'une matrice

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$, on appelle **transposée de la matrice** A, notée A^T, la matrice A' de $\mathscr{M}_{p,n}(\mathbb{K})$ définie pour tout $(i,j) \in [\![1,p]\!] \times [\![1,n]\!]$ par $a'_{ij} = a_{ji}$:

$$\text{si } A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} \quad \text{alors} \quad A^T = \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \vdots \\ a_{1p} & \dots & a_{np} \end{pmatrix}$$

Proposition 2.15: Propriétés transposition

Soient A, B deux matrices de $\mathscr{M}_{m,n}(\mathbb{K})$, C une matrice de $\mathscr{M}_{n,p}(\mathbb{K})$, et soit $\lambda \in \mathbb{K}$ un scalaire, alors :

i) $(\lambda A)^T = \lambda A^T$,

ii) $(A+B)^T = A^T + B^T$,

iii) $(AC)^T = C^T A^T$.

Matrices symétriques

Définition 2.19: Matrices symétriques et anti-symétriques

Soit $A \in \mathscr{M}_n(\mathbb{K})$, alors :

i) A est dite symétrique si $A^T = A$,

ii) A est dite anti-symétrique si $A^T = -A$

En général on note $S_n(\mathbb{K})$ l'ensemble des matrices symétriques d'ordre n sur le corps K et $A_n(\mathbb{K})$ l'ensemble des matrices anti-symétriques d'ordre n sur le corps $\mathbb{K}$. On pourra même montrer que ces deux ensembles forment des sous-espaces vectoriels supplémentaires de $\mathscr{M}_n(\mathbb{K})$, *i.e.* toute matrice carrée peut s'écrire comme la somme unique d'une matrice symétrique et d'une matrice anti-symétrique.

Exemples

Les matrices S et A suivantes sont respectivement des matrices symétriques et anti-symétriques d'ordre 4.

$$S = \begin{pmatrix} a & b & c & d \\ b & e & f & g \\ c & f & h & i \\ d & g & i & j \end{pmatrix} \quad \text{et} \quad A = \begin{pmatrix} 0 & -b & -c & -d \\ b & 0 & -f & -g \\ c & f & 0 & -i \\ d & g & i & 0 \end{pmatrix}.$$

Quid de la matrice nulle à votre avis (celle ne comportant que des 0) ? Noter que le fait d'être anti-symétrique impose nécessairement que la diagonale de la matrice soit nulle.

Représentation des applications linéaires

Après ces quelques rappels sur les matrices, nous allons maintenant pouvoir faire le lien entre les applications linéaires présentées aux sections précédentes et leur représentation matricielle.
Pour cela, nous allons considérer deux espaces vectoriels E et F tout deux de dimension finie p et n respectivement. Les espaces E et F seront également munis des bases $(\mathbf{e}_1, \ldots, \mathbf{e}_p)$ et $(\mathbf{f}_1, \ldots, \mathbf{f}_n)$ respectivement.
Enfin on désignera par u (non plus f comme dans les sections précédentes pour éviter les confusions) une application linéaire de E dans F.
Nous avons précédemment que u est entièrement déterminée par l'image des vecteurs de la base de E dans la base de F

Représentation matricielle

Définition 2.20: Représentation matricielle

Pour tout $j \in [\![1, p]\!]$, notons $(a_{ij})_{i=1}^{n}$ les coordonnées de $u(\mathbf{e}_j)$ dans la base de F, on a donc :

$$u(\mathbf{e}_j) = \sum_{i=1}^{n} a_{ij}\mathbf{f}_i.$$

La matrice $A = (a_{ij})_{i,j=1}^{n,p}$ obtenue est alors appelée **matrice de** u relativement aux bases de E et F. On la note en générale $A = \underset{\mathscr{B}_E, \mathscr{B}_F}{Mat}(u)$ ou encore $Mat(u)$ lorsque le contexte n'est pas ambiguë.

On peut représenter cette matrice de la façon suivante :

$$\underset{\mathscr{B}_E, \mathscr{B}_F}{Mat}(u) = \begin{array}{cc} \begin{array}{ccc} u(\mathbf{e}_1) & \cdots & u(\mathbf{e}_p) \end{array} & \\ \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix} & \begin{array}{c} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_n \end{array} \end{array}.$$

Ainsi la j-ème colonne représente l'image du vecteur $\mathbf{e}_j$ par u dans la base de F.

Interprétation

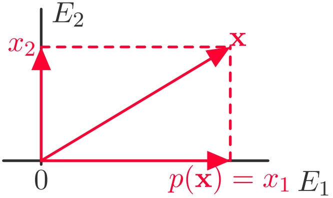
Derrière cette définition, il faut simplement comprendre que derrière chaque matrice se cache une application linéaire.

Regardons maintenant les matrices carrées de plus près, *i.e.* les matrices de $\mathscr{M}_n(\mathbb{K})$, *i.e.* les endomorphismes.

Nous avons vu que certains endomorphismes sont inversibles. Cette inversiblité peut également se caractériser d'un point de vue matriciel, *i.e.* elle implique une certaine propriété sur la matrice.

Exemple I

Soit la projection p qui consiste à conserver uniquement la première composante de ce vecteur, *i.e.* $p(x_1, x_2) = x_1$.



Cette projection peut être représentée par la matrice P suivante :

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

Exemple II

Considérons l'application linéaire u de $\mathbb{R}^3$ dans $\mathbb{R}^2$ définie par

$$f(\mathbf{x}) = (2x_1 + x_3, x_2 - x_3, x_1 - x_2 + 3x_3).$$

Pour déterminer sa représentation matricielle, on calcule l'image des vecteurs de base par l'application u

$$\begin{aligned} u(\mathbf{e}_1) &= (2, 0, 1), \\ u(\mathbf{e}_2) &= (0, 1, -1), \\ u(\mathbf{e}_3) &= (1, -1, 3). \end{aligned}$$

Ainsi la matrice de l'application linéaire u est $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 3 \end{pmatrix}$.

Inverse d'une matrice

Proposition 2.16: Matrice et inverse d'une application linéaire

Soit E un espace vectoriel sur le corps $\mathbb{K}$ de dimension finie $n \in \mathbb{N}^\star$ muni d'une base et soit $u \in \mathscr{L}(E)$. Notons A la représentation matricielle de u, alors u est inversible si et seulement si la matrice A associée est inversible.
De plus, si u est inversible on note $A^{-1} = Mat(u^{-1})$ la matrice inverse de A.

On conserve également les mêmes propriétés que pour les endomorphismes inversibles de E.

On rappellera par la suite comment calculer l'inverse d'une matrice carrée d'ordre n.

Retour famille de vecteurs

La représentation matricielle ne sert pas uniquement à représenter des applications linéaires, elle peut aussi être utilisée pour représenter une famille de vecteurs dans une base.

C'est d'ailleurs cette vision là que nous adoptons lorsque l'on souhaite représenter nos données sous forme de tableaux.

Lien entre base et inversibilité

Proposition 2.17: Base et inversibilité

Soit E un espace vectoriel de dimension $n \in N^\star$ muni d'une base $\mathscr{B}_E$ et considérons une famille de vecteurs $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de E, alors cette famille est une base de E si et seulement si la matrice associée à cette famille est inversible.

Ce résultat est important pour introduire la notion de changement de bases. En effet il est possible que l'on ne souhaite pas forcément travailler avec la base canonique, ce qui est très souvent le cas en analyse de données où l'on préfère regarder les données "sous un autre angle". Il faut alors voir comment faire pour passer d'une base à une autre.

Lien entre base et inversibilité

Définition 2.21: Changement de bases

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie $n \in \mathbb{N}^\star$, et soient $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $\mathscr{B}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ deux bases de E. On appelle **matrice de passage de $\mathscr{B}$ dans $\mathscr{B}'$ ou de $\mathscr{B}$ vers $\mathscr{B}'$** la matrice $P_{\mathscr{B},\mathscr{B}'}$ définie par :

$$P_{\mathscr{B},\mathscr{B}'} = \underset{\mathscr{B}}{Mat}\,(\mathscr{B}') = \underset{\mathscr{B}}{Mat}\,(\mathbf{e}'_1, \ldots, \mathbf{e}'_n).$$

Pour dire les choses plus simplement, les colonnes de la matrice de changement de bases de $\mathscr{B}$ vers $\mathscr{B}'$ sont formées par les coordonnées des vecteurs de $\mathscr{B}'$ dans la base $\mathscr{B}$.

Exemple

On va considérer l'espace vectoriel $E = \mathbb{R}^2$ muni de deux bases différentes :

$$\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2) = ((1,1),(-1,1)) \quad \text{et} \quad \mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2) = ((1,0),(3,1))$$

Pour déterminer la matrice de changement de base, il faut alors exprimer les vecteurs de la base $\mathscr{B}'$ dans $\mathscr{B}$.
On montre alors que la matrice de changement de base est définie par :

$$P_{\mathscr{B},\mathscr{B}'} = \begin{array}{c c} \begin{array}{cc} \mathbf{e}'_1 & \mathbf{e}'_2 \end{array} & \\ \begin{pmatrix} 0.5 & 2 \\ -0.5 & -1 \end{pmatrix} & \begin{array}{c} \mathbf{e}_1 \\ \mathbf{e}_2 \end{array} \end{array}.$$

Pour $\mathbf{e}'_2$ on a bien $\mathbf{e}'_2 = 2\mathbf{e}_1 - \mathbf{e}_2$. On peut ici trouver les coefficients de tête, mais nous verrons plus tard comment faire cela en résolvant ce que l'on appelle des *systèmes linéaires*.

Changement de base pour un vecteur

Proposition 2.18: Changement de base pour un vecteur

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie $n \in \mathbb{N}^\star$. Soient $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et $\mathscr{B}' = (\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ deux bases de E, et soit $P = P_{\mathscr{B},\mathscr{B}'}$.

Considérons un vecteur $\mathbf{x} \in E$, on peut alors écrire

$$\mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{e}_k \quad \text{et} \quad \mathbf{x} = \sum_{k=1}^{n} x'_k \mathbf{e}'_k.$$

Notons alors $\mathbf{x}_{\mathscr{B}} = (x_1, \ldots, x_n)$ et $\mathbf{x}_{\mathscr{B}'} = (x'_1, \ldots, x'_n)$, alors

$$\mathbf{x}_{\mathscr{B}} = P\mathbf{x}_{\mathscr{B}'}.$$

Illustration I

Considérons deux bases $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2)$ et $domB' = (\mathbf{e}'_1, \mathbf{e}'_2)$ d'un espace vectoriel E de dimension 2 et un élément $\mathbf{x}$ de E dont les coordonnées sont respectivement notées $\mathbf{x}_{\mathscr{B}} = (x_1, x_2)$ et $\mathbf{x}_{\mathscr{B}'} = (x'_1, x'_2)$ les coordonnées du vecteur $\mathbf{x}$ dans les bases $\mathscr{B}$ et $\mathscr{B}'$.

Supposons que l'on a également les relations suivantes entre les vecteurs des deux bases

$$\begin{aligned} \mathbf{e}'_1 &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2, \\ \mathbf{e}'_2 &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2. \end{aligned}$$

On va maintenant chercher à trouver notre matrice de passage P de $\mathscr{B}$ vers $\mathscr{B}'$. **Pour cela, on prend l'expression de notre vecteur $\mathbf{x}$ exprimée dans la nouvelle base $\mathscr{B}'$ et on va chercher ces coordonnées dans l'ancienne base $\mathscr{B}$.**

Illustration II

$$
\begin{aligned}
\mathbf{x} &= x_1' \underbrace{\mathbf{e}_1'} + x_2' \underbrace{\mathbf{e}_2'}, \\
&\downarrow \text{ définition de } \mathbf{e}_1' \text{ et } \mathbf{e}_2' \\
&x_1' a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + x_2' a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2, \\
&\downarrow \text{ on factorise} \\
&(a_{11}x_1' + a_{12}x_2')\mathbf{e}_1 + ((a_{21}x_1' + a_{22}x_2')\mathbf{e}_2, \\
&\downarrow \text{ définition de } \mathbf{x} \text{ dans la base } (\mathbf{e}_1, \mathbf{e}_2) \\
&x_1\mathbf{e}_1 + x_2\mathbf{e}_2.
\end{aligned}
$$

En étudiant les deux dernières égalités, nous aboutissons aux relations suivantes :

Illustration III

$$
\begin{aligned}
x_1 &= a_{11}x'_1 + a_{12}x'_2, \\
x_2 &= a_{21}x'_1 + a_{22}x'_2.
\end{aligned}
$$

Ainsi, nous avons la relation $x_{\mathscr{B}} = P x_{\mathscr{B}'}$ où $P = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$.

Changement de base d'une application linéaire

Tout comme il existe une relation permettant le changement de base d'un vecteur, on peut aussi le faire pour une application linéaire. En effet sa représentation est identique à celle d'une famille de vecteurs.

Proposition 2.19: Changement de base pour une application linéaire

Soient E et F des espaces vectoriels sur $\mathbb{K}$ de dimensions finies p et n. Soient $\mathscr{B}_E = (\mathbf{e}_1, \ldots, \mathbf{e}_p)$ et $\mathscr{B}'_E = (\mathbf{e}'_1, \ldots, \mathbf{e}'_p)$ deux bases de E et soit P la matrice de passage de $\mathscr{B}_E$ dans $\mathscr{B}'_E$. De même, soient $\mathscr{B}_F = (\mathbf{f}_1, \ldots, \mathbf{f}_p)$ et $\mathscr{B}'_F = (\mathbf{f}'_1, \ldots, \mathbf{f}'_p)$ deux bases de F et soit Q la matrice de passage de $\mathscr{B}_F$ dans $\mathscr{B}'_F$. Soit $u \in \mathscr{L}(E, F)$ et notons $A = \underset{\mathscr{B}_E, \mathscr{B}_F}{Mat}(u)$ et $A' = \underset{\mathscr{B}'_E, \mathscr{B}'_F}{Mat}(u)$. Alors

$$A' = Q^{-1}AP.$$

Preuve pour comprendre

Démonstration : Soient $\mathbf{x}$ et $\mathbf{x}'$ les représentations d'un vecteur de E dans les bases $\mathscr{B}_E$ et $\mathscr{B}'_E$ et soient $\mathbf{y}$ et $\mathbf{y}'$ les représentations d'un vecteur de E dans les bases $\mathscr{B}_F$ et $\mathscr{B}'_F$.
A est l'unique matrice vérifiant $A\mathbf{x} = \mathbf{y}$ pour tout vecteur $\mathbf{x}, \mathbf{y}$ de $E \times F$.
A' est l'unique matrice vérifiant $A\mathbf{x}' = \mathbf{y}'$ pour tout vecteur $\mathbf{x}, \mathbf{y}$ de $E' \times F'$.
De plus, nous avons les relations $\mathbf{x} = P\mathbf{x}'$ et $\mathbf{y} = Q\mathbf{y}'$.
Donc

$$A\mathbf{x} = \mathbf{y} \iff AP\mathbf{x}' = Q\mathbf{y}' \iff Q^{-1}AP\mathbf{x}' = \mathbf{y}',$$

ainsi $A' = Q^{-1}AP$.

Pour un endomorphisme, on aura la relation $A' = P^{-1}AP$.

Exemple I

Soit u un endomorphisme de $E = \mathbb{R}^2$ dont la matrice associée, notée A, dans la base canonique est donnée par

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

On reconnaît ici l'endomorphisme u qui pour tout $\mathbf{x} \in \mathbb{R}^2$ est défini par

$$u(\mathbf{x}) = (2x_1 + x_2, -x_1 + 3x_3).$$

On souhaite maintenant exprimer cet endomorphisme dans la base formé par les vecteurs $\mathbf{v}_1 = (-1, 1)$ et $\mathbf{v}_2 = (1, 1)$. Dans un premier temps, il nous faut déterminer la matrice de passage P de la base canonique vers la nouvelle base définie par $\mathbf{v}_1$ et $\mathbf{v}_2$.

Exemple II

On rappelle que la matrice de passage P est obtenue en exprimant les vecteurs de la nouvelle base $\mathscr{B}'$en fonction des vecteurs de l'ancienne base $\mathscr{B}$, cette représentation se faisant en colonne :

$$P_{\mathscr{B},\mathscr{B}'} = \begin{array}{c c} \begin{array}{cc} \mathbf{v}_1 & \mathbf{v}_2 \end{array} & \\ \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} & \begin{array}{c} \mathbf{e}_1 \\ \mathbf{e}_2 \end{array} \end{array}.$$

Cette matrice est inversible et son inverse P^{-1} est donnée par

$$P^{-1} = \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}.$$

Il reste alors à calculer le produit $P^{-1}AP$, ce qui nous donne

$$P^{-1}AP = \frac{1}{2}\begin{pmatrix} -1 & 11 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 5 & -1 \\ 3 & 5 \end{pmatrix}$$

Retour sur le rang I

Tout comme nous avons défini le rang d'une application linéaire, on peut également définir le rang d'une matrice.

Définition 2.22: Rang d'une matrice

Soit A une matrice de $\mathscr{M}_{n,p}(\mathbb{K})$, on appelle **rang de** A, noté $rg(A)$, le rang de la famille de vecteurs colonnes de A qui sont des vecteurs de $\mathbb{K}^n$.

Retour sur le rang II

Proposition 2.20: Rang et base

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n muni d'une base $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ et soit $(\mathbf{x}_1, \ldots, \mathbf{x}_p)$ une famille de vecteurs de E, alors :

$$rg(\mathbf{x}_1, \ldots, \mathbf{x}_n) = rg(\underset{\mathscr{B}}{Mat}\,(\mathbf{x}_1, \ldots, \mathbf{x}_n)).$$

De ce résultat on en déduit immédiatement que toute matrice carrée A d'ordre n est inversible si seulement son rang est égale à n. De plus, ce que l'on a vu sur les familles des vecteurs restent valables sur les matrices d'applications linéaires.

Retour sur le rang III

Proposition 2.21: Rang matrice d'une application linéaire

Soient E et F deux espaces vectoriels sur $\mathbb{K}$ de dimensions finies p et n munis de bases $\mathscr{B}_E$ et $\mathscr{B}_F$. Soit u une application linéaire de E dans F dont $A = \underset{\mathscr{B}_E,\mathscr{B}_F}{Mat}(u)$ est la représentation matricielle.

Alors $rg(u) = rg(A)$.

Proposition 2.22: Propriété du rang

Soit $A \in \mathscr{M}_{n,p}(\mathbb{K})$, alors $rg(A^T) = rg(A)$. Le rang d'une matrice est donc invariant par transposition.

Retour sur déterminant I

Une définition à ne pas retenir

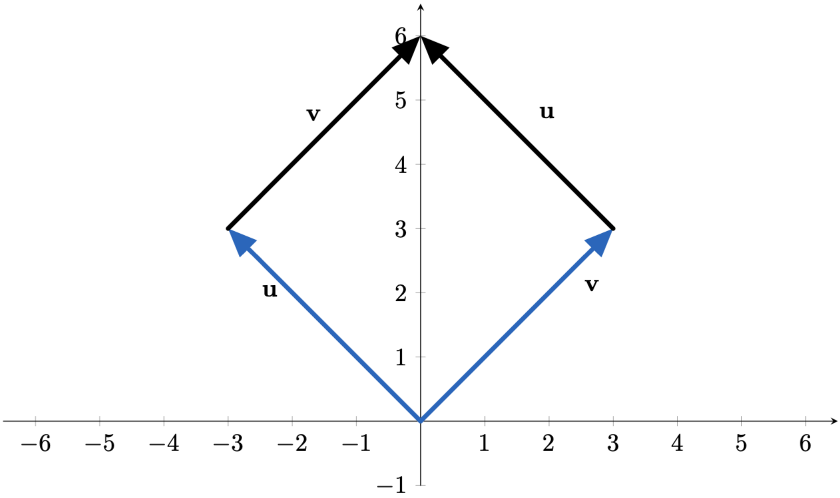
Définition 2.23: Déterminant

Soit A une matrice de $\mathscr{M}_n(\mathbb{R})$ dont les éléments sont notés $(a_{ij})_{i,j=1}^n$.
Soit $\mathfrak{S}_n$ le groupe des permutations de $[\![1, n]\!]$, c'est-à-dire l'ensemble des bijections de $[\![1, n]\!]$ dans $[\![1, n]\!]$ (comme une transposition qui échange de place deux éléments i et j).
On définit le **déterminant** de la matrice A par la relation

$$det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{\sigma(i),i},$$

où $\varepsilon(\sigma)$ est appelé **signature de la permutation** et qui vaut ± 1.

Retour sur déterminant II



Retour sur déterminant III

Dans ce premier exemple la famille de vecteurs $(\mathbf{u}, \mathbf{v})$ est définie par $\mathbf{u} = (-3, 3)$ et $\mathbf{v} = (3, 3)$ ce qui génère un carré avec une surface de 18. Nous aurions donc :

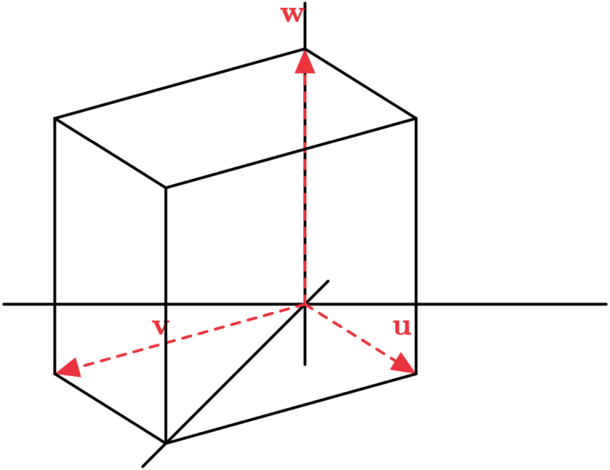
$$det(\mathbf{u}, \mathbf{v}) = \begin{vmatrix} 3 & -3 \\ 3 & 3 \end{vmatrix} = 18.$$

Considérons maintenant les vecteurs $\mathbf{u}, \mathbf{v}, \mathbf{w}$ qui sont définis par

$$\mathbf{u} = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix} \quad \text{et} \quad \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ \sqrt{18} \end{pmatrix}.$$

Ce qui permet de générer le parallélépipède suivant dans un espace de dimension 3, dont le volume qui n'est rien d'autre que le déterminant de la familles des 3 vecteurs, est égal à $18\sqrt{18} = 54\sqrt{2}$.

Retour sur déterminant IV



Retour sur déterminant V

En effet :

$$det(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \begin{vmatrix} 3 & -3 & 0 \\ 3 & 3 & 0 \\ 0 & 0 & \sqrt{18} \end{vmatrix} = 18\sqrt{18}.$$

Nous verrons un peu plus loin comment calculer les déterminants de telles familles de vecteurs. Mais avant cela, regardons quelques propriétés du déterminant.

Retour sur déterminant VI

Proposition 2.23: Caractérisation automorphisme

Soit $u \in \mathscr{L}(E)$, alors u est un automorphisme, *i.e.* un endomorphisme bijectif si et seulement si $det(u) \neq 0$.

Proposition 2.24: Inversibilité d'une matrice

Soit $A \in \mathscr{M}_n(\mathbb{K})$ alors A est inversible si et seulement si $det(A) \neq 0$.

Remarque. Le déterminant n'a de sens que pour les matrices carrées *i.e.* pour les endomorphismes d'espaces vectoriels.

Retour sur déterminant VII

Proposition 2.25: Propriétés du déterminant

oient A et B deux matrices de $\mathscr{M}_n(\mathbb{K})$ alors

$$det(A^T) = det(A) \quad \text{et} \quad det(AB) = det(A)det(B).$$

Si de plus A est inversible, *i.e.* si son déterminant est non nul alors

$$det(A^{-1}) = \frac{1}{det(A)}.$$

En revanche il n'y a pas de propriétés particulières concernant le déterminant de la somme de deux matrices ! En général

$$det(A+B) \neq det(A) + det(B).$$

Inverse d'une matrice I

Il nous reste à définir ce que l'on appelle la comatrice afin de pouvoir fournir une expression de l'inverse d'une matrice $A \in \mathscr{M}_n(\mathbb{K})$.

Définition 2.24: Comatrice

Soit $A = (a_{ij})_{i,j=1}^n \in \mathscr{M}_n(\mathbb{K})$, on appelle **comatrice de** A, notée $Com(A)$, la matrices des cofacteurs de A, c'est-à-dire $Com(A) = (A_{ij})_{i,j=1}^n$.

Un cofacteur A_{ij} est défini comme la déterminant de la matrice A à laquelle on aura supprimé la i-ème ligne et la j-ème colonne.

On verra des exemples concrets à la fin de cette section.

Inverse d'une matrice II

Proposition 2.26: Inverse d'une matrice

Soit $A \in \mathscr{M}_n(\mathbb{K})$ et supposons que A est inversible, alors $Com(A)^T A = A^T Com(A) = det(A) I_n$.

On peut également réécrire le résultat de cette proposition comme suit : si A est inversible alors

$$A^{-1} = \frac{1}{det(A)} Com(A)^T.$$

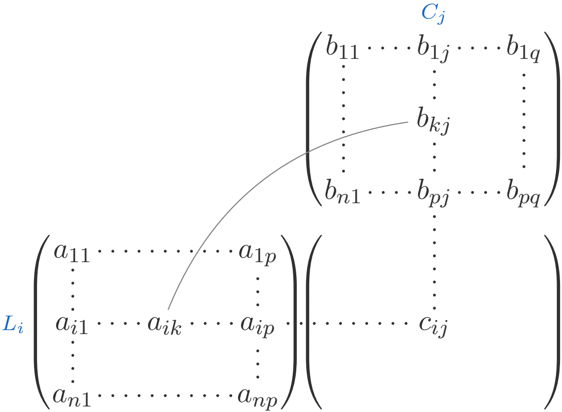
Com(A) est appelée matrice des cofacteurs

Quelques calculs explicites I

C'est le moment de mettre un peu de concret sur ce qui é été développé tout au long de cette section. Nous allons voir comment déterminer le rang, calculer le déterminant et l'inverse d'une matrice en appliquant les définitions, mais aussi en en utilisant quelques petites "*astuces*" pratiques. Avant de faire cela, on rappelle rapidement comment effectuer un produit matriciel.

Quelques calculs explicites II

Schéma illustrant le calcul matriciel $c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$.



Quelques calculs explicites III

Soient $A \in \mathscr{M}_{3,4}(\mathbb{R}),\ B \in \mathscr{M}_{4,2}(\mathbb{R})$ et $\mathbf{x} \in \mathbb{R}^4$ définis par

$$A = \begin{pmatrix} 2 & -1 & 0 & 4 \\ -1 & 0 & 3 & -2 \\ 5 & 0 & 0 & -4 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 4 & 3 \\ 0 & -2 \\ 1 & -1 \end{pmatrix} \quad \text{et} \quad \mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix}.$$

Alors le produit de la matrice A avec le vecteur $\mathbf{x}$ (qui revient donc à calculer l'image du vecteur $\mathbf{x}$ par l'application linéaire associée à la matrice A) est

$$A\mathbf{x} = \begin{pmatrix} 2 & -1 & 0 & 4 \\ -1 & 0 & 3 & -2 \\ 5 & 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \times 1 - 1 \times (-1) + 0 \times 0 + 4 \times 2 \\ \dots \\ \dots \end{pmatrix}$$

Quelques calculs explicites IV

Alors le produit $C = AB$ est

$$C = AB = \begin{pmatrix} 2 & -1 & 0 & 4 \\ -1 & 0 & 3 & -2 \\ 5 & 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 4 & 3 \\ 0 & -2 \\ 1 & -1 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \times 0 - 1 \times 4 + 0 \times 0 + 4 \times 1 & \ldots \\ \ldots & \ldots \\ \ldots & \ldots \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -5 \\ -2 & -5 \\ -4 & 9 \end{pmatrix}.$$

Matrice echelonnée (réduite) I

Définition 2.25: Matrice échelonnée

Une matrice est dite **échelonnée** en lignes si son nombre de zéros précédant la première valeur non d'une ligne augmente ligne par ligne jusqu'à l'éventuelle obtention de lignes ne comportant que des zéros.

Définition 2.26: Matrice échelonnée réduite

Une matrice échelonnée est dite **réduite** si elle est échelonnée et si les premières non nulles de chaque ligne sont égales à 1. De telles valeurs sont appelées des **pivots**.

Matrice echelonnée (réduite) II

Les matrices A, B et C suivantes sont respectivement échelonnée, échelonnée réduite et non échelonnée.

$$A = \begin{pmatrix} 0 & 4 & 3 & 2 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} 1 & 4 & -3 & 5 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ et } C = \begin{pmatrix} 2 & 4 & -3 & 5 \\ 4 & 1 & -2 & 7 \\ 6 & 2 & 4 & 1 \\ 8 & -5 & -2 & 1 \end{pmatrix}.$$

Calcul rang I

Pour déterminer le rang d'une matrice, il n'y a rien de plus simple. Il suffit d'effectuer des opérations élémentaires sur les lignes d'une matrice (ou sur les colonnes) afin d'obtenir une matrice échelonnée ou échelonnée réduite. Le rang de la matrice est alors directement égal au nombre de lignes non nulles de la matrice.
Reprenons les matrices de l'exemple précédent

$$A = \begin{pmatrix} 0 & 4 & 3 & 2 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 4 & -3 & 5 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ et } C = \begin{pmatrix} 2 & 4 & -3 & 5 \\ 4 & 1 & -2 & 7 \\ 6 & 2 & 4 & 1 \\ 8 & -5 & -2 & 1 \end{pmatrix}.$$

On voit directement que les matrices A et B sont de rang 3 et 4.

Calcul rang II

Le rang est cependant moins évident en ce qui concerne la matrice C, on va donc effectuer des opérations élémentaires sur les lignes pour déterminer son rang.

$$C = \begin{pmatrix} 2 & 4 & -3 & 5 \\ 4 & 1 & -2 & 7 \\ 6 & 2 & 4 & 1 \\ 8 & -5 & -2 & 1 \end{pmatrix}$$

$\downarrow$ On va se servir de la valeur 2 en

$\downarrow$ position $(1,1)$ pour annuler les autres lignes

$$= \begin{pmatrix} 2 & 4 & -3 & 5 \\ 0 & -7 & 4 & -3 \\ 0 & -10 & 13 & -18 \\ 0 & -21 & 10 & -19 \end{pmatrix} \begin{array}{l} \\ L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 - 3L_1 \\ L_4 \leftarrow L_4 - 4L_1 \end{array}$$

Calcul rang III

$\downarrow$ on se sert ensuite du -7 en position $(2,2)$ pour

$\downarrow$ faire apparaître des 0 dans les deux dernières lignes

$$= \begin{pmatrix} 2 & 4 & -3 & 5 \\ 0 & -7 & 4 & -3 \\ 0 & 0 & 51 & -96 \\ 0 & 0 & -2 & -10 \end{pmatrix},$$

où $L_3 \leftarrow 7L_3 - 10L_2$ et $L_4 \leftarrow L_4 - 3L_2$.

On peut remarquer que les deux dernières lignes sont indépendantes, la matrice C est donc de rang 4.
Cet exemple montre bien l'intérêt d'utiliser des matrices échelonnées réduites qui permettent de grandement simplifier les calculs.

Calcul déterminant I

Le calcul du déterminant est en général très complexe sauf pour des matrices très particulières. En revanche, ce calcul est très simple dans des espaces à faibles dimensions, comme en dimension 2 et 3 où des formules nous permettent de calculer très facilement le déterminant.

- **En dimension 2 :** le déterminant d'une matrice $A = (a_{ij})_{i,j=1}^{2}$ est donné par :

$$det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Calcul déterminant II

- **En dimension 3 :** le déterminant d'une matrice $A = (a_{ij})_{i,j=1}^3$ est donné par :

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{13}a_{31} + a_{21}a_{32}a_{13}$$
$$- a_{11}a_{13}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Cette règle s'appelle la *règle de Sarrus*.

Calcul déterminant III

- **Matrice diagonale :** soit $A \in \mathscr{M}_n(\mathbb{K})$ une matrice diagonale, *i.e.* $a_{ij} = 0$ pour tout $i \neq j$, alors $det(A) = \prod_{i=1}^{n} a_{ii}$.
  On peut aussi le voir comme :

$$det(A) = \begin{vmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22}\ldots a_{nn} = \prod_{i=1}^{n} a_{ii}.$$

De ce résultat, on voit tout de suite qu'une condition nécessaire et suffisante pour qu'une matrice diagonale soit inversible est que ses éléments diagonaux soient non nuls.

Calcul déterminant IV

- **Matrice triangulaire** soit $A \in \mathscr{M}_n(\mathbb{K})$ une matrice triangulaire supérieure (le résultat reste identique dans le cas des matrices triangulaires supérieures) , *i.e.* $a_{ij} = 0$ pour tout $i > j$, alors $det(A) = \prod_{i=1}^{n} a_{ii}$.
  On peut aussi le voir comme :

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_{(n-1)1} \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22}\dots a_{nn} = \prod_{i=1}^{n} a_{ii}.$$

A nouveau, une matrice triangulaire est inversible si et seulement si ses éléments diagonaux sont non nuls.

Calcul déterminant V

On se propose de calculer le déterminant de la matrice A définie par

$$A = \begin{pmatrix} 3 & 1 & -2 & 4 \\ 0 & 0 & 5 & 1 \\ 0 & 2 & 0 & 0 \\ 3 & 6 & 0 & 1 \end{pmatrix}$$

à l'aide de transformation sur les colonnes et en développant sur les colonnes (ou les lignes). Nous avons donc

$$det(A) = \begin{vmatrix} 3 & 1 & -2 & 4 \\ 0 & 0 & 5 & 1 \\ 0 & 2 & 0 & 0 \\ 3 & 6 & 0 & 1 \end{vmatrix},$$

$\downarrow$ le déterminant reste inchangé par combinaison linéaires de lignes

Calcul déterminant VI

$$= \begin{vmatrix} 3 & 1 & -2 & 4 \\ 0 & 0 & 5 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 5 & 2 & -3 \end{vmatrix}_{L_4 \leftarrow L_4 - L_1},$$

$\downarrow$ on développe selon la première colonne

$$= 3 \begin{vmatrix} 0 & 5 & 1 \\ 2 & 0 & 0 \\ 5 & 2 & -3 \end{vmatrix}$$

$\downarrow$ on développe selon la deuxième ligne

$$= 3 \times (-2) \begin{vmatrix} 5 & 1 \\ 2 & -3 \end{vmatrix},$$

$\downarrow$ on développe le déterminant de taille 2

$$= 3 \times (-2) \times (5 \times (-3) - 1 \times 2) = 102.$$

Calcul de l'inverse d'une matrice I

Une fois que l'on a calculé le déterminant, il reste à déterminer l'expression de la comatrice, *i.e.* la matrice des cofacteurs.

On rappelle que $A^{-1} = \frac{1}{det(A)} Com(A)^T$.

Pour une matrice $A \in \mathscr{M}_2(\mathbb{K})$ inversible nous avons directement :

$$\text{si } A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{alors} \quad A^{-1} = \frac{1}{det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11}. \end{pmatrix}.$$

Pour les matrices carrées de taille 2, la transposée de la matrice des cofacteurs consiste simplement à échanger de place les éléments diagonaux et à changer le signe des éléments anti-diagonaux.

Calcul de l'inverse d'une matrice II

Reprenons notre matrice A précédente.

$$A = \begin{pmatrix} 3 & 1 & -2 & 4 \\ 0 & 0 & 5 & 1 \\ 0 & 2 & 0 & 0 \\ 3 & 6 & 0 & 1 \end{pmatrix}$$

Comme nous avons déjà calculé son déterminant ($det(A) = 102$), il nous reste simplement à déterminer la matrice des cofacteurs A_{ij}. Une fois cette matrice obtenue, il ne faudra pas oublier de la transposer afin de déterminer l'inverse de la matrice A. On se contentera de calculer les valeurs de la première ligne de la comatrice.

Calcul de l'inverse d'une matrice III

- $A_{11} = (-1)^2 \Delta_{11} = \begin{vmatrix} 0 & 5 & 1 \\ 2 & 0 & 0 \\ 6 & 0 & 1 \end{vmatrix} = -2 \times \begin{vmatrix} 5 & 1 \\ 0 & 1 \end{vmatrix} = 10.$

  On aura développer selon la deuxième ligne pour calculer le déterminant de notre matrice carrée de taille 3.

- $A_{12} = (-1)^3 \Delta_{12} = -\begin{vmatrix} 0 & 5 & 1 \\ 0 & 0 & 0 \\ 3 & 0 & 1 \end{vmatrix} = 0.$

  C'est immédiat car notre matrice comprend une ligne avec des zéros uniquement, son déterminant est donc nul.

- $A_{13} = (-1)^4 \Delta_{13} = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 6 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ 3 & 6 \end{vmatrix} = -6.$

  On aura développer selon la première ligne pour obtenir notre nouveau déterminant de taille 2.

Calcul de l'inverse d'une matrice IV

- $A_{14} = (-1)^5 \Delta_{14} = - \begin{vmatrix} 0 & 0 & 5 \\ 0 & 2 & 0 \\ 3 & 6 & 0 \end{vmatrix} = -5 \times \begin{vmatrix} 0 & 2 \\ 3 & 6 \end{vmatrix} = 30.$

  On aura développé selon la première ligne à nouveau.

En calculant tous les coefficients, on montre alors que la matrice des cofacteurs est donnée par :

$$com(A) = \begin{pmatrix} -10 & 0 & -6 & 30 \\ -4 & 0 & 18 & 12 \\ -127 & 51 & -15 & 75 \\ 44 & 0 & 6 & -30 \end{pmatrix}.$$

Algèbre Linéaire

Systèmes linéaires

Algèbre Linéaire

Réduction des endomorphismes

Algèbre Linéaire

Formes quadratiques et espaces euclidiens

Analyses de Données

Généralités et Décomposition en Valeurs Singulières (SVD)

Analyses de Données

Analyse en Composantes Principales (ACP)

Analyses de Données

Généralisation des méthodes

Analyses de Données

Analyse Factorielle des Correspondances (AFC)

Analyses de Données

Analyse factorielle des Correspondances Multiples (ACM)

The End