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 2023

Algèbre Linéaire

Summary

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

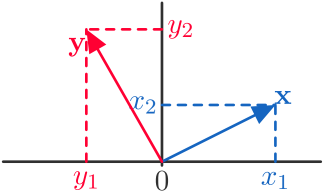
Espaces vectoriels et applications linéaires

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 1.1: 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

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

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

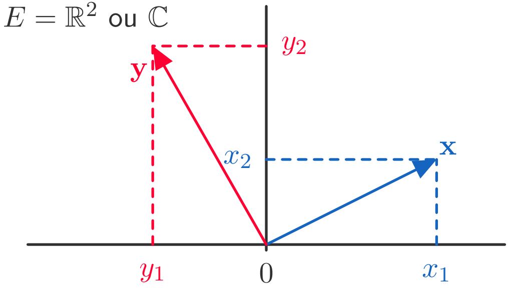
Un groupe abélien est défini par les points suivants :

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.* $\mathscr{M}_2(\mathbb{R})$:

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

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 1.2: 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 1.1

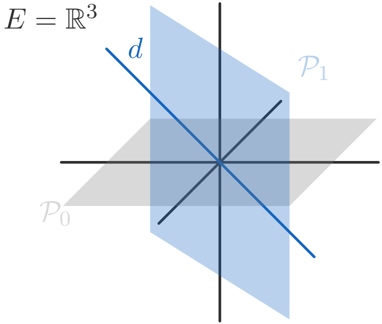
Soit E un $\mathbb{K}$-espace vectoriel et soit F un sous-ensemble de E. F est un **sous-espace** vectoriel de E *si et seulement si* les deux propriétés suivantes sont vraies :

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 1.2: 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 1.3: 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 :

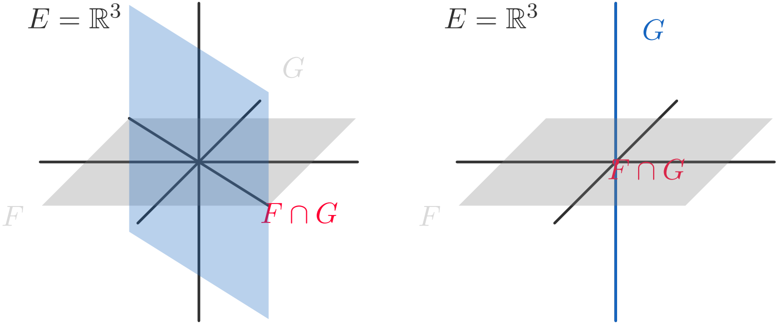
i) $F \cap G = \{0_E\}$,

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 1.3: 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 1.4: 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

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

Noyau et image d'une application linéaire

Définition 1.5: 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 1.6: 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 1.4: 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

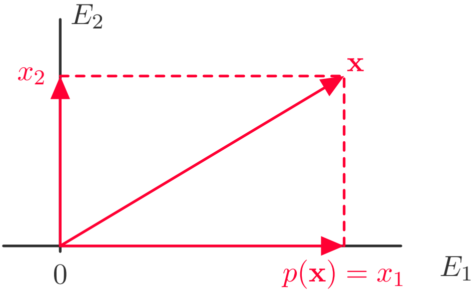
Définition 1.7: Projecteur

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

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

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

Proposition 1.5: 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 1.8: Familles libres, liées

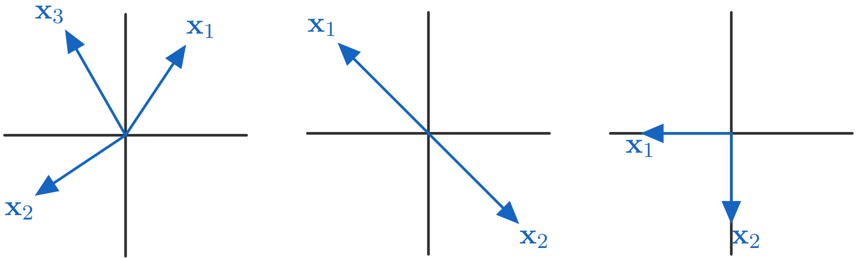
Soit $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$ une famille de vecteurs d'un espace vectoriel E. Cette famille est dite :

- **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 I

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.

Exemples II

On considère la famille de vecteurs $\mathbf{v}_1 = (1, 2, 1, 0)$ et $\mathbf{v}_2 = (-1, 0, -1, 1)$ de $\mathbb{R}^4$.
On montre que cette famille est bien une famille libre. En effet, considérons $\lambda_1, \lambda_2 \in \mathbb{R}$, alors

$$\lambda_1 \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 0 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Si on concentre sur la deuxième et la dernière équation, on trouve directement que $\lambda_1 = \lambda_2 = 0$.

Exemples III

On considère la famille de vecteurs $\mathbf{v}_1 = (1, 1)$ et $\mathbf{v}_2 = (1, -1)$. On On montre qu'il s'agit d'une famille libre. En effet, considérons $\lambda_1, \lambda_2 \in \mathbb{R}$, alors

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

On doit alors résoudre le système formé par les deux équations
$\lambda_1 + \lambda_2 = 0$ et $\lambda_1 - \lambda_2 = 0$.
La somme de ces deux équations implique $\lambda_1 = 0$ et la différence de ces deux équations implique $\lambda_2 = 0$.

Exemples IV

On considère la famille de vecteurs $\mathbf{v}_1 = (1, -1)$, $\mathbf{v}_2 = (3, -2)$ et $\mathbf{v}_3 = (0, 1)$ de $\mathbb{R}^2$.

Que peut-on dire de cette famille de vecteurs ?

Exemples V

On considère la famille de vecteurs $\mathbf{v}_1 = (1, -1),\ \mathbf{v}_2 = (3, -2)$ et $\mathbf{v}_3 = (0, 1)$ de $\mathbb{R}^2$.
Cette famille de vecteur n'est pas libre. En effet, on vérifie aisément que l'on

$$\mathbf{v}_2 = 3\mathbf{v}_1 + \mathbf{v}_3.$$

Nous verrons plus tard que nous aurions pu directement écrire que cette famille de vecteurs est liée sans effectuer de calculs. Pourquoi selon vous ?

Caractérisation des familles liées

Proposition 1.6: 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*).

Exemples I

Reprenons la famille de vecteurs $\mathbf{v}_1 = (1, 1)$ et $\mathbf{v}_2 = (1, -1)$ et montrons qu'ils s'agit d'une famille génératrice de $\mathbb{R}^2$.
Pour cela, on considère un vecteur quelconque $\mathbf{y} = (y_1, y_2)$ de $\mathbb{R}^2$ et on va montrer que ce vecteur peut s'écrire comme une combinaison des vecteurs de $\mathbf{v}_1$ et $\mathbf{v}_2$, *i.e.*

$$\mathbf{y} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2.$$

On cherche donc les valeurs α_1, α_2 telles ques

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

On montre qu'une solution est donnée par

$$\alpha_1 = \frac{y_1 + y_2}{2} \quad \text{et} \quad \alpha_2 = \frac{y_1 - y_2}{2}.$$

Exemples II

Que peut-on dire des familles de vecteurs suivantes ?

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ -4 \end{pmatrix}.$$

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Famille génératrice

Définition 1.9: Famille génératrice

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

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

Exemples I

On pourra par exemple montrer que la famille de vecteurs $\mathbf{v}_1 = (1, 0)$ et $\mathbf{v}_2 = (0, 1)$ est une famille de vecteurs génératrice de $\mathbb{R}^2$.
En effet, pour tout vecteur $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ on a

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Cet exemple est plutôt simple, on va en prendre un autre.

Exemples II

Considérons la famille de vecteurs $\mathbf{v}_1 = (1,1), \mathbf{v}_2 = (1,0)$ et $\mathbf{v}_3 = (-1,1)$.

On peut facilement voir qu'il s'agit d'une famille de vecteurs liée.
On peut aussi montrer qu'il s'agit d'une famille génératrice de $\mathbb{R}^2$.
Considérons à nouveau un vecteur quelconque $\mathbf{x} = (x_1, x2)$ de $\mathbb{R}^2$ et montrons qu'il peut s'exprimer comme une combinaison linéaire des vecteurs de notre famille. Par exemple :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_1 - x_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

ou encore

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 + x_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Exemples III

Nous aurions également pu écrire ce vecteur de la façon suivante :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{x_2 + x_1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{x_2 - x_1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

On a donc trois façons d'exprimer notre vecteur $\mathbf{x}$ à l'aide de vecteurs de cette famille !

Ce n'est donc pas très pratique dans l'absolu, il faut peut-être être plus exigeant sur cette famille de vecteurs afin de simplifier les choses.

Base

Définition 1.10: Base

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

Cette définition va nous servir à introduire la notion qui est coeur de cette section, la notion de *dimension* d'un espace vectoriel.

Exemples I

- 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 II

- 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

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

Exemples III

- 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 1.7: Caractérisation base

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

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

Remarquez bien la différence avec la définition de famille génératrice !
Quelle est cette différence à votre avis ?

Caractérisation des bases

Proposition 1.8: 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 :

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

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

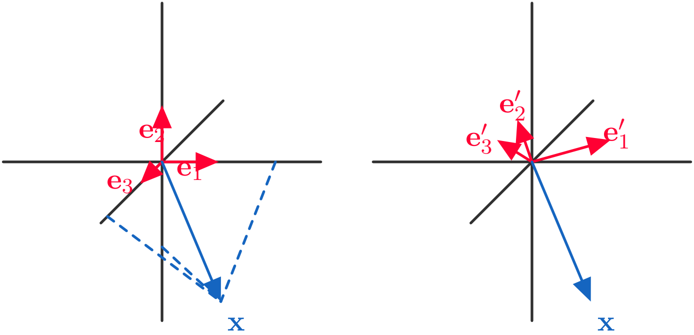
$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \ldots + x_n\mathbf{e}_n.$$

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 1.1: 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 1.11: 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 1.2: 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 1.9: 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 1.10: 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 1.11: 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 1.12: Hyperplan

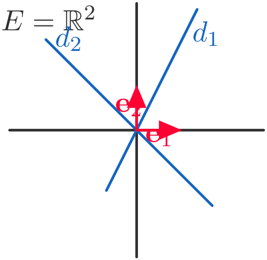
Soit E un espace vectoriel de dimension finie $n \geq 2$. On appelle *hyperplan* de E tout sous-espace vectoriel de E de dimension $n-1$.

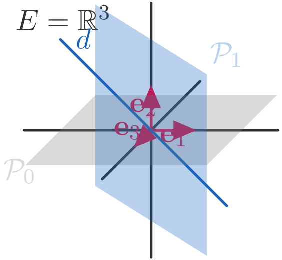
En tant que sous-espace vectoriel, les hyperplans doivent donc nécessairement contenir le vecteur nul, *i.e.* le vecteur 0_E.

On peut montrer que les hyperplans sont les noyaux d'une forme linéaire. Qu'est-ce qu'une forme linéaire et ... pourquoi selon vous ? On peut commencer à répondre à cette question mais on pourra conclure un peu plus tard.

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 1.13: 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 1.12: Propriétés rang d'une famille

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

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

Applications linéaires I

On considère un espace vectoriel E de dimension finie qui admet une base $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$, 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 $(f(\mathbf{x}_1), \ldots, f(\mathbf{x}_n))$ des vecteurs de bases $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de E.

Applications linéaires II

Réciproquement, pour $(\mathbf{y}_1, \ldots, \mathbf{y}_n)$ une famille de vecteurs de F, on définit l'application f de E dans F par : $\forall\, \mathbf{x} \in E$ tel que $\mathbf{x} = \sum_{k=1}^{n} \lambda_k \mathbf{x}_k$, $f(\mathbf{x}) = \sum_{k=1}^{n} \lambda_k \mathbf{y}_k$.

On a alors, $\forall k \in [\![1, n]\!],\ f(\mathbf{x}_k) = \mathbf{y}_k$. De plus f est une application linéaire de E dans F.

Ainsi, on peut définir une application linéaire f en fixant les images $\mathbf{y}_1, \ldots, \mathbf{y}_n$ des vecteurs de la base $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ de E.

Applications linéaires III

On peut résumer tout cela par la proposition suivante :

Proposition 1.13: Application et image d'une base

Soit E un espace vectoriel qui admet une base $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ et soit F un espace vectoriel.
Alors $\forall \ (\mathbf{y}_1, \ldots, \mathbf{y}_n) \in F^n, \ \exists! f \in \mathscr{L}(E, F)$ tel que $\forall k \in [\![1, n]\!], f(\mathbf{x}_k) = \mathbf{y}_k$.

Applications linéaires IV

On peut résumer cette proposition de la façon suivante :

- étant donnée une application linéaire $f \in \mathscr{L}(E, F)$, il suffit de connaître l'image des vecteurs de bases de E par cette application pour la déterminer entièrement.
- ou encore, si l'on connaît n vecteurs de F, il est possible de construire une (unique) application linéaire f de E dans F telle que pour tout $k \in [\![1, n]\!],\ f(\mathbf{x}_k) = \mathbf{y}_k$.

Après avoir vu comment construire une telle application, on va voir que l'on est aussi capable de caractériser les propriétés des morphismes en fonction de l'image des vecteurs de bases.

Caractérisation des applications linéaires

Proposition 1.14: 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 1.14: 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 1.3: Théorème du rang

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

Quelques remarques

- Sa démonstration est relativement compliquée (technique) on ne va donc pas la présenter
- C'est un outils très puissant pour déterminer rapidement la dimension du noyau (resp. de l'espace image) lorsque l'on connait la dimension de l'espace de départ ou de l'image d'une application linéaire (resp. de son noyau).
- Surtout ... on va pouvoir donner une caractérisation super simple des endomorphismes.

Retour sur les Hyperplans

Définition 1.15: 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.

On peut montrer que les hyperplans sont les noyaux d'une forme linéaire. Qu'est-ce qu'une forme linéaire et ... pourquoi selon vous ? On peut y répondre maintenant !

Caractérisation morphismes

Proposition 1.15: 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 1.16: 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 1.17: 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 I

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

Définition 1.18: 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 $(n \times p)$.

Structure de l'espace des matrices II

Quelques remarques anecdotiques

On peut même montrer que l'ensemble des matrices a donc une structure de **groupe abélien** pour la loi additive "$+$" (conséquence du fait d'être un espace vectoriel).
En revanche, il n'y pas de structure de groupe pour la multiplication matricielle ! Pourquoi ?

On peut en revanche montrer que l'ensemble des matrices à une structure d'**anneau unitaire** (ou **unifère**) mais non **intègre**.

- **unifère** : possède un seul et unique élément neutre pour la multiplication matricielle
- **non intègre** : possède des diviseurs de 0, *i.e.* il existe des matrices A et B non nulles telles que $AB = 0$.

Propriétés du produit matriciel

Proposition 1.16: 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 1.19: 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 1.17: 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 1.20: 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.
Quelle est la dimension de ces deux sous-espaces ?

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 1.21: 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

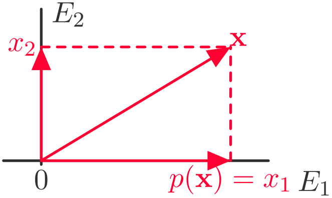
Derrière cette définition, il faut simplement comprendre que derrière chaque matrice se cache une application linéaire.

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 $\mathbb{R}^3$, munis de leur base canonique, 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

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

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 1.18: 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 1.19: 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 1.22: Changement de bases

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie $n \in \mathbb{N}^\star$, et soient $\mathscr{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et $\mathscr{B}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ 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 :

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

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 :

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

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 1.20: Changement de base pour un vecteur

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie $n \in \mathbb{N}^\star$. Soient $\mathscr{B} = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ et $\mathscr{B}' = (\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ 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 $\mathbf{x}_{\mathscr{B}} = (x_1, \dots, x_n)$ et $\mathbf{x}_{\mathscr{B}'} = (x'_1, \dots, x'_n)$, alors

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

Illustration I

Considérons deux bases $\mathscr{B} = (\mathbf{e}_1, \mathbf{e}_2)$ et $\mathscr{B}' = (\mathbf{e}'_1, \mathbf{e}'_2)$ 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 $x_{\mathscr{B}} = Px_{\mathscr{B}'}$ 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 1.21: 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

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

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 :

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

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 1.23: 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 1.22: Rang et base

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

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

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 1.23: 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 1.24: 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 1.24: Déterminant

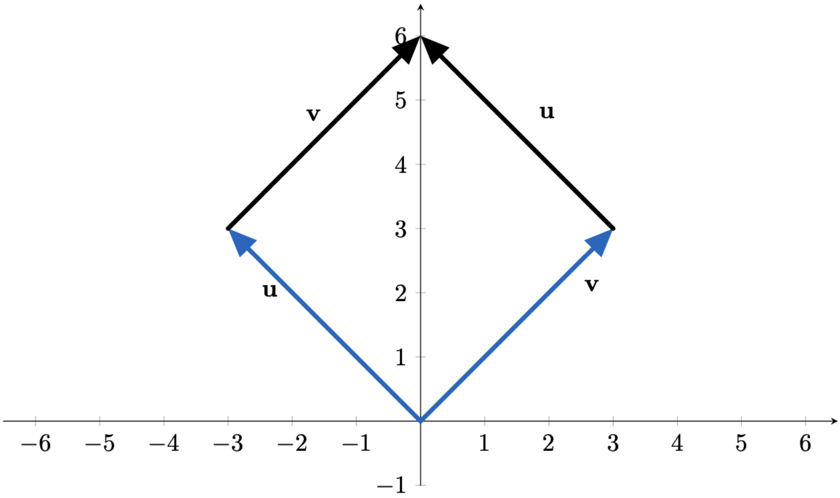
Soit A une matrice de $\mathscr{M}_n(\mathbb{R})$ dont les éléments sont notés $(a_{ij})_{i,j=1}^n$. Soit $\mathfrak{S}_n$ le groupe des permutations de $[\![1,n]\!]$, c'est-à-dire l'ensemble des bijections de $[\![1,n]\!]$ dans $[\![1,n]\!]$ (comme une transposition qui échange de place deux éléments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

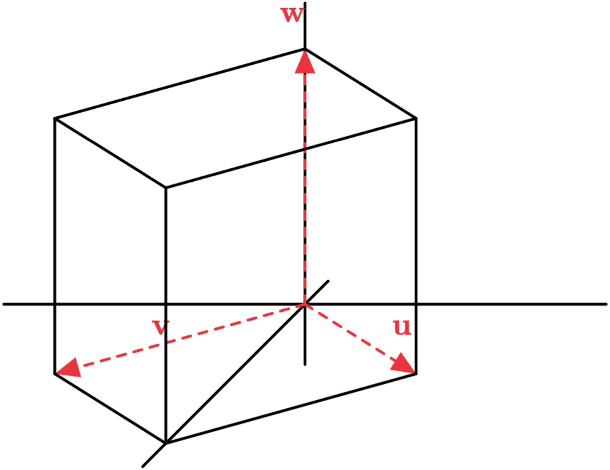
$$det(\mathbf{u}, \mathbf{v}) = \begin{vmatrix} 3 & -3 \\ 3 & 3 \end{vmatrix} = 18.$$

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 1.25: 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 1.26: 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 1.27: Propriétés du déterminant

Soient 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 1.25: Mineure et Cofacteur

Soit $A = (a_{ij})_{i,j=1}^{n} \in \mathscr{M}_n(\mathbb{K})$ alors pour tout couple (i, j)

- on appelle **mineur de A relatif à la i-ème ligne et à la j-ème colonne**, notée Δ_{ij}, le déterminant de la matrice obtenue en supprimant la i-ème ligne et la j-ème colonne de A,
- on appelle **cofacteur de A**, notée A_{ij}, le scalaire défini par $A_{ij} = (-1)^{i+j}\Delta_{ij}$.

Inverse d'une matrice II

Proposition 1.28: Déterminant et mineurs

Soit $A = (a_{ij})_{i,j=1}^{n} \in \mathscr{M}_n(\mathbb{K})$ alors pour tout $1 \leq j \leq n$, $det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \Delta_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij}$.

Inverse d'une matrice III

Définition 1.26: 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 IV

Proposition 1.29: 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 $c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$.

$$
\begin{array}{cc}
 & \begin{array}{c} C_j \\ \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1q} \\ \vdots & & \vdots & & \vdots \\ \vdots & & b_{kj} & & \vdots \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{pj} & \cdots & b_{pq} \end{pmatrix} \end{array} \\
\begin{array}{c} \\ \\ L_i \\ \\ \\ \end{array} \begin{pmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1p} \\ \vdots & & & & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{ip} \\ \vdots & & & & \vdots \\ a_{n1} & \cdots & \cdots & \cdots & a_{np} \end{pmatrix} & \begin{pmatrix} & & \vdots & & \\ & & \vdots & & \\ \cdots & \cdots & c_{ij} & & \\ & & & & \\ & & & & \end{pmatrix}
\end{array}
$$

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

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

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

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

Matrice echelonnée (réduite) I

Définition 1.27: 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 1.28: Matrice échelonnée réduite

Une matrice échelonnée est dite **réduite** si elle est échelonnée et si les premières valeurs 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 rang IV

Les calculs que nous avons effectué sur les lignes de la matrice peuvent également être effectués sur les colonnes de cette dernière !

C'est d'ailleurs le plus naturel ! En effet, on se rappelle que les colonnes d'une matrice d'un endomorphisme représente l'image des vecteurs de base par cet endomorphisme.
Ainsi, pour étudier l'espace image d'un endomorphisme u, noté $Im(u)$, on va alors regarder **le sous-espace engendré** par les colonnes de $Mat(u)$.

On verra un petit peu plus loin une petite "technique" qui permet aussi de calculer directement **une base de l'image et du noyau d'un endomorphisme u simultanément**.

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*.
Elle consiste à faire la *somme du produit des éléments diagonaux* **moins** *la somme du produit des éléments anti-diagonaux*.

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 :

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

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

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 :

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

Calcul de l'inverse d'une matrice V

Il nous restera alors à appliquer la définition

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

où $det(A)$ a été précédemment calculé.
Il faut bien admettre que de passer ainsi par cette définition est loin d'être pratique car elle implique de calculer un nombre important de déterminants ...

On va regarder comment calculer l'inverse d'une matrice de façon beaucoup plus simple en effectuant des opérations sur une matrice **étendue** !

Calcul de l'inverse d'une matrice VI

Il existe aussi une méthode plus pratique pour calculer l'inverse d'une matrice qui repose sur le fait que *les opérations élémentaires sur les lignes et les colonnes sont des automorphismes, ils ne changent donc pas le caractère inversible d'une matrice donnée*.

On peut se servir de cela pour trouver une suite d'opérations élémentaires qui va permettre de transformer une matrice $A \in \mathscr{M}_N(\mathbb{R})$ inversible en la matrice identité. Si on applique, en parallèle, les mêmes transformations à la matrice identité, on sera mesure de déterminer la matrice inverse de la matrice A.

Deux choix s'offrent à nous en terme d'écriture, mais ce choix va conditionner le travail à effectuer et inversement :

Calcul de l'inverse d'une matrice VII

- si on souhaite effectuer des opérations élémentaires sur **les lignes de la matrice de** A, on travaillera sur la *matrice étendue* suivante

$$(A \mid I_n) = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1n} & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & \cdots & 1 \end{array}\right),$$

afin d'aboutir, via **une succession de manipulations sur les lignes** à une matrice étendue de la forme

$$(I_n \mid A') = \left(\begin{array}{ccc|ccc} 1 & \cdots & 0 & a'_{11} & \cdots & a'_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & a'_{n1} & \cdots & a'_{nn} \end{array}\right),$$

où $A' = A^{-1}$ désigne l'inverse de la matrice A.

Calcul de l'inverse d'une matrice VIII

- si on souhaite effectuer des opérations élémentaires sur **les colonnes de la matrice de** A, on travaillera sur la *matrice étendue* suivante

$$\left(\frac{A}{I_n}\right) = \left(\begin{array}{ccc} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \\ \hline 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \end{array}\right).$$

afin d'aboutir, via **une succession de manipulations sur les colonnes** à une matrice étendue de la forme

Calcul de l'inverse d'une matrice IX

$$\left(\frac{I_n}{A'}\right) = \left(\begin{array}{ccc} 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \\ \hline a'_{11} & \cdots & a'_{1n} \\ \vdots & & \vdots \\ a'_{n1} & \cdots & a'_{nn} \end{array}\right),$$

où $A' = A^{-1}$ désigne l'inverse de la matrice A.
On va à nouveau regarder cela sur un exemple pour illustrer la méthode.

Calcul de l'inverse d'une matrice X

Soit $A \in \mathscr{M}_3(\mathbb{R})$ une matrice que l'on admettra inversible (vous pourrez calculer son déterminant afin de vous en convaincre), définie par :

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 0 & 3 & 5 \\ 0 & 0 & 1 \end{pmatrix}.$$

La matrice A est triangulaire supérieure dans cet exemple, on se propose de déterminer son inverse en effectuant des opérations élémentaires sur les lignes. On considère la matrice étendue suivante :

$$\left(\begin{array}{ccc|ccc} 5 & 2 & 2 & 1 & 0 & 0 \\ 0 & 3 & 5 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

Calcul de l'inverse d'une matrice XI

On commence par se servir de la valeur 1 en bas à droite de notre matrice à inverser.
On oubliera pas de répercuter les mêmes transformations sur la matrice identité !

$$\left(\begin{array}{ccc|ccc} 5 & 2 & 0 & 1 & 0 & -2 \\ 0 & 3 & 0 & 0 & 1 & -5 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} L_1 \leftarrow L_1 - 2L_3 \\ L_2 \leftarrow L_2 - 5L_3. \\ \end{array}$$

On va ensuite faire apparaître la valeur 1 dans la deuxième ligne et deuxième colonne. On garde à l'esprit que l'on doit faire apparaître la matrice identité à gauche et on itère ce même type d'opérations.

$$\left(\begin{array}{ccc|ccc} 5 & 2 & 0 & 1 & 0 & -2 \\ 0 & 1 & 0 & 0 & 1/3 & -5/3 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) L_2 \leftarrow L_2/3,$$

Calcul de l'inverse d'une matrice XII

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 1 & -2/3 & 4/3 \\ 0 & 1 & 0 & 0 & 1/3 & -5/3 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)^{L_1 \leftarrow L_1 - 2L_2},$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/5 & -2/15 & 4/15 \\ 0 & 1 & 0 & 0 & 1/3 & -5/3 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right)^{L_1 \leftarrow L_1/5}.$$

On a donc

$$A^{-1} = \begin{pmatrix} 1/5 & -2/15 & 4/15 \\ 0 & 1/3 & -5/3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Bonus : obtention bases noyau et image I

Nous avons vu précédemment que l'on pouvait déterminer une base du noyau de l'image d'une application linéaire $u : \mathbb{R}^n \mapsto \mathbb{R}^p$ par les méthodes suivantes :

- **Noyau :** on cherche les vecteurs $\mathbf{x} \in \mathbb{R}^n$ qui vérifient $u(\mathbf{x}) = \mathbf{0}$, ce qui nous amène à résoudre un système linéaire de n équations à p inconnues (objet de la prochaine section !)
- **Image :** on a vu que l'image d'une application linéaire u est l'espace engendré par les vecteurs colonnes de la matrice de représentation de u. Pour déterminer une base de $Im(u) \subset \mathbb{R}^p$, on va donc chercher une sous famille des n vecteurs de $\mathbb{R}^p$ qui soit libre et engendre l'espace $Im(u)$, *i.e.* la plus grande sous-famille libre des vecteurs colonnes de $Mat(u)$.

Bonus : obtention bases noyau et image II

Cela mène souvent à faire le travail en deux étapes ce qui peut se révéler fastidieux, mais on va voir que l'on peut tout faire d'un coup, sans résoudre de système et en considérant à nouveau une matrice étendue !

En revanche, **cela nécessitera de travailler obligatoire sur les colonnes de la matrice u ! Vu que ce sont ces dernières qui engendrent** $Im(u)$.

Regardons un exemple pour illustrer le principe et considérons

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 11 \end{pmatrix}.$$

On peut montrer que cette matrice n'est pas inversible (calculer son déterminant par exemple), que son noyau est de dimension 1 et que son

Bonus : obtention bases noyau et image III

rang (dimension de l'espace image) est dimension 2 (conséquence du théorème du rang).

On considère la matrice étendue

$$\left(\frac{A}{I_3}\right) = \left(\begin{array}{ccc} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 11 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right),$$

On va ensuite effectuer une succession d'opérations sur les colonnes, pour savoir si certaines colonnes la matrice supérieure peuvent s'annuler ou non.

Bonus : obtention bases noyau et image IV

$$\left(\begin{array}{ccc} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 3 & 7 & 11 \\ \hline 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \underset{\substack{C_2 \leftarrow C_2 - 3C_1 \\ C_3 \leftarrow C_3 - 5C_1}}{\sim} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & -2 & -4 \\ 3 & -2 & -4 \\ \hline 1 & -3 & -5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right) \underset{C_3 \leftarrow C_3 - 2C_2}{\sim} \left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & -2 & 0 \\ 3 & -2 & 0 \\ \hline 1 & -3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right).$$

Et on remarque que les deux premières colonnes de la matrice supérieure sont indépendantes, on va donc s'arrêter là.
In fine, on obtient la matrice étendue suivante

Bonus : obtention bases noyau et image V

$$\left(\begin{array}{ccc} 1 & 0 & 0 \\ 2 & -2 & 0 \\ 3 & -2 & 0 \\ \hline 1 & -3 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{array}\right)$$

qui nous permet de lire directement une base du noyau et une base de l'image de l'application linéaire étudiée.

Bonus : obtention bases noyau et image VI

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

- Rouge : on lit les vecteurs de la base du noyau de l'application linéaire, ici $Ker(u) = Vect(1, -2, 1)$
- Bleu : on lit les vecteurs de la base de l'image de l'application linéaire, ici $Im(u) = Vect\left((1, 2, 3), (0, -2, -2)\right) = Vect\left((1, 2, 3), (3, 4, 7)\right)$

D'ailleurs, les vecteurs de la base de l'image vérifie l'équation définie par le vecteur de la base du noyau $x_1 - 2x_2 + x_3 = 0$ (on étudie ici le noyau d'une forme linéaire !).

Bonus : obtention bases noyau et image VII

Dans l'exemple, nous avons étudié un **endomorphisme** donc une matrice **carrée**, mais la méthode reste valable pour des matrice rectangulaires, *i.e.* pour n'importe quelle application linéaire $u : \mathbb{R}^n \to \mathbb{R}^p$ il suffira de considérer la matrice étendue

$$\left(\frac{A}{I_n}\right) = \left(\begin{array}{ccc} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{p1} & \cdots & a_{pn} \\ \hline 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 1 \end{array}\right).$$

Pour finir

Dans cette section nous avons

- fait des rappels élémentaires les matrices que ce soit par la définition de matrices particulières, leurs structures ou encore les opérations sur les matrices
- introduit la notion de changement de base pour un vecteur ou une famille de vecteurs (important en analyse de données !)
- comment calculer quelques caractéristiques simples d'une matrice

Le contexte matriciel, a permis de s'affranchir dans certains cas de la résolution d'un *système linéaire* bien que cette dernière soit toujours présente de façon implicite.

La prochaine section a pour but de revenir sur les méthodes résolutions de ces systèmes linéaires et complétera les rappels des notions vues l'année passée.

Algèbre Linéaire

Systèmes linéaires

Système linéaire

Définition 1.29: Système linéaire

On appelle *système d'équations linéaires* de n équations à p inconnues tout système S de la forme

$$S : \begin{cases} a_{11}x_1 + \ldots + a_{1p}x_p & = & b_1 \\ a_{21}x_1 + \ldots + a_{2p}x_p & = & b_2 \\ & \vdots & \\ a_{n1}x_1 + \ldots + a_{np}x_p & = & b_n \end{cases}$$

où

- les scalaires $x_1, \ldots, x_p$ sont les inconnues du système d'équations,
- les coefficients $(a_{ij})_{i,j=1}^{n,p}$ sont les coefficients du système,
- le vecteur $\mathbf{b} = (b_1, \ldots, b_n) \in \mathbb{K}^n$ est le second membre du système.

Lien avec les matrices I

Si l'on note $A \in \mathscr{M}_{n,p}(\mathbb{K})$ la matrice associée aux coefficients a_{ij}, $\mathbf{x}$ le vecteur des inconnues x_j et $\mathbf{b}$ le vecteur des scalaires qui constituent le second membre de notre système, nous pouvons réécrire S sous la forme

$$A\mathbf{x} = \mathbf{b}.$$

Tout comme on a introduit la notion de rang pour une matrice, on peut aussi parler de rang pour un système d'équations linéaires. Ainsi, le rang de la matrice A est égal au rang du système S.

Lien avec les matrices II

Dans le cas où le second membre est nul $(=\mathbf{0})$, le système S est dit homogène et s'écrit

$$A\mathbf{x} = 0.$$

Si le second membre n'est pas nul $(\mathbf{b} \neq \mathbf{0})$, le système S_0 défini par $A\mathbf{x} = \mathbf{0}$ est appelé système homogène associé à S. On va voir qu'étudier un tel système va permettre de définir l'ensemble des solutions d'un système d'équations linéaires.

Interprétation I

On se rappelle qu'à toute matrice A est associée une application linéaire, disons u. Ainsi le système S peut aussi s'écrire sous la forme $u(\mathbf{x}) = \mathbf{b}$ et le système d'équation homogène associé S_0 s'écrira $u(\mathbf{x}) = \mathbf{0}$.
Cela permet de faire un premier lien avec l'algèbre linéaire et les solutions d'un système homogène :

Les solutions d'un système homogène S_0 sont exactement les éléments de $Ker(u)$.

Interprétation II

De plus, la nouvelle écriture du système S nous indique que ce dernier admet une solution si et seulement si $\mathbf{b} \in Im(u)$. Si tel est le cas, notons $\mathbf{x}_0$ une solution *particulière* du système S, nous avons donc

$$\begin{aligned} u(\mathbf{x}) = u(\mathbf{x}_0) &\iff u(\mathbf{x} - \mathbf{x}_0) = \mathbf{0} \\ &\iff \mathbf{x} - \mathbf{x}_0 \in Ker(u) \iff \exists K \in Ker(u) \mid \mathbf{x} = \mathbf{x}_0 + K. \end{aligned}$$

On peut donc faire la remarque suivante :

> L'ensemble des solutions du système S est donc $\{\mathbf{x}_0 + K \mid K \in Ker(u)\}$. On peut dire que les solutions de S sont obtenues en trouvant une solution particulière à laquelle on ajoute une solution du système homogène associée.

Remarques anecdotiques

Cette observation est également valable dans d'autres contextes en mathématiques :

- lorsque l'on cherche à déterminer les solutions d'une équation séquentielle de la forme

$$av_{n+1} + bv_n + cv_{n-1} = d(n),$$

- ou encore lorsque l'on cherche à déterminer les solutions d'une équation différentielle (ce qui arrive très souvent en physique) :

$$ay''(t) + by'(t) + cy(t) = f(t).$$

Dimension espace des solutions I

On va maintenant regarder comment déterminer les solutions d'un système linéaire pour des systèmes dits de *Cramer* et pour des systèmes d'équations linéaires quelconques. Pour de tels systèmes, trois cas se présent à nous en ce qui concerne l'espace des solutions :

- Le système peut ne pas avoir de solutions (sauf si c'est un système *homogène*, auquel cas le vecteur nul est toujours solution !)
- le système admet une unique solution
- le système admet une infinité de solutions

Sans informations supplémentaire sur la nature du système, il est difficile de savoir si le système admet ou non des solutions.

On peut cependant être plus précis pour les systèmes homogènes d'**équations linaires indépendantes**. Dans ce cas :

Dimension espace des solutions II

- le système admet une seule et unique solution s'il y a autant d'inconnues que d'équations
- le système admet une infinité de solutions lorsque le nombre d'équations est strictement inférieur au nombre d'inconnues.

Cette dernière remarque est très intéressante car elle permet de déterminer si un système d'équations linéaires homogènes admet des solutions ou non uniquement en regardant le *rang de la matrice A associée au système d'équations linéaires homogènes*.

Il est également possible de caractériser la dimension de l'espace des solutions pour un système linéaire non homogène, c'est le résultat du **Théorème de Rouché-Fontené**.

Dimension espace des solutions III

Théorème 1.4: Théorème de Rouché-Fontené

Soit S un système d'équations linéaires à n variables, de la forme $A\mathbf{x} = \mathbf{b}$. Ce système possède une solution, si et seulement si le rang de la matrice A est égal à celui de la matrice augmenté $\left(\begin{array}{c|c} A & \mathbf{b} \end{array} \right)$.
Si des solutions existent, elles forment un sous-espace **affine** de $\mathbb{R}^n$ de dimension $n - rg(A)$.
De plus :

- si $n = rg(A)$, alors S admet une unique solution
- sinon, il existe une infinité de solutions.

Résolution système de Cramer I

Définition 1.30: Système de Cramer

On appelle **système de Cramer** tout système de n équations à n inconnues de rang n, $n \in \mathbb{N}^\star$.

Un système de Cramer est dons un système S de la forme $A\mathbf{x} = \mathbf{b}$ où $A \in \mathscr{M}_n(\mathbb{K})$ est inversible. Comme la matrice A est inversible, on peut définir son inverse A^{-1} qui vérifie

$$\mathbf{x} = A^{-1}\mathbf{b}.$$

Cette relation montre que tout système de Cramer admet une seule et unique solution définie par la relation précédente. De plus, la matrice A étant inversible, le système homogène associé n'admet pour solution que la solution triviale $\mathbf{x} = \mathbf{0}$.

Résolution système de Cramer II

Nous avons vu comment déterminer A^{-1} dans la section traitant des matrices, ce qui nous suffirait pour résoudre un tel système.
Mais il existe un résultat qui permet aussi de donner immédiatement les solutions d'un tel système. Plus précisément :

Résolution système de Cramer III

Proposition 1.30: Solutions(s) d'un Système de Cramer

Considérons le système de Cramer S suivant :

$$S : \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n & = & b_1 \\ & \vdots & \\ a_{n1}x_1 + \ldots + a_{nn}x_n & = & b_n \end{cases}$$

dont l'écriture matricielle est $A\mathbf{x} = \mathbf{b}$ où $A \in \mathscr{GL}_n(\mathbb{K})$. Alors l'unique solution de S est le n-uplet $(x_1, \ldots, x_n)$ défini par :

$$x_k = \frac{det(A_k)}{det(A)},\ \forall k \in [\![1, n]\!],$$

où A_k est la matrice A dans laquelle on a remplacé la k-ème colonne par le vecteur $\mathbf{b}$.

Exemple d'application I

On considère le système d'équations linéaires suivant que l'on se propose de résoudre.

$$S : \begin{cases} 2x_1 + 3x_2 = & 2 \\ x_1 - 5x_2 = & 1 \end{cases}$$

On commence par montrer que ce système $A\mathbf{x} = \mathbf{b}$ est un système de Cramer donc que la matrice A définie par

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

est inversible. Cela se vérifie facilement en calculant le déterminant qui est ici égal à -13. On peut donc utiliser le résultat de la proposition précédente pour déterminer la solution de ce système.

Exemple d'application II

- $x_1 = \dfrac{det(A_1)}{det(A)} = \dfrac{\begin{vmatrix} 2 & 3 \\ 1 & -5 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -5 \end{vmatrix}} = \dfrac{-13}{-13} = 1,$

- $x_2 = \dfrac{det(A_2)}{det(A)} = \dfrac{\begin{vmatrix} 2 & 2 \\ 1 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 1 & -5 \end{vmatrix}} = \dfrac{0}{-13} = 0.$

Résolution système de Cramer triangulaire I

Il s'agit là d'un cas particulier de Système de Cramer où la matrice A associée est triangulaire (triangulaire supérieure par exemple). Notre système s'écrit alors

$$S : \begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n & = & b_1 \\ a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \vdots & & \\ a_{n-1,1}x_{n-1} + a_{n-1,n}x_n & = & b_{n-1} \\ a_{nn}x_n & = & b_n \end{cases}$$

Résolution système de Cramer triangulaire II

C'est un système particulier qui peut se résoudre "de bas en haut". En effet, il s'agit d'un système de Cramer, la matrice A est donc inversible. Or elle est triangulaire supérieure, cela implique que tous éléments diagonaux de la matrice sont non nuls. La résolution donne :

- $x_n = \dfrac{b_n}{a_{nn}}$
- $x_{n-1} = \dfrac{1}{a_{n-1,n-1}}(b_{n-1} - a_{n-1,n}x_n)$

$\vdots$

-
- $x_1 = \dfrac{1}{a_{11}}(b_1 - a_{12}x_2 - \ldots - a_{1n}x_n)$.

Ce type de système est donc très simple à résoudre car il suffit d'utiliser les lignes du "bas vers le haut" pour déterminer les différentes valeurs de x_i.

Exemple d'application I

Résoudre le système d'équations linéaires suivant :

$$S : \begin{cases} 2x_1 + 3x_2 - x_3 & = & 7 \\ -2x_2 + 3x_3 & = & -1 \\ 7x_3 & = & -7 \end{cases}$$

On va donc déterminer x_3 à l'aide de la troisième équation puis x_2 et enfin x_1, ce qui nous donne :

$$x_3 = -1,$$

$$x_2 = -\frac{1}{2}(-1 - 3x_3) = -\frac{1}{2}(-1 + 3) = -1,$$

$$x_1 = \frac{1}{2}(7 - 3x_2 + x_3) = \frac{1}{2}(7 + 3 - 1) = \frac{9}{2}.$$

Pivot de Gauss I

Considérons un système d'*équations linéaires indépendantes* suivant :

$$S : \begin{cases} a_{11}x_1 + \ldots + a_{1p}x_p & = & b_1 \\ a_{21}x_1 + \ldots + a_{2p}x_p & = & b_2 \\ \vdots & & \\ a_{n1}x_1 + \ldots + a_{np}x_p & = & b_n \end{cases},$$

où $(b_1, \ldots, b_n) \neq 0$. D'après la remarque en début de section nous avons vu que :

- si $p < n$, alors le système n'admet pas de solutions
- si $p = n$, alors le système admet une seule et unique solution
- si $p > n$, alors le système admet une infinité de solutions

Pivot de Gauss II

On va se placer dans le cas où $p = n$ et on va donc essayer de se ramener à un système triangulaire supérieur à l'aide de la **méthode du pivot de Gauss**. On va donc effectuer des opérations élémentaires sur les lignes de la matrice A associée au système :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \longrightarrow \begin{pmatrix} a'_{11} & a'_{12} & \dots & a'_{1n} \\ 0 & a'_{22} & \dots & a'_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a'_{nn} \end{pmatrix}.$$

Pour obtenir un tel système triangulaire, on va servir de a_{11} comme pivot afin d'annuler tous les éléments se trouvant dans la même colonne que a_{11}, *i.e.* tous les éléments de la forme a_{k1}. Concrètement on va effectuer les opérations :

Pivot de Gauss III

$$A \longrightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & \dots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} - \frac{a_{21}}{a_{11}} a_{12} & \dots & a_{nn} - \frac{a_{21}}{a_{11}} a_{1n} \end{pmatrix} \begin{array}{l} \\ L_2 \leftarrow L_2 - \frac{a_{21}}{a_{11}} L_1 \\ \\ L_n \leftarrow L_n - \frac{a_{n1}}{a_{11}} L_1 \end{array} .$$

Notre objectif est faire apparaître des 0 sur tous les éléments de la colonne se trouvant sous le pivot afin de se ramener à une matrice échelonnée.

On va ensuite appliquer ce principe colonne par colonne mais en n'oubliant de descendre d'une ligne à chaque fois. En effet, si nous ne faisons pas cela, nous risquons de refaire apparaître des valeurs non nulles !

Pivot de Gauss IV

Ensuite, on va donc se servir de la première valeur non nulle, *i.e.* $a_{22} - \frac{a_{21}}{a_{11}} a_{12}$, comme nouveau pivot pour annuler toutes les autres valeurs se trouvant dans la même colonne mais pour des indices de lignes supérieurs à la ligne du pivot. Et on continue ainsi de suite jusqu'à l'obtention d'un système triangulaire, ce qui nous donne :

$$A \longrightarrow \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & \dots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & \color{red}{0} & \dots & a_{nn} - \frac{a_{21}}{a_{11}} a_{1n} - \color{red}{\gamma} \end{pmatrix},$$

où $\gamma = \dfrac{a_{n2} - \frac{a_{21}}{a_{11}} a_{12}}{a_{22} - \frac{a_{21}}{a_{11}} a_{12}} \left(a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \right).$

Pivot de Gauss V

Dans le cas où le système n'est pas homogène, les opérations effectuées sur les lignes doivent aussi s'effectuer sur les lignes du vecteur !

Remarque : on remarque que dans la présentation effectuée de la méthode du pivot de Gauss, on essaie de faire intervenir des matrices échelonnées. Il est parfois d'usage de faire apparaître des matrices échelonnées réduites, *i.e.* la première valeur non nulle d'une ligne donnée est un 1. Cela permet de simplifier les calculs dans certaines situations. Cela donnerait :

$$A \longrightarrow \begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} L_1 \leftarrow L_1/a_{11}$$

Pivot de Gauss VI

$\downarrow$ on sert maintenant de la valeur 1 comme pivot

$$\longrightarrow \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & \dots & a_{2n} - \frac{a_{21}}{a_{11}} a_{1n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2} - \frac{a_{21}}{a_{11}} a_{12} & \dots & a_{nn} - \frac{a_{21}}{a_{11}} a_{1n} \end{pmatrix} \begin{array}{l} \\ L_2 \leftarrow L_2 - a_{21} L_1 \\ \\ L_n \leftarrow L_n - a_{n1} L_1 \end{array}$$

Et ainsi de suite ... regardons cela sur quelques exemples

Exemples d'application I

On se propose de résoudre le système S défini par :

$$S : \begin{cases} 2x_1 - 4x_2 - 2x_3 & = & 8 \\ 5x_1 - 2x_2 + 3x_3 & = & -1 \\ -x_1 + 3x_2 + 7x_3 & = & -2 \end{cases}$$

On peut réécrire notre système sous la forme :

$$(A \mid \mathbf{b}) = \left(\begin{array}{ccc|c} 2 & -4 & -2 & 8 \\ 5 & -2 & 3 & -1 \\ -1 & 3 & 7 & -2 \end{array}\right).$$

On peut vérifier qu'il s'agit d'un système d'équations indépendantes de trois équations à trois inconnus, c'est donc bien un système de Cramer. On va appliquer la méthode du pivot de Gauss afin de transformer la matrice A en une matrice triangulaire supérieure

Exemples d'application II

$$A \longrightarrow \left(\begin{array}{ccc|c} 2 & -4 & -2 & 8 \\ 0 & 8 & 8 & -21 \\ 0 & 1 & 6 & 2 \end{array}\right) \begin{array}{l} L_2 \leftarrow L_2 - 2.5L_1 \\ L_3 \leftarrow L_3 + 0.5L_1 \end{array}$$

$\downarrow$ on sert ensuite de la valeur 8 comme pivot

$$\longrightarrow \left(\begin{array}{ccc|c} 2 & -4 & -2 & 8 \\ 0 & 8 & 8 & -21 \\ 0 & 0 & 5 & \frac{37}{8} \end{array}\right) L_3 \leftarrow L_3 - \frac{1}{8}L_2$$

Notre système S est donc équivalent au système

$$S : \begin{cases} 2x_1 - 4x_2 - 2x_3 & = & 8 \\ 8x_2 + 8x_3 & = & -21 \\ 5x_3 & = & \frac{37}{8} \end{cases}$$

Exemples d'application III

que l'on peut alors résoudre " de bas en haut".
Ce qui nous donne le triplet solution suivant

$$S : \begin{cases} x_1 = -\dfrac{435}{200} \\ x_2 = -\dfrac{71}{20} \\ x_3 = \dfrac{37}{40} \end{cases}$$

Dans cet exemple l'espace des solutions est un **point** dans un espace de dimension 3.

Exemples d'application IV

On se propose de résoudre le système S défini par :

$$S : \begin{cases} x_1 - x_2 - 2x_3 & = & -1 \\ 3x_1 - 2x_2 + 3x_3 & = & 2 \end{cases}$$

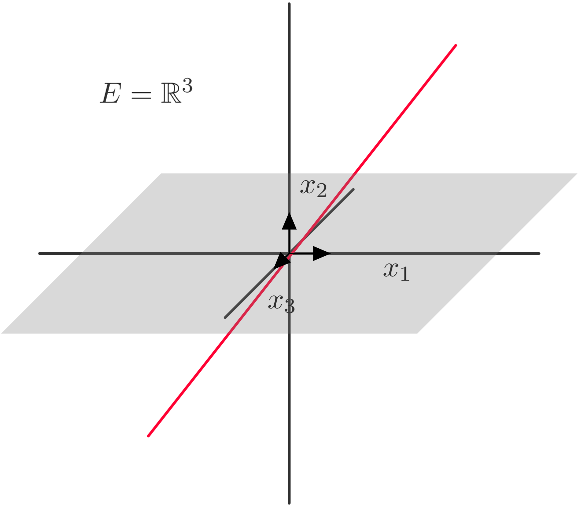
On peut facilement voir que les deux équations sont linéairement indépendantes. On va se servir de la valeur 1 comme pivot et on obtient le système équivalent

$$S : \begin{cases} x_1 - x_2 - 2x_3 & = & -1 \\ \quad x_2 + 9x_3 & = & 5 & : L_2 \leftarrow L_2 - 3L_1 \end{cases}$$

La deuxième équation comporte deux inconnues, on va fixer $x_3 = t \in \mathbb{R}$ qui sera donc un paramètre. On aura donc :

$$S : \begin{cases} x_1 & = & -1 + 2t + x_2 \\ x_2 & = & 5 - 9t \end{cases}, \longrightarrow \begin{cases} x_1 & = & 4 - 7t \\ x_2 & = & 5 - 9t \end{cases}$$

Exemples d'application V



Algèbre Linéaire

Réduction des endomorphismes

Motivations I

Etude des endomorphismes u d'espaces vectoriels E de dimension finie $n \in \mathbb{N}^\star$ muni d'une base $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$.
Représentation matricielle de u, notée $A = Mat(u)$

$$A = \begin{array}{c} \begin{array}{cccc} u(\mathbf{e}_1) & u(\mathbf{e}_2) & \ldots & u(\mathbf{e}_n) \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \end{array}.$$

On a vu comment représenter ce même endomorphisme dans des bases différentes $\mathscr{B}$ et $\mathscr{B}'$ dont les représentations sont notées A et A' respectivement et on note P la matrice de passage de $\mathscr{B}$ vers $\mathscr{B}'$. Alors

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

Motivations II

Dans cette section, nous allons chercher une base $\mathscr{B}'$ dans laquelle la matrice A' se présente sous la forme d'une matrice diagonale, comme suit :

$$A' = \begin{pmatrix} a'_{11} & 0 & \dots & 0 \\ 0 & a'_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a'_{nn} \end{pmatrix}.$$

Valeurs propres et vecteur propres I

Définition 1.31: Vecteur et valeur propres

Soit E un espace vectoriel sur $\mathbb{K}$ de dimension finie n et soit u un endomorphisme de E. On dit qu'un vecteur $\mathbf{x} \in E$ est un **vecteur propre** de u s'il existe $\lambda \in \mathbb{K}$ tel que

$$u(\mathbf{x}) = \lambda \mathbf{x}.$$

Dans ce cas, le scalaire λ est appelé **valeur propre** de l'endomorphisme u.

Si $\mathbf{x}$ est un vecteur propre, il existe une et une seule valeur propre λ associée telle que $u(\mathbf{x}) = \lambda \mathbf{x}$.

Valeurs propres et vecteur propres II

Si λ est une valeur propre, l'ensemble des vecteurs propres associés est l'ensemble des vecteurs non nuls de l'espace $Ker(u - \lambda Id)$. En effet :

$$\lambda \text{ valeur propre } \iff \exists \mathbf{x} \neq 0 \text{ t.q. } u(\mathbf{x}) = \lambda \mathbf{x} \iff \exists \mathbf{x} \neq 0 \text{ t.q. } (u - \lambda Id)(\mathbf{x}) = 0 \iff \mathbf{x} \in Ker(u - \lambda Id).$$

Si les vecteurs propres sont non nuls par définition, rien n'empêche la valeur propre d'être nulle , dans ce cas les vecteurs propres associés à la valeur propre $\lambda = 0$ sont exactement les vecteurs non nuls de $Ker(u)$.

On peut aussi dire que 0 est une valeur propre de u si et seulement si u n'est pas injectif.

Diagonalisation

Définition 1.32: Endomorphisme diagonalisable

Un endomorphisme u d'un espace vectoriel E sur $\mathbb{K}$ de dimension finie n est **diagonalisable** si et seulement s'il existe une base de E formée des vecteurs propres de u.

Si $\mathbf{x}_i$ est un vecteur propre associé à la valeur propre λ_i alors $u(\mathbf{x}_i) = \lambda \mathbf{x}_i$.

$$A = \begin{array}{cc} \begin{array}{cccc} u(\mathbf{x}_1) & u(\mathbf{x}_2) & \dots & u(\mathbf{x}_n) \end{array} & \\ \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} & \begin{array}{c} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{array} \end{array}.$$

Recherche de valeurs propres I

Rappelons que λ est une valeur propre d'un endomorphisme u si l'application $u - \lambda Id$ n'est pas injective, *i.e.* non bijective et son déterminant est donc nul.

C'est ce que nous allons utiliser pour déterminer les valeurs propres de notre application. Ainsi λ est une valeur propre de u si et seulement si

$$det\,(u - \lambda Id) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1n} \\ a_{n1} & \dots & a_{nn-1} & a_{nn} - \lambda \end{vmatrix} = 0.$$

On obtient un polynôme de degré n dont les racines sont les valeurs propres de u.

Recherche de valeurs propres II

Proposition 1.31: Valeurs propres d'un endomorphisme

Soient E un espace vectoriel de dimension finie n et u un endomorphisme de E. Alors les valeurs propres de u sont exactement les racines du polynôme P_u définit par :

$$P_u(\lambda) = det(u - \lambda Id).$$

Ce polynôme est appelé **polynôme caractéristique** de u. Il est de degré n en λ.

Recherche de valeurs propres III

Remarque : on peut également définir le polynôme caractéristique d'un endomorphisme en utilisant sa représentation matricielle A. Ainsi

$$P_u(\lambda) = det(u - \lambda Id) = det(A - \lambda I_n) = P_A(\lambda).$$

Certains auteurs définissent le polynôme caractéristique par :

$$P_u(\lambda) = det(\lambda Id - u) = det(\lambda I_n - A) = P_A(\lambda).$$

Ce qui revient simplement, pour passer de l'une à l'autre des expressions, à multiplier par $(-1)^n$.

Recherche de valeurs propres IV

Considérons l'endomorphisme u dont la représentation dans la base canonique est donnée par :

$$A = \begin{pmatrix} 1 & -1 \\ 3 & 5 \end{pmatrix}.$$

Le polynôme caractéristique $P_A(\lambda)$ est donné par :

$$P_A(\lambda) = det(A - \lambda Id) = \begin{vmatrix} 1-\lambda & -1 \\ 3 & 5-\lambda \end{vmatrix} = (1-\lambda)(5-\lambda)+3 = (\lambda-2)(\lambda-4).$$

Ainsi les valeurs propres de A sont $\lambda_1 = 2$ et $\lambda_2 = 4$.

Recherche de valeurs propres V

Considérons l'endomorphisme u dont la représentation dans la base canonique est donnée par :

$$A = \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}.$$

La matrice est triangulaire supérieure, on peut donc directement lire les valeurs propres sur la diagonale de matrice.

En effet, on se souvient que le déterminant d'une matrice triangulaire est égale au produit des éléments diagonaux.
Ainsi les valeurs propres de A sont $\lambda_1 = 2$ et $\lambda_2 = 2$. **On dira aussi que l'endomorphisme n'admet qu'une seule valeur propre mais de multiplicité** 2.

Recherche de valeurs propres VI

Définition 1.33: Spectre d'une matrice ou d'un endomorphisme

L'ensemble des valeurs propres d'un endomorphisme u est appelé **spectre de** u, noté $Spec(u)$ ou encore $Spec(A)$, si A est la représentation matricielle de u.

Proposition 1.32: Somme et produit des valeurs propres

Soit $A \in \mathscr{M}_n(\mathbb{K})$ et notons $\lambda_1, \ldots, \lambda_n$ ses valeurs propres, c'est-à-dire les racines du polynôme caractéristique associé à A. Alors :

- $tr(A) = \sum_{i=1}^{n} \lambda_i$ et
- $det(A) = \prod_{i=1}^{n} \lambda_i$.

Recherche de valeurs propres VII

La démonstration de ce résultat est compliquée, elle nécessite de repasser par la définition de déterminant et comprendre ce qu'est une permutation. Je ne la présente donc pas, mais vous pourrez la retrouver dans le poly.

Je vais plutôt vous montrer un exemple qui illustre ce résultat en étudiant un endomorphisme de $\mathbb{R}^2$.

Recherche de valeurs propres VIII

Considérons l'endomorphisme u dont la représentation dans la base canonique est donnée par :

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

On pourrait alors utiliser le résultat précédent pour calculer les deux valeurs propres λ_1 et λ_2 !
En effet, d'après la proposition précédente, le polynôme caractéristique de A s'écrit :

$$P_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 = \lambda - tr(A)\lambda + det(A).$$

On a donc le système suivant :

Recherche de valeurs propres IX

$$\begin{cases} \lambda_1 \lambda_2 = det(A) = 4 \\ \lambda_1 + \lambda_2 = tr(A) = 5 \end{cases},$$

dont l'unique solution est donnée par $\lambda_1 = 1$ et $\lambda_2 = 4$.

Recherche de vecteurs propres I

Il s'agit maintenant de déterminer *les vecteurs propres associés à chaque valeur propre*. Cette étape s'effectue donc *après avoir déterminé les valeurs propres d'un endomorphisme*.

Pour cela nous devons déterminer les éléments $\mathbf{x}$ tels que $\mathbf{x} \in Ker(u - \lambda Id)$. Cela peut se faire en résolvant le système d'équations linéaires homogène :

$$(A - \lambda I)\mathbf{x} = 0,$$

pour toutes les valeurs de λ appartenant au spectre de A.

La connaissance de ces vecteurs propres va permettre de définir une base dans laquelle la nouvelle matrice représentant u sera diagonale mais aussi la matrice de passage associée de la base courante vers la base de vecteurs propres.

Recherche de vecteurs propres II

Considérons l'endomorphisme u dont la représentation dans la base canonique est donnée par :

$$A = \begin{pmatrix} 1 & -1 \\ 3 & 5 \end{pmatrix}.$$

Nous avons vu que cet endomorphisme possède deux valeurs propres qui sont $\lambda_1 = 2$ et $\lambda_2 = 4$. Déterminons maintenant les vecteurs propres associés à ces deux valeurs propres en cherchant éléments de $Ker(A - \lambda I)$.

- **Pour** $\lambda = \lambda_1 = 2$ **:**
  Nous devons résoudre le système linéaire $(A - 2I)\mathbf{x} = 0$

$$\begin{cases} -x_1 - x_2 & = & 0 \\ 3x_1 + 3x_2 & = & 0 \end{cases}$$

Recherche de vecteurs propres III

Les deux équations sont dépendantes, on peut donc se concentrer sur la première équation dont une solution est donnée par $(x_1, x_2) = (-1, 1)$.

- **Pour $\lambda = \lambda_2 = 4$:**
  Nous devons résoudre le système linéaire $(A - 4I)\mathbf{x} = 0$

$$\begin{cases} -3x_1 - x_2 & = & 0 \\ 3x_1 + x_2 & = & 0 \end{cases}$$

A nouveau, les deux équations sont dépendantes, on peut donc se concentrer sur la première équation dont une solution est $(x_1, x_2) = (-1, 3)$.

On a donc deux vecteurs propres et on peut vérifier qu'ils forment une base de $E = \mathbb{R}^2$ en calculant le déterminant :

Recherche de vecteurs propres IV

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

Ainsi la matrice A est diagonalisable et elle est semblable à la matrice $A' = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$ et la matrice de passage P vers la vers base des vecteurs propres est défini par $P = \begin{pmatrix} -1 & -1 \\ 1 & 3 \end{pmatrix}$.
On peut ensuite facilement vérifier que $A' = P^{-1}AP$.

A retenir : la matrice de passage est composée des vecteurs propres en colonne !

Recherche de vecteurs propres V

Considérons l'endomorphisme u dont la représentation dans la base canonique est donnée par :

$$A = \begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}.$$

Nous avons vu que cette matrice admet une seule valeur propre $\lambda = 2$ dont la multiplicité est égale à 2. Déterminons les vecteurs propres associés en cherchant les solutions du système $(A - 2I)\mathbf{x} = 0$:

$$\begin{cases} -x_2 & = & 0 \\ 0 & = & 0 \end{cases}$$

Ce système admet pour solution le vecteur de la forme $(x_1, x_2) = (1, 0)$ et ... c'est tout ! On ne trouve ici qu'un seul vecteur propre ... la matrice A n'est donc pas diagonalisable a priori car on ne trouve pas une base de vecteurs propres de $\mathbb{R}^2$.

Caractérisation des endomorphismes diagonalisables I

Définition 1.34: Espace propre

On appelle **sous-espace vectoriel propre** associé à une valeur propre λ, le sous-espace vectoriel

$$E_\lambda(u) = Ker(u - \lambda Id).$$

Un tel sous espace vectoriel est au moins de dimension 1. Si λ n'est pas une valeur propre de u, alors cet espace est de dimension 0 (il est réduit au vecteur nul.

Caractérisation des endomorphismes diagonalisables II

Proposition 1.33: Propriétés des sous-espaces propres

Soient $\lambda_1, \ldots, \lambda_n$ les vecteurs propres, deux à deux distincts, d'un endomorphisme u. Alors la famille des sous-espaces propres $\{E_{\lambda_i}(u)\}_{i=1}^{n}$ est en somme directe.

Ce premier résultat est important car si on dispose de n sous-espaces de dimension au moins égale à 1 qui sont en somme directe dans un espace vectoriel de dimension n. Cela signifie que les vecteurs propres forment une base de cet espace vectoriel, donc l'endomorphisme est diagonalisable.

Caractérisation des endomorphismes diagonalisables III

Théorème 1.5: Diagonalisation et sous-espaces propres

Soit u un endomorphisme d'un espace vectoriel E de dimension n et notons $\lambda_1, \ldots, \lambda_n$ ses valeurs propres (pas nécessairement distinctes). Alors les propriétés suivantes sont équivalentes :

i) u est diagonalisable

ii) E est la somme directe de ses sous-espaces vectoriels propres.

Ce théorème a deux conséquences immédiates :

Caractérisation des endomorphismes diagonalisables IV

Corollaire 1.1: Diagonalisation et valeurs propres

Soit E un espace vectoriel de dimension n et u un endomorphisme de E qui admet $r \leq n$ valeurs propres distinctes $\lambda_1, \dots, \lambda_r$. Alors :

- l'endomorphisme u est diagonalisable si et seulement si la somme $\sum_{k=1}^{r} dim(E_{\lambda_k}(u))$ est égale à n.
- l'endomorphisme u est diagonalisable si et seulement si, pour tout $k \in [\![1, r]\!]$, nous avons

$$dim(E_{\lambda_k}(u)) = m(\lambda_k),$$

où $m(\lambda_k)$ est la multiplicité de la valeur propre λ_k.

Caractérisation des endomorphismes diagonalisables V

Corollaire 1.2: Diagonalisation et polynôme caractéristique

Soit E un espace vectoriel de dimension finie n et u un endomorphisme de E. On notera $A \in \mathscr{M}_n(\mathbb{R})$ la représentation matricielle de cet endomorphisme. Si le polynôme caractéristique de A est scindé sur $\mathbb{R}$ (*i.e.* il peut s'écrire comme le produit de monômes à coefficients dans $\mathbb{R}$) et possède toutes ses racines simples, alors A est diagonalisable dans $\mathscr{M}_n(\mathbb{R})$.

Caractérisation des endomorphismes diagonalisables VI

L'hypothèse *polynôme caractéristique scindé sur* $\mathbb{R}$ est une condition importante pour la diagonalisation. En effet considérons la matrice A définie par

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

dont le polynôme caractéristique $P_A(\lambda)$ est donné par $P_A(\lambda) = \lambda^2 + 1$. Ce polynôme n'est pas scindé dans $\mathbb{R}$ et n'est donc pas diagonalisable car ces deux valeurs propres sont complexes i et $-i$.

D'ailleurs, cette matrice est diagonalisable, en vertu des résultats précédents, mais ... pas dans $\mathbb{R}$, dans $\mathbb{C}$! ! !

Exemple I

Considérons l'endomorphisme u dont la représentation dans la base canonique est donnée par :

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

et montrons que ce dernier diagonalisable.

Exemple II

Son polynôme caractéristique est donné par le déterminant $det(A - \lambda I_3)$:

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

$\downarrow$ on va développer selon la première ligne

$$= -\lambda \begin{vmatrix} 5-\lambda & 2 \\ -3 & -\lambda \end{vmatrix} - 3 \begin{vmatrix} -2 & 2 \\ 2 & -\lambda \end{vmatrix} + 2 \begin{vmatrix} -2 & 5-\lambda \\ 2 & -3 \end{vmatrix},$$

$\downarrow$ on développe chaque déterminant d'ordre 2

$$= \underbrace{-\lambda(-\lambda(5-\lambda)+6)}\,\underbrace{-3(2\lambda-4)}\,\underbrace{+2(6-2(5-\lambda))},$$

$$= -\lambda^3 + 5\lambda^2 - 6\lambda - 6\lambda + 12 + 12 - 20 + 4\lambda,$$

$\downarrow$ factorisation

$$= -\lambda^3 + 5\lambda^2 - 8\lambda + 4,$$

Exemple III

$$
\begin{aligned}
&\downarrow \text{ on observe que 2 est racine du polynôme} \\
=&\ -(\lambda-2)(\lambda^2-3\lambda+2), \\
&\downarrow \text{ on observe à nouveau que 2 et 1 sont racines} \\
=&\ -(\lambda-2)^2(\lambda-1).
\end{aligned}
$$

Le polynôme caractéristique admet donc deux valeurs propres distinctes $\lambda_1 = 1$ et $\lambda_2 = 2$ où $m(\lambda_1) = 1$ et $m(\lambda_2) = 2$.
Pour vérifier si notre endormorphisme est diagonalisable, il faut montrer que sous-espace propre associé à la valeur propre λ_2 est bien de dimension 2. Cependant, on va également chercher à déterminer le sous-espace propre E_{λ_1} afin de définir notre matrice de passage.
On rappelle que les sous-espaces propres sont définis par :

Exemple IV

$$E_\lambda = Ker(u - \lambda Id).$$

- **Espace propre E_{λ_1} :**
  On cherche à résoudre le système suivant en utilisant la méthode du pivot :

$$\begin{cases} -x_1 + 3x_2 + 2x_3 & = & 0 \\ -2x_1 + 4x_2 + 2x_3 & = & 0 \ : \ L_2 \leftarrow L_2 - 2L_1 \\ 2x_1 - 3x_2 - x_3 & = & 0 \ : \ L_3 \leftarrow L_3 + 2L_1 \end{cases}$$

$$\begin{cases} -x_1 + 3x_2 + 2x_3 & = & 0 \\ -2x_2 - 2x_3 & = & 0 \\ 3x_2 + 3x_3 & = & 0 \end{cases}$$

Exemple V

Le système peut donc être réduit au système à deux équations suivants :

$$\begin{cases} -x_1 + 3x_2 + 2x_3 & = & 0 \\ x_2 & = & -x_3 \end{cases}$$

soit

$$\begin{cases} x_1 & = & -x_3 \\ x_2 & = & -x_3 \end{cases}$$

Le sous espace propre E_{λ_1} est alors généré par le vecteur propre
$$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

Exemple VI

- **Espace propre E_{λ_2} :**
  On cherche à résoudre le système suivant en utilisant la méthode du pivot :

$$\begin{cases} -2x_1 + 3x_2 + 2x_3 & = & 0 \\ -2x_1 + 3x_2 + 2x_3 & = & 0 \\ 2x_1 - 3x_2 - 2x_3 & = & 0 \end{cases}$$

  On remarque que le système peut-être réduit à l'équation suivante :

$$-2x_1 + 3x_2 + 2x_3 = 0$$

  dont les solutions sont paramétrés par des paramètres $x_1 = t_1$ et $x_2 = t_2$. Une première solution est alors donnée par le vecteur $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

Exemple VII

en posant $t_2 = 0$ et $t_1 = 1$ et une deuxième solution, indépendante est donnée par le vecteur $\begin{pmatrix} 0 \\ -1/3 \\ 1/2 \end{pmatrix}$ en posant $t_1 = 0$ et $t_2 = -1/3$.

Ainsi le sous-espace propre E_{λ_2} est engendré par les vecteurs $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ -1/3 \\ 1/2 \end{pmatrix}$.

Exemple VIII

Pour chaque valeur propre, la dimension du sous-espace propre est égale à la multiplicité de la valeur propre associée, notre endomorphisme est donc diagonalisable. Dans la base de vecteurs propres, la matrice diagonale est définie par $Diag(1,2,2)$ et la matrice de passage est donnée par :

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & -1/3 \\ -1 & 1 & 1/2 \end{pmatrix}.$$

On pourra vérifier que la matrice $P^{-1}AP$ est bien diagonale.

Une ouverture sur $\mathbb{C}$ I

Dans cette section, nous avons supposé que le corps sur lequel nous avons effectué notre étude est le corps des nombres réels $\mathbb{R}$.

Sur le corps des nombres complexes, la théorie reste inchangée mais on peut cependant remarqué que si des matrices ne sont pas diagonalisables dans $\mathbb{R}$ elle peuvent l'être dans $\mathbb{C}$! Encore mieux, il se peut qu'une matrice n'admette aucune valeur propre réelle mais qu'elle soit diagonalisable (dans $\mathbb{C}$ mais pas dans $\mathbb{R}$ bien évidemment).

Un exemple classique qui permet d'illustrer cela est la matrice de rotation $R(\theta)$ où $\theta \in \mathbb{R}$ est définie par :

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Le polynôme caractéristique de cette matrice est défini par

Une ouverture sur $\mathbb{C}$ II

$$\begin{aligned} P_{R(\theta)}(\lambda) &= (\cos(\theta) - \lambda)^2 + \sin(\theta)^2, \\ &= \lambda^2 - 2\cos(\theta)\lambda + 1. \end{aligned}$$

Si on cherche les racines de ce polynôme, en utilisant le fait que $\cos(\theta) = \frac{1}{2}\left(e^{i\theta} + e^{-i\theta}\right)$, on trouve $\lambda_\pm = \{e^{-i\theta}, e^{i\theta}\}$ où i est tel que $i^2 = -1$. Donc $P_{R(\theta)}(\lambda) = (\lambda - e^{-i\theta})(\lambda - e^{i\theta})$.

Ce polynôme est donc scindé sur $\mathbb{C}$ mais ne l'est pas sur $\mathbb{R}$ sauf dans le cas $\theta = 0$.
Comme le polynôme caractéristique est scindé à racines simples dans $\mathbb{C}$, $R(\theta)$ est diagonalisable dans $\mathbb{C}$. De plus $R(\theta)$ est diagonalisable dans $\mathbb{R}$ si et seulement si $\theta \equiv 0\ [\pi]$.

Pour finir

La réduction des endomorphismes jouera un rôle clef dans l'analyse de données. En effet :

- **vecteurs propres :** ils permettront de définir une base de notre nouvel espace de représentation des données ! Ils vont aussi définir la transformation à appliquer aux données pour les représenter dans cette base.
- **valeurs propres :** elles vont nous permettre de quantifier l'information présente dans les données selon les différents axes définies par les vecteurs propres.
  $\rightarrow$ est-ce que ma projection dans le nouvel espace préserve suffisamment l'information initiale ?

Derniers concepts à voir : **Normes - distances** et **produits scalaires**, bref un peu de géométrie !

Algèbre Linéaire

Formes quadratiques et espaces euclidiens

Motivations I

Nous avons presque tous les outils nécessaires à la présentation des méthodes d'Analyse de Données !

En fait, il reste une dernière étape à franchir avant de se lancer dans les présentations de telles méthodes.

Souvenez vous, on cherchait aussi à projeter les données dans un espace de dimension réduite et en préservant au maximum l'information présente dans les données.

Motivations II

L'information présente dans les données peut s'étudier de deux façons

- En probabilité/statistiques : vous avez souvent parler de la notion de *variance* pour quantifier la variabilité dans vos données.
  Nous verrons cette variance comme une source d'informations dans nos données.
- Analyse de données : on pourra aussi quantifier la quantité d'informations en regardant les distances entre les individus

Le premier point passera par l'étude d'objets comme la matrice de variance-covariance des données (passage d'une matrice rectangulaire à une matrice carrée $\rightarrow$ objet de la prochaine section).

Pour le deuxième point, on va d'avantage se concentrer sur des objets géométriques, comme la notion de norme qui permet de calculer une distance entre deux vecteurs. Ainsi, potentiellement entre deux individus $\mathbf{x}_i$ et $\mathbf{x}_j$.

Motivations III

En outre, comme on cherche à projeter nos données dans un espace de dimension faible, ou sur des axes en particulier (que l'on appellera axe factoriel ou axe principaux), il faut voir comment définir de telles projections (sur une droite ou sur un sous-espace).

On va donc avoir besoin d'étudier les notions de **produit scalaire** (qui est une autre application bilinéaire) - **formes quadratiques** ou encore des notions autour des **espaces euclidiens** (un cas particulier d'espace vectoriel) et enfin des notions **d'orthogonalité**.

Forme bilinéaire I

Définition 1.35: Forme bilinéaire

Soit E un espace vectoriel sur $\mathbb{K}$ et ϕ une application de $E \times E$ dans $\mathbb{K}$. L'application est ϕ est dite **bilinéaire** si elle est linéaire en ses deux arguments, *i.e.* si pour tout $\mathbf{x}, \mathbf{x}', \mathbf{x}'' \in E$ et $\alpha \in \mathbb{K}$ elle est vérifie

- $\phi(\mathbf{x} + \mathbf{x}'', \mathbf{x}') = \phi(\mathbf{x}, \mathbf{x}') + \phi(\mathbf{x}'', \mathbf{x}')$ (linéarité à gauche)
- $\phi(\mathbf{x}, \mathbf{x}' + \mathbf{x}'') = \phi(\mathbf{x}, \mathbf{x}') + \phi(\mathbf{x}, \mathbf{x}'')$ (linéarité à droite)
- $\phi(\lambda \mathbf{x}, \mathbf{x}') = \phi(\mathbf{x}, \lambda \mathbf{x}') = \lambda \phi(\mathbf{x}, \mathbf{x}')$

Forme bilinéaire II

Regardons un peu comment se caractérise une forme bilinéaire si on se place dans un espace E de dimension finie n muni d'une base $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ (par exemple, la base canonique de $\mathbb{R}^n$). Pour cela considérons $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ et $\mathbf{y} = \sum_{i=1}^n y_i \mathbf{e}_i$. Alors

$$\phi(\mathbf{x}, \mathbf{y}) = \phi(\sum_{i=1}^n x_i \mathbf{e}_i, \sum_{i=1}^n y_i \mathbf{e}_i) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \phi(\mathbf{e}_i, \mathbf{e}_j).$$

Ainsi, la forme bilinéaire est entièrement définie par la connaissance des images des vecteurs de base par l'application ϕ.

Forme bilinéaire III

Cette première écriture permet également de montrer que toutes les formes bilinéaires peuvent s'écrire de la forme suivante :

$$\phi(\mathbf{x}, \mathbf{y}) = x_1 y_1 \phi_{11} + x_2 y_1 \phi_{21} + x_1 y_2 \phi_{12} + \ldots + x_i y_j \phi_{ij} + \ldots + x_n y_n \phi_{nn},$$

où $\phi_{ij} = \phi(\mathbf{e}_i, \mathbf{e}_j)$. Une forme bilinéaire se présente donc comme la somme de monômes en $x_i y_j$ où chacun des termes se retrouvent à la puissance 1. Si l'on prend le cas particulier de la base canonique de $\mathbb{R}^n$, nous aurions

$$\phi(\mathbf{x}, \mathbf{y}) = x_1 y_1 \phi_{11} + x_2 y_2 \phi_{21} + \ldots + x_n y_n \phi_{nn},$$

Exemple

Les applications suivantes sont des formes bilinéaires

- $3x_1y_1 + 2x_2y_1 - 3.14x_1y_3$
- $(x_1 - x_2)^2 - (x_1 + x_2)^2$

Les applications suivantes ne sont pas des formes bilinéaires

- $3x_1^2y_2 - x_2y_1 + 5x_3y_3$ (présence d'un terme de degré 2)
- $3x_1 - x_2y_1 + 5x_3y_3$ (présence d'un terme de degré un uniquement)

Forme bilinéaire symétrique I

Définition 1.36: Forme bilinéaire symétrique

Soit ϕ une forme bilinéaire sur un espace vectoriel E. Cette application ϕ est dite bilinéaire **symétrique** si et seulement si pour tout $\mathbf{x}, \mathbf{y} \in E$

$$\phi(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{y}, \mathbf{x}).$$

Une forme bilinéaire symétrique qui nous intéressera plus particulièrement un petit peu plus tard sera *le produit scalaire*. On verra même que cette forme biliénaire dispose de proriétés supplémentaires.

Pour le moment, remarquons qu'une forme bilinéaire est dite symétrique si elle est invariante sous l'action de transposition des vecteurs $\mathbf{x}$ et $\mathbf{y}$.

Exemple

L'application bilinéaire définie par :

$$\phi(\mathbf{x},\mathbf{y}) = x_1y_1 + x_2y_1 + x_1y_2 - 3x_2y3 + 3x_3y_2$$

n'est pas symétrique car la transposition $\mathbf{x} \leftrightarrow \mathbf{y}$ ne laisse pas le résultat invariant comme le montre les deux derniers termes. En revanche l'application

$$\phi(\mathbf{x},\mathbf{y}) = x_1y_1 + x_2y_1 + x_1y_2 - 3x_2y3 - 3x_3y_2$$

est bien symétrique.

Représentation matricielle I

Comme pour les applications linéaires, il est possible de représenter les applications bilinéaires dès lors que l'on connaît l'image des vecteurs de base par notre application.

La différence réside dans le fait que nous devons déterminer l'image par de tous les couples $(\mathbf{e}_i, \mathbf{e}_j)$ par notre application ϕ.

En effet, on se rappelle que pour tout $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ et $\mathbf{y} = \sum_{i=1}^n y_i \mathbf{e}_i$, on a

$$\phi(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \phi(\mathbf{e}_i, \mathbf{e}_j).$$

Représentation matricielle II

Définition 1.37: Matrice d'une forme bilinéaire

Soit ϕ une forme bilinéaire d'un espace vectoriel E de dimension finie n et muni d'une base $\mathscr{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$. La représentation matricielle de la forme bilinéaire s est alors donnée par

$$A_\phi = \begin{pmatrix} \phi(\mathbf{e}_1, \mathbf{e}_1) & \phi(\mathbf{e}_1, \mathbf{e}_2) & \ldots & \phi(\mathbf{e}_1, \mathbf{e}_n) \\ \phi(\mathbf{e}_2, \mathbf{e}_1) & \phi(\mathbf{e}_2, \mathbf{e}_2) & \ldots & \phi(\mathbf{e}_2, \mathbf{e}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\mathbf{e}_n, \mathbf{e}_1) & \phi(\mathbf{e}_n, \mathbf{e}_2) & \ldots & \phi(\mathbf{e}_n, \mathbf{e}_n) \end{pmatrix}$$

On prendra cependant à ne pas confondre cela avec la représentation matricielle d'une application linéaires ! Bien que les objets se ressemblent, ils n'ont pas la même signification !

Représentation matricielle III

Une fois que cette représentation est adoptée, il est possible de réécrire de façon totalement matricielle notre produit scalaire.
En effet, si on considère un espace vectoriel muni d'une base et que l'on note A la représentation matricielle d'une forme bilinéaire ϕ, alors pour tout vecteur $\mathbf{x}$ et $\mathbf{y}$ de E nous avons :

$$\phi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T M \mathbf{y}.$$

Exemple I

Considérons l'application bilinéaire $\phi : E \times E \to \mathbb{R}$, où E est un espace vectoriel sur $\mathbb{R}$ muni de la base canonique de $\mathbb{R}^3$, définie par

$$\phi(\mathbf{x}, \mathbf{y}) = 3x_1y_1 + 6x_2y_1 - 3x_1y_2 - 6x_3y_1 + 2x_1y_3 + 2x_3y_3.$$

La matrice associée à cette forme bilinéaire est alors donnée par

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

La connaissance de cette représentation nous permet par exemple de pouvoir déterminer simplement l'image des vecteurs $\mathbf{x} = (-1, 0, 1)$ et $\mathbf{y} = (0, 1, 1)$

Exemple II

$$\phi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y} = (-1, 0, 1) \begin{pmatrix} 3 & -3 & 2 \\ 6 & 0 & 0 \\ -6 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 3.$$

Exemple III

Considérons l'application bilinéaire $\phi : E \times E \to \mathbb{R}$, où E est un espace vectoriel sur $\mathbb{R}$ muni de la base canonique de $\mathbb{R}^3$, définie par

$$\phi(\mathbf{x}, \mathbf{y}) = 3x_1y_1 + 2x_2y_1 + 2x_1y_2 - 3x_3y_1 + -3x_1y_3 + 2x_3y_3 + 5y_2x_2.$$

La matrice associée à cette forme bilinéaire est alors donnée par

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

On reconnaît ici la matrice d'une forme bilinéaire symétrique. La connaissance de cette représentation nous permet par exemple de pouvoir déterminer simplement l'image des vecteurs $\mathbf{x} = (-1, 0, 1)$ et $\mathbf{y} = (0, 1, 1)$

Exemple IV

$$\phi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T A \mathbf{y} = (-1, 0, 1) \begin{pmatrix} 3 & 2 & -3 \\ 2 & 5 & 0 \\ -3 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 3.$$

On pourra vérifier que $\phi(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \mathbf{y})$.

Forme quadratique I

Dans ce qui précède, nous avons supposé que les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs quelconques de $\mathbb{R}^n$, mais si l'on s'intéresse au cas particulier où $\mathbf{x} = \mathbf{y}$ on associe alors à notre forme bilinéaire ϕ ce que l'on appelle une forme quadratique.

Définition 1.38: Forme quadratique

Soit un E un espace vectoriel sur $\mathbb{K}$ et ϕ une forme bilinéaire symétrique. Alors l'application $q : E \to K$ définie par :

$$q(\mathbf{x}) = \phi(\mathbf{x}, \mathbf{x})$$

est appelée **forme quadratique** sur E associée à la forme bilinéaire ϕ.

Forme quadratique II

En tant que forme **quadratique**, l'application q précédemment définie n'est pas linéaire ! Ainsi, pour tout $\lambda \in \mathbb{K}$ l'égalité $q(\lambda \mathbf{x}) = \lambda q(\mathbf{x})$ est généralement fausse ! En effet, nous avons

$$q(\lambda \mathbf{x}) = \lambda^2 q(\mathbf{x}).$$

Forme quadratique III

Si on note $A = (a_{ij})_{i,j=1}^{n}$ la représentation matricielle de l'application bilinéaire symétrique, alors $A \in S_n(\mathbb{K})$, *i.e. l'ensemble des matrices symétriques à coefficients dans* $\mathbb{K}$, alors la forme quadratique est définie par :

$$q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i>j}^{n} a_{ij}x_i x_j.$$

Cette expression est très pratique lorsque l'on souhaite reconnaître des formes quadratiques. En effet, elle se compose de monômes de degré 2 en les composantes x_i de $\mathbf{x}$.

Exemple

Considérons l'application bilinéaire $\phi : E \times E \to \mathbb{R}$, où E est un espace vectoriel sur $\mathbb{R}$ muni de la base canonique de $\mathbb{R}^3$, définie par

$$\phi(\mathbf{x}, \mathbf{y}) = 3x_1y_1 + 2x_2y_1 + 2x_1y_2 - 3x_3y_1 + -3x_1y_3 + 2x_3y_3 + 5y_2x_2.$$

La matrice associée à cette forme bilinéaire symétrique est alors donnée par

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

La forme quadratique est donnée par
$q(\mathbf{x}) = \phi(\mathbf{x}, \mathbf{x}) = 3x_1^2 + 5x_2^2 + 2x_3^2 + 4x_1x_2 - 6x_1x_3.$

Identités de polarisation I

Si a toute forme bilinéaire symétrique il est possible d'associer une forme quadratique, la réciproque est également vraie.

Identités de polarisation II

Proposition 1.34: Identités de polarisation

Soit $q : E \to \mathbb{K}$ une forme quadratique sur E. Il existe alors une et une seule forme bilinéaire symétrique $\phi : E \times E \to \mathbb{K}$ telle que pour tout $\mathbf{x} \in E$: $q(\mathbf{x}) = \phi(\mathbf{x}, \mathbf{x})$. Elle est donnée par l'une des relations suivantes :

$$\begin{aligned}
\phi(\mathbf{x}, \mathbf{y}) &= \frac{1}{2}\left(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y})\right) \\
\phi(\mathbf{x}, \mathbf{y}) &= \frac{1}{2}\left(q(\mathbf{x}) + q(\mathbf{y}) - q(\mathbf{x}-\mathbf{y})\right) \\
\phi(\mathbf{x}, \mathbf{y}) &= \frac{1}{4}\left(q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}-\mathbf{y})\right).
\end{aligned}$$

L'application ϕ est appelée **forme polaire** de q et les relations précédentes sont appelées **identités de polarisation**.

Identités de polarisation III

Preuve

Nous démontrons uniquement la première relation, le fonctionnement est le même pour les autres identités. On se propose de faire cela de deux façons différentes.

- **A l'aide de la représentation matricielle :**
  Pour cela on va considérer la matrice $A \in S_n(\mathbb{K})$ associée à q et deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ de E. Nous avons :

$$\begin{aligned} & q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}) - q(\mathbf{y}) \\ = & \sum_{i=1}^{n} a_{ii}(x_i + y_i)^2 + 2 \sum_{i>j}^{n} a_{ij}(x_i + y_i)(x_j + y_j) \end{aligned}$$

Identités de polarisation IV

$$- \sum_{i=1}^{n} a_{ii} x_i^2 - 2 \sum_{i>j}^{n} a_{ij} x_i x_j - \sum_{i=1}^{n} a_{ii} y_i^2 - 2 \sum_{i>j}^{n} a_{ij} y_i y_j,$$

$\downarrow$ en développant le premier terme et en simplifiant les termes en bleu

$$= 2 \sum_{i=1}^{n} a_{ii} x_i y_i + 2 \sum_{i>j}^{n} a_{ij} (x_i + y_i)(x_j + y_j)$$

$$- 2 \sum_{i>j}^{n} a_{ij} x_i x_j - 2 \sum_{i>j}^{n} a_{ij} y_i y_j,$$

$\downarrow$ en développant le premier terme en rouge puis en simplifiant

$$= 2 \sum_{i=1}^{n} a_{ii} x_i y_i + 4 \sum_{i>j}^{n} a_{ij} x_i y_j,$$

$$= 2 \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j \right) = 2\phi(\mathbf{x}, \mathbf{y}).$$

Identités de polarisation V

- **A l'aide de la définition :**
  On se rappelle que ϕ est une application bilinéaire symétrique donc

$$\begin{aligned} q(\mathbf{x}+\mathbf{y}) &= \phi(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) \\ &= \phi(\mathbf{x},\mathbf{x}) + 2\phi(\mathbf{x},\mathbf{y}) + \phi(\mathbf{y},\mathbf{y}) \\ &= q(\mathbf{x}) + 2\phi(\mathbf{x},\mathbf{y}) + q(\mathbf{y}) \end{aligned}$$

On obtient la deuxième relation en changeant $\mathbf{y}$ en $-\mathbf{y}$ dans la démonstration précédente.
Enfin, la troisième relation s'obtient en calculant

$$q(\mathbf{x}+\mathbf{y}) - q(\mathbf{x}-\mathbf{y}) = \phi(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y}) - \phi(\mathbf{x}-\mathbf{y},\mathbf{x}-\mathbf{y}).$$

Propriétés des formes quadratiques I

Les formes bilinéaires peuvent être représentées par des matrices, il en est donc de même pour les formes quadratiques qui admettront exactement la même représentation que la forme bilinéaire associée, comme nous avons pu le voir dans les exemples précédentes. On continuera d'appeler cette matrice A et cette dernière sera toujours symétrique !

Comme il s'agit d'une matrice, il est possible de diagonaliser cette matrice est de déterminer ses valeurs propres. Cela se révèle d'un grand intérêt dans le domaine de l'optimisation afin de dériver des propriétés d'un algorithme lorsque ces formes quadratiques sont étudiées en tant que fonction. On va donc rapidement lister quelques résultats intéressants sur ces dernières.

Propriétés des formes quadratiques II

Définition 1.39: Caractérisation des formes quadratiques

Soit E un espace vectoriel sur $\mathbb{K}$ et q une forme quadratique sur E. Notons A sa représentation matricielle. Alors q est dite :

- *positive* (resp. *définie positive*) si pour tout $\mathbf{x} \in E$ non nul :

  $$q(\mathbf{x}) \geq 0 \quad \text{ou} \quad \mathbf{x}^T A\mathbf{x} \geq 0 \quad (\text{resp.} \quad q(\mathbf{x}) > 0 \quad \text{ou} \quad \mathbf{x}^T A\mathbf{x} > 0).$$

  On dit parfois que la matrice associée à q est SDP : Semi-Définie Positive (resp. DP : Définie Positive).

- *négative* (resp. *définie négative*) si pour tout $\mathbf{x} \in E$ non nul :

  $$q(\mathbf{x}) \leq 0 \quad \text{ou} \quad \mathbf{x}^T A\mathbf{x} \leq 0 \quad (\text{resp.} \quad q(\mathbf{x}) < 0 \quad \text{ou} \quad \mathbf{x}^T A\mathbf{x} < 0).$$

  On dit parfois que la matrice associée à q est SDN : Semi-Définie Négative (resp. DP : Définie Négative).

Espace euclidien I

A partir de maintenant on se place uniquement dans des espaces vectoriels réels. Ainsi, E désignera toujours un espace vectoriel sur $\mathbb{R}$.

Définition 1.40: Produit scalaire

Soit E un espace vectoriel sur $\mathbb{K}$. On appelle **produit scalaire** sur E, tout forme bilinéaire ϕ *symétrique et définie positive* sur E :

- **Bilinéaire :** voir Définition 1.35
- **Symétrique :** $\phi(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{y}, \mathbf{x})$
- **Définie positive :** $\phi(\mathbf{x}, \mathbf{x}) > 0$ et $\phi(\mathbf{x}, \mathbf{x}) = 0 \iff \mathbf{x} = 0.$

Regardons un petit exemple pour concrétiser cela.

Espace euclidien II

Soit E un espace vectoriel et montrons que l'application $\phi : E \times E \to \mathbb{R}$ définie par

$$\phi(\mathbf{x}, \mathbf{y}) = x_1y_1 + 3x_1y_2 + 3x_2y_1 + 10x_2y_2,$$

définie un produit scalaire. On va donc vérifier les trois points de la définition.

- L'application est bien symétrique. En effet :

$$\phi(\mathbf{x}, \mathbf{y}) = x_1y_1 + 3x_1y_2 + 3x_2y_1 + 10x_2y_2 = \phi(\mathbf{y}, \mathbf{x}).$$

- Elle est bien définie positive. En effet

$$\phi(\mathbf{x}, \mathbf{x}) = x_1^2 + 6x_1x_2 + 9x_2^2 = (x_1 + 3x_2)^2 + x_2^2.$$

Espace euclidien III

- **Bilinéarité :** Soient $\mathbf{x}, \mathbf{y}$ et $\mathbf{x}$ des éléments de E et $\lambda \in \mathbb{K}$ alors :

$$\begin{aligned}
\phi(\mathbf{x}+\lambda\mathbf{z}, \mathbf{y}) &= (x_1+\lambda z_1)y_1 + 3(x_1+\lambda z_1)y_2 + 3(x_2+\lambda z_2)y_1 \\
&\quad + 9(x_2+\lambda z_2)y_2 \\
&\quad \downarrow \text{ on développe} \\
&= x_1y_1 + 3x_1y_2 + 3x_2y_1 + 9x_2y_2 \\
&\quad + \lambda(z_1y_1 + 3z_1y_2 + 3z_2y_1 + 9z_2y_2), \\
&\quad \downarrow \text{ définition de } \phi \\
&= \phi(\mathbf{x}, \mathbf{y}) + \lambda\phi(\mathbf{z}, \mathbf{y}).
\end{aligned}$$

On conclut qu'elle est au linéaire à droite en utilisant la symétrie.

Espace euclidien IV

Le résultat suivant permet de majorer un produit scalaire entre deux vecteurs en fonction de leurs normes.

Proposition 1.35: Inégalité de Cauchy-Schwartz

Soit E un espace vectoriel. Alors pour tout vecteur $\mathbf{x}$ et $\mathbf{y}$ de E nous avons

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

De plus, l'égalité est atteinte lorsque les vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont colinéaires.

Preuve : Soit $(\mathbf{x}, \mathbf{y}) \in E^2$ et soit $\lambda \in \mathbb{R}$. Comme le produit scalaire est une forme bilinéaire symétrique définie positive, nous avons

Espace euclidien V

$$\langle \mathbf{x} + \lambda \mathbf{y}, \mathbf{x} + \lambda \mathbf{y} \rangle \geq 0.$$

Or

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

Le cas $\mathbf{x}$ ou $\mathbf{y}$ est trivial, on va donc supposer que les deux vecteurs sont nuls dans la suite. L'expression précédente peut donc se voir comme un polynôme de degré 2 en λ. Or $\langle \mathbf{x}, \mathbf{x} \rangle + 2\lambda \langle \mathbf{x}, \mathbf{y} \rangle + \lambda^2 \langle \mathbf{y}, \mathbf{y} \rangle \geq 0$ et $\langle \mathbf{y}, \mathbf{y} \rangle > 0$, donc le discriminant de ce polynôme

$$4\langle \mathbf{x}, \mathbf{y} \rangle^2 - 4\langle \mathbf{y}, \mathbf{y} \rangle \langle \mathbf{x}, \mathbf{x} \rangle$$

est négatif ou nul. On en déduit directement le résultat.

Espace euclidien VI

Définition 1.41: Espace euclidien

On appelle **espace euclidien** tout espace vectoriel réel de dimension finie muni d'un produit scalaire.

Un produit scalaire étant par définition une forme bilinéaire symétrique définie positive, il est donc possible de définir plusieurs produits scalaires sur un même espace vectoriel (une infinité même) vu que l'on peut définir une infinité de forme bilinéaire symétrique définie positive.

Espace euclidien VII

Définition 1.42: Norme Euclidienne

On appelle **norme euclidienne** associée au produit scalaire $\langle \cdot, \cdot \rangle$, notée $\| \cdot \|$, l'application définie de E dans $\mathbb{R}_+$ pour tout $\mathbf{x} \in E$ par

$$\|\mathbf{x}\|_2^2 = \langle \mathbf{x}, \mathbf{x} \rangle.$$

Remarque : La définition précédente ne définit qu'un exemple particulier, mais il existe une définition plus générale de normes que vous serez amenés à rencontrer lorsque l'on fait du calcul matriciel ou encore de la modélisation statistique (*ex :* norme matricielle ou encore les normes L^p pour des méthodes pénalisées).

Espace euclidien VIII

Définition 1.43: Norme

On appelle **norme** sur un espace vectoriel E toute application N définie de E dans $\mathbb{R}_+$ vérifiant les axiomes suivants :

(i) $\forall(\alpha, \mathbf{x}) \in \mathbb{R} \times E,\ N(\alpha\mathbf{x}) = |\alpha| N(\mathbf{x})$ (homogénéité),

(ii) $\forall \mathbf{x} \in E,\ N(\mathbf{x}) = 0 \iff \mathbf{x} = 0$ (séparation),

(iii) $\forall(\mathbf{x}, \mathbf{x}') \in E \times E,\ N(\mathbf{x} + \mathbf{x}') \leq N(\mathbf{x}) + N(\mathbf{x}')$ (inégalité triangulaire).

Espace euclidien IX

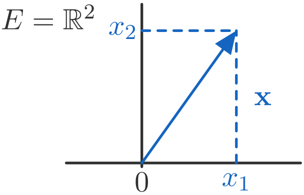
On donne ici quelques exemples de normes :

- L'espace $\mathbb{R}$ muni de la valeur absolue est un espace euclidien, ce qui veut dire dire que l'application $N(\cdot) = |\cdot|$
- La norme euclidienne définie précédemment est une norme, plus précisément, nous avons

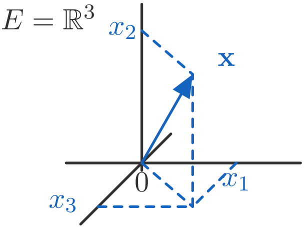
$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Cette norme correspond à un résultat que l'on connaît très bien en dimension 2 et le généralise à la dimension n ... **le théorème de Pythagore !**

Espace euclidien X



$$\|x\| = \sqrt{x_1^2 + x_2^2}$$



$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$$

Vous pouvez essayer de montrer que la norme euclidienne est effectivement une norme en montrant que les trois axiomes sont vérifiés.

Espace euclidien XI

- Plus généralement, toutes les applications $N : \mathbf{x} \mapsto \|\mathbf{x}\|_p$, également appelées p-norme, sont des normes.

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

Un exercice intéressant est de représenter, dans un plan, l'ensemble des vecteurs ayant une norme égale à 1 pour différentes valeurs de p.
C'est ce que l'on appelle la *boule unité* et on peut voir que sa forme varie en fonction de p.

Espace euclidien XII

L'exemple de la norme euclidienne et plus précisément le théorème de Pythagore permet de revenir sur un point important que l'on a déjà abordé précédemment : la notion d'*orthogonalité*.

Définition 1.44: Orthogonalité

Soit E un espace euclidien muni du produit scalaire $\langle \cdot, \cdot \rangle$ et $\mathbf{v}, \mathbf{w}$ deux vecteurs de E. On dit que $\mathbf{v}$ et $\mathbf{w}$ sont orthogonaux si

$$\langle \mathbf{v}, \mathbf{w} \rangle = 0.$$

Espace euclidien XIII

Le résultat suivant permet également de montrer que deux vecteurs sont orthogonaux uniquement en regardant des normes.

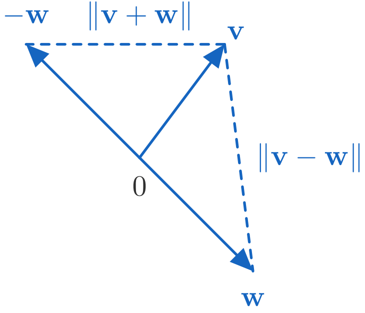
Proposition 1.36: Caractérisation de l'Orthogonalité

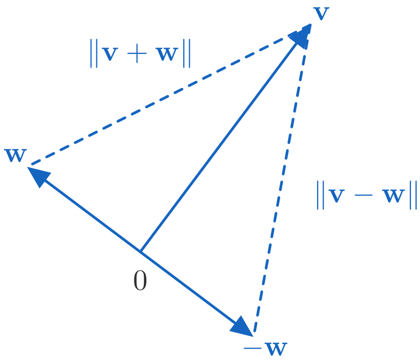
Soit E un espace euclidien muni du produit scalaire $\langle \cdot, \cdot \rangle$, $\| \cdot \|$ la norme associée et $\mathbf{v}, \mathbf{w}$ deux vecteurs de E. Les vecteurs $\mathbf{v}$ et $\mathbf{w}$ sont orthogonaux si et seulement si

$$\|\mathbf{v} - \mathbf{w}\| = \|\mathbf{v} + \mathbf{w}\|.$$

La preuve de ce résultat repose sur le lien entre norme et produit scalaire (Exercice).

Espace euclidien XIV

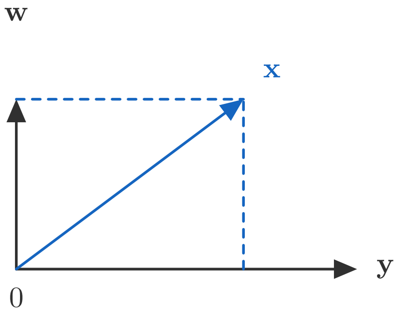




Espace euclidien XV

A partir de cette notion d'orthogonalité, il est également possible de définir l'angle entre deux vecteurs d'un espace euclidien E. Pour cela regardons un petit exemple pour mieux comprendre la construction. Considérons deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ définis dans une certaine base $\mathscr{B}$. On cherche à déterminer l'angle que fait le vecteur $\mathbf{x}$ avec le vecteur $\mathbf{y}$ en utilisant la notion d'orthogonalité vu précédemment.

Espace euclidien XVI



Le vecteur $\mathbf{x}$ peut se décomposer comme la somme de vecteurs orthogonaux, *i.e.* il existe un vecteur $\mathbf{w}$ orthogonal à $\mathbf{y}$ ainsi qu'un scalaire λ tels que :

$$\mathbf{x} = \lambda \mathbf{y} + \mathbf{w}.$$

Espace euclidien XVII

Rappelez-vous, c'est une conséquence des résultats vus précédemment dans cette partie ! Le vecteur $\lambda\mathbf{y}$ est appelé projeté orthogonal de $\mathbf{x}$ sur $\mathbf{y}$. Il nous faut alors déterminer la valeur du scalaire λ. Pour cela, on se rappelle que $\mathbf{y}$ est orthogonal à $\mathbf{w}$, donc $\langle \mathbf{x} - \lambda\mathbf{y}, \mathbf{y} \rangle = 0$, ce qui donne

$$\lambda = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2}.$$

Cette valeur λ qui représente le coefficient de la projection de $\mathbf{x}$ sur $\mathbf{y}$ est fortement liée à l'angle entre les deux vecteurs. Plus précisément

Espace euclidien XVIII

Définition 1.45: Angles entre deux vecteurs

Soient $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs d'un espace euclidien E et notons $\mathbf{x}'$ et $\mathbf{y}'$ les normalisations de ces deux vecteurs, *i.e.*

$$\mathbf{x}' = \frac{\mathbf{x}}{\|\mathbf{x}\|^2} \quad \text{et} \quad \mathbf{y}' = \frac{\mathbf{y}}{\|\mathbf{y}\|^2}.$$

Alors le cosinus de l'angle formé entre les deux vecteurs $\mathbf{x}$ et $\mathbf{y}$, noté $\lambda = \cos(\theta)$, est donné par

$$\lambda = \frac{\langle \mathbf{x}', \mathbf{y}' \rangle}{\|\mathbf{x}'\|^2} = \langle \mathbf{x}', \mathbf{y}' \rangle = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}.$$

C'est donc une valeur comprise entre -1 et 1. En effet, cela est une conséquence de l'inégalité de Cauchy-Schwarz.

Espace euclidien XIX

La définition 1.40 a permis d'introduire le produit scalaire comme une forme **bilinéaire, symétrique et définie positive**. Nous pouvons donc représenter ce produit scalaire à l'aide d'une matrice, si notre espace E est muni d'une base (voir définition 1.37).

Proposition 1.37: Représentation produit scalaire

Soit s un produit scalaire sur un espace euclidien E muni d'une base $(\mathbf{e}_i)_{i\in\mathbb{N}}$ et notons S la représentation de s dans cette base. Considérons deux vecteurs $\mathbf{x}$ et $\mathbf{x}'$ de E alors

$$s(\mathbf{x},\mathbf{y}) = \langle \mathbf{x},\mathbf{y}\rangle_s = \mathbf{x}^T S\mathbf{y}.$$

Espace euclidien XX

En tant que matrice représentante d'un produit scalaire, cette matrice S doit vérifier les mêmes points que la définition de produit scalaire, *i.e.* elle doit être :

- **bilinéaire** ,
- **symétrique** : $S = S^T$,
- **définie positive** : $\forall \mathbf{x} \in E \quad \mathbf{x}^T S \mathbf{x} \geq 0$ et $\mathbf{x}^T S \mathbf{x} = 0 \iff \mathbf{x} = 0$

Remarque. La matrice associée à un produit scalaire est diagonalisable en tant que matrice symétrique réelle.

Espace euclidien XXI

Un premier exemple bien connu de produit scalaire est celui employé jusqu'à présent, c'est-à-dire le produit scalaire canonique de $\mathbb{R}^n$ pour lequel le produit scalaire est communément appelé produit *scalaire euclidien* et qui peut se représenter sous la forme

$$\langle \mathbf{x}, \mathbf{y} \rangle_s = \mathbf{x}^T I \mathbf{x} = \mathbf{x}^T \mathbf{y}, \quad \text{dans ce cas} \quad S = I.$$

Espace euclidien XXII

On peut également définir un produit scalaire sur les matrices, par exemple sur l'ensemble des matrices carrées d'ordre n, $\mathscr{M}_n(\mathbb{R})$. Le produit scalaire ainsi défini s'appelle le *produit scalaire de Frobenius*, noté $\langle \cdot, \cdot \rangle_F$ (auquel on pourra également rattacher une norme, la norme de Frobenius).
Etant données deux matrices carrées A et B d'ordre n, dont les coefficients sont respectivement notées a_{ij} et b_{ij}, alors le produit scalaire de Frobenius est défini par :

$$\langle A, B \rangle_F = Tr(A^T B) = \sum_{i,j=1}^{n} a_{ij} b_{ij}.$$

Orthogonalité I

Définition 1.46: Base orthogonale/orthonormée

Soit E un espace euclidien muni d'une base $\mathscr{B} = (\mathbf{e}_i)_{i=1}^n$. Cette base est dite orthogonale si pour tout i, j tels que $i \neq j$ on a

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = 0.$$

De plus, cette base est dite orthonormée si elle vérifie

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{si i=j,} \\ 0 & \text{sinon.} \end{cases}$$

Orthogonalité II

Dans une base orthonormale, le produit scalaire s'exprime de façon très simple et semblable au produit scalaire canonique. En effet, considérons $(\mathbf{e}_i)_{i=1}^n$ une base orthonormale de E et $\mathbf{x}$ et $\mathbf{y}$ deux vecteurs de E exprimés dans cette base. Alors

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i,j=1}^{n} \langle x_i \mathbf{e}_i, y_j \mathbf{e}_j \rangle = \sum_{i,j=1}^{n} x_i y_j \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \sum_{i=1}^{n} x_i y_i.$$

Orthogonalité III

On peut maintenant retourner à l'étude de nos endomorphismes symétriques réelles dont le résultat suivant va jouer un rôle majeur dans certaines méthodes d'analyses de données.

Proposition 1.38: Endomorphismes symétriques et diagonalisation

Soit ϕ une forme bilinéaire symétrique sur un espace euclidien E, alors il existe une base orthonormée de E dans laquelle la matrice de ϕ est diagonale.

Orthogonalité IV

Bien que nous ne montrons pas ce résultat cela nous permet de justifier l'existence du résultat suivant permettant de caractériser les formes quadratiques relativement au signe de ses valeurs propres.

Proposition 1.39: Caractérisation des formes quadratiques

Soient E un espace vectoriel sur $\mathbb{K}$ et q une forme quadratique sur E dont la représentation matricielle est notée A. Alors

- q est semi-definie positive (*resp.* définie positive) si et seulement si ses valeurs propres sont positives (*resp.* strictement positives)
- q est semi-definie négative (*resp.* définie négative) si et seulement si ses valeurs propres sont négatives (*resp.* strictement négatives).

Orthogonalité V

Définition 1.47: Sous espace orthogonal

Soit E un espace vectoriel euclidien et F un sous-espace quelconque de E strictement inclus dans E. On appelle sous sous espace orthogonal de F dans E, noté $F^{\perp}$, l'ensemble des vecteurs $\mathbf{x}$ de E qui sont orthogonaux à tout vecteur de F

$$F^{\perp} = \{\mathbf{x} \in E \mid \forall \mathbf{z} \in F \langle \mathbf{x}, \mathbf{z} \rangle = 0\}.$$

On peut montrer que $F^{\perp}$ est un sous-espace vectoriel de E.

Orthogonalité VI

Proposition 1.40: Propriétés espaces orthogonaux

Soit E un espace vectoriel euclidien et F un sous espace de E. Alors

- $dim(E) = dim(F) + dim(F^{\perp})$,
- $E = F \oplus F^{\perp}$,
- $\left(F^{\perp}\right)^{\perp} = F$.

Nous avons déjà construits de tels sous-espaces précédemment à l'aide de certaines applications. Souvenez vous des projecteurs orthognonaux étudiez au tout début de cette partie (voir la Définition 1.7 et la Proposition 1.5).

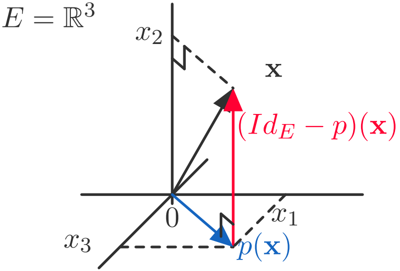
Orthogonalité VII

Pour rappel, ces résultats ont montré que le noyau et l'image d'un projecteur sont supplémentaires dans l'espace de départ. De plus, nous avons

$$Ker(p) = Im(Id_E - p) = (Im(p))^{\perp}\,.$$

On rappelle également que le projecteur est une application idempotente (*i.e.* $p \circ p = p$).

Orthogonalité VIII



On considère un projecteur p sur l'espace engendré par les vecteurs de base $(\mathbf{e}_1, \mathbf{e}_3)$ parallèlement à la droite vectorielle engendrée par le vecteur $\mathbf{e}_2$. Nous avons donc

$$Ker(p) = Vect(\mathbf{e}_2) = Im(Id_E - p) \quad \text{et} \quad Im(p) = Ker(Id_E - p) = Vect(\mathbf{e}_1, \mathbf{e}_3).$$

Orthogonalité IX

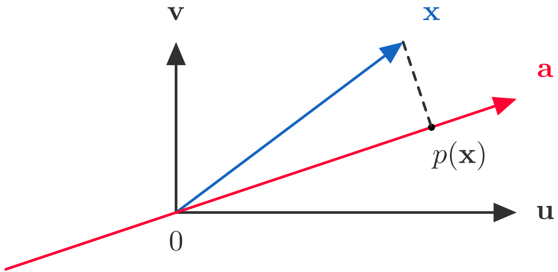
Nous terminons cette section en regardant comment déterminer le projeté orthogonal d'un vecteur sur un sous espace vectoriel d'un espace vectoriel euclidien.

Pour cela, considérons un espace euclidien E de dimension n et F un sous-espace de E de dimension $p < n$.

Soit $\mathbf{x}$ un vecteur de E. Nous allons regarder comment déterminer le projeté de $\mathbf{x}$ sur une droite vectorielle (*i.e.* le cas où F est un espace de dimension 1) puis on traitera le cas plus général où la dimension de F est $2 \leq p < n$.

Orthogonalité X

- **Projection sur une droite :** on considère la droite vectorielle F engendrée par un vecteur $\mathbf{a}$ et on cherche à déterminer le projeté de $\mathbf{x}$ sur $\mathbf{a}$.



$p(\mathbf{x})$ correspond au projeté du vecteur $\mathbf{x}$ sur la droite vectorielle engendrée par $\mathbf{a}$. En ce sens $p(\mathbf{x})$ peut s'interpréter comme le point

Orthogonalité XI

de $Vect(\mathbf{a})$ le plus proche de $\mathbf{x}$, *i.e.* cela reviendrait à chercher les coordonnées α sur le vecteur $\mathbf{a}$ qui soit solution du problème suivant :

$$\min_{\alpha \in \mathbb{R}} \|\mathbf{x} - \alpha \mathbf{a}\|_2^2,$$

où $\alpha \mathbf{a} = p(\mathbf{x})$.
Plutôt que de chercher à résoudre ce problème d'optimisation dont la résolution sera traitée dans un cours dédié, nous allons simplement nous contenter des outils de l'algèbre linéaire.

Pour cela, utilisons simplement le fait que $p(\mathbf{x}) = \alpha \mathbf{a}$ comme énoncé plus tôt (*i.e.* le projeté appartient à la droite vectorielle engendrée par $\mathbf{a}$).

Orthogonalité XII

Dans ce cas, les vecteurs $\mathbf{a}$ et $\mathbf{x} - \alpha\mathbf{a}$ sont orthogonaux, d'où

$$\begin{array}{rcl} \langle \mathbf{a}, \mathbf{x} - \alpha\mathbf{a} \rangle & = & 0, \\ \alpha\langle \mathbf{a}, \mathbf{a} \rangle & = & \langle \mathbf{a}, \mathbf{x} \rangle, \\ \alpha & = & \dfrac{\langle \mathbf{a}, \mathbf{x} \rangle}{\langle \mathbf{a}, \mathbf{a} \rangle}, \\ \alpha & = & \dfrac{\langle \mathbf{a}, \mathbf{x} \rangle}{\|\mathbf{a}\|^2}. \end{array}$$

On aura simplement développer les expressions.

Ainsi $p(\mathbf{x}) = \alpha\mathbf{a} = \dfrac{\langle \mathbf{a}, \mathbf{x} \rangle}{\|\mathbf{a}\|^2}\mathbf{a}$.

Remarquons que seul le vecteur $\mathbf{x}$ a une influence sur l'expression $p(\mathbf{x})$ mais pas le vecteur $\mathbf{a}$ qui ne sert qu'à donner la direction, *i.e.*

Orthogonalité XIII

multiplier le vecteur $\mathbf{a}$ par λ ne change pas l'expression du projeté, mais multiplier $\mathbf{x}$ par λ implique de multiplier l'expression de $p(\mathbf{x})$ par λ.

Dernière remarque, on peut réécrire l'expression de $p(\mathbf{x})$ sous forme matricielle, en remarquant que $p(\mathbf{x}) = \dfrac{\mathbf{a}\mathbf{a}^T}{\|\mathbf{a}\|^2}\mathbf{x} = P\mathbf{x}$ où $P = \dfrac{\mathbf{a}\mathbf{a}^T}{\|\mathbf{a}\|^2}$.
Dans ce cas, la matrice P définie bien une matrice projection (de rang 1) car

$$P^2 = P \quad \text{et} \quad P^T = P.$$

Généralisons cela pour un sous espace de dimension $2 \leq p < n$.

Orthogonalité XIV

- **Projection sur un hyperplan :** Cette fois-ci F est un sous-espace vectoriel de E de dimension $2 \leq p < n$ généré par les vecteurs $\mathbf{a}_1, \cdots, \mathbf{a}_p$ de $\mathbb{R}^n$ qui forment une base de F.

  Dans ce cas, l'espace F est l'espace générée par les colonnes de la matrice $A = (\mathbf{a}_1, \cdots, \mathbf{a}_p)$.

  Nous avons précédemment vu que le projeté de $\mathbf{x}$ sur le sous espace F pouvait s'écrire $p(\mathbf{x}) = P\mathbf{x}$ où P est une certaine matrice, mais surtout, on sait que le projeté va s'écrire comme une combinaison linéaire des vecteurs colonnes de A, *i.e.* $p(\mathbf{x}) = \sum_{j=1}^{p} \alpha_j \mathbf{a}_j = A\boldsymbol{\alpha}$.

  On peut alors utiliser le même raisonnement que précédemment, on sait que le vecteur $\mathbf{x} - p(\mathbf{x})$ est orthogonal à tout vecteur de F, *i.e.*

Orthogonalité XV

$$
\begin{aligned}
\langle \mathbf{a}_1, \mathbf{x} - A\boldsymbol{\alpha} \rangle &= 0, \\
&\vdots \\
\langle \mathbf{a}_p, \mathbf{x} - A\boldsymbol{\alpha} \rangle &= 0.
\end{aligned}
$$

Ce qui peut s'écrire de façon matricielle :

$$\langle A, \mathbf{x} - A\boldsymbol{\alpha} \rangle = A^T(\mathbf{x} - A\boldsymbol{\alpha}) = 0.$$

Notre objectif est de trouver le vecteur $\boldsymbol{\alpha}$, on va donc l'isoler.

Orthogonalité XVI

$$\begin{aligned} A^T(\mathbf{x} - A\boldsymbol{\alpha}) &= 0, \\ A^T A\boldsymbol{\alpha} &= A^T\mathbf{x} \end{aligned}$$

Comme les colonnes de A forment une base de F la matrice $A^T A \in \mathscr{M}_p(\mathbb{R})$ est inversible. On a donc

$$\boldsymbol{\alpha} = \left(A^T A\right)^{-1} A^T\mathbf{x}.$$

Donc le projeté orthogonal de $\mathbf{x}$ sur F est $p(\mathbf{x}) = A\left(A^T A\right)^{-1} A^T\mathbf{x}$.

Orthogonalité XVII

Proposition 1.41: Projection orthogonal sur un sous-espace

Soit E un espace euclidien de dimension n et F un sous espace de E de dimension $p < n$ dont les vecteurs $\mathbf{a}_1, \cdots, \mathbf{a}_p$ forment une base. On note A la matrice dont les colonnes sont formées des vecteurs de cette base.

Alors, pour tout vecteur $\mathbf{x}$ de E, le projeté orthogonal $p(\mathbf{x})$ de $\mathbf{x}$ sur F est donné par

$$p(\mathbf{x}) = A\left(A^T A\right)^{-1} A^T \mathbf{x}$$

et $P = \left(A^T A\right)^{-1} A^T$ est la matrice de projection sur le sous-espace F.

The End ... enfin pour cette partie !