ALGÈBRE III

S. LE BORGNE [1]

1. Ces notes sont basées sur celles de B. Schapira, responsable du cours ces dernières années ; l'image de première page est due à IL.

Ces notes sont provisoires. Elles vous donneront quand même une vision globale du cours. J'y ajouterai sans doute quelques corrigés d'exercices et je compléterai certaines parties. Le plan, lui, ne changera plus avant la fin de l'année. Merci de me signaler les erreurs que vous pourriez y trouver.

Table des matières

1 Matrices, espaces vectoriels, applications linéaires

1.1 Matrices

Définition 1.1. *Une matrice est un tableau rectangulaire de nombres (entiers, réels, complexes).*

Exemples 1.2.

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \quad c = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad b = \begin{pmatrix} 4 \\ 10 \\ 5 \end{pmatrix}.$$

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

Définition 1.3. *La taille d'une matrice est son nombre de lignes et son nombre de colonnes. Si une matrice A a n lignes et m colonnes, on dit que A est de taille $n \times m$. Lorsqu'une matrice n'a qu'une ligne on dit que c'est un vecteur ligne, lorsqu'elle n'a qu'une colonne que c'est un vecteur colonne.*

Exemples 1.4. *Pour les exemples ci-dessus : A est de taille 3×5, C est de taille 2×2, T est de taille 5×5, x et c sont de taille 5×1, b est de taille 3×1. Les matrices x, c, b sont des vecteurs colonnes.*

Définition 1.5. *Une matrice est dite carrée si elle a autant de lignes que de colonnes. Une matrice carrée A est dite diagonale si $a_{ij} = 0$ si $i \neq j$. Elle est dite triangulaire supérieure si $a_{ij} = 0$ si $i > j$. Elle est dite triangulaire inférieure si $a_{ij} = 0$ si $i < j$.*

Exemples 1.6. *Pour les exemples ci-dessus : les matrices C, T et D sont carrées ; T est triangulaire supérieure, D est diagonale.*

Définition 1.7. *Soit A une matrice. La transposée de A est la matrice ${}^tA$ dont les coefficients sont donnés par*

$$({}^tA)_{ij} = A_{ji}.$$

La transposée de A est la matrice dont les lignes sont les colonnes de A.

Exemple 1.8. *La transposée de la matrice A prise en exemple ci-dessus est*

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

Définition 1.9. *Soient A et B deux matrices* **de même taille**. *La somme de A et B est la matrice $A+B$ dont les coefficients sont donnés par*

$$ (A+B)_{ij} = A_{ij} + B_{ij}. $$

Exemple 1.10.

Exemple 1.11. *Soient les deux matrices suivantes :*

$$ C = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}. $$

Leur somme est

$$ C+E = \begin{pmatrix} 0+2 & 1+0 \\ -1+1 & 0+1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}. $$

Proposition 1.12. *L'addition est associative et commutative ; d'autre part la matrice nulle (c'est-à-dire dont les coefficients sont tous nuls) est un élément neutre pour l'addition : si A, B, C sont trois matrices de même taille et 0 est la matrice nulle de cette taille, on a*

$$ (A+B)+C = A+(B+C), \quad A+B = B+A, \quad A+0 = A. $$

Définition 1.13. *Soient A et B deux matrices de tailles $m \times n$ et $n \times p$ respectivement (autrement dit A a autant de colonnes que B a de lignes). Le produit AB de A et B est la matrice dont les coefficients sont donnés par*

$$ (AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}. $$

La matrice AB est de taille $m \times p$.

Exemples 1.14. *Reprenons les deux matrices C et E de l'exemple précédent. Elles sont de taille 2×2. Les deux produits CE et EC sont définis :*

$$ CE = \begin{pmatrix} 0.2+1.1 & 0.0+1.1 \\ (-1).2+0.1 & (-1).0+0.1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2 & 0 \end{pmatrix}, $$

$$ EC = \begin{pmatrix} 2.0+0.(-1) & 2.1+0.0 \\ 1.0+1.(-1) & 1.1+1.0 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix}. $$

On voit en passant que le produit n'est pas commutatif.

Les vecteurs de la base canonique de $\mathbb{R}^5$ sont

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad e_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad e_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Les vecteurs colonnes de A (la matrice prise en exemple plus haut) sont les résultats de la multiplication de A par ces vecteurs de la base canonique. Par exemple

$$A.e_2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

$$A.e_4 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

On peut obtenir de manière analogue les vecteurs ligne en multipliant à gauche par des vecteurs lignes ayant une seule coordonnées 1 et les autres 0 :

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

Soient A et B deux matrices de tailles $m \times n$ et $n \times p$ respectivement. Le produit AB est bien défini (de taille $m \times p$). Appelons $V_1, \ldots, V_p$ les vecteurs colonne de B. Alors le j-ème vecteur colonne de AB est AV_j.

Exemple 1.15. *Considérons les matrices*

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Calculons AB

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

Le deuxième vecteur colonne de B est

$$V_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

et on a

$$A.V_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix} . \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}.$$

Le deuxième vecteur colonne de AB est bien égal à AV_2.

Soit $A = (a_{ij})_{1 \leq i \leq n, 1 \leq j \leq m}$ et $X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} \in \mathbb{R}^m$. Alors on définit

$$A.X = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2m}x_m \\ \dots \\ a_{n1}x_1 + +a_{n2}x_2 + \dots a_{nm}x_m \end{pmatrix}$$

Le vecteur (colonne) $A.X$ est dans $\mathbb{R}^n$.

Définition 1.16. *On appelle matrice identité une matrice carrée diagonale dont les éléments diagonaux sont 1. On note ces matrices I_2, I_3,... (l'indice donne la taille) ou I (si la taille n'est pas ou n'a pas besoin d'être précisée).*

Par exemple, on a

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Proposition 1.17. *La multiplication de matrices est associative et les matrices identité sont des éléments neutres pour la multiplication : si A, B, C sont trois matrices de tailles $m \times n$, $n \times p$ et $p \times q$ (autrement de tailles assurant que les produits écrits ci-dessous soient définis), on a*

$$(AB)C = A(BC), \quad I_m A = A = AI_n.$$

L'associativité permet d'écrire un produit de plusieurs matrices sans utiliser de parenthèses : par exemple ABC (car le résultat est le même si on calcule $(AB)C$ (d'abord AB) ou $A(BC)$ (d'abord BC)).

Proposition 1.18. *La multiplication de matrices est par rapport à l'addition : si A, B, C sont trois matrices de tailles $m \times n$, $n \times p$ et $n \times p$ (autrement de tailles assurant que les produits et sommes écrits ci-dessous soient définis), on a*

$$A(B+C) = AB + AC.$$

Définition 1.19. *On dit qu'une matrice* **carrée** *A est inversible s'il existe une matrice B (de même taille) telle que*

$$AB = BA = I$$

Cette matrice B s'appelle alors l'inverse de A et est notée A^{-1}.

Proposition 1.20. *Soient A et B deux matrices inversibles de même taille. Le produit AB est inversible et*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Pour le voir, il suffit d'écrire le produit

$$AB(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I,$$

la première égalité vient du fait que le produit est associatif; en général on ne peut pas changer l'ordre des termes d'un produit, mais on peut grouper les élément d'un produit comme on le souhaite sans changer le résultat.

Si A est une matrice carrée on peut la multiplier par elle-même autant de fois qu'on veut; on obtient ainsi les puissances de A, notées A^2, A^3, etc... A première vue il est difficile de calculer les puissances d'une matrice. Pour certaines matrices c'est facile, par exemple pour les matrices diagonales :

$$D = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}, \quad D^k = \begin{pmatrix} a^k & 0 & 0 \\ 0 & b^k & 0 \\ 0 & 0 & c^k \end{pmatrix}.$$

On dit qu'une matrice carrée A est diagonalisable si on peut trouver deux matrices P et D de même taille que A telles que P soit inversible, D soit diagonale et

$$D = P^{-1}AP.$$

C'est alors facile de calculer les puissances de A. Remarquons que l'égalité précédente permet d'écrire A en fonction de D. En multipliant à gauche par P et à droite par P^{-1}, on obtient

$$PDP^{-1} = P(P^{-1}AP)P^{-1} = (PP^{-1})A(PP^{-1}) = IAI = A.$$

Calculons maintenant les puissances de A :

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PDDP^{-1} = PD^2P^{-1},$$

$$A^3 = AA^2 = (PDP^{-1})(PD^2P^{-1}) = PD(P^{-1}P)D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1},$$

puis, par récurrence, pour tout entier $k \geq 1$,

$$A^k = PD^kP^{-1}.$$

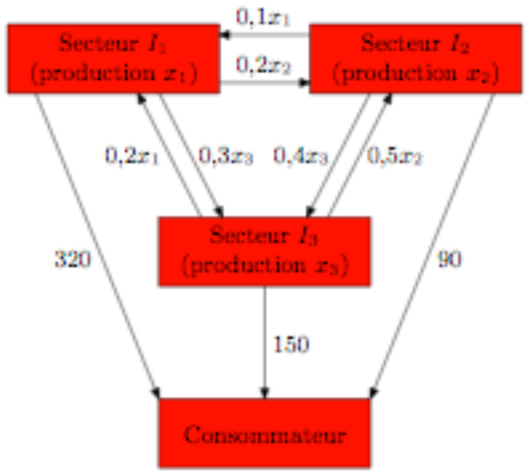
Or on a vu qu'il était facile de calculer les puissances d'une matrice diagonale.

Dans la suite, plutôt que de dire qu'une matrice A est de taille $n \times m$, on dira parfois $A \in M_{n\times m}(\mathbb{R})$. Autrement dit on désignera par $M_{n\times m}(\mathbb{R})$ l'ensemble des matrices A de taille $n \times m$.

1.2 Quelques exemples d'utilisation des matrices

1.2.1 Matrice de Léontief (ou d'entrée-sortie)

1. Considérons une économie avec trois secteurs industriels, I_1 , I_2 , I_3. Quelles quantités x_1, x_2, x_3 doivent-elles produire pour satisfaire à la fois la demande des consommateurs et celle des autres secteurs ? La demande requise de chaque secteur est représentée sur la figure ci-dessous.



2. Lorsqu'on considère un modèle entrée-sortie avec plus de trois secteurs industriels, il devient malaisé de représenter les demandes par un diagramme comme celui de la question précédente. Supposons qu'il y ait des secteurs $I_1, \ldots, I_n$, de productions $x_1, ..., x_n$. Les vecteurs de production x, de demande des consommateurs b et de demande pour le secteur I_j sont

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{pmatrix}, \quad v_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ a_{nj} \end{pmatrix},$$

où b_i est la demande des consommateurs sur le secteur $I_i$et a_{ij} est la demande du secteur I_j au secteur I_i, par euro de production du secteur I_j.

a. Trouver les quatre vecteurs de demandes pour l'économie de la question 1).

b. Quelle est la signification économique de $x_j v_j$?

c. Quelle est la signification économique de $x_1v_1 + x_2v_2 + \ldots + x_nv_n + b$?

d. Quelle est la signification économique de l'équation $x_1v_1+x_2v_2+\ldots+x_nv_n+b = x$?

3. Considérons l'économie d'Israël en 1958. Les trois secteurs industriels considérés sont : I_1 agriculture, I_2 biens manufacturés, I_3 énergie. Production et demande sont mesurés en millions de livres israéliennes, la monnaie d'Israël à cette époque. On nous dit que

$$b = \begin{pmatrix} 13,2 \\ 17,6 \\ 1,8 \end{pmatrix}, \quad v_1 = \begin{pmatrix} 0,293 \\ 0,014 \\ 0,044 \end{pmatrix} \quad v_2 = \begin{pmatrix} 0 \\ 0,207 \\ 0,001 \end{pmatrix} \quad v_3 = \begin{pmatrix} 0 \\ 0,017 \\ 0,216 \end{pmatrix}$$

a. Pourquoi la première composante des vecteurs v_2 et v_3 est-elle nulle ?

b. Trouver les productions x_1, x_2, x_3 qui satisfont la demande.

Considérons les matrices suivantes

$$A = \begin{pmatrix} 0 & 0,2 & 0,3 \\ 0,1 & 0 & 0,4 \\ 0,2 & 0,5 & 0 \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Ce qu'indique la figure est que lorsque les productions des trois secteurs sont x_1, x_2, x_3, les consommations en les produits des trois secteurs sont

$$AX = \begin{pmatrix} 0,2x_2 + 0,3x_3 \\ 0,1x_1 + 0,4x_3 \\ 0,2x_1 + 0,5x_2 \end{pmatrix}.$$

Si les productions des trois secteurs sont x_1, x_2, x_3, alors ce que les trois secteurs offrent aux consommateurs est ce qui reste

$$X - AX = (I - A)X.$$

Cela signifie que si les demandes des consommateurs sont 320, 90, 150 alors le vecteur des productions$X$ doit vérifier l'égalité

$$X - AX = \begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

Cette égalité matricielle est un système linéaire

$$\left\{\begin{array}{cccl} x_1 & -0,2x_2 & -0,3x_3 & = 320 \\ -0,1x_1 & +x_2 & -0,4x_3 & = 90 \\ -0,2x_1 & -0,5x_2 & +x_3 & = 150 \end{array}\right.$$

qu'on peut résoudre comme on le veut. Une façon de le faire est de calculer l'inverse de la matrice $I - A$ car

$$X = (I-A)^{-1}\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

Le calcul d'inverses de petites matrices se fait très bien avec un ordinateur :

$$(I-A)^{-1} \simeq \begin{pmatrix} 1,161 & 0,508 & 0,552 \\ 0,261 & 1,364 & 0,624 \\ 0,363 & 0,784 & 1,422 \end{pmatrix}.$$

On obtient

$$X = \begin{pmatrix} 500 \\ 300 \\ 400 \end{pmatrix}.$$

Voici un autre façon de raisonner. Les trois secteurs veulent répondre à la demande. Chacun produit la quantité souhaité 320, 90, 150 pour chacun des trois secteurs. Mais on constate qu'une partie a été utilisée pour produire ces quantités. La quantité qui reste disponible n'est pas celle demandée mais

$$\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} - A\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

Les trois secteurs décident alors de compléter et produisent ce qui manque. Mais là encore une partie disparaît en consommation intermédiaire. Après cette deuxième étape les trois secteurs ont produit

$$\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} + A\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

et les productions disponibles pour répondre à la demande sont données par

$$\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} - A\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} + A\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} - A^2\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} = \begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} - A^2\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

On peut ainsi continuer à compléter ce qui manque. Au bout de k étapes les trois secteurs ont produit

$$\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} + A\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} + A\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} + \ldots + A^{k-1}\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

et les productions disponibles pour répondre à la demande sont données par

$$\begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix} - A^k \begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

Si A^k tend vers 0 on voit qu'en répétant à l'infini le procédé on "finit" par répondre à la demande en produisant

$$\sum_{k=0}^{\infty} A^k \begin{pmatrix} 320 \\ 90 \\ 150 \end{pmatrix}.$$

Les deux méthodes de calcul proposées doivent aboutir au même résultat ; cela signifie que si A^k tend vers 0 alors

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

On peut le voir aussi de manière formelle. Multiplions $\sum_{k=0}^{\infty} A^k$ par $I - A$, on obtient

$$\begin{aligned} (I - A)(\sum_{k=0}^{\infty} A^k) &= \sum_{k=0}^{\infty} A^k - A \sum_{k=0}^{\infty} A^k \\ &= \sum_{k=0}^{\infty} A^k - \sum_{k=0}^{\infty} A^{k+1} \\ &= I + A + A^2 + A^3 \ldots - (A + A^2 + A^3 \ldots) \\ &= I. \end{aligned}$$

1.2.2 Evolution de parts de marché

Considérons la matrice suivante :

$$A = \begin{pmatrix} 0,6 & 0,15 & 0,05 \\ 0,2 & 0,7 & 0,05 \\ 0,2 & 0,15 & 0,9 \end{pmatrix}$$

Disons que les nombres qu'elle contient ont la signification suivante : trois entreprises se partagent un marché, on suppose que chaque année 60% des clients de l'entreprise 1 (c'est le nombre 0,6) reste avec cette entreprise, 20% passent à l'entreprise 2 et 20% à l'entreprise 3 (ce sont les deux 0,2) ; de manière analogue 70% des clients de l'entreprise 2 restent avec l'entreprise 2 et la moitié du reste (15%) passe à 1 l'autre moitié du reste à 3 ; enfin 90% des clients de l'entreprise 3 restent avec l'entreprise 3 et la moitié du reste (5%) passe à 1 l'autre moitié du reste à 2. On cherche à savoir comment évoluent les parts de marché avec le temps. Imaginons que la répartition soit donnée à une certaine année par les trois nombres positifs ou nuls $p_1^{(0)}$, $p_2^{(0)}$, $p_3^{(0)}$ avec $p_1^{(0)} + p_2^{(0)} + p_3^{(0)} = 1$ ($p_i^{(0)}$ est la part de marché de l'entreprise i à l'année 0). La part de marché à l'année 1 de l'entreprise 1 sera alors :

$$p_1^{(1)} = 0,6.p_1^{(0)} + 0,15.p_2^{(0)} + 0,05.p_3^{(0)}.$$

En effet, la nouvelle part de marché est constituée des clients de 1 qui restent, de ceux qui viennent de 2 et de ceux qui viennent de 3. De façon analogue, on a :

$$p_2^{(1)} = 0,2.p_1^{(0)} + 0,7.p_2^{(0)} + 0,05.p_3^{(0)},$$

$$p_3^{(1)} = 0,2.p_1^{(0)} + 0,15.p_2^{(0)} + 0,9.p_3^{(0)}.$$

Introduisons les vecteurs

$$P^{(0)} = \begin{pmatrix} p_1^{(0)} \\ p_2^{(0)} \\ p_3^{(0)} \end{pmatrix}, \text{ et } P^{(1)} = \begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ p_3^{(1)} \end{pmatrix} \text{ et } P^{(k)} = \begin{pmatrix} p_1^{(k)} \\ p_2^{(k)} \\ p_3^{(k)} \end{pmatrix}$$

donnant les répartitions du marché aux années 0, 1 et k. Les égalités précédentes s'écrivent matriciellement sous la forme

$$P^{(1)} = AP^{(0)}.$$

On a aussi

$$P^{(2)} = AP^{(1)} = AAP^{(0)} = A^2P^{(0)},$$

et, pour tout entier k,

$$P^{(k)} = A^kP^{(0)}.$$

Ce sont donc les puissances de k qui permettent de décrire l'évolution des parts de marchés. Calculons quelques unes de ces puissances :

$$A = \begin{pmatrix} 0,6 & 0,15 & 0,05 \\ 0,2 & 0,7 & 0,05 \\ 0,2 & 0,15 & 0,9 \end{pmatrix} \quad A^2 = \begin{pmatrix} 0,4 & 0,2025 & 0,0825 \\ 0,27 & 0,5275 & 0,09 \\ 0,33 & 0,27 & 0,8275 \end{pmatrix}$$

$$A^{10} = \begin{pmatrix} 0,166 & 0,169 & 0,152 \\ 0,225 & 0,230 & 0,200 \\ 0,609 & 0,600 & 0,648 \end{pmatrix} \quad A^{20} = \begin{pmatrix} 0,158 & 0,158 & 0,158 \\ 0,211 & 0,211 & 0,210 \\ 0,631 & 0,630 & 0,632 \end{pmatrix}$$

$$A^{50} = \begin{pmatrix} 0,15779 & 0,15779 & 0,15779 \\ 0,21053 & 0,21053 & 0,21053 \\ 0,63158 & 0,63158 & 0,63158 \end{pmatrix}$$

(les chiffres sont arrondis au millième pour A^{10} et A^{20}, au cent millième pour A^{50}). On constate que les puissances se stabilisent à une valeur très proche de A^{50}, pour tout $k \geq 50$,

$$A^k \simeq \begin{pmatrix} 0,15779 & 0,15779 & 0,15779 \\ 0,21053 & 0,21053 & 0,21053 \\ 0,63158 & 0,63158 & 0,63158 \end{pmatrix}$$

et que les colonnes de cette matrice sont identiques. Pour $k \geq 50$, on a

$$P^{(k)} = A^k P^{(0)} \simeq \begin{pmatrix} 0,15779 & 0,15779 & 0,15779 \\ 0,21053 & 0,21053 & 0,21053 \\ 0,63158 & 0,63158 & 0,63158 \end{pmatrix} P^{(0)} = \begin{pmatrix} 0,15779 \\ 0,21053 \\ 0,63158 \end{pmatrix}.$$

Autrement dit quelle que soit la répartition du marché à l'année 0, le mécanisme d'évolution aboutit à la même répartition lorsque le temps grandit beaucoup (répartition donnée par les colonnes de la limite de A^k quand k tend vers l'infini.

L'exercice suivant permet d'établir ce comportement limite des puissances de A dans un cadre plus général.

Exercice 1. *Soit A une matrice carrée $n \times n$ $(n \geq 2)$ dont les coefficients sont tous strictement positifs et telle que les sommes des éléments des colonnes soient égales à 1. Notons d le plus petit coefficient de A.*

1. *Montrer que d est inférieur ou égal à $1/2$.*
2. *Pour tout Y vecteur $1 \times n$ dont les coefficients sont positifs ou nuls, notons $m(Y)$ le plus petit coefficient de Y, $M(Y)$ le plus grand. Montrer que, pour tout Y vecteur ligne positif ou nul, on a*

$$M(YA) - m(YA) \leq (1 - 2d)(M(Y) - m(Y)).$$

3. *En déduire que YA^k converge vers un vecteur ligne dont toutes les coordonnées sont égales entre elles (lorsque k tend vers l'infini).*
4. *En déduire que A^k converge vers une matrice dont toutes les colonnes sont égales, puis que l'équation $AX = X$ a une solution X positive.*

Cet exercice a été traité en TD. Voici quelques éléments de réponses supplémentaires.

Commençons par une remarque sur la notion de moyenne. Par exemple, je vous ai dit que votre note finale en Algèbre III serait calculée comme la moyenne de quatre notes n_1, n_2, n_3, n_4 avec les coefficients 1, 3, 3, 6. Autrement dit la note finale sera

$$N = \frac{1}{13}(n_1 + 3n_2 + 3n_3 + 6n_4) = \frac{1}{13}n_1 + \frac{3}{13}n_2 + \frac{3}{13}n_3 + \frac{6}{13}n_4.$$

Imaginons qu'un étudiant vous dise que sa meilleure note est M et sa moins bonne m. Que pouvez-vous dire de sa moyenne ? Vous pouvez dire que sa moyenne est comprise entre m et M. Cela s'écrit facilement

$$N = \frac{1}{13}n_1 + \frac{3}{13}n_2 + \frac{3}{13}n_3 + \frac{6}{13}n_4 \leq \frac{1}{13}M + \frac{3}{13}M + \frac{3}{13}M + \frac{6}{13}n_M = M,$$

car chaque note est majorée par M et les coefficients sont positifs de somme 1. On a de même

$$N = \frac{1}{13}n_1 + \frac{3}{13}n_2 + \frac{3}{13}n_3 + \frac{6}{13}n_4 \geq \frac{1}{13}m + \frac{3}{13}m + \frac{3}{13}m + \frac{6}{13}m = m.$$

Mais vous pouvez dire un peu mieux. La moins bonne moyenne possible est obtenue lorsque M est affectée du plus petit coefficient ($n_1 = M$) et toutes les autres notes sont minimales ($n_2 = n_3 = n_4 = m$), autrement dit

$$N \geq \frac{1}{13}M + \frac{3}{13}m + \frac{3}{13}m + \frac{6}{13}m = \frac{1}{13}M + \frac{12}{13}m.$$

Comment le voir ? Ce qui est sûr c'est que m et M figurent parmi les notes. En remplaçant les deux qui ne sont peut-être pas égales à m ou M par m, on diminue la moyenne. On a donc

$$N \geq \alpha M + (1-\alpha)m = m + \alpha(M-m),$$

où α est le (ou un) coefficient de la note maximale M. Comme $M - m > 0$, la deuxième expression montre que N est minimale quand α est minimal, c'est-à-dire ici quand $\alpha = 1/13$. On montre de la même façon que

$$N \leq \frac{1}{13}m + \frac{12}{13}M.$$

Revenons à l'exercice.

1. Par hypothèse les sommes des coefficients des colonnes de la matrice A sont égales à 1. Pour tout j, on a
$$a_{1j} + a_{2j} + \ldots + a_{nj} = 1.$$
En divisant par n on obtient
$$\frac{1}{n}a_{1j} + \frac{1}{n}a_{2j} + \ldots + \frac{1}{n}a_{nj} = \frac{1}{n}.$$
La moyenne des nombres a_{ij} est égale à $1/n$. Mais la moyenne est supérieure ou égale au plus petit des nombres a_{ij} (ici c'est i qui varie). Cela signifie que chaque colonne comporte au moins un nombre inférieur ou égal à $1/n$. Comme $n \geq 2$, d le plus petit des coefficients de la matrice A est inférieur à $1/2$.

2. Le j-ème coefficient de YA est donnée par la somme
$$(YA)_j = y_1a_{1j}+y_2a_{2j}+\ldots+y_{n-1}a_{(n-1)j}+y_na_{nj} = a_{1j}y_1+a_{2j}y_2+\ldots+a_{(n-1)j}y_{n-1}+a_{nj}y_n.$$
Comme les nombres $a_{1j}, a_{2j}, \ldots, a_{(n-1)j}, a_{nj}$ sont positifs de somme égale à 1, le j-ème coefficient de YA est une moyenne calculée avec les coefficients a_{ij} (i variant). Comme on l'a remarqué plus haut une telle moyenne est inférieure à celle obtenue en prenant la valeur minimale $m(Y)$ avec le plus petit coefficient et toutes les autres valeurs égales à la valeurs maximales $M(Y)$; on obtient
$$(YA)_j \leq (\min_i a_{ij})m(Y) + (1 - \min_i a_{ij})M(Y).$$
Comme $d \leq \min_i a_{ij}$, cela entraîne
$$(YA)_j \leq dm(Y) + (1-d)M(Y).$$

De manière analogue on a

$$(YA)_j \geq dM(Y) + (1-d)m(Y).$$

Tous les coefficients de YA sont donc compris entre $dM(Y)+(1-d)m(Y)$ et $dm(Y)+(1-d)M(Y)$:

$$dM(Y) + (1-d)m(Y) \leq m(YA) \leq M(YA) \leq dm(Y) + (1-d)M(Y),$$

en particulier

$$\begin{aligned} &M(YA) - m(YA) \\ &\quad \leq dm(Y) + (1-d)M(Y) - dM(Y) - (1-d)m(Y) \\ &\quad \leq (1-2d)(M(Y) - m(Y)). \end{aligned}$$

3. Comme dans la question précédente on voit que, pour tout k, les coefficients de YA^{k+1} sont des moyennes des coefficients de YA^k et on en déduit (comme dans la question 2.) que

$$M(YA^{k+1}) - m(YA^{k+1}) \leq (1-2d)(M(YA^k) - m(YA^k)).$$

Par ailleurs on a aussi

$$m(YA^k) \leq m(YA^{k+1}) \leq M(YA^{k+1}) \leq M(YA^k).$$

Par récurrence on montre que pour tout k, on a

$$(M(YA^k) - m(YA^k)) \leq (1-2d)^k(M(YA) - m(YA)).$$

Comme $1-2d$ appartient à $[0,1[$, $(1-2d)^k$ tend vers 0. Les deux suites $(m(YA^k))_k$ et $(M(YA^k))_k$ sont donc adjacentes. Elles convergent vers une limite commune. Tous les coefficients de YA^k étant compris entre $m(YA^k)$ et $M(YA^k)$, le théorème des gendarmes montre que tous les coefficients de YA^k vers une même limite.

4. Les vecteurs ligne de la matrice A^k peuvent s'écrire comme les produits Y_iA^k où Y_i est le vecteur ligne ayant 1 pour i-ème coordonnée et 0 pour les autres ; par exemple

$$Y_1 = (1, 0, 0, \ldots, 0),\ Y_2 = (0, 1, 0, \ldots, 0).$$

La question précédente assure que, pour tout i, il existe une limite l_i telle que

$$Y_iA^k \longrightarrow_{k\to\infty} (l_i, l_i, l_i, \ldots, l_i).$$

Cela signifie (rappelons que Y_iA^k est le i-ème vecteur ligne de la matrice A^k)

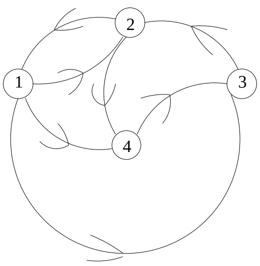
$$A^k \longrightarrow_{k\to\infty} \begin{pmatrix} l_1 & l_1 & l_1 & \cdots & l_1 \\ l_2 & l_2 & l_2 & \cdots & l_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_n & l_n & l_n & \cdots & l_n \end{pmatrix} := L.$$

Pour finir, on remarque que $A^{k+1} = AA^k$ converge vers L et aussi vers AL ; on en déduit $AL = L$. Le premier vecteur colonne (par exemple) de AL est le produit de A par le premier vecteur colonne de L. On en déduit que, si on note X le vecteur colonne de L, on a $AX = X$.

Remarque : le vecteur X obtenu ici n'est pas nul (ses coordonnées sont des nombres positifs ou nuls dont la somme est 1). L'égalité $A0 = 0$ est bien vérifiée mais peu intéressante.

1.2.3 Matrice associée à un graphe

Exercice 2. *Donner le nombre de chemins de longueur 2 allant de 1 à 3 dans le graphe suivant. Combien y-a-t-il de chemins de longueur 3 allant de 1 à 3 ? Combien y-a-t-il de chemins de longueur 8 allant de 1 à 3 ? Combien y-a-t-il de chemins de longueur 5 en tout ? Combien y-a-t-il de chemins de longueur 42 en tout ? Se rendre compte que le calcul matriciel est très utile (la calculatrice aussi) pour répondre à ce genre de question quand les nombres considérés sont grands.*



Je donne quelques indications que je compléterai peut-être plus tard. Décidons de placer dans des matrices 4×4 les nombres de chemins joignant les sommets les uns aux autres :

$$M^{(n)} = \begin{pmatrix} m_{11}^{(n)} & m_{12}^{(n)} & m_{13}^{(n)} & m_{14}^{(n)} \\ m_{21}^{(n)} & m_{22}^{(n)} & m_{23}^{(n)} & m_{24}^{(n)} \\ m_{31}^{(n)} & m_{32}^{(n)} & m_{33}^{(n)} & m_{34}^{(n)} \\ m_{41}^{(n)} & m_{42}^{(n)} & m_{43}^{(n)} & m_{44}^{(n)} \end{pmatrix}.$$

Il faut comprendre donc que $m_{ij}^{(n)}$ désigne le nombre de chemins de longueur n joignant i à j. Gràćce au dessin, on connaît $M^{(1)}$:

$$M^{(n)} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Fixons maintenant un entier n et voyons comment compter le nombre de chemins de longueur $n + 1$ joignant le sommet i au sommet j. On classe ces chemins en fonctions du sommet auquel on se trouve après n pas. L'ensemble des chemins de longueur $n+1$ joignant le sommet i au sommet j est la réunion pour k allant de 1 à 4 des ensembles des chemins de longueur $n+1$ joignant le sommet i au sommet j tels qu'au bout de n pas on se trouve au sommet k. S'il n'y a pas de flèche entre k et j alors il n'y a aucun chemin de longueur $n+1$ joignant le sommet i au sommet j tel qu'au bout de n pas on se trouve au sommet k. S'il a une flèche entre k et j alors il y a autant de chemins de longueur $n+1$ joignant le sommet i au sommet j tel qu'au bout de n pas on se trouve au sommet k qu'il y a de chemins de longueur n joignant le sommet i au sommet k. On peut résumer ces phrases trop longues par l'égalité

$$m_{ij}^{(n+1)} = \sum_{k=1}^{4} m_{ik}^{(n)} m_{kj}^{(1)}.$$

Cela s'écrit encore matriciellement

$$M^{(n+1)} = M^{(n)} M^{(1)}.$$

On montre donc par récurrence qu'on a

$$M^{(n)} = (M^{(1)})^n.$$

Pour calculer les nombres demandés dans l'énoncé de l'exercice il suffit donc de calculer les puissances de $M^{(1)}$:

$$M^{(2)} = (M^{(1)})^2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$M^{(3)} = M^{(1)} M^{(2)} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

$$M^{(5)} = M^{(2)} M^{(3)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 5 & 5 & 5 \\ 2 & 4 & 5 & 4 \\ 2 & 2 & 2 & 3 \\ 1 & 2 & 1 & 1 \end{pmatrix}$$

$$M^{(8)} = M^{(5)} M^{(3)} = \begin{pmatrix} 3 & 5 & 5 & 5 \\ 2 & 4 & 5 & 4 \\ 2 & 2 & 2 & 3 \\ 1 & 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 & 2 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 13 & 21 & 21 & 21 \\ 10 & 17 & 18 & 17 \\ 7 & 10 & 10 & 11 \\ 4 & 7 & 6 & 6 \end{pmatrix}$$

On peut alors répondre aux questions posées : le nombre de chemins de longueur 3 allant de 1 à 3 est 2, le nombre de chemins de longueur 8 allant de 1 à 3 est 21, le nombre total de chemins de longueur 5 est la somme de tous les coefficients de $M^{(5)}$, soit 47. Remarque : prendre la somme de tous les coefficients d'une matrice s'écrit matriciellement. Par exemple ici

$$\begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} M^{(5)} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 47.$$

1.3 Espaces vectoriels

Définition 1.21. *Un ensemble E est un $\mathbb{R}$-espace vectoriel ($\mathbb{C}$-espace vectoriel) si*

1. *il existe une opération sur E, notée $+$ et appelée addition, $(u,v) \in E \times E \mapsto u+v$ telle que*
   (i) $u+v=v+u$
   (ii) $(u+v)+w=u+(v+w)$
   (iii) il existe un unique vecteur 0_E tq $0_E+u=u+0_E=u$
   (iv) Chaque vecteur $u \in E$ a un opposé $-u$ tq $u+(-u)=(-u)+u=0$
2. *Il existe une opération dite "externe" de multiplication d'un vecteur de E par un nombre réel (ou complexe) $(\lambda,u) \in \mathbb{R} \times E \to \lambda \times u$ qui vérifie pour tous $\lambda \in \mathbb{R}$ et $u \in E$ (resp $\lambda \in \mathbb{C}$)*
   (i) $1 \times u = u$
   (ii) $(\lambda+\mu) \times u = \lambda \times u + \lambda \times v$
   (iii) $(\lambda\mu) \times u = \lambda \times (\mu \times u) = \mu \times (\lambda u)$
   (iv) $\lambda(u+v) = \lambda(u+v)$

Voici une liste d'espaces vectoriels.

1. La droite réelle $\mathbb{R}$, le plan usuel $\mathbb{R}^2$ et l'espace usuel $\mathbb{R}^3$ sont des $\mathbb{R}$-espaces vectoriels.
2. Le plan complexe $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel. C'est aussi un $\mathbb{R}$-espace vectoriel.
3. L'ensemble $\mathbb{R}^n$, $n \geq 1$ est un $\mathbb{R}$-espace vectoriel.
4. L'ensemble $\mathbb{C}^n$, $n \geq 1$, est un $\mathbb{C}$-espace vectoriel.
5. L'ensemble $\mathbb{C}^n$, $n \geq 1$, est un $\mathbb{R}$-espace vectoriel.
6. L'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels est un $\mathbb{R}$-espace vectoriel.
7. L'ensemble $\mathbb{R}_n[X]$ des polynômes à coefficients réels. de degré inférieur ou égal à n est un $\mathbb{R}$-eV.
8. L'ensemble $\mathbb{C}_n[X]$ des polynômes à coefficients réels. de degré inférieur ou égal à n est un $\mathbb{C}$-eV.
9. l'ensemble $C(\mathbb{R},\mathbb{R})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un $\mathbb{R}$-eV.
10. L'ensemble $C(\mathbb{R},\mathbb{C})$ des fonctions continues de $\mathbb{R}$ dans $\mathbb{C}$ est un $\mathbb{C}$-espace vectoriel.

11. L'ensemble $M_{n\times m}(\mathbb{R})$ (resp $M_{n\times m}(\mathbb{C})$) des matrices à n lignes et m colonnes remplies de coeff réels (resp complexes) est un $\mathbb{R}$ espace vectoriel (resp $\mathbb{C}$ espace vectoriel).

La liste précédente passe sous silence des éléments très importants de la définition d'un espace vectoriel : l'addition et la multiplication. On omet souvent de les préciser car elles sont considérées comme naturelles. On peut résumer en disant que dans chaque cas, l'addition est l'addition coordonnée à coordonnée et la multiplication par un scalaire la multiplication de chaque coordonnée par ce scalaire.

Définition 1.22. *Un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel E est un sous ensemble $F \subset E$ stable par addition et multiplication :*
(i) si $u, v \in F$ alors $u + v \in F$,
(ii) si $\lambda \in \mathbb{R}$, $u \in F$, alors $\lambda.u \in F$.
En particulier, on doit avoir $0_E \in F$.

Exemples geometriques :

La droite d'équation $y = -2x$ dans $\mathbb{R}^2$. C'est l'ensemble des points de coordonnées (x, y) tq $y = -2x$. C'est un sous-espace vectoriel de $\mathbb{R}^2$.

Le plan d'équation $x + y + z = 0$ dans $\mathbb{R}^3$ est un sous-espace vectoriel.

La droite d'équations $x = y$ et $z = -x$ dans $\mathbb{R}^3$ est un sous-espace vectoriel.

Les sous-espace vectoriel de $\mathbb{R}^2$ sont $\{0\}$, les droites vectorielles, $\mathbb{R}^2$. Les sous-espace vectoriel de $\mathbb{R}^3$ sont $\{0\}$, les droites vectorielles, les plans vectoriels, $\mathbb{R}^3$.

Exemple dans un espace vectoriel de polynomes : Dans $\mathbb{R}_3[X]$ l'ensemble des polynômes vérifiant $2P(X) - XP'(X) + X^2P''(X) = 0$ est un sous-espace vectoriel de $\mathbb{R}_3[X]$.

Proposition 1.23. *Un sous-espace vectoriel d'un espace vectoriel E est un espace vectoriel.*

Définition 1.24. *Soient $u_1 \dots u_n \in E$, $\lambda_1, \dots \lambda_n$ Le vecteur $\vec{u} = \lambda_1\vec{u}_1 + \dots \lambda_n\vec{u}_n$ est appelé combinaison linéaire des u_i. Les λ_i sont les coefficients de la combinaison linéaire. La combinaison linéaire est dite triviale si tous les λ_i sont nuls (et $\vec{u} = 0$).*

Exemple : l'écriture $(2, 3, 2) = 2.(1, 0, 1) + 3.(0, 1, 0)$ peut se lire « $(2, 3, 2)$ est la combinaison linéaire des vecteurs $(1, 0, 1)$ et $(0, 1, 0)$ de coefficients 2 et 3 ».

Autre exemple : Soit $A = (a_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ et $X = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix} \in \mathbb{R}^m$. Alors on définit

$$A.X = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2m}x_m \\ \dots \\ a_{n1}x_1 + a_{n2}x_2 + \dots a_{nm}x_m \end{pmatrix}$$

Le vecteur (colonne) $A.X$ est dans $\mathbb{R}^n$ et on voit qu'il s'écrit

$$A.X = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}.$$

Autrement dit AX est la combinaison linéaire des vecteurs colonne de A avec les coefficients $x_1, \ldots, x_m$. Si on note C_i le i-ème vecteur colonne de A :

$$C_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix},$$

on a

$$AX = x_1C_1 + x_2C_2 + \ldots + x_mC_m = \sum_{i=1}^{m} x_iC_i.$$

Proposition 1.25. *$F \subset E$ est un sous-espace vectoriel de E ssi stable par combinaisons linéaires, i.e. pour toute famille finie de vecteurs $u_1, \ldots, u_k \in F$, et toute famille $\lambda_1, \ldots, \lambda_k$ de nombres on a $\lambda_1 u_1 + \ldots \lambda_k u_k \in F$*

Définition 1.26. *Soit S une partie d'un espace vectoriel E. On appelle sous-espace vectoriel engendré par S, le plus petit sous-espace vectoriel de E contenant S. On le note $Vect(S)$.*

Proposition 1.27. *Soit S une partie d'un espace vectoriel E. Le sous-espace $Vect(S)$ est l'ensemble des combinaisons linéaires d'éléments de S.*

1.4 Applications linéaires

Définition 1.28. *Soient E, F deux $\mathbb{R}$-espace vectoriel (resp $\mathbb{C}$) Une application linéaire de E dans F est une application qui vérifie pour tous les vecteurs u, v de E et tout nombre $\lambda \in \mathbb{R}$ (resp $\mathbb{C}$)*

1. *$f(u + v) = f(u) + f(v)$*
2. *$f(\lambda u) = \lambda f(u)$*

Une forme linéaire est une application lin de E dans $\mathbb{R}$ (resp $E \to \mathbb{C}$)

Un <u>endomorphisme</u> est une application linéaire de E dans E.

Proposition 1.29. *Une application $\varphi : E \to G$ est lin ssi elle préserve les comb lin, i.e. $\varphi(\lambda_1\vec{u}_1 + \ldots \lambda_k\vec{u}_k) = \lambda_1\varphi(\vec{u}_1) + \ldots \lambda_k\varphi(\vec{u}_k)$.*

Exemple projection $p_i : (x_1, \dots, x_n) \in \mathbb{R}^n \to x_i$ est une forme linéaire

Exemple Soient $a_1, \dots a_n$ n nombres reels. Alors $(x_1, \dots, x_n) \in \mathbb{R}^n \to a_1x_1 + a_2x_2 + \dots a_nx_n$ est une forme lineaire.

Proposition 1.30. *Toutes les formes linéaires s'écrivent ainsi*

$$\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n \to a_1x_1 + \cdots + a_nx_n \in \mathbb{R}.$$

Démonstration : Par linéarité, on a :

$$f(x) = f(x_1e_1 + \cdots + x_ne_n) = x_1f(e_1) + \cdots + x_nf(e_n).$$

Les a_i de l'énoncé sont les nombres $f(e_i)$. □

Proposition 1.31. *Soient E, F, G trois $\mathbb{R}$-espace vectoriel. Soient $f : E \to F$ et $g : F \to G$ deux applications linéaires. La composée $g \circ f$ est une application linéaire.*

Soit $A = (a_{ij})_{1 \le i \le n, 1 \le j \le m}$ L'application $\mathbb{R}^m \to \mathbb{R}^n$ définie par

$$X = \begin{pmatrix} x_1 \\ \dots \\ x_m \end{pmatrix} \mapsto A.X = \begin{pmatrix} a_{11}x_1 + \dots a_{1m}x_m \\ \dots \\ a_{n1}x_1 + \dots a_{nm}x_m \end{pmatrix}$$

est une application linéaire.

L'exemple précédent est en fait le cas général : toute application linéaire $\mathbb{R}^m \to \mathbb{R}^n$ peut s'écrire comme un tel produit matriciel.

En effet considérons une application linéaire f de $\mathbb{R}^m$ vers $\mathbb{R}^n$. Appelons $e_1, \dots, e_m$ les vecteurs de la base canonique de $\mathbb{R}^m$:

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots e_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Considérons alors un élément X de $\mathbb{R}^m$ (un vecteur colonne) :

$$X = \sum_{i=1}^{m} x_ie_i.$$

Par linéarité de f on a donc

$$f(X) = f(\sum_{i=1}^{m} x_ie_i) = \sum_{i=1}^{m} x_if(e_i).$$

Les vecteurs (colonne) $f(e_i)$ sont des éléments de $\mathbb{R}^m$. Appelons A la matrice de taille $n \times m$ dont les colonnes sont ces vecteurs $f(e_i)$ (i allant de 1 à m) :

$$f(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \dots \\ a_{ni} \end{pmatrix}.$$

On a vu plus haut (le deuxième exemple après la définition d'une combinaison linéaire) que, pour cette matrice A, AX est la combinaison linéaire des vecteurs colonne de A avec les coefficients x_i, autrement dit :

$$AX = \sum_{i=1}^{m} x_i f(e_i).$$

On a donc bien trouvé une matrice A telle que

$$f(X) = AX.$$

2 Les systèmes linéaires

2.1 Vocabulaire sur les systèmes linéaires

Rappel Soit $E = \mathbb{R}^n$ (resp $\mathbb{C}^n$).

Par exemple $f : (x, y, z) \in \mathbb{R}^3 \to 2x - 3y + \sqrt{2}z$ ou encore $g : (x, y) \in \mathbb{R}^2 \to x$.

Une équation linéaire homogène est une équation du type $a_1x_1 + \dots + a_nx_n = 0$. Autrement dit, c'est une équation du type $f(\vec{x}) = 0$, avec $f : \mathbb{R}^n \to \mathbb{R}$ (ou $f : \mathbb{C}^n \to \mathbb{C}$) une forme linéaire. Le mot homogène signifie que le deuxième membre de l'égalité vaut 0.

Un système linéaire homogène est un système de m équations linéaires homogènes à n inconnues. Soit encore

$$\left\{ \begin{array}{rcl} a_{11}x_1 + \dots a_{1n}x_n & = & 0 \\ a_{21}x_1 + \dots + a_{2n}x_n & = & 0 \\ \dots & = & 0 \\ a_{m1}x_1 + \dots a_{mn}x_n & = & 0 \end{array} \right\}$$

Proposition 2.1. *L'ensemble des solutions d'un système linéaire* **homogène** *est un sous-espace vectoriel de $\mathbb{R}^n$.*

Exemples :

La droite d'équation $y+2x=0$ dans $\mathbb{R}^2$ est un sous-espace vectoriel. La droite d'équations $y+2x=0$ et $z=0$ est un sous-espace vectoriel.

Une équation linéaire est une équation du type $a_1x_1+\cdots+a_nx_n=b$. Autrement dit, c'est une équation du type $f(\vec{x})=b$, avec f une forme linéaire et b un nombre. Un système linéaire est un système de m équations linéaires à n inconnues. Soit encore

$$\left\{\begin{array}{ccc} a_{11}x_1+\ldots a_{1n}x_n & = & b_1 \\ a_{21}x_1+\cdots+a_{2n}x_n & = & b_2 \\ \ldots & = & \ldots \\ a_{m1}x_1+\ldots a_{mn}x_n & = & b_m \end{array}\right\}$$

L'ensemble des solutions d'un système linéaire non homogène n'est PAS un sous-espace vectoriel de $\mathbb{R}^n$. L'ensemble des solutions d'un système linéaire non homogène est appelé <u>sous-espace affine</u>.

Exemple :

La droite d'équation $y+2x=1$ dans $\mathbb{R}^2$ n'est pas un sous-espace vectoriel. La droite d'équations $y+2x=1$ et $z=-2$ n'est pas un sous-espace vectoriel.

Notation matricielle. Soit (S) le système ci-dessus, $A=(a_{ij})_{1\leq i\leq m, 1\leq j\leq n}$ la matrice associée, et $\vec{b}=(b_1,\ldots,b_m)$, ou encore $B=Mat(\vec{b})=\begin{pmatrix} b_1 \\ b_2 \\ \ldots \\ b_m \end{pmatrix}$. Soit $X=\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$ Le système peut se réécrire $A.X=B$.

Autrement dit, matriciellement, l'ensemble des $(x_1,\ldots x_n)$ vérifiant le système d'équations (S) est exactement l'ensemble des $X=\begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$ vérifiant la relation $AX=B$. Ça veut dire la même chose, c'est juste plus rapide à écrire.

2.2 L'algorithme de Gauss-Jordan

La bonne méthode pour résoudre un système linéaire s'appelle l'<u>algorithme de Gauss Jordan</u>.

L'algorithme repose sur quelques remarques.

– Si je change l'ordre d'écriture des équations, je ne change pas les solutions.
– si je multiplie une équation par un nombre non nul, je ne change pas les solutions.

– Si je remplace une équation du système par la somme de cette équation et d'une autre équation du système, je ne change pas les solutions. (Exemple pour illustrer)

Pour comprendre l'algorithme il suffit de l'appliquer à quelques exemples. En voici quelques uns.

$$\left\{\begin{array}{rrrrrcl} & & x_3- & x_4- & x_5 & = & 4 \\ 2x_1+ & 4x_2+ & 2x_3+ & 4x_4+ & 2x_5 & = & 4 \\ 2x_1+ & 4x_2+ & 3x_3+ & 3x_4+ & 3x_5 & = & 4 \\ 3x_1+ & 6x_2+ & 6x_3+ & 3x_4+ & 6x_5 & = & 6 \end{array}\right.$$

Le système ci-dessus peut s'écrire

$$\begin{pmatrix} 0 & 0 & 1 & -1 & -1 \\ 2 & 4 & 2 & 4 & 2 \\ 2 & 4 & 3 & 3 & 3 \\ 3 & 6 & 6 & 3 & 6 \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \\ 6 \end{pmatrix}$$

On pourra faire tous les calculs en omettant les x_i. On notera alors

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

le système, et on appellera cette matrice la Matrice augmentée du système.

Autrement dit, pour un système de n équations à m inconnues, la matrice augmentée a n lignes et $m+1$ colonnes.

Passons maintenant à l'algorithme proprement dit.

Avant de démarrer, je place mon curseur en haut à gauche (sur a_{11} à la première étape).

1. Je veux que ce coefficient soit non nul. S'il est non nul, je passe à l'étape suivante. Sinon, je cherche dans la colonne sous mon curseur un coeff non nul. J'échange alors cette ligne avec celle où est mon curseur. Si ce n'est pas possible (mon coefficient ainsi que tous les coefficients de la colonne sous mon curseur sont nuls) je me déplace d'un cran vers la droite et je recommence.

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

2. Si le coeff est non nul, je divise la ligne où je suis par ce coefficient. J'obtiens un coeff 1, appelé pivot, sur ma ligne pivot.

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

3. Notons $a_{i_0j_0}$ le coefficient sur lequel pointe mon curseur. Je parcours la colonne de mon curseur (colonne pivot). J'ajoute à chaque ligne L_i rencontrée la ligne L_{i_0} multipliée par $(-a_{ij_0})$. Ainsi, à la fin de cette étape, tous les coefficients de la colonne C_{j_0} sauf le pivot sont nuls.

$$\left(\begin{array}{ccccc|c} \mathbf{1} & 2 & 1 & 2 & 1 & 2 \\ 0 & 0 & 1 & -1 & -1 & 4 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 3 & -3 & 3 & 0 \end{array}\right)$$

4. Si c'est possible, je déplace ensuite mon curseur d'un cran vers la droite, puis un cran vers le bas. Je recommence l'algorithme depuis le début. Si ce n'est pas possible l'algorithme s'arrête.

$$\left(\begin{array}{ccccc|c} \mathbf{1} & 2 & 1 & 2 & 1 & 2 \\ 0 & 0 & \mathbf{1} & -1 & -1 & 4 \\ 0 & 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 3 & -3 & 3 & 0 \end{array}\right)$$

$$\left(\begin{array}{ccccc|c} \mathbf{1} & 2 & 0 & 3 & 2 & -2 \\ 0 & 0 & \mathbf{1} & -1 & -1 & 4 \\ 0 & 0 & 0 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 & 6 & -12 \end{array}\right)$$

$$\left(\begin{array}{ccccc|c} \mathbf{1} & 2 & 0 & 3 & 2 & -2 \\ 0 & 0 & \mathbf{1} & -1 & -1 & 4 \\ 0 & 0 & 0 & 0 & 2 & -4 \\ 0 & 0 & 0 & 0 & 6 & -12 \end{array}\right)$$

$$\left(\begin{array}{ccccc|c} \mathbf{1} & 2 & 0 & 3 & 2 & -2 \\ 0 & 0 & \mathbf{1} & -1 & -1 & 4 \\ 0 & 0 & 0 & 0 & \mathbf{1} & -2 \\ 0 & 0 & 0 & 0 & 6 & -12 \end{array}\right)$$

$$\left(\begin{array}{ccccc|c} \mathbf{1} & 2 & 0 & 3 & 0 & 2 \\ 0 & 0 & \mathbf{1} & -1 & 0 & 2 \\ 0 & 0 & 0 & 0 & \mathbf{1} & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Ceci, retraduit en termes de systèmes linéaires, signifie

$$\left\{\begin{array}{ccccc} x_1 & +2x_2 & & +3x_4 & & = 2 \\ & & x_3 - & x_4 & & = 2 \\ & & & & x_5 & = -2 \end{array}\right\}.$$

Ce système a une infinité de solutions. On peut choisir x_2 et x_4 librement, et exprimer $x_1 = -2x_2 - 3x_4 + 2$, $x_3 = x_4 + 2$, $x_5 = -2$.

Dans l'exemple ci-dessus, le rang est 3.

Définition 2.2. La FREL d'une matrice A (forme réduite échelonnée par lignes) est la matrice obtenue après application de l'algorithme de GJ.

Autres exemples (exercice 9 de la feuille 1)

$\star$

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

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

$$\begin{pmatrix} 1 & 1 & -2 & | & 5 \\ 0 & 1 & 8 & | & -8 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -10 & | & 13 \\ 0 & 1 & 8 & | & -8 \end{pmatrix}$$

Conclusion : le système a une infinité de solutions. On peut choisir z arbitrairement ; x et y sont alors donnés par $x = 13 + 2z$, $y = -8 - 8z$.

$\star$

$$\left\{\begin{array}{ccc} x + y & = & 1 \\ 2x - y & = & 5 \\ 3x + 4y & = & 2 \end{array}\right.$$

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

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

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

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

Conclusion : le système a une solution unique donnée par $x = 2$, $y = -1$.

⋆

$$\begin{cases} x_3 + x_4 & = & 0 \\ x_2 + x_3 & = & 0 \\ x_1 + x_2 & = & 0 \\ x_1 + x_4 & = & 0 \end{cases}$$

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

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

$$\begin{pmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & -1 & 0 & 1 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 1 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 1 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & 1 & | & 0 \\ 0 & 1 & 0 & -1 & | & 0 \\ 0 & 0 & 1 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}$$

Conclusion : le système a une infinité de solutions. On peut choisir x_4 arbitrairement ; les autres x_i sont alors donnés par $x_1 = -x_4$, $x_2 = x_4$, $x_3 = -x_4$.

⋆

$$\begin{cases} 4x_1+3x_2+2x_3-x_4 &= 4 \\ 5x_1+4x_2+3x_3-x_4 &= 4 \\ -2x_1-2x_2-x_3+2x_4 &= -3 \\ 11x_1+6x_2+4x_3+x_4 &= 11 \end{cases}$$

$$\begin{pmatrix} 4 & 3 & 2 & -1 & | & 4 \\ 5 & 4 & 3 & -1 & | & 4 \\ -2 & -2 & -1 & 2 & | & -3 \\ 11 & 6 & 4 & 1 & | & 11 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/4 & 1/2 & -1/4 & | & 1 \\ 5 & 4 & 3 & -1 & | & 4 \\ -2 & -2 & -1 & 2 & | & -3 \\ 11 & 6 & 4 & 1 & | & 11 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/4 & 1/2 & -1/4 & | & 1 \\ 0 & 1/4 & 1/2 & 1/4 & | & -1 \\ 0 & -1/2 & 0 & 3/2 & | & -1 \\ 0 & -9/4 & -3/2 & 15/4 & | & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3/4 & 1/2 & -1/4 & | & 1 \\ 0 & 1 & 2 & 1 & | & -4 \\ 0 & -1/2 & 0 & 3/2 & | & -1 \\ 0 & -9/4 & -3/2 & 15/4 & | & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 & -1 & | & 4 \\ 0 & 1 & 2 & 1 & | & -4 \\ 0 & 0 & 1 & 2 & | & -3 \\ 0 & 0 & 3 & 6 & | & -9 \end{pmatrix}$$

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

Conclusion : le système a une infinité de solutions. On peut choisir x_4 arbitrairement ; les autres x_i sont alors donnés par $x_1 = 1 - x_4$, $x_2 = 2 + 3x_4$, $x_3 = -3 - 2x_4$.

$\star$

$$\begin{cases} 3x+11y+19z &= 22 \\ 7x+23y+39z &= 10 \\ -4x-3y-2z &= 6 \end{cases}$$

$$\begin{pmatrix} 3 & 11 & 19 & | & 22 \\ 7 & 23 & 39 & | & 10 \\ -4 & -3 & -2 & | & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 11/3 & 19/3 & | & 22/3 \\ 7 & 23 & 39 & | & 10 \\ -4 & -3 & -2 & | & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 11/3 & 19/3 & | & 22/3 \\ 0 & -8/3 & -16/3 & | & -124/3 \\ 0 & 35/3 & 70/3 & | & 106/3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 11/3 & 19/3 & | & 22/3 \\ 0 & 1 & 2 & | & 31/2 \\ 0 & 35 & 70 & | & 106 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 11/3 & 19/3 & | & 22/3 \\ 0 & 1 & 2 & | & 31/2 \\ 0 & 0 & 0 & | & 106 - 35.31/2 < 0 \end{pmatrix}$$

Conclusion : le système n'a pas de solutions, on dit qu'il est incompatible.

⋆

$$\begin{cases} 3x + 6y + 14z & = & 22 \\ 7x + 14y + 30z & = & 46 \\ 4x + 8y + 7z & = & 6 \end{cases}$$

$$\begin{pmatrix} 3 & 6 & 14 & | & 22 \\ 7 & 14 & 30 & | & 46 \\ 4 & 8 & 7 & | & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 14/3 & | & 22/3 \\ 7 & 14 & 30 & | & 46 \\ 4 & 8 & 7 & | & 6 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 14/3 & | & 22/3 \\ 0 & 0 & -8/3 & | & -16/3 \\ 0 & 0 & -35/3 & | & -70/3 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 14/3 & | & 22/3 \\ 0 & 0 & 1 & | & 2 \\ 0 & 0 & 1 & | & 2 \end{pmatrix}$$

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

Conclusion : le système a une infinité de solutions. On peut choisir y arbitrairement ; x et z sont alors donnés par $x = -2 - 2y$, $z = 2$.

⋆

$$\begin{cases} 3x+5y+3z &= 25 \\ 7x+9y+19z &= 65 \\ -4x+5y+11z &= 5 \end{cases}$$

$$\left(\begin{array}{ccc|c} 3 & 5 & 3 & 25 \\ 7 & 9 & 19 & 65 \\ -4 & 5 & 11 & 5 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 5/3 & 1 & 25/3 \\ 0 & -8/3 & 12 & 20/3 \\ 0 & 35/3 & 15 & 115/3 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 5/3 & 1 & 25/3 \\ 0 & 1 & -9/2 & -5/2 \\ 0 & 7 & 9 & 23 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 17/2 & 25/2 \\ 0 & 1 & -9/2 & -5/2 \\ 0 & 0 & 81/2 & 81/2 \end{array}\right)$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 17/2 & 25/2 \\ 0 & 1 & -9/2 & -5/2 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

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

Conclusion : le système a une solution unique donnée par $x = 4$, $y = 2$, $z = 1$.

2.3 Calcul de l'inverse d'une matrice par GJ

Rappelons qu'une matrice $A \in M_{n\times n}(\mathbb{R})$ est dite inversible s'il existe une matrice B telle que $A.B = B.A = I_n$. On note cette matrice A^{-1}, et on l'appelle l'inverse de A.

Exemple $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{pmatrix}$ est-elle inversible ? Ceci revient à dire que le système

$$\begin{cases} x_1 + x_2 + x_3 = y_1 \\ 2x_1 + 3x_2 + 2x_3 = y_2 \\ 3x_1 + 8x_2 + 2x_3 = y_3 \end{cases}$$

admet une unique solution. On résout par GJ. On obtient

$$\begin{cases} x_1 + x_2 + x_3 = y_1 \\ 0 + x_2 = -2y_1 + y_2 \\ 5x_2 - x_3 = -3y_1 + y_3 \end{cases}$$

puis

$$\left\{\begin{array}{ccccccccc} x_1 & & + & x_3 & = & 3y_1 & -y_2 & \\ & x_2 & & & = & -2y_1 & +y_2 & \\ & & - & x_3 & = & 7y_1 & -5y_2 & +y_3 \end{array}\right.$$

puis

$$\left\{\begin{array}{cccccccc} x_1 & & & = & 10y_1 & -6y_2 & +y_3 \\ & x_2 & & = & -2y_1 & +y_2 & \\ & & x_3 & = & -7y_1 & +5y_2 & -y_3 \end{array}\right.$$

Ainsi,

$$X = \begin{pmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{pmatrix} .Y$$

ce qui signifie que

$$A^{-1} = \begin{pmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{pmatrix} .$$

On présente souvent de manière un peu différente ce calcul (en n'écrivant pas les x_i et y_i). On applique l'algorithme de Gauss-Jordan à $(A|I_3)$ et on obtient $(I_3|A^{-1})$.

Dans le cas précédent, cela donne

$$\begin{pmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 \\ 2 & 3 & 2 & | & 0 & 1 & 0 \\ 3 & 8 & 2 & | & 0 & 0 & 1 \end{pmatrix}$$

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

$$\begin{pmatrix} 1 & 0 & 1 & | & 3 & -1 & 0 \\ 0 & 1 & 0 & | & -2 & 1 & 0 \\ 0 & 0 & -1 & | & 7 & -5 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 & | & 10 & -6 & 1 \\ 0 & 1 & 0 & | & -2 & 1 & 0 \\ 0 & 0 & 1 & | & -7 & 5 & -1 \end{pmatrix}$$

C'est vrai en toute dimension. Si A est inversible, pour trouver A^{-1}, on applique GJ à la matrice étendue $(A|I_n)$ et on obtient $(I_n|A^{-1})$.

Un exemple : inverser la matrice

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

(c'est la deuxième matrice de l'exercice 2 de la feuille 2). On applique la méthode de Gauss-Jordan à la matrice augmentée par I_3 :

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

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

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

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

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

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

Conclusion

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & -1 \\ 2 & -1 & 2 \end{pmatrix}$$

2.4 Structure de l'ensemble des solutions d'un système linéaire

Un système linéaire $AX = B$ est dit compatible ou consistant s'il a au moins une solution, incompatible ou inconsistant sinon.

Théorème 2.3. *Soit A la matrice des coeff d'un système linéaire* **homogène** $AX = \vec{0}$.

- *L'ensemble des solutions du système linéaire* **homogène** *est un seV.*
- *Un système linéaire homogène a toujours des solutions*
- *Le système a une unique solution lorsque $Frel(A)$ a exactement un pivot par colonne (autant de pivots que d'inconnues).*
- *L'ensemble des solutions est infini, de dimension $nbcolonnes - nbpivots$, lorsqu'une colonne de $Frel(A)$ au moins ne contient pas de pivot (moins de pivot que d'inconnues).*

Soit maintenant (S) un système linéaire avec second membre $AX = B$. Observons que si X_0 est une solution particulière de (S), et X est une autre solution, alors $A(X - X_0) = \vec{0}$. Autrement dit, $X - X_0$ est solution du système homogène. Les solutions s'écrivent toutes sous la forme $X = X_0 + X'$ où $X' = X - X_0$ appartient au sev défini par l'équation homogène $AY = 0$. On dit aussi que l'ensemble des solutions est un sous-espace affine.

Définition 2.4. *Appelons rang d'une matrice A le nombre de pivots de la forme échelonnée obtenue après application de l'algorithme de Gauss-Jordan.*

Ce n'est pas la définition naturelle que nous donnerons plus tard. Nous verrons que les deux définitions coàŕncident.

Théorème 2.5. *Soit A la matrice des coeff d'un système linéaire (S) $AX = B$, et $C = (A|B)$ la matrice augmentée.*
- **L'ensemble des solutions d'un système linéaire est un seV uniquement lorsque le système est homogène (second membre nul)**
- *Le système n'a pas de solution ssi la Frel(C) contient la ligne $0\dots0|1$.*
- *Si le système est consistant, il a une unique solution lorsque $Frel(A)$ a exactement un pivot par colonne (autant de pivots que d'inconnues).*
- *Si le système est consistant, il a une infinité de solutions lorsque une colonne de $Frel(A)$ au moins ne contient pas de pivot (moins de pivot que d'inconnues).*

Théorème 2.6. *Soit $A \in M_{n\times m}(\mathbb{R})$. On a $rg(A) \leq n$ et $rg(A) \leq m$.*
Si $rgA = n$ (nb d'équations) le système est compatible.
Si $rg(A) = m$ (nb d'inconnues) le système a au plus une solution
Si $rg(A) < m$ (nb d'inconnues) alors le système a une infinité ou aucune solution.
Le système a une unique solution lorsque $rg(A) = n = m$ et $Frel(A) = I_n$.

Théorème 2.7. *Soit $A \in M_{n\times n}(\mathbb{R})$ une matrice carrée. Les assertions suivantes sont équivalentes :*

1. *A est inversible,*
2. *$rgA = n$,*
3. *$AX = 0 \implies X = 0$,*
4. *$X \neq 0 \implies AX \neq 0$,*
5. *$\forall Y\ \exists X\ AX = Y$,*

6. $\forall Y \; \exists! X \; AX = Y$.

Explications

3. et 4. sont contraposées l'une de l'autre donc équivalentes.
6. entraîne 5. : elle est plus précise.
1. entraîne 3. : en effet si A est inversible et $AX = 0$ alors

$$A^{-1}(AX) = A^{-1}0 = 0 \text{ et } A^{-1}(AX) = (A^{-1}A)X = IX = X,$$

autrement dit $X = 0$.
1. entraîne 5. : en effet supposons A inversible et donnons nous un vecteur Y ; en posant $X = A^{-1}Y$ on a

$$AX = A(A^{-1}Y) = (AA^{-1})Y = IY = Y,$$

on a donc bien l'existence de X tel que $AX = Y$.
5. entraîne 1. : considérons les vecteurs de la base canonique e_i ; d'après 5. il existe des vecteurs colonne C_i tels que $AC_i = e_i$. Considérons alors la matrice B dont les vecteurs colonnes sont les C_i. Alors le i-ème vecteur colonne du produit AB est $AC_i = e_i$. Cela signifie que $AB = I$ et A est inversible. Pourquoi a-t-on $BA = I$ aussi ?
5. entraîne 6. : soit X et X' deux solutions

$$A(X - X') = Y - Y = 0.$$

Comme 5. entraîne 1. qui entraîne 3., on a $X - X' = 0$, soit $X = X'$.
5. entraîne 2. : si $rang(A) < n$ alors on peut trouver des Y tels que le système $AX = Y$ soit incompatible ; autrement dit si 2. est faux, 5. l'est aussi.
3. entraîne 2. : si $rang(A) < n$ alors le système homogène a une infinité de solution, autrement dit si 2. est faux alors 3. est faux (3. signifie que le système homogène a une seule solution).

3 Dimensions dans les espaces vectoriels, dimensions associées à une matrice

3.1 Injectivité, surjectivité

Définition 3.1. *Une application (linéaire) $f : E \to F$ est injective si deux vecteurs distincts ont des images distinctes. Autrement dit, si $x \neq y$ alors $f(x) \neq f(y)$. Ou encore, si $f(x) = f(y)$ alors $x = y$.*

Elle est dite surjective si pour tout $z \in F$, il existe au moins un $x \in E$ tq $f(x) = z$.

Elle est bijective si elle est injective et surjective.

Définition 3.2. *Le noyau d'une application* **linéaire** *de E dans F est l'ensemble* $Ker(f) = \{\vec{x} \in E, f(\vec{x}) = 0_F\}$.

C'est un sous ev de E (exo, à faire.)

Proposition 3.3. *Si E, F sont deux ev et $f : E \to F$ est linéaire, alors f est injective ssi $Ker(f) = \{0_E\}$*

Autrement dit pour une application linéaire il suffit de vérifier l'injectivité pour 0.

Démonstration : En effet, $f(x) = f(y)$ équivaut **par linéarité** à $f(x-y) = 0_F$, autrement dit $x - y \in Ker(f)$.

Dire que $Ker(f) = \{0_E\}$ signifie donc que $x - y = 0_E$ soit encore $x = y$, donc f est injective. Réciproquement, si f est injective, et si $x \in Ker(f)$, alors $f(x) = 0_F = f(0_E)$. Donc $x = 0_E$. Cela dit bien que $Ker(f) = \{0_E\}$. □

Définition 3.4. *Soit $f : E \to F$. L'image de f, notée $Im(f)$, est l'ensemble $Im(f) = \{z \in F, \exists x \in E, f(x) = z\}$. C'est un sev de F*

Proposition 3.5. *Soient E, F deux espaces vectoriels et $f : E \to F$ une application linéaire. Si G est un sev de E, alors $f(G) = \{y \in F, \exists x \in E, y = f(x)\}$ est un sev de F. Si H est un sev de F alors $f^{-1}(H) = \{x \in E, f(x) \in H\}$ est un sev*

Attention à la notation

Exemple $Im(f) = f(E)$ et $Ker(f) = f^{-1}(\{0\})$.

Démonstration : A faire pour manipuler □

Exemple

$$f : (x, y, z, y) \in \mathbb{R}^4 \mapsto (x, x + y, y + z, z - x, y - x) \in \mathbb{R}^5$$

Alors $Ker(f) = \{(0, 0, 0, t), t \in \mathbb{R}\}$ et $Im(f) = \{x(1, 1, 0, -1, -1) + y(0, 1, 1, 0, 1) + z(0, 0, 1, 1, 0)\}$.

Exemple tiré de l'analyse. Sur $E = C^\infty([0, 1], \mathbb{R})$, $D : f \to f'$ et $I : f \to \int_0^1 f(t)dt$, $I : E \to \mathbb{R}$. Alors $KerD = \{cstes\}$ et $Im(D) = E$.

Soit A une matrice $n \times p$. Notons aussi l'application linéaire donnée par la multiplication par A. Trouver le noyau de l'application A c'est résoudre le système linéaire $AX = 0$:

$$ker(A) = \{X \ / \ AX = 0\}.$$

Trouver l'ensemble $Im(A)$ c'est trouver les seconds membres Y tels que le système $AX = Y$ soit compatible.

3.2 Isomorphismes

Lorsque f est bijective, pour tout $z \in F$ il existe un et un seul $x \in E$ tq $f(x) = z$. on note f^{-1} l'inverse (réciproque) de f i.e. l'application qui à $z \in F$ associe cet unique x. Par construction, on vérifie (vérifiez le!!) que $f \circ f^{-1}(z) = z$ et $f^{-1} \circ f(x) = x$.

Soit A une matrice $n \times p$. Notons aussi l'application linéaire donnée par la multiplication par A. Si $n < p$, A ne peut pas être injective (le système $AX = 0$ a une infinité de solutions). Si $n > p$, A ne peut pas être surjective (nous y reviendrons; le voir pour une matrice échelonnée). Si $n = p$, A est bijective si et seulement si A est inversible.

Proposition 3.6. *Si $f : E \to F$ est une application linéaire bijective, alors son inverse $f^{-1} : F \to E$ est elle aussi linéaire et bijective.*

Une application linéaire bijective est souvent appelée un <u>isomorphisme</u>

<u>Démonstration :</u> Soient $z, z' \in F$ et $x = g(z), x' = g(z')$. Alors $g(z+z') = g(f(x)+f(x')) = g(f(x + x')) = x + x' = g(z) + g(z')$ par linéarité de f. De même $g(\lambda z) = g(\lambda f(x)) = g(f(\lambda x)) = \lambda x = \lambda g(z)$.

□

Proposition 3.7. *Soit $f, g : E \to E$ linéaires. Si f, g bijectives, alors $f \circ g$ est bijective et $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$*

<u>Démonstration :</u> Soit $z \in E$. Il existe un unique y tq $f(y) = z$ et il existe un unique x tq $g(x) = y$. Donc $f \circ g(x) = z$. Et s'il existe un autre x' tq $f \circ g(x') = z$, comme f est injective, on aurait $g(x') = y = g(x)$ et comme g injective, on aurait $x' = x$. Donc x est unique donc $f \circ g$ bijective. Et $x = g^{-1}(y)$ avec $y = f^{-1}(z)$ donc $x = g^{-1} \circ f^{-1}(z) = (f \circ g)^{-1}(z)$.
□

Proposition 3.8. *Si $f : \mathbb{R}^n \to \mathbb{R}^n$ est un isomorphisme, alors*

$$Mat(f) \times Mat(f^{-1}) = Mat(Id) = Mat(f^{-1}) \circ Mat(f) = I_n$$

3.3 Familles libres et génératrices

Définition 3.9. *On dit qu'une partie S est génératrice si $Vect(S) = E$.*

Exemple $E = \mathbb{R}^2$, $S = \{(1, 0)\}$ n'est pas génératrice $S = \mathbb{R}^2$ est génératrice. $S = \{(1, 1), (2, 1), (1, -1)\}$ est génératrice. $S = \{(1, 1), (-1, -1)\}$ n'est pas génératrice.

Exemple Toute forme linéaire $\varphi : \mathbb{R}^n \to \mathbb{R}$ s'écrit de la manière suivante $\varphi(x) = a_1x_1 + \dots a_nx_n$. Notons $p_i : x \to x_i$. Alors $\varphi = a_1p_1 + \dots a_np_n$. Autrement dit la famille $\{p_1, \dots, p_n\}$ est génératrice de l'ensemble $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ des formes linéaires.

Soit A une matrice $n \times p$. Notons aussi A l'application linéaire associée. Dire que A est surjective revient à dire que le sev engendré par les vecteurs colonne de A est $\mathbb{R}^n$, ou encore que la famille des vecteurs colonne de A est génératrice.

Remarque 3.10. Si $S \subset G$ et S génératrice, alors G aussi.

Définition 3.11. *Une famille est linéairement dépendante ssi il existe une combinaison linéaire $\lambda_1\vec{u}_1 + \dots \lambda_n\vec{u}_n = \vec{0}$ avec au moins l'un des $\lambda_i \neq 0$. On dit qu'il existe une combinaison linéaire nulle non triviale des (u_i).*

Proposition 3.12. *Soit E ev. Dans une famille $\vec{u}_1, \dots, \vec{u}_p$, le vecteur $\vec{u}_k$, $1 \leq k \leq p$ est dit redondant s'il s'écrit comme combinaison linéaire des vecteurs précédents $\vec{u}_1, \dots, \vec{u}_{k-1}$. Une famille $\vec{u}_1, \dots, \vec{u}_n$ est dite linéairement dépendante ou liée s'il existe (au moins) un vecteur redondant dans la famille.*

Démonstration : Supposons que $\vec{u}_1, \dots, \vec{u}_n$ soit liée. Il existe une combinaison linéaire $\lambda_1\vec{u}_1 + \dots \lambda_n\vec{u}_n = \vec{0}$ avec au moins l'un des $\lambda_i \neq 0$. Prenons i_0 le plus grand indice tel que $\lambda_{i_0} \neq 0$. On a alors

$$\sum_{j=1}^{i_0} \lambda_i u_i = 0, \text{ ou encore } \lambda_{i_0}u_{i_0} = \sum_{j=1}^{i_0-1} (-\lambda_i)u_i,$$

ce qui donne

$$u_{i_0} = \sum_{j=1}^{i_0-1} (-\lambda_i/\lambda_{i_0})u_i.$$

□

En particulier une famille $\vec{u}_1, \dots, \vec{u}_n$ est liée si l'un des vecteurs de la famille est combinaison linéaire des autres.

Définition 3.13. *Soit E ev. Une famille qui n'est pas linéairement dépendante (i.e. ne contient pas de vecteur redondant) est dite libre.*

La famille $\vec{u}_1, \dots, \vec{u}_n$ est libre si on a l'implication

$$\lambda_1\vec{u}_1 + \dots \lambda_n\vec{u}_n = \vec{0} \Rightarrow \forall i, \lambda_i = 0.$$

Exemple dans $\mathbb{R}$, deux vecteurs non nuls sont toujours liés car $u_2u_1 - u_1u_2 = 0$.

Exemple dans $\mathbb{R}^2$: $(1,0)$ et $(0,1)$ forment une famille libre, $(1,1)$, $(1,0)$ et $(0,1)$ forment une famille liée.

Dans $\mathbb{R}^4$, $((1,1,1,0),(1,0,1,0),(0,1,0,0),(0,0,0,1))$ est liée.

Soit A une matrice $n \times p$. Notons aussi A l'application linéaire associée. Dire que A est injective revient à dire que la famille des vecteurs colonne de A est libre.

Proposition 3.14. *Soit E ev Si $S \subset L$ et L est libre alors S aussi.*

Démonstration : Si S n'est pas libre, elle contient un vecteur redondant. Donc L aussi... $\square$

Proposition 3.15. *Soit $(u_1,\ldots,u_k)$ des vecteurs de E.*
- *La famille $(u_1,\ldots,u_k)$ est libre ssi $\varphi : \mathbb{R}^k \to E$ définie par $(\lambda_1,\ldots,\lambda_k) \to \lambda_1 u_1 + \ldots \lambda_k u_k$ est injective.*
- *La famille $(u_1,\ldots,u_k)$ est génératrice ssi $\varphi : \mathbb{R}^k \to E$ définie par $(\lambda_1,\ldots,\lambda_k) \to \lambda_1 u_1 + \ldots \lambda_k u_k$ est surjective.*

Proposition 3.16. *Soit $(u_1,\ldots,u_k)$ des vecteurs de E.*
- *Si $(u_1,\ldots u_k)$ est libre et $f : E \to F$ est injective, alors $(f(u_1),\ldots,f(u_k))$ est libre*
- *Si $(u_1,\ldots u_k)$ est génératrice et $f : E \to F$ est surjective, alors $(f(u_1),\ldots,f(u_k))$ est génératrice.*

Démonstration : On a $f(\sum \lambda_i u_i) = \sum \lambda_i f(u_i)$ par linéarité. Si $\sum \lambda_i f(u_i) = 0$ alors $f(\sum \lambda_i u_i) = 0$ et comme f injective, $\sum \lambda_i u_i = 0$, donc tous les λ_i sont nuls. Donc les $f(u_i)$ sont libres.

De même f est surjective donc tout $z \in F$ s'écrit $z = f(x)$. La famille (u_i) est génératrice donc $x = \sum \lambda_i u_i$ donc $z = \sum \lambda_i f(u_i)$. $\square$

Remarque : l'image par une application linéaire d'une famille liée est une famille liée ; par contraposée, si f est une application linéaire et $f(u_1),\ldots,f(u_k)$ est libre alors $u_1,\ldots,u_k$ est libre.

3.4 Bases

Définition 3.17. *Soit E ev. Une famille $(u_1,\ldots,u_k)$ est une base ssi elle est libre et génératrice. Elle est libre si $\sum \lambda_i u_i = 0$ implique que tous les λ_i sont nuls. Elle est génératrice si $Vect(u_1,\ldots,u_k) = E$.*

Exemples de bases.

Dans $\mathbb{R}^n$ $(1,0,\ldots,0),(0,1,0\ldots0),\ldots(0,\ldots,0,1)$.

Dans $\mathbb{C}^n$ comme $\mathbb{C}$-ev $(1,0,\ldots,0),(0,1,0\ldots0),\ldots(0,\ldots,0,1)$.

Dans $\mathbb{C}^n$ comme $\mathbb{R}$-ev $(1,0,\ldots,0),(i,0,\ldots,0),(0,1,0\ldots0),(0,i,0\ldots0)\ldots$

Dans $\mathbb{R}_n[X]$, $(1,X,\ldots,X^n)$

Dans $\mathcal{L}(\mathbb{R}^n,\mathbb{R})$ $p_1,\ldots p_n$

Soit A une matrice $n\times n$. Notons aussi A l'application linéaire associée. Dire que A est bijective revient à dire que la famille des vecteurs colonne de A forme une base de $\mathbb{R}^n$.

Proposition 3.18. *Soit (u_i) une base de E. Alors tout $u\in E$ s'écrit de manière unique comme $u=\sum\lambda_i u_i$*

Démonstration : existence car génératrice, unicité car libre □

Proposition 3.19. *Soit (u_i) base de E et $f:E\to F$ linéaire bijective. Alors $(f(u_i))$ est une base de F.*

Proposition 3.20. *Soit $(u_1,\ldots,u_k)$ des vecteurs de E. La famille $(u_1,\ldots,u_k)$ est une base ssi $\varphi:\mathbb{R}^k\to E$ définie par $(\lambda_1,\ldots,\lambda_k)\to\lambda_1u_1+\ldots\lambda_ku_k$ est bijective.*

Théorème 3.21 (Base incomplète)**.** *Soit E ev. Si $(u_i)_{i\in I}\subset(u_i)_{i\in J}$, que la première famille est libre et la deuxième est génératrice, il existe une famille intermédiaire qui est une base.*

Exemple dans $\mathbb{R}^3$, $\{(1,0,1),(1,1,1)\}\subset\{(1,0,1),(1,1,1),(1,3,4),(2,1,0)\}$

Corollaire 3.22. *Tout ev possède une base. Toute famille libre incluse dans une base. Toute famille génératrice contient une base.*

Explications pour des familles finies.

Supposons que soit donnée une famille $(u_1,\ldots,u_n)$ génératrice d'un espace vectoriel $E\neq\{0\}$. Pourquoi peut-on en extraire une base ? Si la famille $(u_1,\ldots,u_n)$ est libre, alors c'est déjà une base. Sinon l'un des vecteurs est combinaison linéaire des autres. Disons par exemple que u_n est combinaison linéaire des autres. Alors l'espace vectoriel engendré par $(u_1,\ldots,u_n)$ est égal à l'espace vectoriel engendré par $(u_1,\ldots,u_{n_1})$. Cela signifie que $(u_1,\ldots,u_{n-1})$ est génératrice. Si elle est libre, c'est une base. Sinon l'un de ses vecteurs est combinaison linéaire des autres. On peut l'enlever tout en conservant une famille génératrice... On enlève ainsi des vecteurs redondants jusqu'à obtenir une famille libre (toujours génératrice) donc une base.

Supposons maintenant que soit donnée une famille $(u_1,\ldots,u_n)$ libre dans un espace vectoriel $E\neq\{0\}$. Si cette famille est génératrice alors c'est une base. Sinon il existe un vecteur

u_{n+1} de E qui n'est pas combinaison linéaire de $u_1, \ldots, u_n$ et la famille $(u_1, \ldots, u_{n+1})$ est libre. Si elle est génératrice alors c'est une base. Sinon on peut lui ajouter un vecteur etc... Lorsque E est de dimension finie, on obtient une base au bout d'un certain nombre de répétitions. Si E est de dimension infinie le résultat est encore vrai mais il faut procéder autrement pour le démontrer.

3.5 Dimension

Théorème 3.23. *Toutes les bases de E ont même cardinal, appelé dimension de E*

Démonstration :

Soit $(e_1, \ldots, e_n)$ une base d'un espace vectoriel. Soit $(u_1, \ldots, u_n)$ une famille libre de ce même espace vectoriel. Écrivons les coordonnées de chaque vecteur u_i dans la base $(e_1, \ldots, e_n)$. On obtient n vecteur colonne qu'on peut ranger les un à côté des autres dans une matrice A. La matrice A est carrée (de taille $n \times n$) et ses vecteurs colonne forment une famille libre (car les u_i forment une famille libre). Comme on l'a vu cela s'exprime par l'implication : $AX = 0 \implies X = 0$. Or on a vu aussi que cela entraînait que A est inversible et que la famille des vecteurs colonne de A est une base. Autrement dit la famille $(u_1, \ldots, u_n)$ est une base.

Cela montre que les bases ont toutes le même nombre d'éléments. En effet supposons que $(e_1, \ldots, e_n)$ $(f_1, \ldots, f_p)$ soient deux bases de cardinaux différents. Par exemple on a $p < n$. Comme $(e_1, \ldots, e_n)$ est une base, c'est une famille libre, donc $(e_1, \ldots, e_p)$ est aussi une famille libre. D'après ce qu'on a dit plus haut $(e_1, \ldots, e_p)$ est une base. Mais alors e_{p+1} par exemple est une combinaison linéaire des vecteurs $e_1, \ldots, e_p$. Cela contredit le fait que la famille $(e_1, \ldots, e_n)$ est une base (donc libre).

□

Dans toute la suite on suppose que E est un ev dont les bases ont un nb fini d'éléments, c'est la dimension.

Proposition 3.24. *Si E est de dimension n toute famille génératrice a au moins n éléments et toute famille libre au plus n éléments.*

Proposition 3.25. *Soit E de dim n. Une famille $(u_i)_{1 \leq i \leq n}$ est libre ssi elle est génératrice ssi elle est une base.*

Proposition 3.26. *Si F sev de E,* $\dim F \leq \dim E$.

Proposition 3.27. *Soit $f : E \to F$ linéaire,*

1. *si f injective,* $\dim F \geq \dim E$,

2. *si f surjective* $\dim F \leq \dim E$*,*
3. *si f bijective* $\dim F = \dim E$*.*

Soit A une matrice $n \times p$. Notons aussi A l'application linéaire associée. Si $n < p$ A ne peut pas être injective, si $n > p$ A ne peut pas être surjective.

Proposition 3.28. *Si $f : \mathbb{R}^n \to \mathbb{R}^m$ est une application linéaire bijective, alors $n = m$.*

Exemple $(x, y, z, t) \in \mathbb{R}^4 \to (3x + 5y, 7y + 11z, 3x + z + t)$ non injective.

$\mathbb{R}^4 \to \mathbb{R}^5$ non surjective

Proposition 3.29. *Si E et F ont même dimension on peut construire un isomorphisme de E dans F.*

Définition 3.30. *Soit $F \subset E$ un sev. Un paramétrage linéaire de F est une application linéaire bijective $\varphi : \mathbb{R}^k \to F$ pour un certain $k \geq 0$.*

Par exemple $F = \{(x, y, z) \in \mathbb{R}^3, x + y + z = 0\}$. Alors $\varphi : \mathbb{R}^2 \to F$ définie par $(\lambda, \mu) \to (\lambda, \mu, -\lambda - \mu)$ est une param. linéaire de F.

Exercice 1. Soient $u_1 = (1, 2, 3)$ et $u_2 = (2, -1, 1)$, $v_1 = (1, 0, 1)$ et $v_2 = (2, -1, 1)$. Montrer que le sous-espace engendré par u_1 et u_2 coïncide avec le sous-espace vectoriel engendré par v_1 et v_2, de deux manières.

1. En écrivant u_1 et u_2 chacun comme combinaison linéaire de v_1 et v_2.
2. En trouvant des équations pour chacun de ces deux sous-espaces.
3. Soient P_1 le plan paramétré par $P_1 = \{(\lambda, \mu, \lambda + \mu), \lambda \in \mathbb{R}, \mu \in \mathbb{R}\} \subset \mathbb{R}^3$ et $P_2 = \{(\lambda + 2\mu, 2\lambda - \mu, 3\lambda + \mu), \lambda \in \mathbb{R}, \mu \in \mathbb{R}\} \subset \mathbb{R}^3$. Comparer ces deux plans.

C'est l'exercice 7 de la feuille 3.

1. On cherche α_1, α_2 tels que $u_1 = \alpha_1 v_1 + \alpha_2 v_2$ et β_1, β_2 tels que $u_2 = \beta_1 v_1 + \beta_2 v_2$. Cela revient à résoudre les deux systèmes linéaires

$$\left\{\begin{array}{ccc} \alpha_1 & +2\alpha_2 & = & 1 \\ 0.\alpha_1 & -\alpha_2 & = & 2 \\ \alpha_1 & +\alpha_2 & = & 3 \end{array}\right. \quad \left\{\begin{array}{ccc} \beta_1 & +2\beta_2 & = & 2 \\ 0.\beta_1 & -\beta_2 & = & -1 \\ \beta_1 & +\beta_2 & = & 1 \end{array}\right.$$

Remarque : a priori il n'y a pas de raison que ces systèmes aient des solutions; l'énoncé semble indiquer que oui. On peut résoudre les deux en même temps par la méthode de Gauss-Jordan puisque les matrices des premiers membres sont les mêmes :

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

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

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

$$\left(\begin{array}{cc|cc} 1 & 0 & 5 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Conclusion les deux systèmes ont chacun une solution unique. On obtient :

$$u_1 = 5v_1 - 2v_2 \text{ et } u_2 = v_2$$

Oui, on aurait pu remarquer que u_2 et v_2 étaient égaux.

Comme u_1 et u_2 sont des combinaisons linéaires de v_1 et v_2, on a

$$Vect(u_1, u_2) \subset Vect(v_1, v_2).$$

Mais u_1 et u_2 ne sont pas colinéaires. En effet, comme ni u_1 ni u_2 n'est nul, ils seraient colinéaires s'il existait un nombre λ tel que $u_1 = \lambda u_2$. On aurait en particulier $1 = 2\lambda$ et $2 = -\lambda$, ce qui est impossible.

Le sous-espace vectoriel $Vect(u_1, u_2)$ est donc un plan. Comme $Vect(v_1, v_2)$ est au plus de dimension 2 et contient le plan $Vect(u_1, u_2)$ il est de dimension 2 et coïncide avec $Vect(u_1, u_2)$.

2. Pour trouver une équation de $Vect(u_1, u_2)$ écrivons que le vecteur $w = (x, y, z)$ est combinaison linéaire de u_1 et u_2 : il existe λ_1 et λ_2 tels que

$$w = \lambda_1 u_1 + \lambda_2 u_2$$

soit

$$\left\{\begin{array}{rrcl} \lambda_1 & +2\lambda_2 & = & x \\ 2\lambda_1 & -\lambda_2 & = & y \\ 3\lambda_1 & +\lambda_2 & = & z \end{array}\right.$$

Appliquons l'algorithme de Gauss-Jordan :

$$\left(\begin{array}{cc|c} 1 & 2 & x \\ 2 & -1 & y \\ 3 & 1 & z \end{array}\right)$$

$$\left(\begin{array}{cc|c} 1 & 2 & x \\ 0 & -5 & y - 2x \\ 0 & -5 & z - 3x \end{array}\right)$$

$$\left(\begin{array}{cc|c} 1 & 0 & 1/5.x+2/5.y \\ 0 & 1 & 2/5.x-1/5.y \\ 0 & 0 & z-x-y \end{array}\right)$$

Le système n'a de solution (alors unique) que si $z-x-y=0$. C'est l'équation de $Vect(u_1,u_2)$ recherchée.

Passons à $Vect(v_1,v_2)$. On procède de la même façon. Pour trouver une équation de $Vect(u_1,u_2)$ écrivons que le vecteur $w=(x,y,z)$ est combinaison linéaire de v_1 et v_2 : il existe λ_1 et λ_2 tels que

$$w=\lambda_1 v_1+\lambda_2 v_2$$

soit

$$\left\{\begin{array}{cccc} \lambda_1 & +2\lambda_2 & = & x \\ 0.\lambda_1 & -\lambda_2 & = & y \\ \lambda_1 & +\lambda_2 & = & z \end{array}\right.$$

Appliquons l'algorithme de Gauss-Jordan :

$$\left(\begin{array}{cc|c} 1 & 2 & x \\ 0 & -1 & y \\ 1 & 1 & z \end{array}\right)$$

$$\left(\begin{array}{cc|c} 1 & 2 & x \\ 0 & -1 & y \\ 0 & -1 & z-x \end{array}\right)$$

$$\left(\begin{array}{cc|c} 1 & 0 & x+2y \\ 0 & 1 & -y \\ 0 & 0 & z-x-y \end{array}\right)$$

Le système n'a de solution (alors unique) que si $z-x-y=0$. C'est l'équation de $Vect(v_1,v_2)$ recherchée.

Les deux plans $Vect(u_1,u_2)$ et $Vect(v_1,v_2)$ ont même équation ; ils sont égaux (il aurait suffit qu'ils aient des équations proportionnelles).

3. On remarque que
$$P_2=Vect(u_1,u_2)$$
et que si on pose $w_1=(1,0,1)$ et $w_2=(0,1,1)$,
$$P_1=Vect(w_1,w_2).$$
On vérifie que les coordonnées de w_1 et w_2 satisfont l'équation du plan $Vect(u_1,u_2)$. Par ailleurs w_1 et w_2 ne sont pas colinéaires. Conclusion : $P_1=P_2$.

Exercice 2. Montrer que F est un sous-espace vectoriel de $\mathbf{R}^n$, donner une base de F et sa dimension dans les cas suivants :

1. $F=\{(x_1,\dots,x_n)\in\mathbf{R}^n | x_1+x_2+\ldots+x_n=0\}$
2. $F=\{(x_1,\dots,x_n)\in\mathbf{R}^n | x_1+x_n=x_2+x_n=\ldots+x_n=x_{n-1}+x_n=0\}$

3.6 Théorème du rang et applications

Définition 3.31. *Le rang d'une application linéaire est* $rang(f) = \dim Im(f)$. *Le rang d'une matrice est* $rang(A) = rang(f_A) = \dim Im(f_A)$.

Autrement dit, comme l'image de f_A *est engendrée par les vecteurs colonnes de* A*, le rang d'une matrice est la dimension du sous-espace vectoriel engendré par ses colonnes. Le rang d'une famille de vecteurs est*

$$rang(u_1, \dots, u_k) = \dim Vect(u_1, \dots, u_k)\,.$$

C'est le rang de la matrice dont les colonnes sont $u_1, \dots, u_k$

Si S *syst linéaire, de matrice* A*, le rang de* S *est le rang de* A.

Proposition 3.32. *Si* f *surjective,* $rg(g \circ f) = rg(g)$ *Si* g *inj,* $rg(g \circ f) = rg(f)$*. Matriciellement, si* $A \in M_n(\mathbb{R})$ *matrice et* $P \in M_n(\mathbb{R})$ *inversible alors*

$$rg(A) = rg(PA) = rg(AP)\,.$$

Démonstration : Si f surjective , alors $Im(g \circ f) = Im(g)$. Si g inj, g est bijective de $Im(f)$ dns $g(Imf) = Im(g \circ f)$. Donc $\dim(Im(g \circ f)) = rg(f)$ □

Corollaire 3.33. *Le rang d'une matrice est le nombre de pivots dans sa forme réduite échelonnée.*

Comme appliquer l'algorithme de Gauss-Jordan qui mène à la forme réduite échelonnée revient à multiplier à gauche par une matrice inversible une matrice a même rang que sa matrice réduite échelonnée. En faisant des opérations sur les colonnes, ce qui revient à multiplier à droite par une matrice inversible, on obtient une matrice de la forme (quitte à permuter des colonnes ce qui revient aussi à multiplier à droite par une matrice inversible)

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

qui a même rang que la matrice de départ. Les vecteurs colonne d'une telle matrice sont des éléments distincts de la base canonique ou bien 0. Son rang est donc bien égal au nombre de 1 contenu dans la forme réduite échelonnée. On remarque sur la forme donnée ci-dessus que la dimension de l'espace engendré par les lignes coïncide avec la dimension de l'espace engendré par les colonnes (le nombre de 1 dans les deux cas). On en déduit la proposition suivante :

Corollaire 3.34. *Le rang d'une matrice est égal au rang de sa transposée.*

Théorème 3.35. *Soit $A \in M_{n\times n}(\mathbb{R})$ une matrice carrée. Les assertions suivantes sont équivalentes :*

1. *A est inversible,*
2. *$rgA = n$,*
3. *$AX = 0 \implies X = 0$,*
4. *$X \neq 0 \implies AX \neq 0$,*
5. *$\forall Y\ \exists X\ AX = Y$,*
6. *$\forall Y\ \exists! X\ AX = Y$,*
7. *les vecteurs colonne de A forment une famille libre dans $\mathbb{R}^n$,*
8. *les vecteurs colonne de A forment une famille génératrice dans $\mathbb{R}^n$,*
9. *les vecteurs colonne de A forment une base de $\mathbb{R}^n$,*
10. *les vecteurs ligne de A forment une famille libre dans $\mathbb{R}^n$,*
11. *les vecteurs ligne de A forment une famille génératrice dans $\mathbb{R}^n$,*
12. *les vecteurs ligne de A forment une base de $\mathbb{R}^n$.*

Théorème 3.36 (Théorème du rang)**.** *Soit $f : E \to F$ linéaire avec E de dimension finie. Alors*

$$\dim Ker f + \dim Im(f) = \dim E$$

Démonstration : Le noyau $Kerf$ est un sous-espace vectoriel de E. C'est donc un espace vectoriel de dimension finie, notons k cette dimension. Soit $(u_1, \ldots, u_k)$ une base de $Ker(f)$. C'est une famille libre dans E. On peut la compléter en une base $(u_1, \ldots, u_n)$ de E (où on a noté n la dimension de E ; $k \leq n$). Appelons G le sous-espace vectoriel engendré par $(u_{k+1}, \ldots, u_n)$: $G = Vect(u_{k+1}, \ldots, u_n)$. On a

$$G \cap Ker(f) = \{0\}.$$

En effet, soit $v \in G \cap Ker(f)$ alors il existe $\lambda_1, \ldots, \lambda_k$ et $\mu_1, \ldots, \mu_{n-k}$ tels que

$$v = \sum_{i=1}^{k} \lambda_i u_i = \sum_{j=1}^{n-k} \mu_j u_{k+j}.$$

En particulier on a

$$\sum_{i=1}^{k} \lambda_i u_i + \sum_{j=1}^{n-k} (-\mu_j) u_{k+j} = 0.$$

Mais la famille $(u_1, \ldots, u_n)$ est une base. Elle est donc libre et l'égalité précédente entraîne que tous les λ_i et tous les μ_j sont nuls. Le vecteur v est nécessairement nul. Considérons maintenant la restriction de f à G :

$$f_g\ :\ G \to F,\ v \mapsto f(v).$$

Le noyau de f_G est $G \cap Ker(f)$ (en effet $v \in G$ est dans le noyau de f_G si $f_G(v) = 0$ mais $F_G(v) = f(v)$ donc il faut et il suffit que v soit dans $Kerf$). Comme $G \cap Ker(f) = \{0\}$, cela signifie que f_G est injective. Par ailleurs l'image de f_G est égale à celle de f. En effet on a évidemment $Im(f_G) \subset Im(f)$ car $G \subset E$. Et par ailleurs prenons $v \in E$: on peut écrire $v = \sum_{i=1}^n \lambda_i u_i$ et, comme $f(u_i) = 0$ si $i \leq k$, on a

$$f(v) = f(\sum_{i=k+1}^{n} \lambda_i u_i).$$

Cela signifie que $f(v)$ a la même image qu'un élément de G. On a donc aussi $Im(f) \subset Im(f_G)$ et $Im(f_G) = Im(f)$. On déduit de ce qui précède que f_G est un isomorphisme de G sur $Im(f_G) = Im(f)$. Cela implique en particulier que G et $Im(f)$ ont même dimension : $n - k$. Autrement dit on a montré que $dim(Im(f)) = n - k = dim(E) - dim(Ker(f))$ soit encore

$$dim(Ker(f)) + dim(Im(f)) = dim(E).$$

□

3.7 Intersection, produit, somme de sous-espaces vectoriels

Soient E_1 et E_2 deux $\mathbb{R}$-ev. Leur produit est défini ainsi

$$E_1 \times E_2 = \{(u_1, u_2), u_1 \in E_1, u_2 \in E_2\}.$$

Exemple $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$, $\mathbb{R}^2 \times \mathbb{R} = \mathbb{R}^3$, ...

$E_1 \times E_2$ est un ev. Somme $(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$. Produit $\lambda(u_1, u_2) = (\lambda u_1, \lambda u_2)$.

On dispose de deux projections $p_1 : E_1 \times E_2 \to E_1$ tq $p_1(u_1, u_2) = u_1$ et $p_2 : E_1 \times E_2 \to E_2$ tq $p_2(u_1, u_2) = u_2$

Soient F_1 et F_2 deux sev d'un même ev E. Leur somme

$$F_1 + F_2 = \{u_1 + u_2, u_1 \in F_1, u_2 \in F_2\}$$

est un sev de E qui contient F_1 et F_2. Dessin dans $\mathbb{R}^3$, somme de deux droites. (droite d'équations $y = x$ et $z = 0$ et droite d'équations $z = x, x = y$.

Soient F_1, F_2 deux sev d'un même ev E. Leur intersection est l'ensemble

$$E_1 \cap E_2 = \{u \in E, u \in E_1 \text{ et } u \in E_2\}$$

C'est un sev de E.

Attention $E_1 \cup E_2$ n'est pas un sev en général (dessin)

Soient E_1, E_2 deux sev de E. On peut définir leur somme et leur intersection, mais ce sont donc aussi des ev, on peut donc définir leur produit. Soit $\Phi : E_1 \times E_2 \to E_1 + E_2$ définie par

$$(u_1, u_2) \to u_1 + u_2$$

C'est une application linéaire qui vérifie $Im(\Phi) = E_1 + E_2$ et $Ker(\Phi) = \{(u, -u), u \in E_1 \cap E_2\}$.

Proposition 3.37. *Soient E_1, E_2 deux sev de E. Alors les prop suiv sont equiv*

1. *$E = E_1 + E_2$ et $E_1 \cap E_2 = \{0\}$ (on note $E = E_1 \oplus E_2$)*
2. *tout vecteur $u \in E$ s'écrit de manière unique $u = u_1 + u_2$ avec $u_i \in E_i$*
3. *$\Phi : E_1 \times E_2 \to E$ est un isom*

Démonstration : $2 \to 1$ suffit de l'écrire.

$1 \to 3$ ecrire $Ker\Phi$ et $Im(\Phi)$

$3 \to 2$ $u = u_1 + u_2$ avec $(u_1, u_2) = \Phi^{-1}(u)$.

□

Exemple $\mathbb{R}^5 = \mathbb{R}^2 \times \mathbb{R}^3 = \mathbb{R}^2 \times \{0\} \oplus \{0\} \times \mathbb{R}^3$

Exemple

$$M_n(\mathbb{R}) = S_n(\mathbb{R}) \oplus A_n(\mathbb{R})$$

Ecrire dans $M_3(\mathbb{R})$ $S_3 = \{M, m_{ji} = m_{ij}\}$ exemple $A_3 = \{M, m_{ji} = -m_{ij}\}$ exemple $M = (M + {}^tM) + (M - {}^tM)$

Proposition 3.38. *Si $E_1 \oplus E_2 = E$ alors $\dim E_1 + \dim E_2 = \dim E$.*
$\dim E_1 \times E_2 = \dim E_1 + \dim E_2$.

Proposition 3.39. *Soit E ev, et $H \subset E$ sev. Alors H admet un supplémentaire.*

Proposition 3.40. *On a toujours*

$$\dim(E_1 + E_2) = \dim E_1 + \dim E_2 - \dim(E_1 \cap E_2)$$

En particulier $E_1 \oplus E_2 = E$ ssi $E_1 \cap E_2 = \{0\}$ et $\dim E_1 + \dim E_2 = \dim E$.

Démonstration : Regarder l'application $E_1 \times E_2 \to E_1 + E_2$ définie par $\Psi(u_1, u_2) = u_1 + u_2$. alors par le thm du rang, $\dim E_1 + E_2 + \dim E_1 \cap E_2 = \dim E_1 + \dim E_2$. □

Définition 3.41. *Soit E un espace vectoriel de dimension finie n. Un sous-espace vectoriel de E de dimension $n-1$ est appelé hyperplan (vectoriel) de E.*

Si $\varphi : E \to \mathbb{R}$ forme linéaire non nulle. L'image de φ est $\mathbb{R}$. Le théorème du rang appliqué à φ donne

$$dim(ker(\varphi)) + dim(Im(\varphi)) = dim(E) = n, \text{ soit } dim(ker(\varphi)) + 1 = n,$$

$$\text{ou encore } dim(ker(\varphi)) = n - 1.$$

Le noyau d'une forme linéaire non nulle est un hyperplan.

Remarque : Toute droite vectorielle non incluse dans un hyperplan H est un supplémentaire de H.

Prenons φ et ψ deux formes linéaires linéairement indépendantes (en particulier aucune des deux n'est nulle). Quelle est la dimension de l'intersection de leur noyau ? Notons E_1 le noyau de φ, E_2 le noyau de ψ. On a

$$dim(E_1 + E_2) + \dim(E_1 \cap E_2) = dim(E_1) + dim(E_2).$$

Le sous-espace vectoriel $E_1 + E_2$ contient E_1 et E_2 et est inclus dans E. Sa dimension ne peut être que $n-1$ ou n. Si c'était $n-1$ alors on aurait $E_1 + E_2 = E_1 = E_2$. Dans ce cas considérons un vecteur $u \notin E_1$. Alors pour tout v dans E on peut trouver $w \in E_1$ et $\lambda \in \mathbb{R}$ tels que

$$v = w + \lambda.u.$$

On a alors

$$\varphi(v) = \varphi(w + \lambda.u) = \varphi(w) + \lambda\varphi(u) = \lambda\varphi(u)$$

$$\psi(v) = \psi(w + \lambda.u) = \psi(w) + \lambda\psi(u) = \lambda\psi(u)$$

Mais cela signifie que pour tout v on a

$$\varphi(v) = (\varphi(u)/\psi(u))\psi(v),$$

u étant un vecteur fixé, cela signifie que φ et ψ sont proportionnelles, ce qui est contraire à l'hypothèse. Cela signifie que E_1 et E_2 ne sont pas égaux et donc qu'on a $E_1 + E_2 = E$. L'égalité sur les dimensions ci-dessus est donc

$$n + \dim(E_1 \cap E_2) = (n-1) + (n-1),$$

soit

$$\dim(E_1 \cap E_2) = n - 2.$$

Proposition 3.42. *Soient φ_1, φ_2,...φ_k k formes linéaires indépendantes sur un espace E, H_1, H_2,..., H_k leurs noyaux et ψ une autre forme linéaire. Si le noyau de ψ contient l'intersection des H_i alors ψ est une combinaison linéaire de φ_1, φ_2,...φ_k.*

Explication pour $E = \mathbb{R}^n$. Les formes linéaires φ_i et ψ sont de la forme suivante :

$$\varphi_i \ : \ (x_1, \ldots, x_n) \mapsto a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n,$$

$$\psi \ : \ (x_1, \ldots, x_n) \mapsto c_1x_1 + c_2x_2 + \ldots + c_nx_n.$$

Dire que $x = (x_1, \ldots, x_n)$ appartient au noyau de φ_i c'est dire que $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0$. Dire que $x = (x_1, \ldots, x_n)$ appartient à l'intersection des noyaux des φ_i c'est dire que $(x_1, \ldots, x_n)$ est solution du système linéaire :

$$\left\{\begin{array}{cccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & 0 \\ a_{21}x_1 & +\cdots+ & a_{2n}x_n & = & 0 \\ \vdots & \ldots & \vdots & = & 0 \\ a_{k1}x_1 & +\cdots+ & a_{kn}x_n & = & 0 \end{array}\right.$$

Cela peut se dire sous une autre forme : le vecteur colonne dont les coordonnées sont les x_i appartient au noyau de l'application

$$f_A \ : \ X \mapsto AX,$$

où A est la matrice $k \times n$ suivante :

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ldots & \vdots \\ a_{k1} & \ldots & a_{kn} \end{pmatrix}.$$

Dire que l'intersection des noyaux des φ_i est contenue dans celui de ψ c'est dire que si le système précédent est satisfait alors

$$c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0.$$

Cela signifie que les deux systèmes

$$\left\{\begin{array}{ccccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & 0 \\ a_{21}x_1 & +\cdots+ & a_{2n}x_n & = & 0 \\ \vdots & \ldots & \vdots & = & 0 \\ a_{k1}x_1 & +\cdots+ & a_{kn}x_n & = & 0 \end{array}\right. \quad \text{et} \quad \left\{\begin{array}{ccccc} a_{11}x_1 & +\cdots+ & a_{1n}x_n & = & 0 \\ a_{21}x_1 & +\cdots+ & a_{2n}x_n & = & 0 \\ \vdots & \ldots & \vdots & = & 0 \\ a_{k1}x_1 & +\cdots+ & a_{kn}x_n & = & 0 \\ c_1x_1 & +\cdots+ & c_nx_n & = & 0 \end{array}\right.$$

ont les mêmes ensembles de solutions. Cela signifie que les formes réduites échelonnées des deux matrices

$$\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ldots & \vdots \\ a_{k1} & \ldots & a_{kn} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ldots & \vdots \\ a_{k1} & \ldots & a_{kn} \\ c_1 & \ldots & c_n \end{pmatrix}$$

ont le même nombre de pivots. En particulier la dernière ligne de la forme réduite échelonnée de la deuxième matrice est nulle. Cela veut dire que le vecteur ligne $(c_1, \ldots, c_n)$ est combinaison linéaire des lignes de A ou encore que ψ est combinaison linéaire des φ_i.

Cette proposition permet de voir d'une autre façon le théorème du rang.

Remarquons d'abord la chose suivante : si F est un sous-espace vectoriel d'un espace E de dimension finie et si H est un hyperplan de E alors ou bien $F \subset H$, ou bien $dim(F \cap H) = dim(F) - 1$. Pourquoi ? On a

$$dim(F + H) + dim(F \cap H) = dim(F) + dim(H).$$

Si $F \not\subset H$, $F + H = E$ et la formule précédente devient :

$$dim(E) + dim(F \cap H) = dim(F) + dim(E) - 1,$$

ce qui donne bien

$$dim(F \cap H) = dim(F) - 1.$$

On peut donc dire la chose suivante quand on prend l'intersection d'un sev avec un hyperplan vectoriel on obtient soit le même sev (s'il est inclus dans l'hyperplan) soit un sev d'une dimension inférieure d'une unité.

Revenons au théorème du rang. On se donne une matrice

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ldots & \vdots \\ a_{k1} & \ldots & a_{kn} \end{pmatrix}.$$

Un vecteur colonne X est dans son noyau s'il est dans l'intersection des noyaux des formes linéaires définies par les lignes de la matrices. Notons r le rang de A. On peut trouver r lignes de A formant une famille libre et $k - r$ lignes combinaisons linéaires des r autres. Le noyau de A est alors l'intersection des noyaux des r lignes formant une famille libre. Sa dimension est $n - r$: la dimension de l'intersection de r hyperplans indépendants ; à chaque intersection la dimension diminue de 1, on part de n et on prend l'intersection r fois.

4 Repérage dans $\mathbb{R}^d$, matrices et transformations géométriques

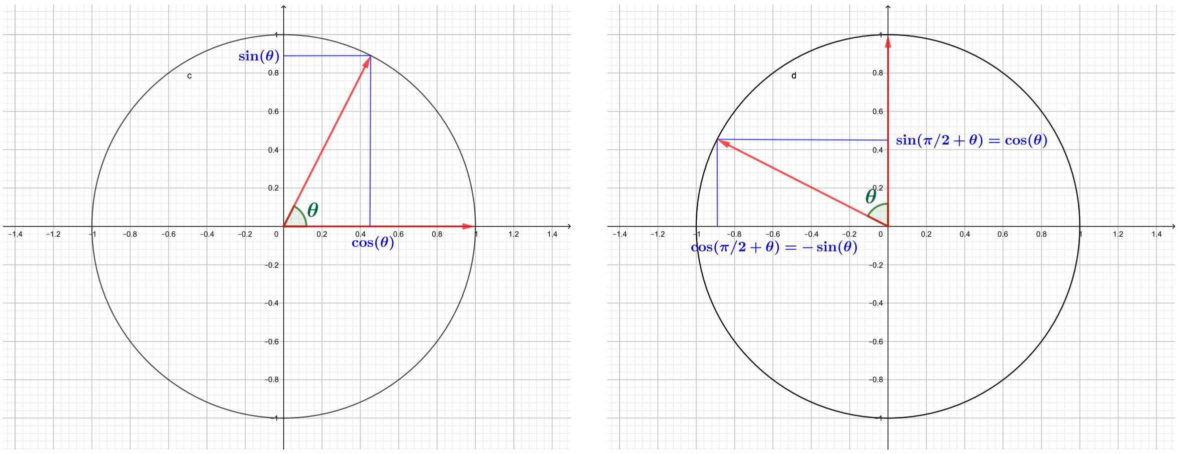
4.1 Les homothéties

Ce sont les applications linéaires qui dilatent / contractent tout. L'homothétie de rapport α est l'application $h_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ définie par $h_\alpha : \vec{x} \to \alpha\vec{x}$

On peut définir des homothéties de $\mathbb{C}^n$ dans $\mathbb{C}^n$ aussi bien sûr. Ce sont les applications les plus simples. Elles préservent les droites, elles ne font que changer l'échelle des figures.

4.2 Les rotations du plan et de l'espace

La rotation $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ qui fait tourner tous les vecteurs d'un angle θ est linéaire. Pour trouver sa matrice dans la base canonique on cherche les coordonnées des images $(1, 0)$ et de $(0, 1)$. Les deux dessins ci-dessous montrent ce qu'on obtient.



La rotation R_θ est l'application linéaire associée à la matrice $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ (première colonne : coordonnées de l'image de $(1, 0)$; deuxième colonne : coordonnées de l'image de $(0, 1)$).

La rotation d'angle 2π est l'identité.

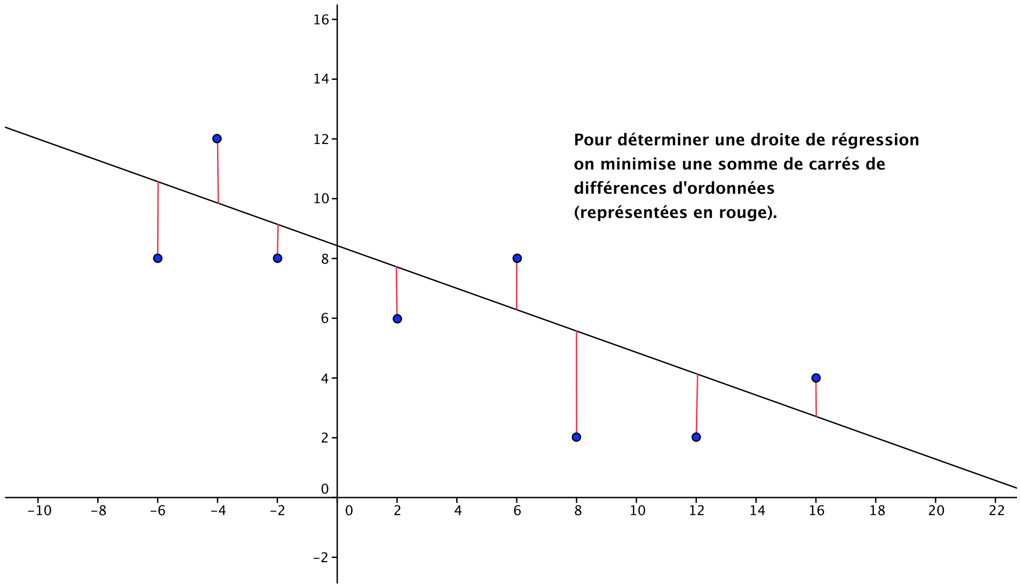
La rotation d'angle π est l'homothétie de rapport -1, encore notée $-Id$.

Une rotation de $\mathbb{R}^3$ est une application linéaire qui, dans la direction d'un certain vecteur $\vec{u}$ (l'axe de la rotation), ne change rien, et dans le plan orthogonal à la droite $\mathbb{R}\vec{u}$, est une rotation plane.

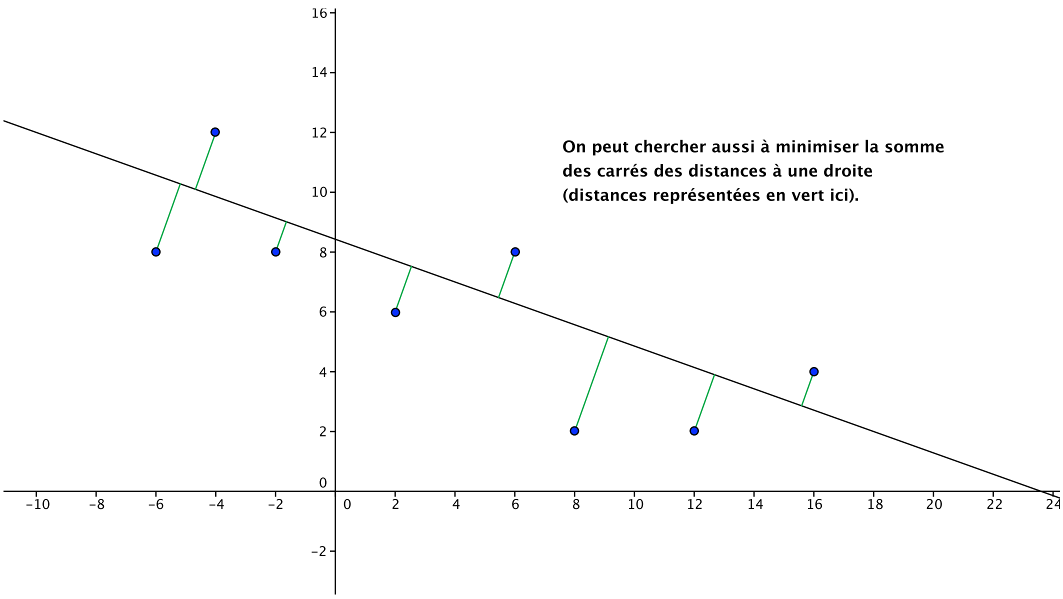
4.3 Les projections

Une projection est une application $p : \mathbb{R}^n \to \mathbb{R}^n$ telle que $p \circ p = p$. Par exemple, dans $\mathbb{R}^2$, la projection orthogonale sur l'axe horizontal, définie par $p(x, y) = (x, 0)$ est une projection au sens ci-dessus. Dans $\mathbb{R}^3$ la projection orthogonale sur le plan d'équation $x = 0$, définie comme $(x, y, z) \to (0, y, z)$, est également une projection au sens ci-dessus.

Les projections sont utilisées en statistique descriptive.



C'est une projection dans la direction de l'axe des ordonnées sur une droite qu'on cherche.



C'est une projection orthogonale sur une droite qu'on cherche.

Attention : ce sont des projections affines (sur des droites qui ne passent généralement pas par 0) mais de telles projection ont une partie linéaire qui est une projection vectorielle.

Question : quelle est la matrice de la projection orthogonale de $\mathbb{R}^2$ sur la droite d'équation $x = y$? Le vecteur $(1, 0)$ est envoyé sur $(1/2, 1/2)$ et le vecteur $(0, 1)$ aussi. Réponse :

$$\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

On pourra vérifier que cette matrice est égale à son carré.

Pour voir qu'une matrice est la matrice d'une projection il suffit de voir qu'elle est égale à son carré. Mais cela ne donne pas des renseignements qu'on aimerait parfois avoir : projection sur quel sous-espace parallèlement à quel sous espace ? Considérons par exemple cette matrice

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

On peut vérifier qu'elle est égale à son carré. C'est bien la matrice d'une projection. Mais comment voir que c'est la projection sur l'hyperplan d'équation $x+y+z=0$ parallèlement à $Vect((1,1,1))$? C'est assez simple. Si p est une projection, c'est une projection sur $Im(p)$ parallèlement à $ker(p)$. Pour décrire géométriquement p, il suffit de déterminer son image et son noyau. Si une projection est donnée par une matrice $n \times n$, il faut déterminer son noyau (donc résoudre un système linéaire) et déterminer le sous-espace vectoriel engendré par ses colonnes.

Soit p une projection définie dans un espace vectoriel E. On a

$$E = Im(p) \oplus ker(p).$$

Soit v un élément de E. On peut écrire

$$v = p(v) + v - p(v).$$

Par définition $p(v)$ appartient à $Im(p)$. D'autre part on a $p(v - p(v)) = p(v) - p(p(v)) = p(v) - p^2(v) = p(v) - p(v) = 0$ (car $p^2(v) = p(v)$ puisque p est une projection), donc $v - p(v)$ appartient au noyau de p. On a donc bien

$$E = Im(p) + ker(p).$$

Reste à voir que la somme est directe c'est-à-dire que $Im(p) \cap ker(p) = \{0\}$. Prenons $v \in Im(p) \cap ker(p)$. Comme v est dans l'image de p, il existe u tel que $v = p(u)$, comme v est dans le noyau de p on a $p(v) = 0$. On a donc $p(p(u)) = 0$, mais $p(p(u)) = p(u)$ (puisque p est un projecteur. On a donc $v = 0$. Autrement dit tout élément de $Im(p) \cap ker(p)$ est nul.

4.4 Les symétries

Une <u>symétrie</u> est une application linéaire $s : \mathbb{R}^n \to \mathbb{R}^n$ telle que $s \circ s = Id$. En particulier, s est bijective, d'inverse $s^{-1} = s$.

L'identité est une symétrie.

L'application $-Id$, encore appelée symétrie centrale en géométrie, est une symétrie.

Dans le plan $\mathbb{R}^2$, la réflexion orthogonale d'axe horizontal, définie par $(x, y) \to (x, -y)$, est une symétrie.

La réflexion d'axe d'équation $y = x$ est aussi une symétrie ; C'est l'application $(x, y) \to (y, x)$ qui échange les coordonnées.

Proposition 4.1. *Soit $s : E \to E$ une symétrie $s \circ s = id$. Alors $E = Ker(s - id) \oplus Ker(s + id)$*

Soit v un élément de E. On peut écrire

$$v = (v - s(v))/2 + (v + s(v))/2.$$

Or

$$(s + id)(v - s(v)) = s(v) - s^2(v) + v - s(v) = s(v) - v + v - s(v) = 0,$$

car $s^2(v) = v$. De même

$$(s - id)(v + s(v)) = s(v) + s^2(v) - v - s(v) = s(v) + v - v - s(v) = 0.$$

Les deux vecteurs $(v - s(v))/2$ et $(v + s(v))/2$ appartiennent à $Ker(s + id)$ et $Ker(s - id)$ respectivement et leur somme est v. On a donc bien

$$E = Ker(s - id) + Ker(s + id).$$

Reste à voir que la somme est directe c'est-à-dire que $Ker(s - id) \cap Ker(s + id) = \{0\}$. Prenons $v \in Ker(s - id) \cap Ker(s + id)$. Comme v est dans le noyau de $s - id$ on a $v = s(v)$, comme il es dans le noyau de $s + id$, on a aussi $v = -s(v)$. On en déduit $v = -v$ donc $v = 0$. Autrement dit tout élément de $Ker(s - id) \cap Ker(s + id)$ est nul.

Pour montrer qu'une matrice A de taille $n \times n$ est la matrice d'une symétrie il faut montrer

$$\mathbb{R}^n = Ker(A - Id) \oplus Ker(A + Id).$$

Si on sait que A est une matrice de symétrie on peut trouver $Ker(A - Id)$ et $Ker(A + Id)$ en résolvant deux systèmes linéaires. On décrit alors l'action de A comme la symétrie par rapport à $Ker(A - Id)$ parallèlement à $Ker(A + Id)$.

4.5 Bases et coordonnées

Jusqu'à maintenant j'ai identifié un vecteur de $\mathbb{R}^n$ et ses coordonnées dans la base canonique (le plus souvent placées en colonnes). Prenons un vecteur X de $\mathbb{R}^n$ et considérons une autre base de $\mathbb{R}^n$: $(b_1, \ldots, b_n)$. On peut écrire X dans la base $(b_1, \ldots, b_n)$:

$$X = \alpha_1 b_1 + \ldots + \alpha_n b_n.$$

On a alors deux représentations de X : X lui-même, la matrice colonne de ses coordonnées dans la base canonique et la matrice de ses coordonnée dans la base $(b_1, \ldots, b_n)$. Il ne faut pas les confondre.

Dans un espace vectoriel E muni d'une base $\mathcal{B} = (b_1, \dots, b_n)$. Si $u = u_1\vec{b}_1 + \dots u_n\vec{b}_n$ on écrit

$$Mat_{\mathcal{B}}(u) = \begin{pmatrix} u_1 \\ \dots \\ u_n \end{pmatrix}$$

Par exemple considérons le polynôme $P = X^3 + X^2 - 2X + 3$ comme élément de l'espace vectoriel $\mathbb{R}_3[X]$. Ses coordonnées dans la base $1, X, X^2, X^3$ sont (en ligne) $(3, -2, 1, 1)$. On peut montrer que $1, X+1, (X+1)^2, (X+1)^3$ est aussi une base de $\mathbb{R}_3[X]$. Quels sont les coordonnées de P dans cette base ? On peut utiliser ce qu'on sait sur les polynômes et procéder de la façon suivante. On a

$$(X+1)^3 = X^3 + 3X^2 + 3X + 1$$

On voit que $P - (X+1)^3$ est de degré 2 :

$$P - (X+1)^3 = X^3 + X^2 - 2X + 3 - X^3 - 3X^2 - 3X - 1 = -2X^2 - 5X + 2$$

Pour éliminer le terme de degré 2 il nous faut ajouter $2(X+1)^2$:

$$P - (X+1)^3 + 2(X+1)^2 = -2X^2 - 5X + 2 + 2(X^2 + 2X + 1) = -X + 4.$$

On en déduit alors l'égalité

$$P - (X+1)^3 + 2(X+1)^2 + (X+1) = 5,$$

soit

$$P = (X+1)^3 - 2(X+1)^2 - (X+1) + 5.$$

Les coordonnées de P dans la base $1, X+1, (X+1)^2, (X+1)^3$ sont (en ligne) $(5, -1, -2, 1)$. Dans cet exemple on est moins tenté de confondre le polynôme avec l'un ou l'autre vecteur de ses coordonnées que dans le cas de $\mathbb{R}^4$. Probablement grâce à la notation et au fait qu'on se représente souvent plutôt un polynôme comme une fonction.

Quoi qu'il en soit il faut souvent bien distinguer un vecteur des vecteurs de ses coordonnées dans différentes bases.

Reprenons l'exemple précédent. On peut présenter le problème de manière légèrement différente : trouver α_1, α_2, α_3, α_4, tels que

$$P = \alpha_1.1 + \alpha_2.(X+1) + \alpha_3.(X+1)^2 + \alpha_4.(X+1)^3.$$

Ecrivons (en colonne) les coordonnées de chaque vecteur de la base $1, X+1, (X+1)^2, (X+1)^3$ dans la base $1, X, X^2, X^3$:

$$1 \to \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad X+1 \to \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad (X+1)^2 \to \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} \quad (X+1)^3 \to \begin{pmatrix} 1 \\ 3 \\ 3 \\ 1 \end{pmatrix}$$

En écrivant les coordonnées des polynômes apparaissant dans l'équation précédente dans la base $1, X, X^2, X^3$, on voit que l'équation est équivalente au système

$$\begin{cases} 1.\alpha_1 & +1.\alpha_2 & +1.\alpha_3+ & +1.\alpha_4 & = 3 \\ 0.\alpha_1 & +1.\alpha_2 & +2.\alpha_3+ & +3.\alpha_4 & = -2 \\ 0.\alpha_1 & +0.\alpha_2 & +1.\alpha_3+ & +3.\alpha_4 & = 1 \\ 0.\alpha_1 & +0.\alpha_2 & +0.\alpha_3+ & +1.\alpha_4 & = 1 \end{cases}$$

On peut résoudre ce système par la méthode de Gauss-Jordan

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 3 \\ 0 & 1 & 2 & 3 & -2 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

$$\left(\begin{array}{cccc|c} 1 & 0 & -1 & -2 & 5 \\ 0 & 1 & 2 & 3 & -2 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

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

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

Conclusion : les coordonnées de P dans la base $1, X+1, (X+1)^2, (X+1)^3$ sont (en ligne) $(1, -1, -2, 1)$. On a bien trouvé la même chose. Pour trouver le résultat, on a utilisé la matrice

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

obtenue en rassemblant les colonnes des coordonnées des éléments de la nouvelle base dans l'ancienne base.

Exemple dans $\mathbb{R}^4$ on regarde E l'hyperplan d'équation $x+y+z+t = 0$. Alors une base est $(1, -1, 0, 0)$, $(1, 0, -1, 0)$, $(1, 0, 0, -1)$. Dans cette base la matrice du vecteur $(5, 3, -7, -1)$ est $\begin{pmatrix} -3 \\ 7 \\ 1 \end{pmatrix}$.

Comme nous l'avons déjà remarquéi $\mathcal{B} = (\vec{b}_1, \dots \vec{b}_n)$ base de E ev, alors $\Psi_{\mathcal{B}} : \mathbb{R}^n \to E$ définie par

$$(x_1, \dots, x_n) \to x_1\vec{b}_1 + \dots x_n\vec{b}_n$$

est un isomorphisme d'ev.

Définition 4.2. *Si $f : E \to F$ est une application linéaire, $\mathcal{B} = (\vec{b}_1, \dots \vec{b}_n)$ base de E et $\mathcal{C} = (\vec{c}_1, \dots \vec{c}_n)$ base de F alors on appelle matrice de f par rapport aux bases $\mathcal{B}$ et $\mathcal{C}$, la matrice*

$$Mat_{\mathcal{B},\mathcal{C}}(f) = (Mat_{\mathcal{C}}(f(\vec{b}_1))| \dots |Mat_{\mathcal{C}}(f(\vec{b}_n))),$$

c'est-à-dire la matrice dont les colonnes sont les vecteurs des coordonnées des images par f des éléments de $\mathcal{B}$ dans la base $\mathcal{C}$.

Définition 4.3. *Si $f : E \to E$ est un endomorphisme, $\mathcal{B} = (\vec{b}_1, \dots \vec{b}_n)$ une base de E , on appelle matrice de f dans la base $\mathcal{B}$, la matrice*

$$Mat_{\mathcal{B}}(f) = (Mat_{\mathcal{B}}(f(\vec{b}_1))| \dots |Mat_{\mathcal{B}}(f(\vec{b}_n))),$$

c'est-à-dire la matrice dont les colonnes sont les vecteurs des coordonnées des images par f des éléments de $\mathcal{B}$ dans la base $\mathcal{B}$.

Prenons deux exemples p une projection et s une symétrie dans $\mathbb{R}^n$. On sait que

$$\mathbb{R}^n = ker(p) \oplus Im(p).$$

Prenons une $\mathcal{B}$ base de $\mathbb{R}^n$ constituée par la réunion d'une base de $Im(p)$ et d'une base de $ker(p)$. Alors la matrice de p dans la base $\mathcal{B}$ est de la forme

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0\dots & 0 \\ 0 & 1 & \dots & 0 & 0\dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0\dots & 0 \\ 0 & 0 & \dots & 0 & 0\dots & 0 \\ 0 & 0 & \dots & 0 & 0\dots & 0 \end{pmatrix}.$$

On sait que

$$\mathbb{R}^n = ker(s - id) \oplus ker(s + id).$$

Prenons une $\mathcal{C}$ base de $\mathbb{R}^n$ constituée par la réunion d'une base de $ker(s - id)$ et d'une base de $ker(s + id)$. Alors la matrice de s dans la base $\mathcal{C}$ est de la forme

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & -1 \end{pmatrix}.$$

Exercice 3. Soit f l'endomorphisme de $\mathbf{R}^3$ dont la matrice dans la base canonique est $M = \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$. Soient $v_1 = (1,-1,1), v_2 = (1,1,0), v_3 = (-1,1,2)$.

1. Montrer que $\mathcal{B} = (v_1, v_2, v_3)$ est une base de $\mathbf{R}^3$.
2. Déterminer la matrice associée à f dans la base $\mathcal{B}$.
3. En déduire que f est une symétrie par rapport à un plan vectoriel que l'on déterminera.

$$f(v_1) <-> \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$f(v_2) <-> \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$f(v_3) <-> \frac{1}{3}\begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

On a donc $f(v_1) = -v_1$, $f(v_2) = v_2$, $f(v_3) = v_3$. La matrice de f dans la base (v_1, v_2, v_3) est

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

L'application f est la symétrie par rapport au plan $Vect(v_2, v_3)$ dans la direction v_1 (orthogonale).

Proposition 4.4. *Si $f : E \to F$ et $g : F \to G$ sont deux applications linéaires, est une application linéaire, $\mathcal{B}_1$, $\mathcal{B}_1$, $\mathcal{B}_1$ des bases de E, F, G respectivement, alors*

$$Mat_{\mathcal{B}_1,\mathcal{B}_3}(g \circ f) = Mat_{\mathcal{B}_2,\mathcal{B}_3}(g) \times Mat_{\mathcal{B}_1,\mathcal{B}_2}(f)$$

Rappel : $\mathcal{L}(E, F)$ est l'ens des appl linéaires de E dans F. C'est un ev.

Proposition 4.5. *Soient E, F deux $\mathbb{R}$-ev, soit $\mathcal{B} = (b_1, \dots b_m)$ base de E, et $\mathcal{C} = (c_1, \dots c_n)$ base de F. Alors l'application $g \in \mathcal{L}(E, F) \to Mat_{\mathcal{B},\mathcal{C}}(g) \in M_{n\times m}(\mathbb{R})$ est un isom d'ev.*

Démonstration : C'est linéaire. C'est bijectif car si $M \in M_{n\times m}$ on définit $g_M : \vec{x} = x_1 b_1 + \dots x_m b_m = y_1 c_1 + \dots y_n c_n$ avec $Y = MX$.

Autrement dit, $g_M = \psi_{\mathcal{C}} \circ f_M \circ \psi_{\mathcal{B}}^{-1}$ Dessin diagramme $\square$

Remarque : une base de $M_{n\times m}(\mathbb{R})$ est la famille des matrices E_{ij} Donc $\dim M_{n\times m}(\mathbb{R}) = nm$

Corollaire 4.6. $\dim \mathcal{L}(E,F) = \dim E \times \dim F$.

En particulier, $\dim \mathcal{L}(E,\mathbb{R}) = \dim E$.

4.6 Matrices de passage

Exemple : E hyperplan d'équation $x+y+z-t = 0$ dans $\mathbb{R}^4$. Base $(1,-1,0,0),(1,0,-1,0),(1,0,0,1)$ ou encore $(1,-1,0,0),(0,1,-1,0),(0,0,1,1)$

Laquelle choisir ? Y en a-t-il une meilleure ? Plus naturelle ? Comment passer de l'une à l'autre ?

Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E, $\mathcal{C}$ et $\mathcal{C}'$ deux bases de F. On veut comparer $Mat_{\mathcal{B},\mathcal{C}}(f)$ et $Mat_{\mathcal{B}',\mathcal{C}'}(f)$.

Surtout, on n'apprend pas de formules par coeur avec des notations arbitraires.

On commence par faire un diagramme

$$f : (E,\mathcal{B}) \to (F,\mathcal{C})$$

$$f : (E,\mathcal{B}') \to (F,\mathcal{C}')$$

On introduit l'application identité $Id_E : (E,\mathcal{B}) \to (E,\mathcal{B}')$ et $Id_F : (F,\mathcal{C}) \to (F,\mathcal{C}')$.

On écrit ensuite que $f = Id_F \circ f \circ Id_E$. On en déduit donc (d'après les propositions ci-dessus) que

$$Mat_{\mathcal{B}',\mathcal{C}'}(f) = Mat_{\mathcal{C},\mathcal{C}'}(Id_F) \times Mat_{\mathcal{B},\mathcal{C}}(f) \times Mat_{\mathcal{B}',\mathcal{B}}(Id_E)$$

On appelle matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$ la matrice

$$P = Mat_{\mathcal{B},\mathcal{B}'}(Id_E)$$

Proposition 4.7. *Une matrice de passage est inversible. De plus*

$$Mat_{\mathcal{B},\mathcal{B}'}(Id_E) \times Mat_{\mathcal{B}',\mathcal{B}}(Id_E) = I_n.$$

Les colonnes de cette matrice sont les vecteurs de la base $(b_1,\dots b_n)$ décomposés dans la base $\mathcal{B}'$. En pratique, c'est plutôt le contraire qu'on a naturellement. On connait la première base $\mathcal{B}$ et l'expression des vecteurs $(b'_1,\dots b'_n)$ dans la base $\mathcal{B}$. Ainsi la matrice $(b'_1|\dots b'_n)$ est la matrice de passage de $\mathcal{B}'$ vers $\mathcal{B}$.

Observez que si $A' = Mat_{\mathcal{B}',\mathcal{C}'}(f)$, $A = Mat_{\mathcal{B},\mathcal{C}}(f)$, on a alors $rg(A) = rg(A')$.

La relation ci-dessus se résume parfois en

$$A' = QAP^{-1}$$

avec $A' = Mat_{\mathcal{B}',\mathcal{C}'}(f)$, $A = Mat_{\mathcal{B},\mathcal{C}}(f)$, $P = Mat_{\mathcal{B},\mathcal{B}'}(Id_E)$ et $Q = Mat_{\mathcal{C},\mathcal{C}'}(Id_F)$.

Exemple dans $\mathbb{R}^2$, $\mathcal{B} = \{(1,0),(0,1)\}$ et $\mathcal{B}' = \{(1,2),(3,-1)\}$ Alors $Mat_{\mathcal{B}',\mathcal{B}}(Id_{\mathbb{R}^2}) = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}$.

Définition 4.8. *Si $A' = Q^{-1}AP$ avec P, Q inversibles, alors on dit que A et A' sont équivalentes.*

Cas particulier si $E = F$, $\mathcal{B} = \mathcal{C}$ et $\mathcal{B}' = \mathcal{C}'$ Alors

$$Mat_{\mathcal{B}'}(f) = Mat_{\mathcal{B},\mathcal{B}'}(Id_E) \times Mat_{\mathcal{B}}(f) \times Mat_{\mathcal{B}',\mathcal{B}}(Id_E)\,.$$

Avec les notations ci-dessus, on a

$$A' = PAP^{-1}$$

On dit que A et A' sont semblables.

Exemple : $f : \mathbb{R}^2 \to \mathbb{R}^2$ défini par $f(x,y) = (x+2y, 2x+y)$. $\mathcal{B} = \{(1,0),(0,1)\}$ et $\mathcal{B}' = \{(1,1),(1,-1)\}$

$$Mat_{\mathcal{B}',\mathcal{B}}(Id_E) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Et $f(b'_1) = 3b' - 1$, $f(b'_2) = -b - 2')$ donc $A' = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}$

Intérêt ? ? $A'^n = \begin{pmatrix} 3^n & 0 \\ 0 & (-1)^n \end{pmatrix}$ Donc

$$A^n = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3^n & 0 \\ 0 & (-1)^n \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

Exemple : étude de $u_{n+1} = u_n + 2v_n$ et $v_{n+1} = 2u_n + v_n$. Alors

$$\begin{pmatrix} u_{n+1} \\ v_{n+1} \end{pmatrix} = A \begin{pmatrix} u_n \\ v_n \end{pmatrix}$$

Donc $\begin{pmatrix} u_n \\ v_n \end{pmatrix} = A^n \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$. Donc $u_n = \frac{5.3^n + (-1)^n}{2}$ et $v_n = \frac{5.3^n - (-1)^n}{2}$.

4.7 Diagonalisation

Soit $f : E \to E$ un endomorphisme d'un espace vectoriel E de dimension finie..

Définition 4.9. *S'il existe un vecteur* **non nul** $\vec{u} \neq 0$ *et* λ *un réel (ou complexe) tq* $f(\vec{u}) = \lambda\vec{u}$ *alors on dit que* $\vec{u}$ *est un vecteur propre de* f *(*$\vec{vp}$*) et* λ *une valeur propre de* f *(vp).*

Dans la direction de la droite vectorielle $\mathbb{R}\vec{u}$ *(ou* $\mathbb{C}\vec{u}$*) tous les autres vecteurs* $\vec{v} = \alpha\vec{u}$ *vérifient* $f(\vec{v}) = \lambda\vec{v}$*. Autrement dit* $f_{|Vect(\vec{u}}$ *est une homothétie de rapport* λ*.*

Pour $\lambda \in \mathbb{R}$ *ou* $\mathbb{C}$ *on note* $E_\lambda = \{\vec{w} \in E,\ f(\vec{w}) = \lambda\vec{w}\}$ *. Alors* λ *vp ssi* $E_\lambda \neq \{0\}$*. Dans ce cas on dit que* E_λ *est l'esp propre associé à la vp* λ*.*

Géométriquement, si $\vec{u}$ vecteur propre, alors $f(\mathbb{R}\vec{u}) \subset \mathbb{R}\vec{u}$ (= si vp non nulle) ($f(\mathbb{C}u) \subset \mathbb{C}u$ pour un $\mathbb{C}$-ev). Autrement dit, f préserve les droites propres.

Remarque 4.10. E_λ est un sev.
Si $\lambda = 0$ est vp $E_0 = Ker(f)$. En particulier, f bij ssi $E_0 = \{0\}$.

Exemples.

1. $h : \mathbb{R}^3 \to \mathbb{R}^3$ homothétie de rapport 5. 5 est valeur propre, $E_5 = \mathbb{R}^3$
2. S symétrie de $\mathbb{R}^3$ orthogonale par rapport au plan d'équation $z = 0$. Alors 1 est vp, espace propre $E_1 = \{(x, y, 0)\}$ et -1 vp, $E_{-1} = \{(0, 0, z), z \in \mathbb{R}\}$.
3. p la projection sur le plan $x = 0$. vp $0, 1$. Si $u = (0, y, z)$, $p(u) = u$ vect propre pour vp 1, si $u = (x, 0, 0)$ alors $p(u) = \vec{0}$, donc $\vec{u} \in E_0 = Ker(p)$.
4. dans $\mathbb{R}^2$ $f(x, y) = (-y, x)$ rotation, matrice $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ Pas de vp ni $\vec{vp}$
5. $f : \mathbb{C}^2 \to \mathbb{C}^2$ définie par $f(x, y) = (-y, x)$. Vp $\pm i$, $(i, 1)$ est vect propre pour $\lambda = i$ et $(-i, 1)$ est vect propre pour $\lambda = -i$

Proposition 4.11. *Si*λ *vp de* f*, alors* λ^k *vp de* f^k

Proposition 4.12. Soit E ev, $\lambda \neq \mu$ deux vp. Alors $E_\lambda \cap E_\mu = \{0\}$

Démonstration : $\vec{u} \in E_\lambda \cap E_\mu$ alors $f(u) = \lambda u = \mu u$ donc $(\mu - \lambda)u = \vec{0}$, donc $\vec{u} = \vec{0}$. □

Corollaire 4.13. *Si* E *est un* $\mathbb{R}$*-ev de dim* n *alors il y a au plus* n *valeurs propres distinctes.*

Démonstration : Si $\lambda_1, \dots \lambda_k$ sont les vp, $E \supset E_{\lambda_1} + \cdots + E_{\lambda_k} = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$ et chaque E_{λ_i} est de dimension au moins 1. □

Définition 4.14. *Soit E un $\mathbb{R}$-ev ($\mathbb{C}$-ev). $f : E \to E$ lin est dite diagonalisable si il existe une base de E tq la matrice de f dans cette base soit diagonale. C'est équivalent à dire qu'il existe une base $(b_1, \dots b_n)$ dont les vecteurs sont des vecteurs propres de f. Idem dans $\mathbb{C}$. C'est équivalent à dire que $E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \dots E_{\lambda_k}$, pour E_{λ_i} les espaces propres.*

Soit A une matrice de taille $n \times n$. Comment savoir si elle est diagonalisable ? Comment la diagonaliser ? Supposons qu'on ait une base de $\mathbb{R}^n$ $(V_1, \dots, V_n)$ formée de vecteurs propres de A :

$$AV_i = \lambda_i V_i.$$

Construisons la matrice P en rassemblant les vecteurs colonne V_i :

$$P = \left(\; V_1 \; \middle| \; V_2 \; \middle| \; \cdots \; \middle| \; V_n \; \right).$$

La matrice P est inversible puisque ses vecteurs colonne forment une base de $\mathbb{R}^n$. La matrice inverse de P vérifie, pour tout i,

$$P^{-1} V_i = e_i,$$

car $Pe_i = V_i$. Formons le produit $P^{-1}AP$:

$$P^{-1}AP = P^{-1}A \left(\; V_1 \; \middle| \; V_2 \; \middle| \; \cdots \; \middle| \; V_n \; \right) = P^{-1} \left(\; AV_1 \; \middle| \; AV_2 \; \middle| \; \cdots \; \middle| \; AV_n \; \right).$$

Comme les V_i sont des vecteurs propres de A, on obtient

$$P^{-1}AP = P^{-1} \left(\; \lambda_1 V_1 \; \middle| \; \lambda_2 V_2 \; \middle| \; \cdots \; \middle| \; \lambda_n V_n \; \right),$$

puis

$$P^{-1}AP = \left(\; \lambda_1 P^{-1} V_1 \; \middle| \; \lambda_2 P^{-1} V_2 \; \middle| \; \cdots \; \middle| \; \lambda_n P^{-1} V_n \; \right) = \left(\; \lambda_1 e_1 \; \middle| \; \lambda_2 e_2 \; \middle| \; \cdots \; \middle| \; \lambda_n e_n \; \right).$$

Autrement dit

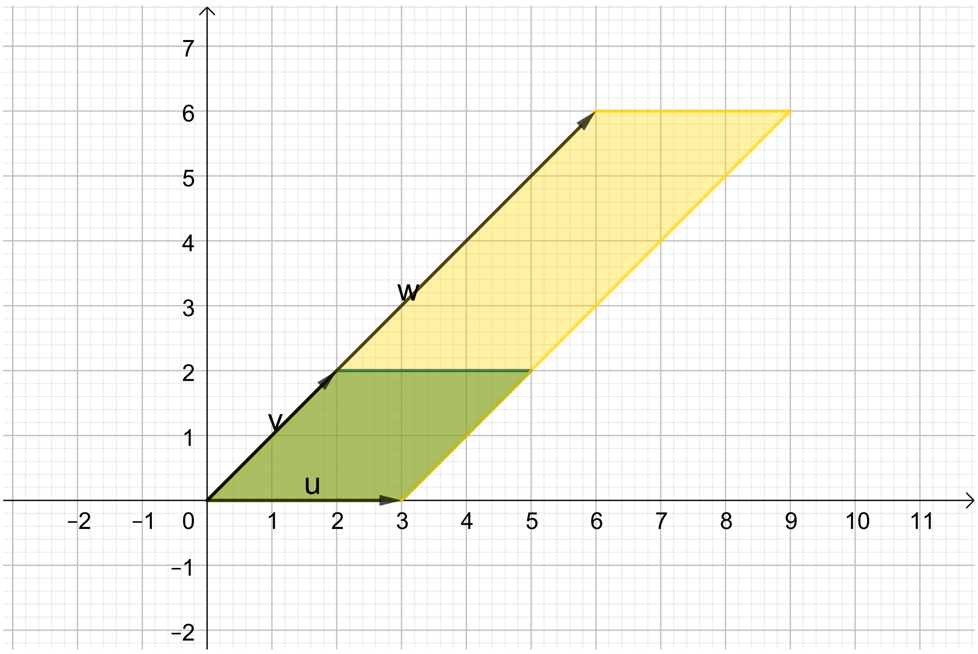
$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

5 Déterminants

Il faut avoir en tête qu'un déterminant est une aire en dimension 2, un volume en dimension 3 et ce qu'on appelle encore un volume en dimension supérieur, mais une aire ou un volume avec un signe.

5.1 Aires et déterminant dans $\mathbb{R}^2$

Dans le plan l'aire d'un parallélogramme est égal au produit base fois hauteur. Un parallélogramme est défini par deux vecteurs issus d'un même sommet et portés par deux côtés. Par exemple, sur la figure ci-dessous le parallélogramme vert est défini par les deux vecteurs u et v, le parallélogramme jaune est défini par les deux vecteurs u et w (le jaune contient le vert).



Comme w est le triple de v, le parallélogramme vert a une hauteur trois fois supérieure au parallélogramme vert et, comme les deux ont même base, l'aire du parallélogramme jaune est trois fois celle du parallélogramme vert.

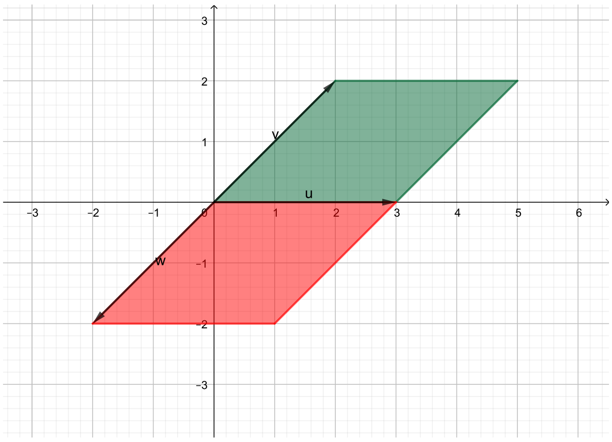
Désignons par $\mathcal{A}(u, v)$ l'aire d'un parallélogramme construit deux vecteurs u et v. Ce qu'on vient de remarquer sur l'exemple indique qu'on a

$$\mathcal{A}(u, 3v) = 3\mathcal{A}(u, v).$$

Plus généralement, pour tout nombre réel positif λ, on a

$$\mathcal{A}(u, \lambda v) = \lambda\mathcal{A}(u, v).$$

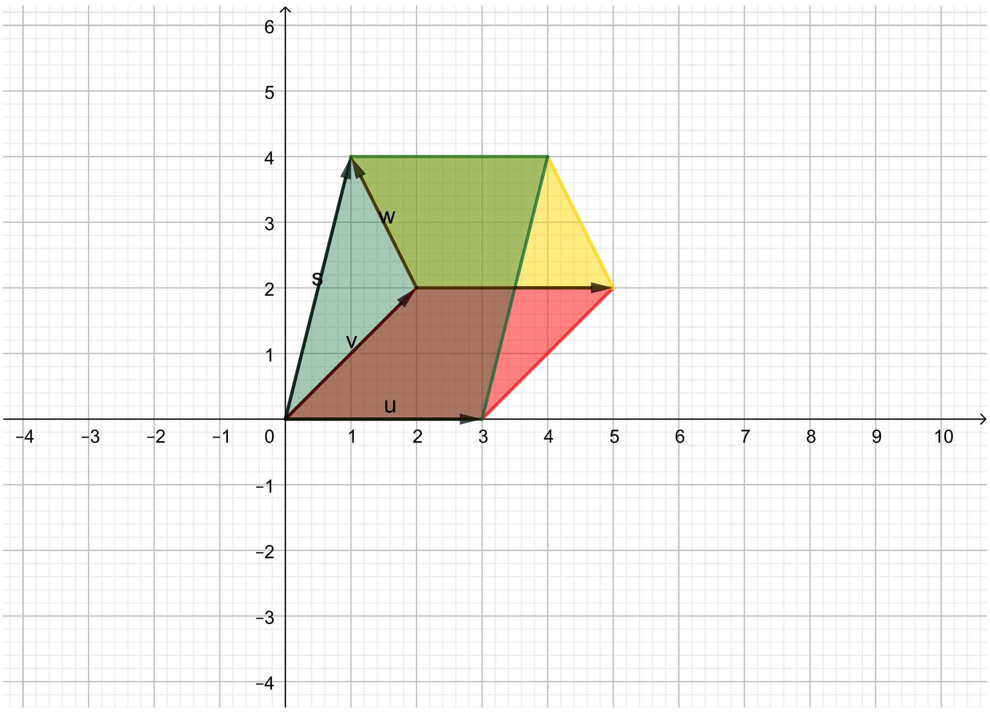
Si λ est négatif, il faut prendre la valeur absolue car il n'y a pas d'aire négative. La figure ci-dessous montre le cas $w = -v$.



Pour la base on peut faire les mêmes remarques. On a donc, pour tout nombre réel λ

$$\mathcal{A}(u, \lambda v) = |\lambda| \mathcal{A}(u, v),\ \mathcal{A}(\lambda u, v) = |\lambda| \mathcal{A}(u, v).$$

La figure suivante montre ce qui se passe lorsque l'un des vecteurs s'écrit comme une somme (ici $s = v + w$).



Sur le dessin précédent la hauteur du parallélogramme vert est la somme des parallélogrammes jaune et rouge (et les trois ont même base). On en déduit

$$\mathcal{A}(u, v+w) = \mathcal{A}(u, v) + \mathcal{A}(u, w).$$

Cette formule marche bien dans le cas dessiné mais pas en général. Que se passe-t-il par exemple si $w = -v$? On n'a pas

$$\mathcal{A}(u, v-v) = \mathcal{A}(u, v) + \mathcal{A}(u, -v) = 2\mathcal{A}(u, v),$$

car $\mathcal{A}(u, v-v) = \mathcal{A}(u, 0) = 0$ (c'est l'aire d'un triangle aplati). La formule marcherait si on écrivait $\mathcal{A}(u, -v) = -\mathcal{A}(u, v)$, mais ce ne serait plus une aire !

Prenons un autre exemple. L'aire d'un parallélogramme défini par deux vecteurs colinéaires est nulle (c'est un parallélogramme aplati). Prenons deux vecteur u et v et considérons le parallélogramme aplati construit sur $u+v$ et $u+v$ et essayons d'exprimer son aire (nulle) en utilisant la "règle de la somme" :

$$0 = \mathcal{A}(u+v, u+v) = \mathcal{A}(u, u+v) + \mathcal{A}(v, u+v) = \mathcal{A}(u, u) + \mathcal{A}(u, v) + \mathcal{A}(v, u) + \mathcal{A}(v, v).$$

Comme $\mathcal{A}(u, u)$ et $\mathcal{A}(v, v)$ sont nuls cela donne

$$0 = \mathcal{A}(u, v) + \mathcal{A}(v, u).$$

Si on veut pouvoir appliquer la "règle de la somme" il faut introduire introduire un signe. Le déterminant de deux vecteurs dans le plan est justement cette aire avec un signe.

On définit donc un objet qui a un signe et qui vérifie les propriétés voulues

Définition 5.1. *Le déterminant, noté* det *est une application de* $\mathbb{R}^2 \times \mathbb{R}^2$ *dans* $\mathbb{R}$ *qui vérifie*
- $|\det(u, v)| = \mathcal{A}(u, v)$ *et* $\det(u, v) = \mathcal{A}(u, v)$ *si la base* $(\vec{u}, \vec{v})$ *est directe.*
- $\det(e_1, e_2) = 1$ *(base canonique)*
- det *est linéaire en chaque variable (2-linéaire)*
- $\det(u, v) = -\det(v, u)$ *(alternée, antisymétrique)*

Proposition 5.2. *Une telle application existe et est unique.*

Démonstration : Cela découle du calcul suivant, qui montre qu'il n'y a qu'une seule possibilité. Décomposons u et v dans la base canonique $u = u_1e_1 + u_2e_2$, $v = v_1e_1 + v_2e_2$. On a alors

$$\begin{aligned} det(u, v) &= det(u_1e_1 + u_2e_2, v_1e_1 + v_2e_2) \\ &= u_1 det(e_1, v_1e_1 + v_2e_2) + u_2 det(e_2, v_1e_1 + v_2e_2) \\ &= u_1v_1 det(e_1, e_1) + u_1v_2 det(e_1, e_2) + u_2v_1 det(e_2, e_1) + u_2v_2 det(e_2, e_2) \end{aligned}$$

Mais $det(e_1, e_1) = det(e_2, e_2) = 0$, $det(e_1, e_2) = 1$ et $det(e_2, e_1) = -det(e_1, e_2)$. On a donc

$$det(u, v) = u_1 v_2 - u_2 v_1.$$

Cette formule définit une fonction qui vérifie les propriétés ci-dessus. Cela montre la proposition. □

5.2 Volume et déterminant dans $\mathbb{R}^3$

Si w orthogonal au plan$Vect(u, v)$ on a $Vol(u, v, w) = \mathcal{A}(u, v).|w|$. On aimerait définir un det qui représenterait ce que l'on attendrait d'un volume avec un signe.

Définition 5.3. *Le déterminant, noté* det *est une application de* $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ *dans* $\mathbb{R}$ *qui vérifie*
- $|\det(u, v, w)| = \mathrm{vol}(u, v, w)$ *et* $\det(u, v, w) = \mathrm{vol}(u, v, w)$ *si la base* $(\vec{u}, \vec{v}, \vec{w})$ *est directe.*
- $\det(u, v, w) = \|w\|\mathcal{A}(u, v)$ *si la base* (u, v, w) *est directe et* w *orthogonal au plan* $Vect(u, v)$
- $\det(e_1, e_2, e_3) = 1$ *(base canonique)*
- det *est linéaire en chaque variable (3-linéaire)*
- $\det(u, v, w) = -\det(v, u, w)$ *(alternée, antisymétrique) et idem avec les autres paires de variables*
- $\det(u, v, w) = 0$ *ssi la famille* (u, v, w) *est liée*

Proposition 5.4. *Le determinant de* $\mathbb{R}^3$ *existe et est unique.*

En cas d'ambiguité, on notera $\det_2$ le det de $\mathbb{R}^2$ et $\det_3$ celui de $\mathbb{R}^3$

<u>Démonstration :</u> Il se définit de même par un calcul. En effet, on écrit $\vec{u}_i = u_{1i}\vec{e}_1 + u_{2i}\vec{e}_2 + u_{3i}\vec{e}_3$. Les propriétés demandées au déterminant imposent alors

$$\begin{aligned}
\det(\vec{u}_1, \vec{u}_2, \vec{u}_3) &= u_{11}u_{22}u_{33}\det(e_1, e_2, e_3) + u_{11}u_{32}u_{23}\det(e_1, e_3, e_2) \\
&+ u_{21}u_{12}u_{33}\det(e_2, e_1, e_3) + u_{21}u_{32}u_{13}\det(e_2, e_3, e_1) \\
&+ u_{31}u_{12}u_{23}\det(e_3, e_1, e_2) + u_{31}u_{22}u_{13}\det(e_3, e_2, e_1) \\
&= u_{11}u_{22}u_{33} - u_{11}u_{32}u_{23} - u_{21}u_{12}u_{33} \\
&+ u_{21}u_{32}u_{13} + u_{31}u_{12}u_{23} - u_{31}u_{22}u_{13}
\end{aligned}$$

Ecrire la règle de Sarrus mnémotechnique.

Par ailleurs, cette formule définit une fonction de $\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$ dans $\mathbb{R}$dont on vérifie aisément (fastidieux...) qu'elle vérifie les propriétés souhaitées. □

5.3 Déterminant dans $\mathbb{R}^n$

Qu'est ce que c'est le volume ? Le déterminant ? Finalement, ayant compris les cas de dim 2 et 3 on définit un déterminant dans $\mathbb{R}^n$ et on définit le volume n-dimensionnel comme sa valeur absolue.

Le déterminant se définit en quelque sorte par récurrence, à l'aide du déterminant de $\mathbb{R}^{n-1}$

Définition 5.5. *Le déterminant, noté* det *est une application définie sur les n-uplets de vecteurs de* $\mathbb{R}^n$ *à valeurs dans* $\mathbb{R}$ *qui vérifie*

- *Soient* $u_1, \dots u_{n-1}$ *des vecteurs de* $\mathbb{R}^{n-1}$ *identifié à* $\mathbb{R}^{n-1} \times \{0\} \subset \mathbb{R}^n$*. Et soit* $\vec{u}_n = \lambda_n \vec{e_n}$*. Alors* $\det_{\mathbb{R}^n}(u_1, \dots, u_{n-1}, u_n) = \det_{\mathbb{R}^{n-1}}(u_1, \dots u_{n-1}) \times \lambda_n$*.*
- $\det(e_1, e_2, \dots e_n) = 1$ *(base canonique)*
- det *est linéaire en chaque variable (n-linéaire)*
- $\det(u_1 \dots u_n) = -\det(u_2, u_1, u_3 \dots u_n)$ *(alternée, antisymétrique) et idem avec les autres paires de variables*
- $\det(u_1, \dots u_n) = 0$ *ssi la famille* $(u_1, \dots u_n)$ *est liée*

Proposition 5.6. *Le determinant de* $\mathbb{R}^n$ *existe et est unique.*

Démonstration :

Comme plus haut, cela vient du fait que les conditions ci-dessus impliquent qu'il doit être défini par une certaine formule. Cette formule est la suivante

$$\det(\vec{u}_1, \dots, \vec{u}_n) = \sum_{\sigma \in \mathcal{S}_n} (-1)^{\epsilon(\sigma)} u_{1\sigma(1)}.u_{2\sigma(2)} \dots u_{n\sigma(n)}$$

Ici σ est une permutation des indices $\{1, \dots, n\}$. Une telle permutation peut être décrite comme une succession d'échanges de variables comme au 4ème point de la définition ci-dessus. Ainsi, si on fait un nombre pair d'échanges de variables, les -1 se compensent. Alors que si on fait un nombre impair d'échanges de variables, il y a un -1 qui sort. Si σ est une permutation, $\epsilon(\sigma)$ désigne le nombre d'échanges de variables qu'il faut faire pour arriver à σ. Ce qui est important est qu'il soit pair ou impair.

Cette formule, vérifiée dans les cas $n = 2$ et $n = 3$, se justifie par récurrence. Supposons que nous ayons défini un déterminant dans $\mathbb{R}^{n-1}$. Alors les propriétés de la définition ci-dessus impliquent que

$$\det(\vec{u}_1, \dots, \vec{u}_n) \quad = \quad u_{1n} \det_{n-1}() + u_{2n} \det_{n-1}() + \dots u_{nn} \det_{n-1}()$$

L'hypothèse de récurrence appliquée à ces déterminants donne le résultat voulu. ADMIS.

□

5.4 Développement suivant une ligne ou une colonne

Concrètement, un déterminant de taille au moins 4 (ou même trois) ne se calcule pas avec ces formules, mais suivant la procédure de dév suivant une ligne ou une colonne.

$$\begin{aligned}
\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,n-1} & a_{1n} \\ a_{21} & \dots & \dots & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & \dots & a_{nn} \end{vmatrix} &= a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots & a_{2,n-1} & a_{2n} \\ a_{32} & \dots & \dots & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n2} & \dots & \dots & \dots & a_{nn} \end{vmatrix} \\
&- a_{12} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2,n-1} & a_{2n} \\ a_{31} & a_{33} & \dots & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n3} & \dots & \dots & a_{nn} \end{vmatrix} \\
&+ \quad \dots \\
&+ (-1)^n a_{1n} \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2,n-2} & a_{2,n-1} \\ a_{21} & \dots & \dots & \dots & a_{2,n-1} \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & \dots & a_{n,n-1} \end{vmatrix}
\end{aligned}$$

On note a_{ij} le coeff de la i-ème ligne et la j-ème colonne. A_{ij} est le mineur d'ordre ij, i.e. la matrice A moins la $i$ème col et la j-ème ligne. Ainsi, on a

$$\det(\vec{u}_1, \dots, \vec{u}_n) = a_{11} \det A_{11} - a_{12} \det A_{12} + \dots + (-1)^n a_{1n} \det A_{1n}$$

Ce qu'on a fait pour la première ligne peut se faire pour n'importe quelle ligne ou colonne.

Ajouter exemples en dim 3 pour comprendre d'où ça vient ;

5.5 Déterminant d'une matrice et déterminant d'un endomorphisme de $\mathbb{R}^n$

Définition 5.7. *Le déterminant d'une matrice est le déterminant de ses vecteurs colonnes. Le déterminant d'un endomorphisme de $\mathbb{R}^n$ est le déterminant de sa matrice dans la base canonique.*

On utilise des barres pour noter les déterminants. Par exemple si A est la matrice suivante

$$A = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 4 & -1 \\ 3 & 1 & -3 \end{pmatrix},$$

son déterminant est noté

$$det(A) = \begin{vmatrix} 1 & 0 & -3 \\ 0 & 4 & -1 \\ 3 & 1 & -3 \end{vmatrix}.$$

Proposition 5.8. *Soit φ une application linéaire. Alors elle multiplie tous les volumes par $|det(\varphi)|$. Autrement dit, pour toute partie A raisonnable de E, $vol(\varphi(A)) = |\det\varphi|.vol(A)$*

Démonstration : On va démontrer cette proposition sans trop de rigueur. Imaginons qu'on découpe A en petits cubes. $\varphi(A)$ est découpée en les images des petits cubes par φ. Il suffit donc de comprendre comment est transformé un petit cube. Par définition du déterminant, le volume de $\varphi(e_1, \dots, e_n)$ vaut $|\det\varphi|$. Si on regarde un cube C_l de coté l, son volume est $1/l^n = \det(\frac{1}{l}e_1, \dots, \frac{1}{l}e_n)$. Et son image $\varphi(C_l)$ a pour volume $\frac{1}{l^n}\det(\varphi(e_1), \dots \varphi(e_n)) = \det\varphi \times vol(C_l)$

Toute forme A raisonnablement approximé par des petits cubes voit donc son volume multiplié par $\det\varphi$: $vol(\varphi(A)) = \det\varphi \times vol(A)$ □

Théorème 5.9. *Soit A et B deux matrices carrées. On a*

$$det(AB) = det(A)det(B)$$

De même si f et g sont deux endomorphismes de $\mathbb{R}^n$ alors

$$\det g \circ f = \det g \times \det f.$$

Démonstration : Soit A un matrice. Considérons l'application

$$(V_1, \dots, V_n) \mapsto det(AV_1, \dots, AV_n).$$

En écrivant les coordonnées des vecteurs V_i dans la base canonique on peut écrire

$$V_i = \sum_{j=1}^{n} v_{ji}e_j$$

$$det(AV_1, \dots, AV_n) = det(A\sum_{j=1}^{n} v_{j1}e_j, \dots, A\sum_{j=1}^{n} v_{jn}e_j) = det(\sum_{j=1}^{n} v_{j1}Ae_j, \dots, \sum_{j=1}^{n} v_{jn}Ae_j)$$

En appliquent les mêmes règles que plus haut on trouve

$$det(AV_1, \dots, AV_n) = det(V_1, \dots, V_n)det(Ae_1, \dots, Ae_n) = det(V_1, \dots, V_n)det(A)$$

Prenons maintenant deux matrices A et B. On a

$$det((AB)V_1, \dots, (AB)V_n) = det(V_1, \dots, V_n)det(AB),$$

mais aussi

$$\begin{aligned} det((AB)V_1, \ldots, (AB)V_n) &= det(A(BV_1), \ldots, A(BV_n)) \\ &= det(BV_1, \ldots, BV_n)det(A) \\ &= det(V_1, \ldots, V_n)det(B)det(A). \end{aligned}$$

D'où l'égalité $det(AB) = det(B)det(A)$. □

Une autre façon de comprendre pourquoi le résultat précédent est vrai (ne tenant pas compte du signe). Multiplier par B multiplie les volumes par $|det(B)|$, multiplier par A multiplie les volumes par $|det(A)|$, donc multiplier par B puis par A multiplie les volumes par $|det(A)||det(B)|$. Mais par ailleurs, multiplier par AB multiplie les volumes par $|det(AB)|$. On en déduit

$$|det(AB)| = |det(A)||det(B)|.$$

Remarque $f : \mathbb{R}^n \to \mathbb{R}^n$ bij ssi inj ssi surj ssi les vecteurs $f(e_1), \ldots, f(e_n)$ forment une base ssi $\det(f(e_1), \ldots, f(e_n)) \neq 0$

Proposition 5.10. *Un endomorphisme f de $\mathbb{R}^n$ est bijectif si et seulement si* $\det(f) \neq 0$

Démonstration : Si $\det(f) \neq 0$ alors le volume du parallélépipède construit sur les vecteurs colonne de la matrice A de f dans la base canonique n'est pas 0. Cela signifie que les vecteurs colonne en question forment une base de $\mathbb{R}^n$ (sinon le parallélépipède serait aplati car les vecteurs colonne seraient tous dans un même sous-espace vectoriel de dimension strictement inférieure à n). Nous avons vu que cela entraîne que la matrice A est inversible (donc f aussi).

Si f est inversible, sa matrice A dans la base canonique l'est aussi. On a

$$1 = det(Id) = det(AA^{-1}) = det(A)det(A^{-1}).$$

Le déterminant de f n'est donc pas nul. □

Définition 5.11. *déterminant d'une matrice : déterminant de ses vecteurs colonnes.*

Proposition 5.12. – $\det {}^tA = \det A$ *et*
– $\det AB = \det A \det B$.
– $\det I_n = 1$
– $\det A^{-1} = 1/\det A$

En particulier, si P inversible, $\det A = \det P^{-1}AP$.

On peut ajouter une caractérisation de matrice inversible à notre liste.

Théorème 5.13. *Soit $A \in M_{n\times n}(\mathbb{R})$ une matrice carrée. Les assertions suivantes sont équivalentes :*

1. *A est inversible,*
2. *$rgA = n$,*
3. *$AX = 0 \implies X = 0$,*
4. *$X \neq 0 \implies AX \neq 0$,*
5. *$\forall Y\ \exists X\ AX = Y$,*
6. *$\forall Y\ \exists! X\ AX = Y$,*
7. *les vecteurs colonne de A forment une famille libre dans $\mathbb{R}^n$,*
8. *les vecteurs colonne de A forment une famille génératrice dans $\mathbb{R}^n$,*
9. *les vecteurs colonne de A forment une base de $\mathbb{R}^n$,*
10. *les vecteurs ligne de A forment une famille libre dans $\mathbb{R}^n$,*
11. *les vecteurs ligne de A forment une famille génératrice dans $\mathbb{R}^n$,*
12. *les vecteurs ligne de A forment une base de $\mathbb{R}^n$,*
13. *$det(A) \neq 0$.*

Cela donne aussi une nouvelle façon de calculer le rang d'une matrice.

Proposition 5.14. *Le rang d'une matrice est le plus grand nombre k tel qu'on puisse extraire un déterminant $k \times k$ différent de 0 de la matrice.*

Extraire signifie ici obtenir en effaçant un certain nombre de lignes et un certain nombre de colonnes. Un tel déterminant extrait d'une matrice s'appelle un mineur de la matrice.

5.6 Polynôme caractéristique

Définition 5.15. *Soit E un $\mathbb{R}$ ev (ou $\mathbb{C}$ ev). Soit $f : E \to E$ linéaire. On définit $\chi_f(X) = \det(X.Id - f) \in \mathbb{R}[X]$ (ou $\mathbb{C}[X]$). C'est un polynôme en X à coeff dans $\mathbb{R}$ ou $\mathbb{C}$.*

Proposition 5.16. $\chi_A(X) = X^n - Tr(A)X^{n-1} + \cdots + (-1)^n \det A$

Démonstration : Calcul matriciel □

Exemple $f(x,y) = (-y,x)$ dans $\mathbb{R}^2$ alors $\chi_f(X) = \det\begin{pmatrix} X & 1 \\ -1 & X \end{pmatrix} = X^2 + 1$.

Exemple $f(x,y) = (-y,x)$ dans $\mathbb{C}^2$ alors $\chi_f(X) = \det\begin{pmatrix} X & 1 \\ -1 & X \end{pmatrix} = X^2 + 1$.

Exemple f_A associée à la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ Alors $\chi_f(X) = (X-1)^2(X+1)$

Proposition 5.17. *Soit E un $\mathbb{R}$ ev. Si $\lambda \in \mathbb{R}$ vp de f alors λ racine réelle de χ_f. Idem sur $\mathbb{C}$*

Démonstration : $e_1 = \vec{u}$ vect propre pour λ. G suppl de $Vect(e_1)$ Soit $e_2, \dots e_n$ base de G. Alors $Mat_{(e_1,\dots,e_n)}(X.Id - f) = \begin{pmatrix} X-\lambda & * & * \\ 0 & * & * \\ 0 & * & * \\ 0 & * & * \\ \dots & * & * \end{pmatrix}$ Alors en développant suivant 1e colonne, on trouve $\det(XId - f) = (X - \lambda) \times Q(X)$ où $Q(X)$ est le déterminant de la matrice obtenue en rayant 1e ligne et 1e colonne. Donc λ est racine de χ_f. $\square$

Proposition 5.18. *Si λ racine réelle de χ_f alors λ vp*

Démonstration : λ racine de P ssi $f - \lambda Id$ pas injective ssi il existe $\vec{u} \neq \vec{0}$ dans $Ker(f - \lambda Id)$, ssi il existe $\vec{u} \neq \vec{0}$ tq $f(\vec{u}) - \lambda\vec{u} = \vec{0}$ ssi λ vp de f $\square$

Théorème 5.19. *Les vp de f sont les racines de χ_f.*

Corollaire 5.20. *$f : E \to E$ avec E $\mathbb{R}$-ev de dim n. Alors f admet au plus n vp et f a n vp distinctes réelles ssi χ_f a n racines distinctes réelles. Idem dans $\mathbb{C}$.*

Proposition 5.21. *Soit E un ev de dim n (sur $\mathbb{R}$ ou $\mathbb{C}$) Si χ_f a n racines distinctes alors f est diagonalisable.*

Démonstration : Soient $\lambda_1, \dots, \lambda_n$ les vp. Alors $\dim E_{\lambda_i} \geq 1$. Et $E_{\lambda_1} \oplus E_{\lambda_2} \dots \oplus E_{\lambda_n} \subset E$. Le terme de gauche est de dim au moins n et E de dim n. Donc tous les E_{λ_i} sont de dim 1. On prend $u_i \neq 0$ dans le E_{λ_i} et on obtient une famille libre de n vecteurs propres, donc une base. $\square$

Remarque 5.22. Ce n'est pas nécessaire. Exemple S la symétrie par rapport au plan d'équation $y = 0$ dans $\mathbb{R}^3$ a pour vp ± 1, $\chi_S(X) = (X-1)^2(X+1)$ Sa matrice dans la base canonique est $\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Elle est diagonale donc S diagonalisable. Pourtant que 2 vp.

Ce critère est bien pratique, il est "souvent" vérifié en pratique.

Voir semestre prochain pour la diagonalisation des matrices symétriques.

Proposition 5.23. *Si χ_f n'a que $0 \leq k < n$ racines (donc vp) alors*

$$k \leq \dim(E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}) \leq \dim E = n$$

et tous les cas intermédiaires peuvent arriver.

Un ex important $f(x,y) = (2x+y, 2y)$. Alors $\chi_f(X) = (X-2)^2$ une seule vp 2. $Mat(f) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ Si f était diagonalisable, alors f serait egal à $2.Id$. Or ce n'est pas le cas .Donc f pas diagonalisable.

Il peut n'y avoir aucune vp. Ex $f(x,y) = (-y,x)$ de $\mathbb{R}^2$ dans $\mathbb{R}^2$ alors $\chi_f(X) = X^2+1$ n'a pas de racine réelle et f (rotation d'angle $\pi/2$) n'a pas de vp

Attention $g : \mathbb{C}^2 \to \mathbb{C}^2$ définie par $g(x,y) = (-y,x)$ a deux vp, $\pm i$, car $\chi_g(X) = \chi_f(X)$ a deux racines complexes.

Proposition 5.24. *Si E $\mathbb{R}$-ev de dim impaire ou bien E $\mathbb{C}$-ev, alors au moins une vp.*

Démonstration : Il y a au moins une racine de χ_f. □

Remarque : en pratique, étant donnée une application linéaire f / une matrice A , il arrive fréquemment qu'on ait besoin de calculer ses puissances.

Bien sûr, si A est une matrice diagonale, $A = Diag(\lambda_1, \ldots, \lambda_k)$, alors $A^n = Diag(\lambda_1^n, \ldots, \lambda_k^n)$, c'est ce qui est le mieux pour calculer les puissances.

Cela étant dit, même si A n'est pas diagonalisable, s'il y a une relation du type $A^3 = A^2 + A - 1$, alors on peut calculer les puissances en écrivant $A^4 = A.A^3 = A^3 + A^2 - A = A^2 + A - 1 + A^2 - A = 2A^2 - 1$, puis $A^5 = 2A^3 - A = 2A^2 + 2A - 2 - A) = 2A^2 + A - 2$, etc

Ce genre de relation est donc très utile, même et surtout quand A n'est pas diagonalisable.

Définition 5.25. *Soit $P \in \mathbb{R}[X]$ ou $P \in \mathbb{C}[X]$ un polynôme, ie $P(X) = \sum_{i=0}^N a_i X^i$. Soit $f : E \to E$ une application linéaire (E $\mathbb{R}$-ev ou $\mathbb{C}$-ev)s L'application $P(f)$ est définie comme*

$$P(f) = a_0 Id + a_1 f + a_2 f \circ f + \ldots a_N f \circ f \circ \ldots f = \sum_{i=0}^N a_i f^i$$

Le théorème suivant s'appelle le théorème de Cayley Hamilton.

Théorème 5.26. *$\chi_f(f) = 0$ en tant qu'application linéaire.*

Exemple $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, $\chi_f(X) = (X-2)(X-3) = X^2 - 5X + 6$, alors $f^2 - 5f + 6Id = 0$.

6 Deux applications en statistique descriptive

6.1 Régression linéaire

Sur une collection d'individus on observe des variables y_i et des variables explicatives (ou régresseurs) x_i, $i = 1, ...n$, chaque paire (y_i, x_i) représentant une expérience. On les arrange dans un tableau de la façon suivante :

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 & x_{12} & \dots & x_{1p} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n2} & \dots & x_{np} \end{pmatrix}.$$

x_i est donc un vecteur ligne. On notera $x_{.j}$ la j-ème colonne. On convient généralement que le premier régresseur est la constante, mais ça n'est pas obligatoire. Pour tout vecteur β, on définit l'erreur quadratique moyenne d'ajustement (ou erreur résiduelle) $S(\beta)$ comme

$$S(\beta)^2 = \frac{1}{n}\|y - X\beta\|^2 = \frac{1}{n}\sum_i (y_i - x_i\beta)^2.$$

Le vecteur des coefficients de regression calculé aux moindres carrés ordinaires est

$$\hat{\beta} = \arg\min_{\beta} S(\beta)$$

Proposition 6.1. *Si la matrice ${}^tXX$ est inversible, alors $\hat{\beta}$ est donné par*

$$\hat{\beta} = ({}^tXX)^{-1}.{}^tXy.$$

D'abord nous allons montrer que S a un point critique en $\hat{\beta}$ puis qu'elle a un minimum global en $\hat{\beta}$.

Calculons les dérivées partielles de S :

$$\begin{aligned}
\frac{\partial S}{\partial \beta_j} &= \frac{1}{n}\frac{\partial}{\partial \beta_j}\|y - X\beta\|^2 \\
&= \frac{1}{n}\frac{\partial}{\partial \beta_j}\sum_i (y_i - x_i\beta)^2 \\
&= \frac{1}{n}\frac{\partial}{\partial \beta_j}\sum_i (y_i - \sum_k x_{ik}\beta_k)^2 \\
&= \frac{1}{n}\sum_i \frac{\partial}{\partial \beta_j}(y_i - \sum_k x_{ik}\beta_k)^2 \\
&= \frac{1}{n}\sum_i (-2x_{ij})(y_i - \sum_k x_{ik}\beta_k) \\
&= \frac{-2}{n}\sum_i y_i x_{ij} + \frac{2}{n}\sum_i\sum_k x_{ij}x_{ik}\beta_k \\
&= \frac{-2}{n}({}^tXy)_j + \frac{2}{n}\sum_k(\sum_i x_{ij}x_{ik})\beta_k \\
&= \frac{-2}{n}({}^tXy)_j + \frac{2}{n}\sum_k ({}^tXX)_{jk}\beta_k \\
&= \frac{2}{n}(({}^tXX\beta)_j - ({}^tXy)_j) \\
&= \frac{2}{n}(({}^tXX\beta - {}^tXy)_j).
\end{aligned}$$

Autrement dit

$$\nabla S = \frac{2}{n}({}^tXX\beta -{}^t Xy).$$

Le gradient s'annule lorsque ${}^tXX\beta - {}^tXy = 0$ c'est-à-dire $\beta = ({}^tXX)^{-1}({}^tXy)$.

Pour montrer que S a un minimum en son point critique on peut montrer que S est coercive (comme il y a une base de $\mathbb{R}^p$ formée de vecteurs ligne de X, si la norme de β est très grande au moins un produit de l'un de ces vecteurs ligne par β est très grand ce qui entraîne que $S(\beta)$ est très grand).

On peut aussi calculer la hessienne de S.

$$\begin{aligned}
\frac{\partial^2 S}{\partial \beta_l \partial \beta_j} &= \frac{-2}{n}\sum_i x_{ij}\frac{\partial}{\partial \beta_l}(y_i - \sum_k x_{ik}\beta_k) \\
&= \frac{2}{n}\sum_i x_{ij}x_{il} \\
&= \frac{2}{n}({}^tXX)_{jl}.
\end{aligned}$$

Cela signifie que la hessienne de S est $2\ {}^tXX/n$, matrice définie positive. La fonction S est donc strictement convexe ; elle a un minimum en son unique point critique.

Autre façon de faire le calcul.

$$\begin{aligned} S(\beta) &= \frac{1}{n}\|y - X\beta\|^2 \\ &= \frac{1}{n}\langle y - X\beta, y - X\beta\rangle \\ &= \frac{1}{n}\left(\|y\|^2 - 2\langle y, X\beta\rangle + \langle X\beta, X\beta\rangle\right) \\ &= \frac{1}{n}\left(\|y\|^2 - 2\langle {}^tXy, \beta\rangle + \langle {}^tXX\beta, \beta\rangle\right) \end{aligned}$$

On en déduite que le gradient de S en β est $\frac{2}{n}\left({}^tXX\beta - {}^tXy\right)$ et sa hessienne $2\ {}^tXX/n$. Le gradient est nul lorsque

$${}^tXX\beta - {}^tXy = 0, \text{ soit } \beta = ({}^tXX)^{-1}({}^tXy),$$

et comme la hessienne est constante définie positive S est strictement convexe. Elle a donc un minimum global en son unique point critique.

6.2 Analyse en composantes principales

On se donne n vecteurs $x_1, \dots, x_n$ dans $\mathbb{R}^d$. Ces vecteurs représentent des données recueillies sur n individus (d nombres par individus). Ces n vecteurs définissent un nuage de points dans $\mathbb{R}^d$. On veut projeter ce nuage de points sur une droite ou un plan en perdant le moins d'information possible. Nous supposons que les données sont centrées c'est-à-dire que la moyenne des vecteurs x_i est le vecteur nul.

Commençons par chercher une droite Δ telle que la somme des distances de chaque x_i à Δ au carré soit minimale. Appelons p_Δ la projection sur Δ. On cherche à minimiser

$$\sum_{i=1}^n \|x_i - p_\Delta(x_i)\|^2.$$

Le théorème de Pythagore donne

$$\|x_i\|^2 = \|x_i - p_\Delta(x_i)\|^2 + \|p_\Delta(x_i)\|^2.$$

On a donc

$$\sum_{i=1}^n \|x_i - p_\Delta(x_i)\|^2 = \sum_{i=1}^n \|x_i\|^2 - \sum_{i=1}^n \|p_\Delta(x_i)\|^2,$$

et, comme $\sum_{i=1}^n \|x_i\|^2$ ne dépend pas de Δ, minimiser $\sum_{i=1}^n \|x_i - p_\Delta(x_i)\|^2$ revient à maximiser $\sum_{i=1}^n \|p_\Delta(x_i)\|^2$. Ecrivons qu'une droite est déterminée par un vecteur directeur unitaire v : $\Delta = \mathbb{R}v$. Cela permet de donner une expression analytique à la quantité que nous cherchons à maximiser car $p_\Delta(x_i) = \langle x_i, v\rangle v$. On cherche v de norme 1 tel que

$$\phi(v) = \sum_{i=1}^n \langle x_i, v\rangle^2.$$

C'est un problème d'optimisation sous contrainte ($\|v\|^2 = 1$). La fonction ϕ est un forme quadratique. Pour obtenir la matrice symétrique (positive par définition de ϕ (somme de carrés)) associée à ϕ, il suffit d'écrire le produits scalaire sous forme de produit de matrices :

$$\langle x_i, v\rangle^2 = {}^t v x_i {}^t x_i v.$$

Cela donne

$$\phi(v) = \sum_{i=1}^n \langle x_i, v\rangle^2 = \sum_{i=1}^n {}^t v x_i {}^t x_i v = {}^t v \left(\sum_{i=1}^n x_i {}^t x_i\right) v,$$

soit, en posant $A = \sum_{i=1}^n x_i {}^t x_i$:

$$\phi(v) = \langle v, Av\rangle.$$

Remarque : si nous appelons X la matrice $d \times n$ dont les vecteurs colonnes sont les x_i. On peut vérifier que l'on a

$$A = X.{}^t X,$$

qui est une matrice définie positive si X est de rang d (ce qui nécéssite $n \geq d$).

Les gradients de ϕ et de la contrainte au point v sont

$$\nabla\phi(v) = 2Av, \;\; \nabla\|v\|^2 = 2v.$$

Le gradient de la contrainte ne s'annule pas sur l'ensemble $\|v\|^2 = 1$, on peut donc appliquer le théorème des extrema liés : en un point où ϕ est maximale sous la contrainte, il existe λ tel que

$$\nabla\phi(v) = \lambda\nabla\|v\|^2 \text{ soit } Av = \lambda v.$$

(Remarque : grâce au calcul précédent on peut démontrer que les matrices réelles symétriques sont diagonalisables. Résultat qui apparaît alors comme une conséquence du théorème des extrema liés)

Ce que nous avons obtenu est que ϕ ne peut être maximale qu'en un vecteur propre de A. La valeur de ϕ en un vecteur propre de A unitaire est la valeur propre associée. On en déduit que $\phi(v)$ est maximale si v est un vecteur propre de A associé à la plus grande valeur propre de A.

Nous pouvons alors énoncer la proposition suivante.

Proposition 6.2. *Les droites qui approchent le mieux le nuage de points (au sens des moindres carrés pour une projection orthogonale) sont les droites propres associées à la plus grande valeur propre de la matrice $A = X.{}^t X$. Si la dimension de l'espace propre associé à la plus grande valeur propre est 1 alors une seule droite donne la meilleure approximation.*

Intéressons-nous maintenant à la meilleure approximation par un plan. Nous allons montrer la proposition suivante.

Proposition 6.3. *Les plans qui approchent le mieux le nuage de points (au sens des moindres carrés pour une projection orthogonale) sont les plans obtenus comme somme de deux droites propres associées aux deux plus grandes valeurs propres de la matrice* $A = X.^tX$. *Si la multiplicité de la plus grande valeur propre est au moins 2 alors on prend deux droites propres associées à cette plus grande valeur propre. Si la somme des dimensions des espaces propres associés aux deux plus grandes valeurs propre est 2 alors un seul plan donne la meilleure approximation.*

On cherche deux vecteurs unitaires v_1 v_2 orthogonaux tels que la somme

$$\sum_{i=1}^{n} \|x_i - p(x_i)\|^2,$$

où p est la projection orthogonale sur $Vect(v_1, v_2)$. On montre que

$$p(x_i) = \langle x_i, v_1\rangle v_1 + \langle x_i, v_2\rangle v_2.$$

Là encore le théorème de Pythagore donne

$$\|x_i\|^2 = \|x_i - p(x_i)\|^2 + \|p(x_i)\|^2,$$

et minimiser $\sum_{i=1}^{n} \|x_i - p(x_i)\|^2$ revient à maximiser $\sum_{i=1}^{n} \|p(x_i)\|^2$. La dépendance en v_1 et v_2 est cachée dans le p :

$$\psi(v_1, v_2) = \sum_{i=1}^{n} \|p(x_i)\|^2 = \sum_{i=1}^{n} \|\langle x_i, v_1\rangle v_1 + \langle x_i, v_2\rangle v_2\|^2 = \sum_{i=1}^{n} \langle x_i, v_1\rangle^2 + \sum_{i=1}^{n} \langle x_i, v_2\rangle^2.$$

On cherche donc à maximiser $\psi(v_1, v_2) = \phi(v_1) + \phi(v_2)$ sous la contrainte v_1 et v_2 sont orthogonaux. En prenant pour v_1 et v_2 deux vecteurs propres de A unitaires associés aux deux plus grandes valeurs propres de A on obtient la somme de ces deux plus grandes valeurs propres pour $\psi(v_1, v_2)$. On ne peut pas faire mieux. Comment le voir ?

Numérotons les valeurs propres de A par valeurs descendantes : $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d$ et fixons une famille $u_1, \ldots, u_d$ de vecteurs propres unitaires associés à ces valeurs propres ($Au_i = \lambda_i u_i$). Le maximum de $\phi(v) = \langle v, Av\rangle$ sur l'hyperplan $u_1^{\perp}$ est λ_2. Considérons alors deux vecteurs v_1 et v_2 unitaires et orthogonaux. Le plan $Vect(v_1, v_2)$ et l'hyperplan $u_1^{\perp}$ s'intersectent au moins le long d'une droite. Prenons une base orthonormée w_1, w_2 de $Vect(v_1, v_2)$ telle que w_1 appartiennent à $u_1^{\perp}$. On a alors $\phi(w_1) \leq \lambda_2$ et

$$\psi(v_1, v_2) = \psi(w_1, w_2) = \phi(w_1) + \phi(w_2) \leq \lambda_2 + \phi(w_2) \leq \lambda_2 + \lambda_1,$$

car λ_1 est le maximum de ϕ sur la sphère unité.

On montre de manière analogue que, pour tout $k = 1, \ldots, d-1$, le sous-espace vectoriel de dimension k qui approche le mieux le nuage de points est $Vect(u_1, \ldots, u_k)$. Les droites $\mathbb{R}u_i$ s'appellent les composantes principales ou les axes principaux.