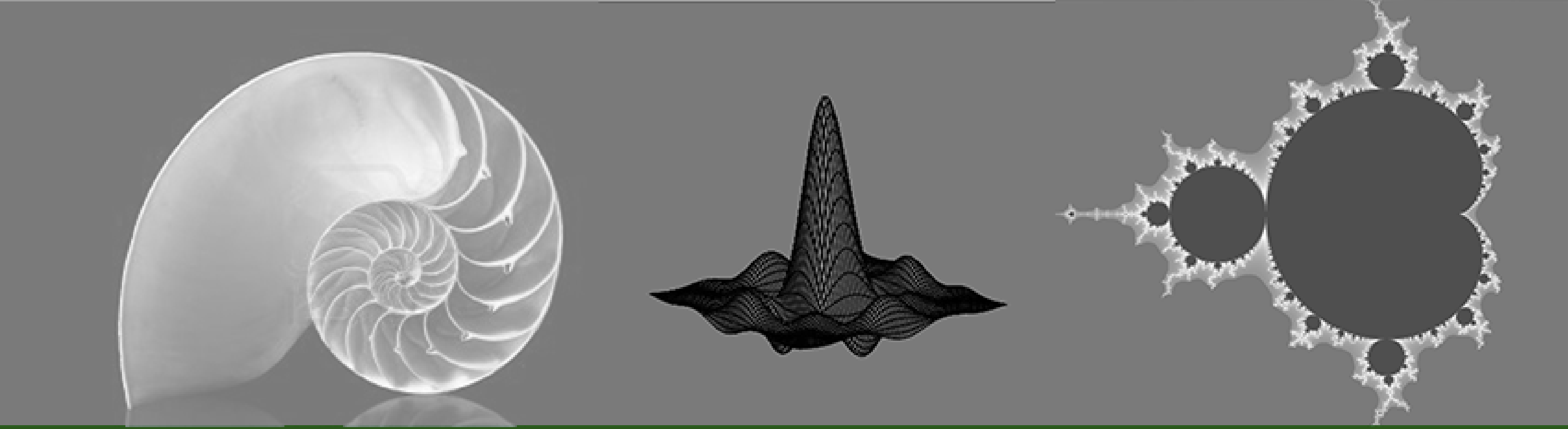
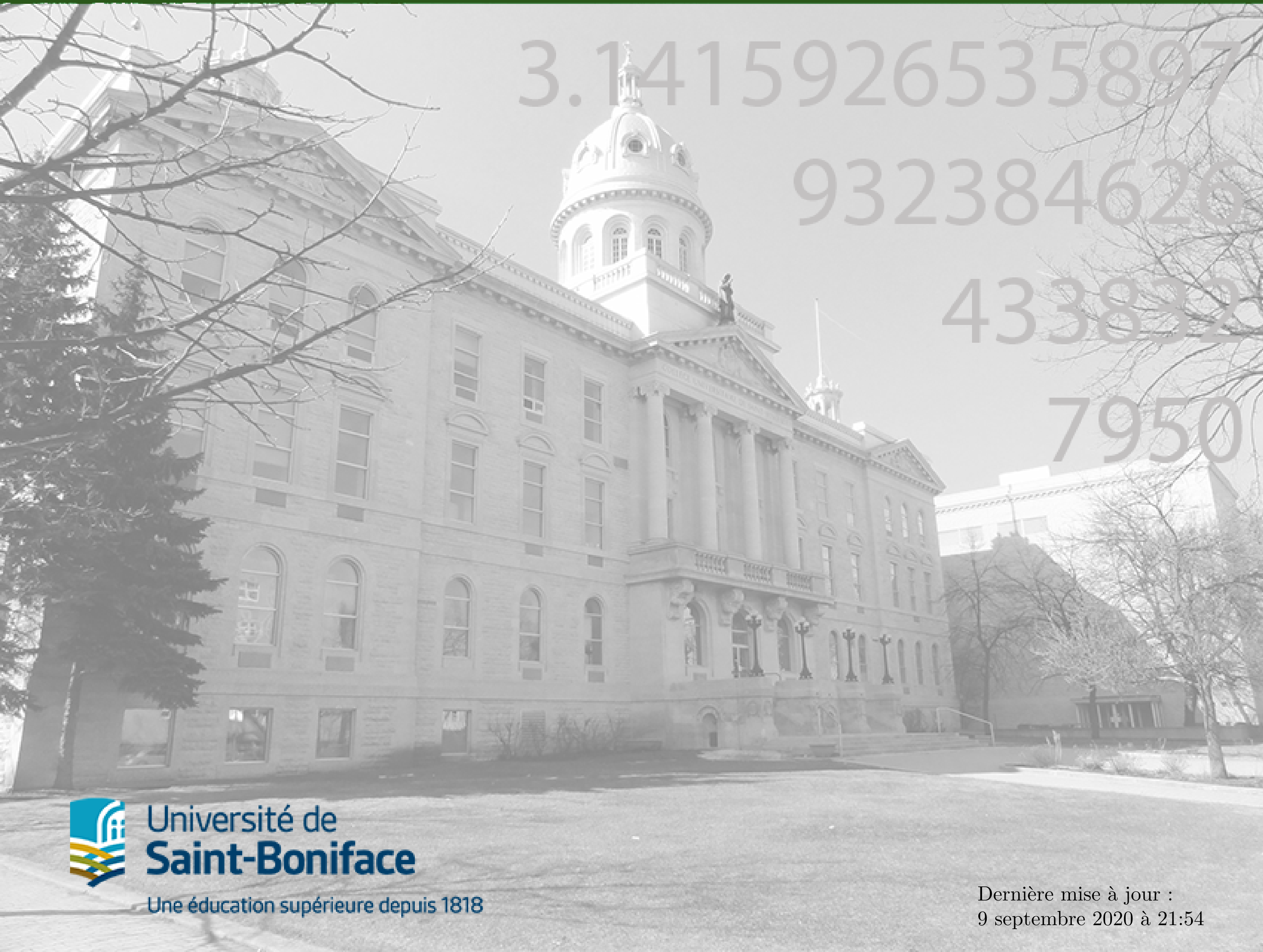


Algèbre linéaire 2

Nicolas Bouffard

Automne 2020



Université de Saint-Boniface
Une éducation supérieure depuis 1818

Dernière mise à jour :
9 septembre 2020 à 21:54

Table des matières

Chapitre 1

Les espaces vectoriels

En algèbre linéaire I, vous avez appris qu'un vecteur est un élément de $\mathbb{R}^n$. La réalité est en fait beaucoup plus générale que cela. Un vecteur n'est en fait rien d'autre qu'un élément d'un espace vectoriel. Le cas de $\mathbb{R}^n$ n'est qu'un cas particulier d'espace vectoriel. Un espace vectoriel est la structure de base sur laquelle il est possible de faire de l'algèbre linéaire. Nous allons cependant voir dans ce chapitre que lorsqu'on travaille en dimension finie, tous les espaces vectoriels ressemblent et quelque sorte à $\mathbb{R}^n$, donc bien qu'on ne vous ait pas dit toute l'histoire en algèbre linéaire I, on ne vous a pas non plus vraiment menti.

Dans ce chapitre, nous allons donc commencer par apprendre ce qu'est un espace vectoriel, puis nous allons en déduire certaines propriétés concernant entre autre les bases et la dimension.

1.1 Les espaces vectoriels

Definition 1.1.1. Un espace vectoriel (réel) est un ensemble V muni d'un addition $\oplus$ et d'une multiplication par un scalaire $\otimes$ et qui respecte les 10 propriétés ci dessous :

1. $u \oplus v \in V$ pour tout $u, v \in V$
2. $u \oplus v = v \oplus u$ pour tout $u, v \in V$
3. $u \oplus (v \oplus w) = (u \oplus v) \oplus w$ pour tout $u, v, w \in V$
4. Il existe un élément $0 \in V$ tel que $u \oplus 0 = 0 \oplus u = u$ pour tout $u \in V$
5. Pour tout $u \in V$, il existe $(-u) \in V$ tel que $u \oplus (-u) = (-u) \oplus u = 0$
6. $\lambda \otimes u \in V$ pour tout $\lambda \in \mathbb{R}$ et $u \in V$
7. $\lambda \otimes (\mu \otimes u) = (\lambda\mu) \otimes u$ pour tout $\lambda, \mu \in \mathbb{R}$ et $u \in V$.
8. Il existe un élément $1 \in \mathbb{R}$ tel que $1 \otimes u = u$ pour tout $u \in V$
9. $\lambda \otimes (u \oplus v) = (\lambda \otimes u) \oplus (\lambda \otimes v)$ pour tout $\lambda \in \mathbb{R}$ et pour tout $u, v \in V$
10. $(\lambda + \mu) \otimes u = (\lambda \otimes u) \oplus (\mu \otimes u)$ pour tout $\lambda, \mu \in \mathbb{R}$ et $u \in V$

On définit un espace vectoriel complexe de la même manière que pour le cas réel, en remplaçant $\mathbb{R}$ par $\mathbb{C}$.

Exemple 1.1.1. Il existe beaucoup d'exemple d'espace vectoriel qui nous sont familier :

1. $\mathbb{R}^n$ et $\mathbb{C}^n$ l'ensemble des vecteurs (flèche) de dimension n
2. $\mathbb{P}_n(\mathbb{R})$ et $\mathbb{P}_n(\mathbb{C})$ l'ensemble des polynômes de degré au plus n
3. $M_{m\times n}(\mathbb{R})$ et $M_{m\times n}(\mathbb{C})$ l'ensemble des matrices de dimension $m \times n$
4. $\mathcal{F}(\mathbb{R})$ et $\mathcal{F}(\mathbb{C})$ l'ensemble de toutes les fonctions réelles ou complexes

Nous n'allons pas démontrer en détail chacun de ces exemples (vous devriez pouvoir le faire en exercice), mais pour vous donner un aperçu de comment faire, nous allons démontrer que l'ensemble $\mathbb{P}_1(\mathbb{R}) = \{ax+b : a, b \in \mathbb{R}\}$ forme bien un espace vectoriel. Nous allons pour ce faire vérifier chacune des 10 propriétés une par une, dans le même ordre que nous les avons énoncés, mais avant, nous allons définir correctement nos deux opérations :

$$(ax + b) \oplus (cx + d) = (a + c)x + (b + d)$$

$$\lambda \otimes (ax+b) = (\lambda a)x + (\lambda b)$$

Vérifions maintenant les 10 propriétés :

1. Si $(ax+b)$ et $(cx+d)$ sont des éléments de $\mathbb{P}_1(\mathbb{R})$, alors
$$(ax+b) \oplus (cx+d) = (a+c)x + (b+d) \in \mathbb{P}_1(\mathbb{R})$$
2. Si $(ax+b)$ et $(cx+d)$ sont des éléments de $\mathbb{P}_1(\mathbb{R})$, alors
$$(ax+b) \oplus (cx+d) = (a+c)x + (b+d) = (c+a)x + (d+b) = (cx+d) \oplus (ax+b)$$
Notez que pour obtenir la deuxième égalité, on a utilisé les propriétés des nombres réels que nous connaissons déjà $(a+c=c+a)$.
3. Si $(ax+b)$, $(cx+d)$ et $ex+f$ sont des éléments de $\mathbb{P}_1(\mathbb{R})$, alors on a :
$$\begin{array}{rcl} ((ax+b)\oplus(cx+d))\oplus(ex+f) & = & ((a+c)x+(b+d))\oplus(ex+f) \\ & = & ((a+c)+e)x+((b+d)+f) \\ & = & (a+(c+e))x+(b+(d+f)) \\ & = & (ax+b)\oplus((c+e)x+(d+f)) \\ & = & (ax+b)\oplus((cx+d)\oplus(ex+f)) \end{array}$$
Ici, la troisième égalité a été possible grace à la propriété d'associativité des nombres réels. Toutes les autres égalité proviennent de la définition de l'addition de vecteurs $\oplus$.
4. Nous allons montrer que le vecteur 0 est donné par $0x+0$. En effet, si $(ax+b) \in \mathbb{P}_1(\mathbb{R})$, on a :
$$(0x+0) \oplus (ax+b) = (0+a)x + (0+b) = ax+b$$
L'autre égalité est automatiquement satisfaite par la propriété 2 que nous avons déjà démontré.
5. Si $u = (ax+b) \in \mathbb{P}_1(\mathbb{R})$, on veut montrer que $-u = (-a)x + (-b)$. Pour ce faire, on a :
$$(ax+b) \oplus ((-a)x + (-b)) = (a+(-a))x + (b+(-b)) = (a-a)x + (b-b) = 0x+0 = 0$$
6. Si $\lambda \in \mathbb{R}$ et $(ax+b) \in \mathbb{P}_1(\mathbb{R})$, alors on a :
$$\lambda \otimes (ax+b) = (\lambda a)x + (\lambda b) \in \mathbb{P}_1(\mathbb{R})$$
7. Si $\lambda, \mu \in \mathbb{R}$, et $(ax+b) \in \mathbb{P}_1(\mathbb{R})$, alors on a :
$$\begin{array}{rcl} \lambda \otimes (\mu \otimes (ax+b)) & = & \lambda \otimes ((\mu a)x + (\mu b)) \\ & = & (\lambda(\mu a))x + (\lambda(\mu b)) \\ & = & ((\lambda\mu)a)x + ((\lambda\mu)b) \\ & = & (\lambda\mu) \otimes (ax+b) \end{array}$$
8. On veut montrer que l'élément 1 est bel et bien le nombre 1. Si $(ax+b) \in \mathbb{P}_1(\mathbb{R})$, on a donc :
$$1 \otimes (ax+b) = (1a)x + (1b) = ax+b$$
9. Si $\lambda \in \mathbb{R}$ et $(ax+b)$ et $(cx+d)$ sont des éléments de $\mathbb{P}_1(\mathbb{R})$, alors on a :
$$\begin{array}{rcl} \lambda \otimes ((ax+b)\oplus(cx+d)) & = & \lambda \otimes ((a+c)x+(b+d)) \\ & = & (\lambda(a+c))x + (\lambda(b+d)) \\ & = & (\lambda a + \lambda c)x + (\lambda b + \lambda d) \\ & = & ((\lambda a)x + (\lambda b)) \oplus ((\lambda c)x + (\lambda d)) \\ & = & (\lambda \otimes (ax+b)) \oplus (\lambda \otimes (cx+d)) \end{array}$$

10. Si $\lambda, \mu \in \mathbb{R}$ et $(ax+b) \in \mathbb{P}_1(\mathbb{R})$, alors on a :

$$\begin{aligned}
(\lambda+\mu)\otimes(ax+b) &= ((\lambda+\mu)a)x+((\lambda+\mu)b) \\
&= (\lambda a+\mu a)x+(\lambda b+\mu b) \\
&= ((\lambda a)x+(\lambda b))\oplus((\mu a)x+(\mu b)) \\
&= (\lambda\otimes(ax+b))\oplus(\mu\otimes(ax+b))
\end{aligned}$$

On remarque qu'il est relativement long de vérifier qu'un ensemble est un espace vectoriel. Dans certain cas, une propriété peut sembler complètement évidente, mais il reste tout de même essentielle de prendre le temps de vérifier chacune des propriétés une par une, en supposons connu uniquement la définition de $\oplus$ et $\otimes$, ainsi que les propriétés des nombres réels. Autrement, plusieurs erreurs pourrait survenir. Nous verrons cependant dans la prochaine section, que dans bien des cas il est possible de vérifier seulement 2 des 10 propriétés pour s'assurer qu'un ensemble forme bien un espace vectoriel.

Exercice 1.1.1. Démontrer que chacun des exemples précédent forment bien des espaces vectoriels.

Nous allons maintenant voir dans le théorème suivant qu'en supposons vraie uniquement les 10 propriétés d'un espace vectoriel qui nous servent d'axiome, on peut en déduire d'autre propriétés qui nous sont familière pour le cas de $\mathbb{R}^n$. Ces propriétés seront donc vraie pour n'importe quel espace vectoriel.

Théorème 1.1.1. **(Propriétés des espaces vectoriels)** Si V est un espace vectoriel, u un vecteur de V et $k \in \mathbb{R}$ (ou $\mathbb{C}$). Alors on a :

1. Le vecteur 0 définie à l'axiome 4 est unique
2. Le vecteur $-u$ définie à l'axiome 5 est unique
3. $0 \otimes u = 0$
4. $k \otimes 0 = 0$
5. $(-1) \otimes u = -u$
6. Si $k \otimes u = 0$ alors $k = 0$ ou $u = 0$

Démonstration.

1. Supposons qu'il y a deux vecteurs qui satisfont l'axiome 4. Appelons ces vecteurs 0_1 et 0_2. Alors par l'axiome 4 nous avons
$$0_1 \oplus 0_2 = 0_1 \text{ et } 0_1 \oplus 0_2 = 0_2$$
on doit donc avoir
$$0_1 = 0_2$$
ce qui confirme que le vecteur est unique.

2. Prenons $u \in V$ et supposons que $(-u)_1$ et $(-u)_2$ sont tous deux des vecteurs qui satisfont l'axiome 5. Alors on a :
$$\begin{aligned}
& u\oplus(-u)_1 = 0 \text{ et } u\oplus(-u)_2 = 0 \text{ par l'axiome 5} \\
\Rightarrow\ & u\oplus(-u)_1 = u\oplus(-u)_2 \\
\Rightarrow\ & (-u)_1\oplus u\oplus(-u)_1 = (-u)_1\oplus u\oplus(-u)_2 \text{ en additionnant à gauche } (-u)_2 \\
\Rightarrow\ & 0+(-u)_1 = 0+(-u)_2 \text{ par l'axiome 5} \\
\Rightarrow\ & (-u)_1 = (-u)_2 \text{ par l'axiome 4}
\end{aligned}$$
l'inverse additif d'un vecteur est donc unique.

3. Prenons $u \in V$. Par l'axiome 5, il existe un vecteur $[-(0 \otimes u)]$ tel que

$$(0 \times u) \oplus [-(0 \otimes u)] = 0$$

Ce qui nous donne :

$$\begin{aligned} & 0 \otimes u = (0+0) \otimes u = (0 \otimes u) \oplus (0 \otimes u) \text{ axiome 10} \\ \Rightarrow \quad & (0 \otimes u) \oplus [-(0 \otimes u)] = (0 \otimes u) \oplus (0 \otimes u) \oplus [-(0 \otimes u)] \\ \Rightarrow \quad & 0 = (0 \otimes u) \text{ axiome 4} \end{aligned}$$

4. Prenons $k \in \mathbb{R}$, alors on a :

$$\begin{aligned} & k \otimes 0 = k \otimes (0 \oplus 0) \text{ axiome 4} \\ \Rightarrow \quad & k \otimes 0 = (k \otimes 0) \oplus (k \otimes 0) \text{ axiome 9} \\ \Rightarrow \quad & (k \otimes 0) \oplus -(k \otimes 0) = (k \otimes 0) \oplus (k \otimes 0) \oplus -(k \otimes u) \\ \Rightarrow \quad & 0 = k \otimes 0 \text{ axiome 4} \end{aligned}$$

5. Prenons $u \in V$, alors on a :

$$\begin{aligned} 0 &= 0 \otimes u \\ &= (1-1) \otimes u \\ &= (1 \otimes u) \oplus ((-1) \otimes u) \text{ axiome 10} \\ &= u \oplus [(-1) \otimes u] \text{ axiome 8} \end{aligned}$$

Par l'unicité de l'inverse additif, nous avons donc que $(-1) \otimes u = -u$.

6. Supposons que $k \otimes u = 0$ et $k \neq 0$, alors on a :

$$\begin{aligned} k \otimes u = 0 \quad & \Rightarrow \quad k^{-1} \otimes (k \otimes u) = k^{-1} \otimes 0 \text{ car } k \neq 0 \\ & \Rightarrow \quad (k^{-1}k) \otimes u = 0 \text{ axiome 7 et partie 3 du théorème} \\ & \Rightarrow \quad 1 \otimes u = 0 \\ & \Rightarrow \quad u = 0 \text{ axiome 8} \end{aligned}$$

Donc si $k \otimes u = 0$ alors $k = 0$ ou $u = 0$.

$\square$

1.2 Les sous-espaces vectoriels

La plupart des espaces vectoriels qui apparaissent en pratique sont en fait rien d'autre qu'un sous-ensemble d'un espace vectoriel qui nous est déjà familier. Dans ce cas, il n'est pas nécessaire de vérifier chacune des 10 propriétés (ce qui serait particulièrement long), car la plupart des propriétés pourront être obtenu comme simple conséquence qu'elles sont satisfaites dans le plus grand espace vectoriel. C'est ce que nous allons voir dans cette section.

Definition 1.2.1. W est un sous-espace vectoriel d'un espace vectoriel V si W est un sous ensemble de V et W est un espace vectoriel.

Exemple 1.2.1. Nous avons vu dans la section précédente que l'ensemble de toutes les fonctions réelles $\mathcal{F}(\mathbb{R})$, et l'ensemble de tous les polynômes de degré au plus 2 dénoté $\mathbb{P}_2(\mathbb{R})$ sont tous deux des espaces vectoriels. Par contre, il est facile de voir que tout les polynômes sont en fait des fonctions, et que l'addition et multiplication par un scalaire sont définie de la même façon. On peut donc affirmer que $\mathbb{P}_2(\mathbb{R})$ est un sous espace vectoriel de $\mathcal{F}(\mathbb{R})$.

Théorème 1.2.1. Si W est un sous-ensemble non-vide d'un espace vectoriel V, alors W est un sous-espace vectoriel de V si et seulement si les deux propriétés suivantes sont satisfaites :

1. $u \oplus v \in W$ pour tout $u, v \in W$
2. $\lambda \otimes u \in W$ pour tout $\lambda \in \mathbb{R}$ et $u \in W$

Démonstration. La démonstration de ce théorème n'est pas très difficile et vous est laissé en exercice. $\square$

Exemple 1.2.2. Démontrez que l'ensemble

$$W = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$$

Muni des opérations d'addition et de multiplication par un scalaire habituel pour les matrices est un sous-espace vectoriel de $V = M_{2\times 2}(\mathbb{R})$. Comme W est bien un sous ensemble de V, il suffit de vérifier les 2 propriétés d'un sous-espace vectoriel :

1. $\begin{pmatrix} a & b \\ b & a \end{pmatrix} \oplus \begin{pmatrix} c & d \\ d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ b+d & a+c \end{pmatrix} \in W$
2. $\lambda \otimes \begin{pmatrix} a & b \\ b & a \end{pmatrix} = \begin{pmatrix} \lambda a & \lambda b \\ \lambda b & \lambda a \end{pmatrix} \in W$

Comme les deux propriétés sont satisfaites, il s'agit bien d'un sous-espace vectoriel.

Remarque : Pour simplifier la notation, à partir de maintenant, le symbole $\oplus$ sera remplacer par le symbole habituel +, et le symbole $\otimes$ sera tout simplement omis. Vous devez cependant vous assurer de bien comprendre la distinction entre les opérations sur les scalaires, et les opérations sur les vecteurs.

Definition 1.2.2. Si V est un espace vectoriel et $S = \{u_1, u_2, ..., u_n\}$ est un sous-ensemble de V, alors on définie l'espace généré par S (aussi appellé span) comme étant :

$$span\{u_1, u_2, ..., u_n\} = \left\{ \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n : \alpha_1, \alpha_2, ..., \alpha_n \in \mathbb{R} \right\}$$

Exemple 1.2.3. Est-ce que le vecteur $\begin{pmatrix} 1 \\ 14 \end{pmatrix}$ appartient à $W = span\left\{ \begin{pmatrix} 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ -3 \end{pmatrix} \right\}$? Pour le savoir, on doit résoudre le système d'équation linéaire suivant :

$$\lambda_1 \begin{pmatrix} 5 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \lambda_3 \begin{pmatrix} 4 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 14 \end{pmatrix} \qquad \Longrightarrow \qquad \left\{ \begin{array}{l} 5\lambda_1 + 3\lambda_2 + 4\lambda_3 = 1 \\ 2\lambda_1 + 1\lambda_2 - 3\lambda_3 = 14 \end{array} \right.$$

En applicant la méthode de Gauss, on obtient donc :

$$\left(\begin{array}{ccc|c} 5 & 3 & 4 & 1 \\ 2 & 1 & -3 & 14 \end{array} \right) \sim^{2L_1 - 5L_2 \to L_2} \left(\begin{array}{ccc|c} 5 & 3 & 4 & 1 \\ 0 & 1 & 23 & -68 \end{array} \right)$$

On remarque donc que le système d'équations linéaires possède une infinité de solutions, en particulier il en possède au moins une. On peut donc conclure que le vecteur appartient bien à W.

Exemple 1.2.4. Dans l'espace vectoriel $\mathbb{P}_2(\mathbb{R})$, est-ce que le vecteur $3x^2 + 1$ appartient à $W = span\{2x^2 + 2x + 1, 4x^2 - x + 5\}$? Pour ce faire, on doit résoudre le système d'équations linéaires suivant :

$$\lambda_1(2x^2 + 2x + 1) + \lambda_2(4x^2 - x + 5) = 3x^2 + 1 \qquad \Longrightarrow \qquad \left\{ \begin{array}{l} 2\lambda_1 + 4\lambda_2 = 3 \\ 2\lambda_1 - 1\lambda_2 = 0 \\ 1\lambda_1 + 5\lambda_2 = 1 \end{array} \right.$$

En applicant la méthode de Gauss, on obtient donc :

$$\left(\begin{array}{cc|c} 2 & 4 & 3 \\ 2 & -1 & 0 \\ 1 & 5 & 1 \end{array}\right) \sim_{\substack{L_1-L_2\to L_2 \\ L_1-2L_2\to L_2}} \left(\begin{array}{cc|c} 2 & 4 & 3 \\ 0 & 5 & 3 \\ 0 & -6 & 1 \end{array}\right) \sim^{6L_2+5L_3\to L_3} \left(\begin{array}{cc|c} 2 & 4 & 3 \\ 0 & 5 & 3 \\ 0 & 0 & 23 \end{array}\right)$$

On remarque que ce système d'équations linéaires a donc aucune solution. On peut donc conclure que le vecteur n'appartient pas à W.

Nous somme maintenant amener à énoncer un théorème simple, mais particulièrement important pour le reste du texte. Le théorème est en fait l'une des raisons pourquoi nous somme intéressé par le span d'un ensemble de vecteurs.

Théorème 1.2.2. Si V est un espace vectoriel et $S = \{u_1, u_2, ..., u_n\}$ est un sous-ensemble de V, alors $W = span\{u_1, u_2, ..., u_n\}$ est un sous-espace vectoriel de V.

Démonstration. Supposons que V est un espace vectoriel, et prenons $S = \{u_1, u_2, ..., u_n\}$. Alors on a :

1. $(\alpha_1 u_1 + \alpha_2 u_2 + ...\alpha_n u_n) + (\beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n) = (\alpha_1 + \beta_1)u_1 + (\alpha_2 + \beta_2)u_2 + ... + (\alpha_n + \beta_n)u_n \in span(S)$
2. $\lambda(\alpha_1 u_1 + \alpha_2 u_2 + ...\alpha_n u_n) = \lambda\alpha_1 u_1 + \lambda\alpha_2 u_2 + ...\lambda\alpha_n u_n \in span(S)$

Par le théorème, il s'agit donc bien d'un sous-espace vectoriel de V. □

1.3 Les bases et la dimension

Les espaces vectoriels contiennent règle générale beaucoup de vecteurs. Il est donc difficile de les étudier directement. Dans cette section, nous allons voir qu'il est toujours possible de définir une base sur un espace vectoriel. Une base n'est rien d'autre qu'un ensemble de vecteur, le plus petit possible, nous permettant d'étudier l'espace vectoriel au complet. Il nous sera donc possible d'étudier un espace vectoriel à partir de seulement quelques vecteurs.

Definition 1.3.1 (Ensemble générateur)**.** Si V est un espace vectoriel et $S = \{u_1, u_2, ..., u_n\}$ sont des vecteurs de V, alors on dit que les vecteurs de S sont générateur de V si $span(S) = V$. C'est à dire, si pour tout $v \in V$, l'équation

$$\alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n = v$$

a au moins une solution. De plus, on dit que V est un espace vectoriel de dimension finie, s'il existe un ensemble S ayant un nombre fini d'élément et qui est générateur de V. Autrement, on dit que V est de dimension infinie.

Remarquez que dans la définition précédente, lorsqu'on travaille avec des espaces vectoriels réels, on va supposer que les α_i sont des nombres réels. Lorsqu'on travaille avec des espaces vectoriels complexes, on va supposer que les α_i sont des nombres complexes. Notez aussi que bien qu'occasionnellement nous feront référence aux espaces vectoriels de dimension infinie, c'est dernier sont considérablement plus difficile à étudier. C'est pour cela que dans presque tous les cas durant le reste du cours, nous nous concentrerons uniquement sur les espaces vectoriels de dimension finie.

Exemple 1.3.1. On veut montrer que l'ensemble

$$\mathcal{S} = \left\{ \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix} \right\}$$

est générateur de $\mathbb{R}^3$. Pour ce faire, on doit vérifier que n'importe quel vecteur de $\mathbb{R}^3$ peut s'écrire comme combinaison linéaire des vecteurs de l'ensemble. C'est à dire qu'on doit vérifier que le système d'équations linéaires suivant possède toujours au moins une solution :

$$\begin{cases} 2\lambda_1 + 0\lambda_2 + 1\lambda_3 = a \\ 5\lambda_1 + 3\lambda_2 - 2\lambda_3 = b \\ 1\lambda_1 - 1\lambda_2 + 4\lambda_3 = c \end{cases}$$

En applicant la méthode de Gauss, on obtient :

$$\left(\begin{array}{ccc|c} 2 & 0 & 1 & a \\ 5 & 3 & -2 & b \\ 1 & -1 & 4 & c \end{array}\right) \sim^{5L_1-2L_2\to L_2}_{L_1-2L_3\to L_3} \left(\begin{array}{ccc|c} 2 & 0 & 1 & a \\ 0 & -6 & 9 & 5a-2b \\ 0 & 2 & -7 & a-2c \end{array}\right) \sim^{L_2+3L_3\to L_3} \left(\begin{array}{ccc|c} 2 & 0 & 1 & a \\ 0 & -6 & 9 & 5a-2b \\ 0 & 0 & -12 & 8a-2b-6c \end{array}\right)$$

Comme ce système possède toujours exactement une solution, il en possède toujours au moins une. On peut donc conclure que l'ensemble est générateur de $\mathbb{R}^3$.

Definition 1.3.2 (Ensemble linéairement indépendant)**.** Si V est un espace vectoriel et $S = \{u_1, u_2, ..., u_n\}$ sont des vecteurs de V, alors on dit que les vecteurs de S sont linéairement indépendant si l'équation

$$\alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n = 0$$

a une unique solution. Cette solution sera alors $\alpha_1 = \alpha_2 = ... = \alpha_n = 0$.

Remarquez que comme pour les ensembles générateurs, dans la définition précédente lorsqu'on travaille avec des espaces vectoriels réels, on va supposer que les α_i sont des nombres réels. Lorsqu'on travaille avec des espaces vectoriels complexes, on va supposer que les α_i sont des nombres complexes.

Théorème 1.3.1. Supposons que V est un espace vectoriel de dimension finie, et $\mathcal{B}$ est une famille de vecteur de V qui ne sont pas linéairement indépendant, alors l'un des vecteurs de $\mathcal{B}$ peut s'écrire comme combinaison linéaire des autres vecteurs de $\mathcal{B}$.

Démonstration. Supposons que $\mathcal{B} = \{u_1, u_2, ..., u_n\}$, et supposons que $\mathcal{B}$ n'est pas linéairement indépendant. Alors, il existe $\lambda_1, \lambda_2, ..., \lambda_n \in \mathbb{R}$ qui ne sont pas tous 0 tel que :

$$\lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_n u_n = 0$$

Comme les λ_i ne sont pas tous zéro, on va supposer sans perte de généralité que $\lambda_1 \neq 0$. On a donc :

$$u_1 = \frac{-\lambda_2}{\lambda_1} u_2 + \frac{-\lambda_3}{\lambda} u_3 + ... + \frac{-\lambda_n}{\lambda_1} u_n$$

Donc u_1 peut s'écrire comme combinaison linéaire des autre vecteurs de $\mathcal{B}$. □

Exemple 1.3.2. Dans l'espace vectoriel $\mathbb{R}^2$, vérifier si les vecteurs $\left\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}5\\6\end{pmatrix}\right\}$ sont linéairement indépendant. S'ils ne sont pas linéairement indépendant, on veut écrire l'un des vecteurs comme combinaison linéaire des autres. Pour ce faire, nous allons commencer par résoudre le système d'équations linéaires suivant :

$$\alpha_1 \begin{pmatrix}1\\2\end{pmatrix} + \alpha_2 \begin{pmatrix}3\\4\end{pmatrix} + \alpha_3 \begin{pmatrix}5\\6\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

En applicant la méthode de Gauss, on obtient donc :

$$\left(\begin{array}{ccc|c} 1 & 3 & 5 & 0 \\ 2 & 4 & 6 & 0 \end{array}\right) \sim^{2L_1-L_2\to L_2} \left(\begin{array}{ccc|c} 1 & 3 & 5 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right)$$

On remarque donc facilement que le système d'équations linéaires possède une infinités de solutions. Les vecteurs ne sont donc pas linéairement indépendant. Pour écrire l'un des vecteurs comme combinaison linéaire des autres, on commence par trouver l'ensemble des solutions du système, ce qui nous est donné par :

$$\begin{cases} \alpha_1 = s \\ \alpha_2 = -2s \\ \alpha_3 = s \end{cases}$$

On choisit ensuite l'une de ces solutions qui est différente de 0. On peut donc prendre par exemple $s = 1$. En remplaçant dans l'équation 1.3.2 on obtient donc :

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

Ce qui nous permet ensuite d'écrire :

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

Nous avons donc écrit l'un des vecteurs de l'ensemble comme combinaison linéaire des autres.

Definition 1.3.3 (Une base)**.** Si V est un espace vectoriel de dimension finie et $\mathcal{B}$ est un sous-ensemble de V, alors on dit que $\mathcal{B}$ est une base s'il s'agit d'un ensemble linéairement indépendant et générateur de V.

Il est aussi possible de parler de base dans le cas d'espaces vectoriels de dimension infinie. Par contre, dans ce cas certaine difficulté se pose, et il est nécessaire de modifier légèrement notre définition. Nous ne traiterons pas de ce cas dans ce cours.

L'importance des bases en algèbre linéaire vient du fait qu'il nous est possible de représenter de façon unique n'importe quel vecteur de l'espace vectoriel à partir d'élément de notre base, ce qui signifie qu'il nous est possible d'étudier un espace vectoriel à partir uniquement (ou presque) des éléments de notre base. Le théorème suivant est un premier pas dans cette direction et démontre l'unicité de la représentation.

Théorème 1.3.2. **(Unicité de la représentation)** Si V est un espace vectoriel de dimension finie et $\mathcal{B} = \{u_1, u_2, u_3, ..., u_n\}$ est une base de V, alors l'écriture de tout vecteur $x \in V$ dans la base $\mathcal{B}$ est unique. C'est à dire que si :

$$x = \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n = \beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n$$

alors $\alpha_i = \beta_i$ pour tout i.

Démonstration. Supposons que

$$x = \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n = \beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n$$

alors on a :

$$\begin{aligned} 0 &= x - x \\ &= \Big(\alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n\Big) - \Big(\beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n\Big) \\ &= (\alpha_1 - \beta_1)u_1 + (\alpha_2 - \beta_2)u_2 + ... + (\alpha_n - \beta_n)u_n \end{aligned}$$

par indépendance linéaire, on a donc que $\alpha_i - \beta_i = 0$ pour tout i, ou de façon équivalente :

$$\alpha_i = \beta_i, \qquad \forall i$$

ce qui démontre que l'écriture dans la base $\mathcal{B}$ est unique. □

Definition 1.3.4. Si V est un espace vectoriel, et $\mathcal{B} = \{u_1, u_2, ..., u_n\}$ est une base de V, alors on écrira :

$$u = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ ... \\ \alpha_n \end{pmatrix}_{\mathcal{B}}$$

Pour signifier que $u = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + ... + \alpha_n u_n$. Lorsque aucun indice n'est écrit, alors on supposera qu'il s'agit de la base habituelle.

Exemple 1.3.3. On veut vérifier que l'ensemble

$$\mathcal{B} = \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

forme une base de $\mathbb{R}^3$ et nous voulons ensuite écrire le vecteur

$$u = \begin{pmatrix} 8 \\ 13 \\ 1 \end{pmatrix}$$

dans cette base. Pour ce faire, commençons par montrer que l'ensemble $\mathcal{B}$ est linéairement indépendant :

$$\lambda_1 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

ce qui nous donne le système d'équations linéaires suivant :

$$\begin{cases} 2\lambda_1 + 3\lambda_2 + \lambda_3 = 0 \\ \lambda_1 + 2\lambda_3 = 0 \\ 4\lambda_2 + \lambda_3 = 0 \end{cases}$$

En appliquant la méthode de Gauss, nous obtenons donc :

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 0 \\ 1 & 0 & 2 & 0 \\ 0 & 4 & 1 & 0 \end{array}\right) \underset{\sim}{L_1 - 2L_2 \to L_2} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 4 & 1 & 0 \end{array}\right) \underset{\sim}{4L_2 - 3L_3 \to L_3} \left(\begin{array}{ccc|c} 2 & 3 & 1 & 0 \\ 0 & 3 & -3 & 0 \\ 0 & 0 & -15 & 0 \end{array}\right)$$

Comme ce système a une unique solution, l'ensemble est linéairement indépendant. On veut maintenant montrer qu'il est générateur. Pour ce faire, on veut résoudre le système d'équations linéaires suivant :

$$\lambda_1 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

ce qui nous donne le système d'équations linéaires suivant :

$$\begin{cases} 2\lambda_1 + 3\lambda_2 + \lambda_3 = x \\ \lambda_1 + 2\lambda_3 = y \\ 4\lambda_2 + \lambda_3 = z \end{cases}$$

En appliquant la méthode de Gauss, nous obtenons donc :

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & x \\ 1 & 0 & 2 & y \\ 0 & 4 & 1 & z \end{array}\right) \underset{\sim}{L_1 - 2L_2 \to L_2} \left(\begin{array}{ccc|c} 2 & 3 & 1 & x \\ 0 & 3 & -3 & x - 2y \\ 0 & 4 & 1 & z \end{array}\right) \underset{\sim}{4L_2 - 3L_3 \to L_3} \left(\begin{array}{ccc|c} 2 & 3 & 1 & x \\ 0 & 3 & -3 & x - 2y \\ 0 & 0 & -15 & 4x - 8y - 3z \end{array}\right)$$

Comme ce système a encore une fois une unique solution, on peut donc conclure que les vecteurs sont générateurs de $\mathbb{R}^3$. Il s'agit donc bien d'une base. Finalement pour écrire le vecteur u dans la base $\mathcal{B}$, on remplace les valeurs de x, y, z par $8, 13, 1$ et on complète la résolution du système d'équation précédent.

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 8 \\ 0 & 3 & -3 & 8-2(13) \\ 0 & 0 & -15 & 4(8)-8(13)-3(1) \end{array}\right) = \left(\begin{array}{ccc|c} 2 & 3 & 1 & 8 \\ 0 & 3 & -3 & -18 \\ 0 & 0 & -15 & -75 \end{array}\right)$$

On obtient donc que

$$\lambda_3 = \frac{-75}{-15} = 5$$
$$\lambda_2 = \frac{-18+3(5)}{3} = \frac{-3}{3} = -1$$
$$\lambda_1 = \frac{8-5-3(-1)}{2} = \frac{6}{2} = 3$$

Le vecteur u peut donc s'écrire sous la forme :

$$u = \begin{pmatrix} 3 \\ -1 \\ 5 \end{pmatrix}_{\mathcal{B}}$$

Théorème 1.3.3. Si V est un espace vectoriel, et $\mathcal{B} = \{u_1, u_2, ..., u_n\}$, alors

1. si $\mathcal{B}$ est un ensemble générateur de V, alors tout ensemble $\mathcal{B}'$ qui contient $\mathcal{B}$ est aussi générateur de V.
2. si $\mathcal{B}$ est un ensemble linéairement indépendant de V, alors tout ensemble $\mathcal{B}'$ contenu dans $\mathcal{B}$ est aussi linéairement indépendant.

Démonstration.

1. Supposons que $\mathcal{B}' = \{u_1, u_2, ..., u_n, w_1, w_2, ..., w_m\}$ et prenons $x \in V$. Alors comme $\mathcal{B}$ est générateur, on peut écrire
$$x = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_n u_n$$
maintenant, on remarque que x peut aussi s'écrire comme combinaison linéaire d'éléments de $\mathcal{B}'$. En effet, remarquons que :
$$x = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_n u_n + 0w_1 + 0w_2 + ... + 0w_m$$
l'ensemble $\mathcal{B}'$ est donc générateur de V.
2. Prenons $\mathcal{B} = \{u_1, u_2, ..., u_n, w_1, w_2, ..., w_m\}$ un ensemble linéairement indépendant, et $\mathcal{B}' = \{u_1, u_2, ..., u_n\}$. Supposons que $\mathcal{B}'$ n'est pas linéairement indépendant, c'est à dire qu'il existe des constante $\lambda_1, \lambda_2, ..., \lambda_n$ qui ne sont pas toute nulle tel que :
$$\lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_n u_n = 0$$
Dans ce cas, si on pose $\mu_1, \mu_2, ..., \mu_m = 0$, on a donc :
$$\lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_n u_n + \mu_1 w_1 + \mu_2 w_2 + ... + \mu_m w_m = 0$$
et pour laquelle au moins l'un des coefficients n'est pas nul, ce qui contredit l'indépendance linéaire de $\mathcal{B}$. On peut donc conclure que $\mathcal{B}'$ est un ensemble linéairement indépendant.

$\square$

Théorème 1.3.4. Tous les espaces vectoriels V de dimension finie admettent au moins une base. De plus, toutes les bases de V ont le même nombre d'éléments.

Démonstration. Commençons par démontrer l'existence d'une base. Si V est un espace vectoriel de dimension fini, alors il existe un ensemble $\mathcal{S} = \{u_1, u_2, ..., u_n\}$ qui est générateur de V, c'est à dire que $span(\mathcal{S}) = V$. Sans perte de généralité, nous pouvons supposer qu'aucun des u_i n'est égal au vecteur nul. Nous savons que $\mathcal{S}_1 = \{u_1\}$ est un ensemble linéairement indépendant. S'il est aussi générateur, on a terminé (on a trouvé une base). Dans le cas contraire, il existe un vecteur $u_* \in \mathcal{S}$, pour lequel l'ensemble $\{u_1, u_*\}$ est encore linéairement indépendant. Sans perte de généralité, on peut supposer que $u_* = u_2$. Si l'ensemble $\mathcal{S}_2 = \{u_1, u_2\}$ est générateur, alors on a terminé. Autrement, il existe un vecteur u_* dans $\mathcal{S}$ qui est linéairement indépendant avec $\{u_1, u_2\}$. Sans perte de généralité, on peut supposer que $u_* = u_3$. Si l'ensemble $\mathcal{S}_3 = \{u_1, u_2, u_3\}$ est générateur, alors on a terminé. Autrement, on continue de la même manière en prenant un autre élément u_*. Comme l'ensemble $\mathcal{S}$ est fini, le processus doit obligatoirement s'arrêter. On peut donc conclure qu'en dimension fini tous les espaces vectoriel admettent au moins une base.

Nous voulons maintenant montrer que toutes les bases de V ont le même nombre d'élément. Supposons que $\mathcal{B}_1 = \{u_1, u_2, ..., u_n\}$ et $\mathcal{B}_2 = \{w_1, w_2, ..., w_m\}$ sont des bases de V tel que $m > n$. Alors on a :

$$w_1 = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_n u_n$$

Comme w_1 ne peut pas être 0, au moins l'un des λ_i n'est pas 0. Sans perte de généralité, on va supposer que $\lambda_1 \neq 0$, ce qui nous donne :

$$u_1 = \frac{1}{\lambda_1} w_1 + \frac{-\lambda_2}{\lambda_1} u_2 + \frac{-\lambda_3}{\lambda_1} u_3 + ... + \frac{-\lambda_n}{\lambda_1} u_n$$

Alors l'ensemble $\{w_1, u_2, u_3, ..., u_n\}$ est aussi un ensemble générateur de V. On a donc :

$$w_2 = \lambda_1 w_1 + \lambda_2 u_2 + \lambda_3 u_3 + ... + \lambda_n u_n$$

De plus, au moins un des λ_i, $i > 1$ n'est pas 0. Sans perte de généralité, on va supposer que $\lambda_2 \neq 0$. On a donc :

$$u_2 = \frac{-\lambda_1}{\lambda_2} w_1 + \frac{1}{\lambda_2} w_2 - \frac{\lambda_3}{\lambda_1} u_3 - ... - \frac{\lambda_n}{\lambda_1} u_n$$

Alors l'ensemble $\{w_1, w_2, u_3, ..., u_n\}$ est aussi un ensemble générateur de V. En continuant de la même façon, on obtient que $\{w_1, w_2, ..., w_n\}$ est aussi un ensemble générateur de V. On obtient donc que w_m peut s'écrire comme combinaison linéaire des vecteurs $\{w_1, w_2, w_3, ..., w_n\}$. L'ensemble $\mathcal{B}_2$ n'est donc pas linéairement indépendant, et donc $\mathcal{B}_2$ n'est donc pas une base de V, ce qui est une contradiction. Toutes les bases de V doivent donc avoir le même nombre d'élément. □

Definition 1.3.5. Si V est un espace vectoriel, alors la dimension de V est le nombre d'élément dans une base de V. On dénote la dimension de V par $dim(V)$.

Exemple 1.3.4. Voici les bases habituelles de certain espaces vectoriels (réels) courant, ainsi que leur dimension :

1. L'espace vectoriel $\mathbb{R}^n$ a comme base les vecteurs :

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, ..., \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}$$

   et sa dimension est n

2. L'espace vectoriel $\mathbb{P}_n(\mathbb{R})$ a comme base les vecteurs :

$$\mathcal{B} = \{1, x, x^2, x^3, ..., x^n\}$$

   et a comme dimension $n + 1$.
3. L'espace vectoriel $M_{n\times n}(\mathbb{R})$ a comme base l'ensemble :

$$\mathcal{B} = \{A_{ij}\}_{i,j=1}^{n}$$

   Où les matrices A_{ij} sont des matrices de dimension $n \times n$ ayant des zéros partout sauf un 1 à la i-eme colonne et la j-eme ligne. Sa dimension est n^2.

Théorème 1.3.5. Si V est un espace vectoriel de dimension finie :

1. Si $\mathcal{S}$ est un sous-ensemble linéairement indépendant de V, alors on peut ajouter des vecteur à $\mathcal{S}$ de sorte que l'ensemble devienne une base de V.
2. Si $\mathcal{S}$ est un sous-ensemble générateur de V, alors on peut enlever des vecteurs à $\mathcal{S}$ de sorte que l'ensemble devienne une base de V.
3. Si $dim(V) = n$, alors tout ensemble contenant plus de n vecteurs n'est pas linéairement indépendant.
4. Si $dim(V) = n$, alors tout ensemble contenant moins de n vecteurs n'est pas générateur de V.

Démonstration.

1. Premièrement, rappelons nous que comme V est par hypothèse un espace vectoriel de dimension finie, alors il existe un ensemble fini $E \subseteq V$ qui est générateur de V. Supposons que $\mathcal{S}$ est un ensemble linéairement indépendant. Si $\mathcal{S}$ est générateur, alors $\mathcal{S}$ est déjà une base et il n'y a rien à faire. Autrement, si $\mathcal{S}$ n'est pas générateur, alors $span(\mathcal{S}) \neq span(E) = V$. Il existe donc un vecteur $u \in E \backslash span(\mathcal{S})$ qui est différent de 0. L'ensemble $\mathcal{S} \cup \{u\}$ est alors linéairement indépendant, car autrement u serait dans $span(\mathcal{S})$. Si $\mathcal{S} \cup \{u\}$ est générateur, alors on a obtenu une base. Autrement, on répète la même étape tant et aussi longtemps que l'ensemble obtenu n'est pas générateur. Le processus doit obligatoirement s'arrêter éventuellement car l'ensemble E est par hypothèse générateur de V.
2. Supposons que $\mathcal{S}$ est un ensemble générateur de V qui n'est pas linéairement indépendant. Alors l'un des vecteurs de $\mathcal{S}$ peut s'écrire comme combinaison linéaire des autres vecteurs. Il s'agit donc de retirer ce vecteur de $\mathcal{S}$ ce qui ne changera pas le span étant donné qu'il était déjà possible de l'écrire à partir des autres vecteurs. Si l'ensemble ainsi obtenu est linéairement indépendant, on a donc obtenu un base. Autrement on répète la même étape jusqu'à ce qu'on obtienne un ensemble linéairement indépendant. Cette méthode doit éventuellement s'arrêter car dans le pire des cas il ne nous restera plus qu'un seul vecteur et ce dernier sera donc linéairement indépendant.
3. Supposons que $dim(V) = n$ et que $\mathcal{S}$ est un ensemble linéairement indépendant contenant m vecteurs avec $m > n$. Par la première partie du théorème, nous pouvons donc ajouter des vecteurs à l'ensemble $\mathcal{S}$ pour obtenir un nouvel ensemble $\mathcal{S}'$ contenant k vecteurs, qui formera une base. Notez que $k \geq m > n$. On a donc obtenu une base de V qui contient k élément. Par contre comme toute les bases doivent avoir le même nombre d'élément, ceci est une contradiction. L'ensemble $\mathcal{S}$ ne pouvait donc pas être linéairement indépendant.
4. Cette partie est semblable à la précédente et est laissé en exercice.

□

Exemple 1.3.5. Considérez l'ensemble de vecteur suivant :

$$S = \left\{ \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} \right\}$$

Cet ensemble est générateur de $\mathbb{R}^3$. Par la troisième partie du théorème, on peut cependant voir que cet ensemble ne peut pas être linéairement indépendant car $\mathbb{R}^3$ est un espace vectoriel de dimension 3. Nous voulons donc retirer des vecteurs à cet ensemble jusqu'à ce qu'on obtienne un ensemble qui sera linéairement indépendant. Nous allons maintenant appliquer la méthode de Gauss en plaçant ces vecteurs à la vertical et en ignorant les constantes qui devrait se trouver à droite.

$$\begin{pmatrix} 3 & 4 & 2 & 5 \\ 2 & 1 & 3 & 2 \\ 1 & 3 & -1 & 0 \end{pmatrix} \sim_{\substack{2L_1-3L_2\to L_2 \\ 3L_3-L_1\to L_3}} \begin{pmatrix} 3 & 4 & 2 & 5 \\ 0 & 5 & -5 & 4 \\ 0 & 5 & -5 & -5 \end{pmatrix} \sim^{L_2-L_3\to L_3} \begin{pmatrix} 3 & 4 & 2 & 5 \\ 0 & 5 & -5 & 4 \\ 0 & 0 & 0 & 9 \end{pmatrix}$$

La forme échelon de la matrice contient donc des pivots dans les colonnes 1, 2 et 4. On remarque donc que si la 3e colonne aurait été omise durant tout le calcul, on aurait obtenu un système d'équation qui aurait eu exactement une solution (et donc une base...). Si on enlève le 3e vecteurs de S, on obtient donc une base de $\mathbb{R}^3$. La base est donc :

$$S' = \left\{ \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} \right\}$$

Exemple 1.3.6. Considérons l'ensemble suivant :

$$S = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \right\}$$

cet ensemble est clairement linéairement indépendant dans $\mathbb{R}^3$, mais ne peux pas être générateur de $\mathbb{R}^3$ car il contient moins de 3 vecteurs. Le théorème nous affirme cependant qu'on peut ajouter des vecteurs à S de sorte qu'on obtienne une base. La question est donc comment faire pour ajouter des vecteurs à S pour que l'ensemble devienne générateur. Pour ce faire, remarquons que nous connaissons déjà la base suivante de $\mathbb{R}^3$:

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

en ajoutant ces vecteurs à S, on obtient donc un ensemble qui est générateur, mais qui n'est plus linéairement indépendant. On peut donc appliquer la méthode de l'exemple précédent pour obtenir une base de $\mathbb{R}^3$ qui contient tous les vecteurs de l'ensemble S :

$$\begin{pmatrix} 1 & 4 & 1 & 0 & 0 \\ 2 & 5 & 0 & 1 & 0 \\ 3 & 6 & 0 & 0 & 1 \end{pmatrix} \sim_{\substack{2L_1-L_2\to L_2 \\ 3L_1-L_3\to L_3}} \begin{pmatrix} 1 & 4 & 1 & 0 & 0 \\ 0 & 3 & 2 & -1 & 0 \\ 0 & 6 & 3 & 0 & -1 \end{pmatrix} \sim^{2L_2-L_3\to L_3} \begin{pmatrix} 1 & 4 & 1 & 0 & 0 \\ 0 & 3 & 2 & -1 & 0 \\ 0 & 0 & 1 & -2 & 1 \end{pmatrix}$$

Les colonnes contenant les pivots sont donc les colonnes 1,2 et 3. On obtient donc la base suivante :

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

qui contient tous les éléments de S.

Comme application de ce que nous venons de voir nous allons chercher une base du span d'un ensemble de vecteurs. Il existe deux manières de procéder. La première est celle qui a été utilisé dans le premier cours d'algèbre linéaire. C'est à dire en plaçant les vecteurs horizontalement dans une matrice et en applicant la méthode de Gauss. Les vecteurs (lignes) non nul ainsi obtenu forme une base du span. La seconde méthode consiste à appliquer la théorie que nous avons vu dans l'exemple précédent. Les vecteurs de l'ensemble par définition sont générateur du span. La question revient donc à savoir comment retirer des vecteurs de sorte que l'ensemble devienne linéairement indépendant.

Exemple 1.3.7. Si $V = \mathbb{R}^3$, trouvez une base de

$$span\left\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix},\begin{pmatrix}7\\8\\9\end{pmatrix}\right\}$$

Première méthode : On place les vecteurs sous forme de ligne et on applique la méthode de Gauss :

$$\begin{pmatrix}1 & 2 & 3\\4 & 5 & 6\\7 & 8 & 9\end{pmatrix} \sim_{\substack{4L_1-L_2\to L_2\\7L_1-L_3\to L_3}} \begin{pmatrix}1 & 2 & 3\\0 & 3 & 6\\0 & 6 & 12\end{pmatrix} \sim^{2L_2-L_3\to L_3} \begin{pmatrix}1 & 2 & 3\\0 & 3 & 6\\0 & 0 & 0\end{pmatrix}$$

Les lignes non nulle forme donc une base. On a donc la base suivante :

$$\mathcal{B} = \left\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}0\\3\\6\end{pmatrix}\right\}$$

Deuxième méthode : On place les vecteurs sous forme de colonne et on applique la méthode de Gauss :

$$\begin{pmatrix}1 & 4 & 7\\2 & 5 & 8\\3 & 6 & 9\end{pmatrix} \sim_{\substack{2L_1-L_2\to L_2\\3L_1-L_3\to L_3}} \begin{pmatrix}1 & 4 & 7\\0 & 3 & 6\\0 & 6 & 12\end{pmatrix} \sim^{2L_2-L_3\to L_3} \begin{pmatrix}1 & 4 & 7\\0 & 3 & 6\\0 & 0 & 0\end{pmatrix}$$

Comme les pivots ce trouve dans les colonnes 1 et 2, on obtient donc la base suivante :

$$\mathcal{B}' = \left\{\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix}\right\}$$

Théorème 1.3.6. Si V est un espace vectoriel de dimension finie et W est un sous-espace vectoriel de V, alors $dim(W) \leq dim(V)$.

Démonstration. Supposons que $\mathcal{B}$ est une base de W. Alors, tous les éléments de $\mathcal{B}$ sont aussi des éléments de V, et $\mathcal{B}$ est un ensemble linéairement indépendant dans V. Donc $\mathcal{B}$ peut être complété en une base de V. Comme toute les bases d'un espace vectoriel ont le même nombre d'élément, on peut donc conclure que :

$$\dim(W) \leq \dim(V)$$

□

Exercice 1.3.1. Utilisez le théorème précédent pour démontrer que l'espace vectoriel $\mathcal{F}(\mathbb{R})$ est de dimension infinie. Indice : comparer cet espace avec l'espace vectoriel $\mathbb{P}_n(\mathbb{R})$.

1.4 Le théorème de la matrice inverse

Nous avons remarquez dans la section précédente que l'étude des bases d'un espace vectoriel se ramène à étudier des systèmes d'équations linéaires, et en particulier à vérifier l'unicité d'une solution. Il semble donc approprié de réviser certaine propriété des systèmes d'équations linéaires, et en particulier de caractériser les systèmes d'équations linéaires ayant exactement une solution. Nous allons donc commencer par rappeler un résultat simple, mais important.

Théorème 1.4.1. Si $Ax = b$ est un système d'équations linéaires, alors l'un des trois cas ci-dessous est vrai :

1. Le système n'a aucune solution.
2. Le système a exactement une solution.
3. Le système a une infinité de solutions.

Démonstration. Pour démontrer le théorème, il suffit de démontrer que si un système d'équations linéaires possède au moins deux solutions, alors il doit obligatoirement en avoir une infinité. Pour ce faire, supposons que x_1 et x_2 sont des solutions du système d'équations linéaires. C'est à dire que $Ax_1 = b$ et $Ax_2 = b$. Si $\lambda \in \mathbb{R}$, alors on a :

$$A(\lambda x_1 + (1-\lambda)x_2) = \lambda Ax_1 + (1-\lambda)Ax_2 = \lambda b + (1-\lambda)b = b$$

Donc $\lambda x_1 + (1-\lambda)x_2$ est une solution du système d'équations linéaire pour chaque λ. On a donc une infinité de solution. □

Nous somme maintenant prêt à caractériser les systèmes d'équations linéaires ayant exactement une solution. Ceci est fait à l'aide du théorème de la matrice inverse. Il s'agit d'un théorème absolument fondamental en algèbre linéaire, et nous y ajouterons certain élément dans les prochains chapitres.

Théorème 1.4.2. **(Théorème de la matrice inverse)** Si A est une matrice de dimension $n \times n$, alors les énoncés suivant sont équivalents :

1. A est une matrice inversible
2. $det(A) \neq 0$
3. Le système d'équations linéaires $Ax = 0$ a une unique solution
4. Le système d'équations linéaires $Ax = b$ a une unique solution pour tout $b \in \mathbb{R}^n$.
5. Les colonnes de A forment une base de $\mathbb{R}^n$
6. Les lignes de A forment une base de $\mathbb{R}^n$.

Nous allons maintenant démontrer le théorème, mais en ignorant la seconde partie qui est malheureusement un peu trop technique pour notre cours.

Démonstration.
(1) ⇒ (3) Premièrement, on remarque qu'un système d'équations linéaires homogène doit obligatoirement avoir $x = 0$ comme solution. Maintenant, si x est une solution de $Ax = 0$, alors $AA^{-1}x = A^{-1}0$, ce qui nous donne finalement $x = 0$. Donc il y a exactement une solution.

(3) ⇒ (4) On va commencer par montrer qu'il doit y avoir au moins une solution. Pour ce faire, supposons au contraire que $Ax = b$ ne possède aucune solution. Si $\{e_1, e_2, ..., e_n\}$ est une base de $\mathbb{R}^n$, alors $\{Ae_1, Ae_2, ..., Ae_n\}$ ne peux pas être une base, car cet ensemble n'est pas générateur (Son span ne contient pas b). On peut donc conclure que les vecteurs de cet ensemble ne peuvent pas non plus être linéairement indépendant. Il existe donc des scalaires $\lambda_1, \lambda_2, ..., \lambda_n$ qui ne sont pas tous nul et tel que $\lambda_1 Ae_1 + \lambda_2 Ae_2 + ... + \lambda_n Ae_n = 0$. En factorisant le A, on obtient donc :

$$A(\lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n) = 0$$

Par hypothèse, ce système possède exactement une solution (la solution 0), on obtient donc :

$$\lambda_1 e_1 + \lambda_2 e_2 + ... + \lambda_n e_n = 0$$

Comme les λ_i ne sont pas tous nul, on obtient donc que l'ensemble $\{e_1, e_2, ..., e_n\}$ n'est pas linéairement, ce qui est une contradiction. On peut donc conclure que le système d'équations linéaires $Ax = b$ doit avoir au moins une solution.

On veut maintenant montrer que le système d'équations linéaires ne peut pas avoir plus qu'une solution. Pour ce faire, supposons au contraire que x_1 et x_2 sont des solutions. C'est à dire, on suppose que $Ax_1 = b$ et $Ax_2 = b$. En soustrayant les deux équations, on obtient donc :

$$Ax_1 - Ax_2 = 0 \quad \Rightarrow \quad A(x_1 - x_2) = 0$$

Par l'hypothèse (3), on peut donc conclure que $x_1 - x_2 = 0$, c'est à dire $x_1 = x_2$. Le système ne peut donc pas avoir plus qu'une solution.

(4) ⇒ (1) Prenons $\{e_1, e_2, ..., e_n\}$ la base habituelle de $\mathbb{R}^n$. Par hypothèse, chacun des systèmes d'équations linéaires $Ax = e_i$ possède exactement une solution. On va dénoté ces solutions par x_i de sorte que $Ax_i = e_i$ pour tout i. Maintenant, on construit une matrice en plaçant les vecteurs x_i en colonne :

$$A^{-1} = [x_1|x_2|...|x_n]$$

Il est alors facile de vérifier que $AA^{-1} = I$, ce qui confirme qu'il s'agit bien de l'inverse de la matrice A. En particulier, la matrice A est inversible.

(3 ⇔ (5) La démonstration repose sur une réécriture du système d'équations linéaires $Ax = 0$. Si on écrit :

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \qquad x = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$$

Donc le système d'équation linéaire $Ax = 0$ devient :

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \qquad \Leftrightarrow \qquad \begin{cases} a_{11}\lambda_1 + a_{12}\lambda_2 + \dots + a_{1n}\lambda_n = 0 \\ a_{21}\lambda_1 + a_{22}\lambda_2 + \dots + a_{2n}\lambda_n = 0 \\ \dots \\ a_{n1}\lambda_1 + a_{n2}\lambda_2 + \dots + a_{nn}\lambda_n = 0 \end{cases}$$

$$\Leftrightarrow \qquad \lambda_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + \lambda_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \dots + \lambda_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Ce qui nous permet d'affirmer que le système d'équations linéaires $Ax = 0$ possède exactement une solution si et seulement les colonnes de la matrice A sont linéairement indépendante. Finalement, pour compléter la démonstration, il suffit de remarquer que dans un espace vectoriel de dimension n, tout ensemble linéairement indépendant de n vecteur doit obligatoirement être générateur, et donc il doit s'agit d'une base.

(5 ⇔ (6) Cette partie est relativement simple. Il suffit de remarquer que le système d'équations linéaires $Ax = 0$ est équivalent en prenant la transposé de chaque côté au système d'équations linéaires $x^T A^T = 0$. Ce dernier est en fait une réécriture du système d'équations linéaires en inversant le rôle de ligne et des colonnes. □

Exemple 1.4.1. On veut vérifier que l'ensemble ci-dessous forme une base de $\mathbb{R}^3$

$$\mathcal{B} = \left\{ \begin{pmatrix} 5 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \\ -1 \end{pmatrix} \right\}$$

Pour ce faire, en utilisant le théorème de la matrice inverse, on remarque qu'il est suffisant de calculer le déterminant de la matrice obtenu en plaçant les vecteurs en colonne (ou en ligne) dans une matrice. On a donc :

$$\begin{vmatrix} 5 & 0 & 4 \\ 2 & 1 & 0 \\ -3 & 5 & -1 \end{vmatrix} = 5 \begin{vmatrix} 1 & 0 \\ 5 & -1 \end{vmatrix} - 0 \begin{vmatrix} 2 & 0 \\ -3 & -1 \end{vmatrix} + 4 \begin{vmatrix} 2 & 1 \\ -3 & 5 \end{vmatrix} = 5(-1) - 0 + 4(13) = 47$$

Comme le déterminant n'est pas 0, on peut donc conclure qu'il s'agit bien d'une base.

1.5 Bases et déterminant

La question est maintenant de savoir comment faire pour utiliser le déterminant pour savoir si un ensemble de vecteurs forment une base dans un espace vectoriel autre que $\mathbb{R}^n$. De manière générale, il n'y a pas de truc magique. Par contre, lorsque l'on connaît déjà une base de l'espace, il est possible de vérifier qu'un ensemble est une base avec une méthode très semblable à ce que nous avons fait dans la section précédente.

Théorème 1.5.1. Si V est un espace vectoriel de dimension n, $\mathcal{B}$ est une base de V, et

$$S = \{u_1, u_2, ..., u_n\}$$

sont des vecteurs de V, alors les vecteurs de S forment une base de V si et seulement le déterminant de la matrice ayant pour colonne les vecteurs de S écrit dans la base $\mathcal{B}$ n'est pas zéro.

Démonstration. Il s'agit d'une application du théorème de la matrice inverse □

Attention, bien que le théorème soit valable de manière tout à fait général, il y a un détail important qu'il ne faut pas oublier. Les opérations d'addition et de multiplication par un scalaire peuvent être redéfini de manière arbitraire tant qu'elles satisfont les 10 propriétés. Donc pour pouvoir calculer le déterminant comme nous le faisons à l'habitude, il faut que les opérations de V correspondent aux opérations habituelles de $\mathbb{R}^n$ lorsque l'on écrit les vecteurs de V dans la base $\mathcal{B}$. Bien que ceci puisse sembler être un problème majeur, on peut cependant remarquer que les espaces vectoriels $\mathbb{P}_n(\mathbb{R})$ et $M_{n\times n}(\mathbb{R})$ muni des opérations habituelles satisfont cette condition.

Exemple 1.5.1. démontrez que

$$S = \left\{ \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \right\}$$

est une base de l'espace vectoriel

$$W = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}$$

muni des opérations d'addition et de multiplication par un scalaire habituel pour les matrices. Comme première méthode, nous allons montrez que c'est vecteur sont linéairement indépendant et générateur :

1. Pour montrer que les vecteurs de S sont linéairement indépendant, on doit résoudre le système d'équations linéaires suivant :

$$\alpha \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

   Ce qui peut s'écrire sous la forme :

$$\begin{cases} \alpha + \beta = 0 \\ \alpha - \beta = 0 \\ \alpha - \beta = 0 \\ \alpha + \beta = 0 \end{cases}$$

   En résolvant ce système, on obtient que la seule solution à ce système est $\alpha = \beta = 0$. Les vecteurs sont donc linéairement indépendant.

2. Pour montrer que les vecteurs de S sont générateur de W, on doit résoudre le système d'équations linéaires suivant :

$$\alpha \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

   Ce qui peut s'écrire sous la forme :

$$\begin{cases} \alpha + \beta = a \\ \alpha - \beta = b \\ \alpha - \beta = b \\ \alpha + \beta = a \end{cases}$$

On résout ce système en utilisant la méthode de Gauss :

$$\left(\begin{array}{cc|c} 1 & 1 & a \\ 1 & -1 & b \\ 1 & -1 & b \\ 1 & 1 & a \end{array}\right) \begin{array}{c} L_1 - L_2 \to L_2 \\ L_1 - L_3 \to L_3 \\ \sim \\ L_1 - L_4 \to L_4 \end{array} \left(\begin{array}{cc|c} 1 & 1 & a \\ 0 & 2 & a-b \\ 0 & 2 & a-b \\ 0 & 0 & 0 \end{array}\right) \begin{array}{c} L_2 - L_3 \to L_3 \\ \sim \end{array} \left(\begin{array}{cc|c} 1 & 1 & a \\ 0 & 2 & a-b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Comme ce système possède toujours exactement une solution, les vecteurs de S sont générateurs de W.

On a donc montrer que les vecteurs de S forment bien une base de W. On va maintenant montrer à nouveau que les vecteurs de S sont une base, mais cette fois en utilisant la méthode du déterminant. Pour ce faire on doit premièrement identifier une base "évidente" de W. Celle ci est relativement facile à trouver, et nous est donné par les vecteurs suivant :

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\}$$

On va ensuite écrire les vecteurs de S dans la base $\mathcal{B}$:

$$\begin{aligned} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{\mathcal{B}} \\ \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ -1 \end{pmatrix}_{\mathcal{B}} \end{aligned}$$

Et donc tout ce que nous avons à faire est de calculer le déterminant suivant :

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

Comme le déterminant n'est pas zéro, il s'agit donc bien d'une base.

1.6 Changement de bases pour les vecteurs

Si V est un espace vectoriel, on a vu que V admet toujours au moins une base. Nous avons aussi vu que dans une base donné, l'écriture d'un vecteur est unique. Cependant, il est tout à fait possible (et c'est pratiquement toujours le cas) que l'espace V admette plus d'une base. Dans ce cas, un vecteur pourra s'écrire de plusieurs façon, c'est à dire d'une façon (habituellement) différente pour chacune des bases. Dans cette section, nous allons maintenant voir comment faire pour convertir un vecteur écrit dans une base $\mathcal{B}_1$ en un vecteur écrit dans une base $\mathcal{B}_2$.

Théorème 1.6.1. Si V est un espace vectoriel, et $\mathcal{B}_1, \mathcal{B}_2$ sont des bases sur V, alors il existe une matrice $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ qui permet de convertir un vecteur écrit dans la base $\mathcal{B}_1$, en un vecteur écrit dans la base $\mathcal{B}_2$. C'est à dire que si v est un vecteur de V, et $[v]_{\mathcal{B}_1}$ et $[v]_{\mathcal{B}_2}$ sont les représentation de v dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement, alors :

$$[v]_{\mathcal{B}_2} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} [v]_{\mathcal{B}_1}$$

De plus, les colonnes de la matrice $P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1}$ sont formés des vecteurs de la base $\mathcal{B}_1$ écrit dans la base $\mathcal{B}_2$.

Bien que le théorème précédent nous permette de résoudre entièrement le problème du changement de base de manière théorique, trouver la matrice de passage peut dans certain cas devenir un peu long. Un truc souvent pratique pour simplifier les calculs est l'utilisation d'une troisième base qui va nous servir de base intermédiaire. C'est le théorème ci-dessous qui va nous permettre de simplifier (dans certain cas) nos calcul. Le théorème nous donne aussi de l'information concernant l'inverse d'une matrice de passage.

Théorème 1.6.2. Si V est un espace vectoriel, et $\mathcal{B}_1, \mathcal{B}_2$ et $\mathcal{B}_3$ sont des bases, alors on a les relations suivantes entre les matrices de passage :

1. $P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2} = (P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1})^{-1}$
2. $(P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2})(P_{\mathcal{B}_2 \leftarrow \mathcal{B}_3}) = P_{\mathcal{B}_1 \leftarrow \mathcal{B}_3}$

Exemple 1.6.1. Si $V = \mathbb{P}_2(\mathbb{R})$, on veut utiliser la méthode du changement de base pour trouver les valeurs de $a, b, c \in \mathbb{R}$ tel que :

$$2(x^2 + x + 1) + 3(x^2 + x - 1) + 1(x^2 - x + 1) = a(x^2 + 1) + b(x^2 + x) + c(x + 1)$$

On va donc commencer par montrer que les deux ensembles suivants sont des bases :

$$\mathcal{B}_1 = \{(x^2 + x + 1), (x^2 + x - 1), (x^2 - x + 1)\}$$

$$\mathcal{B}_2 = \{(x^2 + 1), (x^2 + x), (x + 1)\}$$

Pour ce faire, on va utiliser la base $\varepsilon = \{x^2, x, 1\}$, qui est la base standard de $\mathbb{P}_2(\mathbb{R})$. On doit donc écrire les vecteurs de $\mathcal{B}_1$ et $\mathcal{B}_1$ dans la base ε :

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$$

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Et donc on doit tout simplement calculez les déterminants suivant pour vérifier qu'ils s'agit de base :

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{vmatrix} = -4$$

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

Comme les déterminants ne sont pas zéro, il s'agit bien de base. On va maintenant trouver les matrices de passage :

$$P_{\varepsilon \leftarrow \mathcal{B}_1} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$P_{\varepsilon \leftarrow \mathcal{B}_2} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

On a donc l'équation suivante qui nous permet de calculer $a, b, c \in \mathbb{R}$:

$$\begin{aligned} \begin{pmatrix} a \\ b \\ c \end{pmatrix} &= P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = P_{\mathcal{B}_2 \leftarrow \varepsilon} P_{\varepsilon \leftarrow \mathcal{B}_1} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = (P_{\varepsilon \leftarrow \mathcal{B}_2})^{-1} P_{\varepsilon \leftarrow \mathcal{B}_1} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 5 \\ -1 \end{pmatrix} \end{aligned}$$

On a donc l'égalité suivante :

$$2(x^2+x+1)+3(x^2+x-1)+1(x^2-x+1)=1(x^2+1)+5(x^2+x)-(x+1)$$

Exemple 1.6.2. Considérez l'espace vectoriel $\mathbb{R}^2$ et les bases :

$$\mathcal{B}_1=\left\{\begin{pmatrix}1\\1\end{pmatrix},\begin{pmatrix}1\\-1\end{pmatrix}\right\}$$

$$\mathcal{B}_2=\left\{\begin{pmatrix}2\\5\end{pmatrix},\begin{pmatrix}1\\3\end{pmatrix}\right\}$$

On veut trouver les matrices de passage $P_{B_2\leftarrow B_1}$ et $P_{B_1\leftarrow B_2}$

Méthode 1 : Pour trouver la matrice de passage $P_{B_2\leftarrow B_1}$, il s'agit d'écrire les vecteurs de B_1 dans la base B_2. Pour ce faire, on doit donc résoudre les systèmes d'équations linéaires suivants :

$$\begin{pmatrix}1\\1\end{pmatrix}=a\begin{pmatrix}2\\5\end{pmatrix}+b\begin{pmatrix}1\\3\end{pmatrix}$$

$$\begin{pmatrix}1\\-1\end{pmatrix}=c\begin{pmatrix}2\\5\end{pmatrix}+d\begin{pmatrix}1\\3\end{pmatrix}$$

Ce qui nous donne :

$$\left(\begin{array}{cc|c}2&1&1\\5&3&1\end{array}\right)\sim^{5L_1-2L_2\to L_2}\left(\begin{array}{cc|c}2&1&1\\0&-1&3\end{array}\right)\sim^{L_1+L_2\to L_1}\left(\begin{array}{cc|c}2&0&4\\0&-1&3\end{array}\right)\sim^{\frac{1}{2}L_1\to L_1}_{-L_2\to L_2}\left(\begin{array}{cc|c}1&0&2\\0&1&-3\end{array}\right)$$

Donc $a=2$ et $b=-3$ ce qui nous donne

$$\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}2\\-3\end{pmatrix}_{B_2}$$

Pour le deuxième système d'équations linéaires, on a :

$$\left(\begin{array}{cc|c}2&1&1\\5&3&-1\end{array}\right)\sim_{5L_1-2L_2\to L_2}\left(\begin{array}{cc|c}2&1&1\\0&-1&7\end{array}\right)\sim^{L_1+L_2\to L_1}\left(\begin{array}{cc|c}2&0&8\\0&-1&7\end{array}\right)\sim^{\frac{1}{2}L_1\to L_1}_{-L_2\to L_2}\left(\begin{array}{cc|c}1&0&4\\0&1&-7\end{array}\right)$$

Donc $c=4$ et $b=-7$ ce qui nous donne

$$\begin{pmatrix}1\\-1\end{pmatrix}=\begin{pmatrix}4\\-7\end{pmatrix}_{B_2}$$

On obtient donc la matrice de passage suivante :

$$P_{B_2\leftarrow B_1}=\begin{pmatrix}2&4\\-3&-7\end{pmatrix}$$

Et finalement, en utilisant le théorème, on obtient :

$$P_{B_1\leftarrow B_2}=(P_{B_2\leftarrow B_1})^{-1}=\begin{pmatrix}2&4\\-3&-7\end{pmatrix}^{-1}=\begin{pmatrix}7/2&2\\-3/2&-1\end{pmatrix}$$

Méthode 2 : Comme dans $\mathbb{R}^2$ on a une base évidente, c'est à dire la base canonique :

$$\varepsilon=\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$$

On a donc les matrices de passage suivante :

$$P_{\varepsilon\leftarrow B_1}=\begin{pmatrix}1&1\\1&-1\end{pmatrix}$$

$$P_{\varepsilon \leftarrow B_2} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}$$

On obtient donc les matrices de passage suivante :

$$\begin{aligned} P_{B_2 \leftarrow B_1} &= P_{B_2 \leftarrow \varepsilon} P_{\varepsilon \leftarrow B_1} = (P_{\varepsilon \leftarrow B_2})^{-1} P_{\varepsilon \leftarrow B_1} = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ -3 & -7 \end{pmatrix} \end{aligned}$$

Pour trouver l'autre matrice de passage, on calcul l'inverse ce qui nous donne :

$$P_{B_1 \leftarrow B_2} = (P_{B_2 \leftarrow B_1})^{-1} = \begin{pmatrix} 2 & 4 \\ -3 & -7 \end{pmatrix}^{-1} = \begin{pmatrix} 7/2 & 2 \\ -3/2 & -1 \end{pmatrix}$$

Chapitre 2

Les applications linéaires

Nous avons vu dans le chapitre précédent ce qu'est un espace vectoriel. Nous allons maintenant étudier les transformations linéaires entre deux espaces vectoriels. Lorsque ces dernier sont de dimension fini, une telle transformation sera représenté par une matrice.

2.1 Matrices et applications linéaires

Definition 2.1.1. Si V et W sont des espaces vectoriels et $f : V \to W$, alors on dit que f est une application linéaire si :

1. $f(u+v) = f(u) + f(v)$ pour tout $u, v \in V$
2. $f(\lambda u) = \lambda f(u)$ pour tout $\lambda \in \mathbb{R}$ et $u \in V$

Exemple 2.1.1. Il existe beaucoup d'exemples d'applications linéaires. En voici quelques unes :

1. Si on considère l'espace vectoriel $\mathbb{R}$, alors les applications linéaires sont les fonctions de la forme $f(x) = ax$ où a est une constante. Il faut cependant faire attention, bien qu'il soit tentant d'affirmer que toute les fonctions de la forme $g(x) = ax + b$ soit linéaire, ceci est vrai seulement si $b = 0$. Une fonction de la forme $g(x) = ax + b$ devrait plutôt être appelé fonction affine.
2. Dans l'espace vectoriel $\mathbb{R}^2$, c'est à dire dans le plan, les symétries, rotation et homothétie sont des transformations linéaires. Par contre, les translations ne le sont pas.
3. On peut aussi considérer des applications linéaires sur des espaces vectoriels de dimension infinie [1]. Si on prend l'espace vectoriel $V = C^1(\mathbb{R})$ de toutes les fonctions dérivables sur $\mathbb{R}$, alors on peut définir la transformation suivante :
$$f(g) = \frac{d}{dx}(g), \quad \forall g \in C(\mathbb{R})$$
Cette transformation est linéaire, car les propriétés de la dérivée nous affirme que :
$$\frac{d}{dx}(\lambda f) = \lambda \frac{d}{dx}(f), \quad \forall \lambda \in \mathbb{R}, \ \forall f \in C^1(\mathbb{R})$$
$$\frac{d}{dx}(f+g) = \frac{d}{dx}(f) + \frac{d}{dx}(g), \quad \forall f, g \in C^1(\mathbb{R})$$
Bien que ce dernier exemple soit particulièrement important, nous n'en diront cependant pas plus sur les applications linéaires dans des espaces vectoriels de dimension infinie.

Definition 2.1.2. Si V, W sont des espaces vectoriels et $f : V \to W$ est une application linéaire, alors on définit l'image de f comme étant :
$$Im(f) = \{f(x) : x \in V\}$$

1. L'étude des applications linéaires sur des espaces vectoriels de dimension infinie fait habituellement partie d'un sujet appelé l'analyse fonctionnelle, dans lequel on combine les techniques de l'algèbre linéaire avec ceux de l'analyse.

et on définit le noyau de f comme étant :

$$ker(f) = \{x \in V : f(x) = 0\}$$

Les deux concepts que nous venons de définir sont en fait l'image d'une fonction et l'ensemble des zéros d'une fonction.

Théorème 2.1.1. Si V et W sont des espaces vectoriels et $f : V \to W$ est une application linéaire, alors

1. $Ker(f)$ est un sous espace vectoriel de V
2. $Im(f)$ est un sous espace vectoriel de W

Démonstration. La démonstration est laissé en exercice. □

Théorème 2.1.2. Si V et W sont des espaces vectoriels de dimension fini et $f : V \to W$ est une application linéaire, alors on peut trouver une matrice A tel que :

$$f(u) = Au \text{ pour tout } u \in V$$

Démonstration. Prenons $\mathcal{B} = \{e_1, e_2, ..., e_n\}$ une base de V. Alors si $u \in V$, alors :

$$u = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 + ... + \lambda_n e_n$$

alors on a :

$$\begin{aligned}
f(u) &= f(\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 + ... + \lambda_n e_n) \\
&= \lambda_1 f(e_1) + \lambda_2 f(e_2) + \lambda_3 f(e_3) + ... + \lambda_n f(e_n) \\
&= \begin{pmatrix} [\lambda_1 f(e_1) + \lambda_2 f(e_2) + \lambda_3 f(e_3) + ... + \lambda_n f(e_n)]_1 \\ [\lambda_1 f(e_1) + \lambda_2 f(e_2) + \lambda_3 f(e_3) + ... + \lambda_n f(e_n)]_2 \\ \vdots \\ [\lambda_1 f(e_1) + \lambda_2 f(e_2) + \lambda_3 f(e_3) + ... + \lambda_n f(e_n)]_m \end{pmatrix} \\
&= \begin{pmatrix} [f(e_1)]_1 & [f(e_2)]_1 & ... & [f(e_n)]_1 \\ [f(e_1)]_2 & [f(e_2)]_2 & ... & [f(e_n)]_2 \\ \vdots & \vdots & & \vdots \\ [f(e_1)]_m & [f(e_2)]_m & ... & [f(e_n)]_m \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \\
&= Au
\end{aligned}$$

□

Remarquez que la matrice A du théorème précédent dépend du choix de base. Si on change de base, la matrice changera.

Definition 2.1.3. Des matrices A et B sont dites semblable si elles représentent la même application linéaire, mais dans des bases différentes.

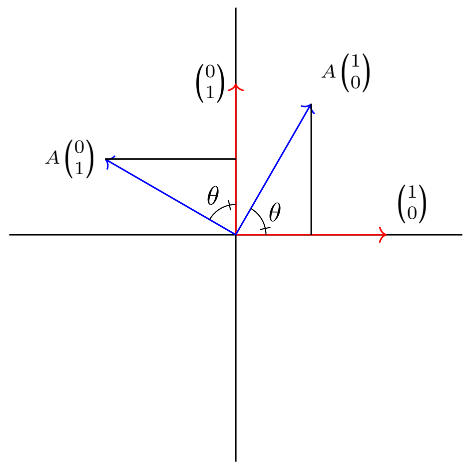
Une question importante est d'essayer de trouver des critères nous permettant de savoir si deux matrices données sont semblables. Nous allons revenir à quelques reprise sur cette question durant le reste du chapitre et dans le chapitre suivant.

2.2 Les transformations du plan

Il existe plusieurs matrices standard représentant des applications linéaires telles que des symétries, rotations et homotéties dans le plan, ou même plus généralement dans $\mathbb{R}^n$. Bien que nous allons mettre l'emphase sur $\mathbb{R}^2$, la théorie que nous allons développer ici s'applique un des espaces de n'importe quel dimension. Pour trouver la matrice A d'une transformation linéaire f écrit dans une base $\mathcal{B} = \{u_1, u_2, ..., u_n\}$ sur un espace vectoriel V, on place les vecteurs $f(u_1), f(u_2), ..., f(u_n)$ en colonne dans une matrice.

Exemple 2.2.1. On veut trouver la matrice A représentant une rotation d'angle θ dans $\mathbb{R}^2$ en utilisant la base habituelle. Pour ce faire, commençons par représenter les vecteurs $\vec{i}$ et $\vec{j}$ dans le plan, ainsi que leur image selon la rotation. Sachant que la longueur de chacun des vecteurs que nous avons dessiner est de

FIGURE 2.1 – Rotation d'angle θ



longueur 1, un peu de trigonométrie nous permet d'obtenir :

$$A\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix} \qquad A\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-\sin(\theta)\\ \cos(\theta)\end{pmatrix}$$

Ce qui nous permet d'écrire la matrice A comme étant :

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

Exemple 2.2.2. Dans $\mathbb{R}^2$, on veut trouver la matrice représentant une symétrie selon l'axe $y = x$. Utiliser premièrement la base habituelle, puis utiliser la base $\mathcal{B} = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\right\}$. En faisant un schéma comme nous l'avons fait pour l'exemple précédent, on remarque que :

$$[A]_\varepsilon \begin{pmatrix}1\\0\end{pmatrix}_\varepsilon = \begin{pmatrix}0\\1\end{pmatrix}_\varepsilon, \qquad [A]_\varepsilon \begin{pmatrix}0\\1\end{pmatrix}_\varepsilon = \begin{pmatrix}1\\0\end{pmatrix}_\varepsilon \qquad \Rightarrow \qquad [A]_\varepsilon = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}_\varepsilon$$

De la même façon, avec la base $\mathcal{B}$ on obtient :

$$[A]_\mathcal{B} \begin{pmatrix}1\\0\end{pmatrix}_\mathcal{B} = \begin{pmatrix}1\\0\end{pmatrix}_\mathcal{B}, \qquad [A]_\mathcal{B} \begin{pmatrix}0\\1\end{pmatrix}_\mathcal{B} = \begin{pmatrix}0\\-1\end{pmatrix}_\mathcal{B} \qquad \Rightarrow \qquad [A]_\mathcal{B} = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}_\mathcal{B}$$

De plus, comme les deux matrices représente la même application linéaire, mais dans des bases différentes, on peut donc affirmer que les deux matrices sont semblable.

Table 2.1 – Transformations importantes du plan

Matrice	Signification géométrique
$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$	Grossissement (ou rapetissement) de facteur a horizontalement et de facteur b verticalement. Ici on suppose $a > 0$ et $b > 0$.
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	Symétrie selon l'axe des x
$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$	Symétrie selon l'axe des y
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	Projection orthogonale sur l'axe des x
$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$	Projection orthogonale sur l'axe des y
$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$	Rotation d'angle θ dans le sens anti-horaire.

La table 2.1 représente quelques transformations simples de $\mathbb{R}^2$ que vous devriez normalement être capable de retrouver facilement (sinon il faut les connaitre par coeur).

Bien entendu, dans bien des cas nous somme intéresser à travailler avec des transformation linéaire qui sont plus compliquer que celle énoncé dans le tableau ci-dessus. Par contre, la plupart du temps elles peuvent être décomposer en plusieurs applications linéaires simple. Le théorème ci-dessous nous sera donc d'une grande utilité. Remarquez que dans le premier cours d'algèbre linéaire, nous avions utiliser ce théorème pour définir notre produit de matrice. Le théorème ne devrait donc pas vous surprendre.

Théorème 2.2.1. La composition des applications linéaires correspond au produit de matrices.

Exemple 2.2.3. On cherche la matrice correspondant à une symétrie selon l'axe des x suivi d'une rotation de 45 degrés dans $\mathbb{R}^2$. En utilisant les matrices vu précédemment, combiné avec le théorème, on a donc :

$$\begin{pmatrix} \cos(45) & -\sin(45) \\ \sin(45) & \cos(45) \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$$

2.3 Changement de base pour les matrices

Nous avons déjà vu dans le chapitre précédent qu'un vecteur a une représentation différente pour chaque base, et nous avons vu comment passer d'une base à l'autre. Dans ce chapitre, nous avons étudier les applications linéaires, que nous avons écrite sous forme de matrice. Comme c'était le cas pour les vecteurs, la matrice correspondant à une application linéaire dépend de la base choisi. La question est donc la même. Si on connait la matrice d'une application linéaire dans une base donné, comment faire pour trouvé la matrice qui correspond à la même transformation linéaire mais dans une base différente.

Théorème 2.3.1. Si V est un espace vectoriel et f est une application linéaire $f : V \to V$. Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont des bases de V, et $A_{\mathcal{B}_1}, A_{\mathcal{B}_2}$ sont les matrices représentant l'application linéaire f dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement. Alors

$$A_{\mathcal{B}_2} = P_{\mathcal{B}_2 \leftarrow \mathcal{B}_1} A_{\mathcal{B}_1} P_{\mathcal{B}_1 \leftarrow \mathcal{B}_2}$$

Corollaire 2.3.1. Des matrices A et B sont semblable (i.e. qui représente la même application linéaire, mais dans des bases différentes) si et seulement si il existe une matrice P inversible tel que $A = PBP^{-1}$

Exemple 2.3.1. Pour illustrer l'utilisation du théorème, nous allons chercher à retrouver la formule permettant de calculer la projection d'un vecteur sur un autre, mais plutôt que de le faire géométriquement comme dans le premier cours d'algèbre linéaire, nous allons cette fois le faire à l'aide d'un changement de base. Rappellons du premier cours d'algèbre linéaire que si x et u sont des vecteurs de $\mathbb{R}^n$, alors la projection du x sur le vecteur u est donné par :

$$x_u = \frac{x \cdot u}{||u||^2} u$$

Pour cet exercice, nous allons travailler uniquement dans $\mathbb{R}^2$, mais le même raisonnement pourrait s'appliquer aux autres dimension. Premièrement, supposons que $u = \begin{pmatrix} a \\ b \end{pmatrix}$. On peut donc créer la base suivante de $\mathbb{R}^2$:

$$\mathcal{B} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -b \\ a \end{pmatrix} \right\}$$

qui contient le vecteur u et un vecteur qui lui est orthogonal (perpendiculaire). Il n'est pas très difficile de vérifier que si le vecteur u n'est pas 0, il s'agit bien d'une base. Si on dénote par A la matrice qui représente la projection sur le vecteur u, on obtient donc :

$$[A]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Cette matrice est cependant écrite dans la base $\mathcal{B}$ ce qui n'est pas l'idéal. Nous allons donc effectuer un changement de base pour la remettre dans la base habituelle :

$$[A]_{\varepsilon} = P_{\varepsilon \leftarrow \mathcal{B}} [A]_{\mathcal{B}} P_{\mathcal{B} \leftarrow \varepsilon} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix}^{-1} = \frac{1}{a^2 + b^2} \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = \frac{1}{a^2 + b^2} \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}$$

En multipliant la matrice ainsi obtenu par un vecteur x quelconque, on obtient donc la projection x_u. Question d'obtenir exactement la formule du premier cours d'algèbre linéaire, posons $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. On obtient donc :

$$\begin{aligned} x_u &= [A]_{\varepsilon} x = \frac{1}{a^2 + b^2} \begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{||u||^2} \begin{pmatrix} a^2 x_1 + abx_2 \\ abx_1 + b^2 x_2 \end{pmatrix} = \frac{1}{||u||^2} \begin{pmatrix} (ax_1 + bx_2)a \\ (ax_1 + bx_2)b \end{pmatrix} \\ &= \frac{ax_1 + bx_2}{||u||^2} \begin{pmatrix} a \\ b \end{pmatrix} = \frac{x \cdot u}{||u||^2} u \end{aligned}$$

Ce qui est exactement la formule que nous avons vu dans le premier cours.

Exemple 2.3.2. On veut trouver la matrice (dans la base habituel) qui décrit une symétrie par rapport à l'axe $y = 5x$. Pour ce faire, on va commencer par résoudre le problème dans la base

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ -1 \end{pmatrix} \right\}$$

Nous avons choisi cette base de sorte que le premier vecteur soit dans la direction de la droite, et le second perpendiculaire. Dans cette base, il est facile de trouver la matrice correspondant au problème :

$$[A]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

On doit maintenant réécrire la matrice dans la base habituelle :

$$\begin{aligned}[A]_\varepsilon &= P_{\varepsilon\leftarrow\mathcal{B}}[A]_\mathcal{B}P_{\mathcal{B}\leftarrow\varepsilon}\\ &= \begin{pmatrix}1 & 5\\ 5 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1 & 5\\ 5 & -1\end{pmatrix}^{-1}\\ &= \begin{pmatrix}-12/13 & 5/13\\ 5/13 & 12/13\end{pmatrix}\end{aligned}$$

À l'aide de cette matrice, on peut maintenant trouver l'image de n'importe quel point de $\mathbb{R}^2$ écrit dans la base habituelle.

Exemple 2.3.3. Dans $V = \mathbb{P}_2(\mathbb{R})$, on a la transformation linéaire suivante :

$$T(ax^2 + bx + c) = ax^2 + c$$

On veut maintenant appliquer la même transformation, mais pour des vecteurs écrit sous la forme :

$$\alpha(x^2 + x) + \beta(x + 1) + \gamma(x^2 + 1)$$

Sans réécrire les vecteurs sous la forme (base) habituelle. Pour ce faire, on va commencer par montrer que l'ensemble suivante est une base :

$$\mathcal{B} = \{x^2 + x, x + 1, x^2 + 1\} = \left\{ \begin{pmatrix}1\\1\\0\end{pmatrix}_\varepsilon, \begin{pmatrix}0\\1\\1\end{pmatrix}_\varepsilon, \begin{pmatrix}1\\0\\1\end{pmatrix}_\varepsilon \right\}$$

Pour ce faire, on calcul le déterminant :

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

Il s'agit donc bien d'une base. On a donc que la matrice qui représente cette transformation linéaire dans la base $\mathcal{B}$ est donné par :

$$\begin{aligned}[T]_\mathcal{B} &= P_{\mathcal{B}\leftarrow\varepsilon}[T]_\varepsilon P_{\varepsilon\leftarrow\mathcal{B}}\\ &= \begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 0 & 1 & 1\end{pmatrix}\\ &= \begin{pmatrix}1/2 & -1/2 & 0\\ -1/2 & 1/2 & 0\\ 1/2 & 1/2 & 1\end{pmatrix}\end{aligned}$$

On peut donc utiliser cette matrice pour calculer par exemple l'image du vecteur $2(x^2+x)+4(x+1)+6(x^2+1)$:

$$\begin{aligned}T\Big(2(x^2 + x) + 4(x + 1) + 6(x^2 + 1)\Big) &= \begin{pmatrix}1/2 & -1/2 & 0\\ -1/2 & 1/2 & 0\\ 1/2 & 1/2 & 1\end{pmatrix}\begin{pmatrix}2\\4\\6\end{pmatrix} = \begin{pmatrix}-1\\1\\9\end{pmatrix}\\ &= -(x^2 + x) + (x + 1) + 9(x^2 + 1)\end{aligned}$$

2.4 Les 4 espaces fondamentaux

Dans le chapitre précédent, nous avons étudier les sous-espaces vectoriels dans un contexte général. Nous allons maintenant étudier 4 sous-espace important qui sont associé à une matrice. Les trois premiers : l'espace colonne, l'espace ligne et l'espace nul sont relativement facile à définir et à comprendre. Le 4e est cependant un peu plus obscure, mais tout aussi important. Il s'agit de l'espace nul de la matrice transposé.

Definition 2.4.1. Si A est une matrice, alors on définit $col(A)$ comme étant l'espace généré (span) par les vecteurs colonnes de A. De la même façon, on définit $row(A)$ comme étant l'espace généré (span) par les vecteurs lignes de la matrice A.

Théorème 2.4.1. Les espaces lignes et colonnes d'une matrice ont les propriétés suivantes :

1. Si A est une matrice de dimension $m \times n$, alors $col(A)$ est un sous-espace vectoriel de $\mathbb{R}^n$ et $row(A)$ est un sous-espace vectoriel de $\mathbb{R}^m$.
2. Si V et W sont des espaces vectoriel et $f : V \to W$ est un application linéaire qui peut être représenté par la matrice A, c'est à dire que $f(v) = Av$, alors :
$$col(A) = Im(f)$$
3. Si $Ax = b$ est un système d'équations linéaires, alors le système est consistent (i.e. il existe au moins une solution) si et seulement si $b \in col(A)$.

Démonstration.

1. Il s'agit d'une conséquence que le span d'un ensemble de vecteur forme toujours un sous-espace vectoriel.
2. On se rappelle que
$$Im(f) = \{f(u) : u \in V\}$$
Prenons $e_1, e_2, ..., e_n$ une base de V, alors on obtient :
$$Im(f) = \{\lambda_1 f(e_1) + \lambda_2 f(e_2) + ... + \lambda_n f(e_n) : \lambda_i \in \mathbb{R}\} = span\{f(e_1), f(e_2), ..., f(e_n)\}$$
Maintenant, comme les $f(e_1), f(e_2), ..., f(e_n)$ sont les colonnes de la matrice A, alors :
$$Im(f) = col(A)$$
3. Il s'agit d'une conséquence directe de la partie précédent. L'image d'une application linéaire étant par définition l'ensemble des valeurs de b pour lesquels le système d'équations linéaire $Ax = b$ possède une solution.

$\square$

Nous venons de voir que l'espace colonne d'une matrice est en fait la même chose que l'image de l'application linéaire représenté par cette matrice. La question que vous devriez maintenant vous poser est que représente l'espace ligne ? L'idée est relativement simple, l'espace ligne d'une matrice A est en fait la même chose que l'espace colonne de la matrice A^T. Donc si V et W sont des espaces vectoriels, et $f : V \to W$ est une application linéaire, alors on peut trouver une matrice A tel que $f(v) = Av$ pour tout $v \in V$. Regardons maintenant ce qui ce passe en prenant la transposer de la matrice A. Si on définit $g(w) = A^T w$, alors on obtient une transformation linéaire $g : W \to V$. Attention, bien qu'il y ait certaine similarité avec l'inverse d'une matrice, dans la grande majorité des cas il ne s'agit absolument pas de l'inverse de la matrice (l'exception étant les matrices orthogonales). L'espace ligne de la matrice A devient l'espace colonne de la matrice A^T. On peut donc conclure que l'espace ligne de la matrice A est en fait l'image de l'application linéaire g. Remarquez finalement qu'on peut réécrire l'application linéaire g comme étant $g(w) = A^T w = (w^T A)^T$. Ce qui nous donne un indice de ce que sera le 4e espace fondamental que nous aurons besoin.

Definition 2.4.2. Si A est une matrice, alors l'espace nul de A, dénoté $null(A)$ est l'ensemble des solutions du système d'équations linéaires $Ax = 0$.

Théorème 2.4.2. L'espace null d'une matrice a les propriétés suivantes :

1. Si V et W sont des espaces vectoriel et $f : V \to W$ est une application linéaire pouvant être représenté par la matrice A, c'est à dire $f(v) = Av$ pour tout $v \in V$, alors on a
$$null(A) = \ker(f)$$
2. Si A est une matrice de dimension $m \times n$, alors $null(A)$ est un sous-espace vectoriel de $\mathbb{R}^n$.

Démonstration.

1. Il s'agit de comparer la définition du noyau d'un application linéaire avec celui de l'espace null d'une matrice.
2. Il s'agit d'une conséquence que $\ker(f)$ est un sous-espace vectoriel si f est un application linéaire.

□

Nous somme maintenant prêt à définir notre 4e espace fondamental. Premièrement, nous avons déjà remarquer qu'il y a un lien important entre l'espace ligne et l'espace colonne. En effet, l'espace ligne n'est rien d'autre que l'espace colonne de la matrice transposé, et inversement. Donc prendre la transposé de la matrice A nous permet de faire le lien entre les deux premiers espaces fondamentaux. La même idée s'applique si on compare l'espace null d'une matrice A avec l'espace null de la matrice A^T. On appelle souvent ce quatrième espace le noyau à gauche d'une matrice, et on peut le définir comme étant :

$$coker(f) = null(A^T) = \{x^T : x^T A = 0\}$$

Nous allons maintenant illustrer les 4 espaces fondamentaux sous forme d'un diagramme [2]. Ce dernier est pour le moment encore incomplet. Il nous faudra attendre au chapitre 4 pour pouvoir le compléter. Le diagramme va tout de même nous permettre de mieux visualiser l'interaction entre ces 4 espaces. Si V et W sont des espaces vectoriels, et $f : V \to W$ est une application linéaire, alors on peut représenter f sous forme d'une matrice A, ceci nous permet d'obtenir le diagramme suivant :

Pour trouver une base de l'espace colonne, ligne et nul, on commence par appliquer la méthode de Gauss à la matrice. Les lignes non nulle de la matrice échelon forme une base de l'espace ligne. Les colonnes contenant les pivots dans la matrice échelon correspondent aux colonnes de la matrice originale qui forme une base de l'espace colonne. Finalement, pour trouver une base de l'espace nul, on résolut le système d'équation $Ax = 0$, puis pour obtenir la base on pose chacun des paramètres libre comme étant 1 pendant que les autres sont 0.

Exemple 2.4.1. Trouver des bases pour $col(A), row(A)$ et $null(A)$ pour la matrice

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

On commence par appliquer la méthode de Gauss pour mettre la matrice sous forme échelon.

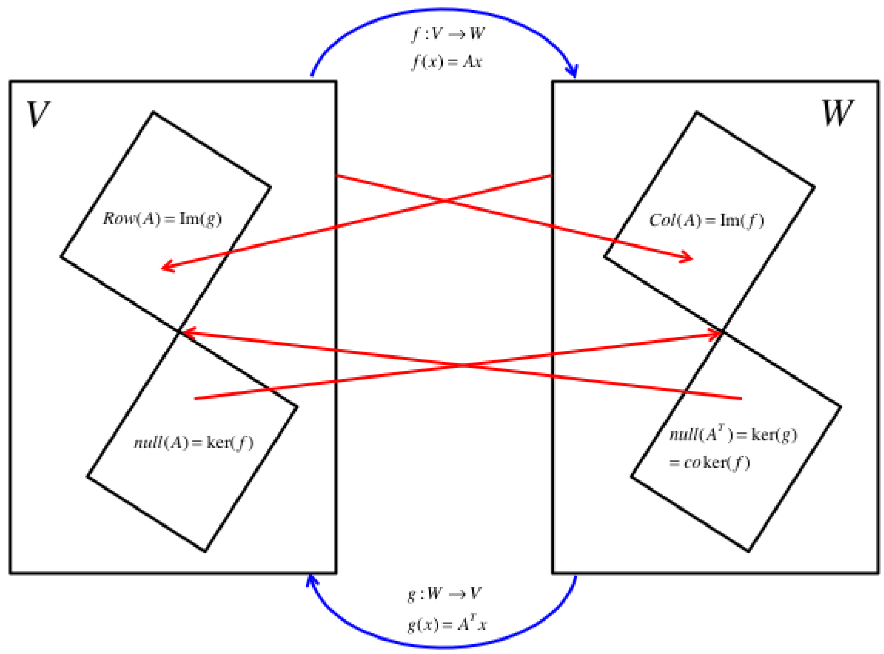
$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \sim^{\substack{4L_1 - L_2 \to L_2 \\ 7L_1 - L_3 \to L_3}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \\ 0 & 6 & 12 \end{pmatrix} \sim^{2L_2 - L_3 \to L_3} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{pmatrix} \sim^{\frac{1}{3}L_2 \to L_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

On obtient donc comme base pour $col(A)$:

$$\left\{ \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} \right\}$$

2. Notez que le diagramme, bien que légèrement différent, est fortement inspirer de l'article de Strang [6]

Figure 2.2 – Les 4 espaces fondamentaux



comme base pour $row(A)$ on a :

$$\left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\}$$

Il ne nous reste plus qu'à trouver une base pour $null(A)$. Pour ce faire, on va résoudre le système d'équation $Ax = 0$:

$$\left\{ \begin{array}{l} z = t \\ y = -2z = -2t \\ x = -2y - 3z = 4t - 3t = t \end{array} \right. \qquad \Rightarrow \qquad null(A) = \left\{ \begin{pmatrix} t \\ -2t \\ t \end{pmatrix} : t \in \mathbb{R} \right\}$$

Une base de $null(A)$ est donc donné par :

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\}$$

Remarquez finalement que bien que nous ayons fait attention de faire la différence entre les espaces associé à un application linéaire f, et les espaces fondamentaux associé à une matrice A, cette différence est un peu artificielle. Par abus de language, il sera donc courant de parler de l'image de la matrice A plutôt que de l'espace colonne, et de parler du noyau de la matrice A plutôt que de l'espace nulle. D'ailleurs, plusieurs auteur ne font pas la différence entre les deux.

2.5 La théorie du rang

Dans la section précédente, nous avons définie les 4 espaces fondamentaux associés à une matrice et nous avons vu comment trouver une base pour chacun d'eux. En fait, pour être précis nous avons vu comment trouver une base de 3 de ces 4 espaces fondamentaux, mais il n'est pas difficile de comprendre comment trouver une base pour le 4e lorsqu'on a compris comment trouver une base de l'espace null (ou noyau). Il est maintenant temps de regarder la dimension de chacun de ces espaces. Dans ce cas, des liens importants apparaissent.

Definition 2.5.1. Si A est une matrice, alors on définit le rang de A comme étant la dimension de l'espace colonne $col(A)$, c'est à dire la dimension de $Im(f)$ où f est l'application linéaire $f(x) = Ax$. De plus, on définit la nullité de A comme étant la dimension de l'espace nul de A, c'est à dire la dimension de $Ker(f)$.

Exemple 2.5.1. Dans l'exemple que nous avons fait à la section précédente avec la matrice

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

on a donc que le $rang(A) = 2$ et $nullite(A) = 1$.

Théorème 2.5.1. Si A et P sont des matrices pour lesquels les produits ci-dessous sont définies, et si P est inversible, alors on a :

1. $rang(A) = rang(PA)$
2. $rang(A) = rang(A^T)$
3. $rang(A) = rang(AP)$

Démonstration.

1. Supposons que $\{x_1, x_2, ..., x_n\}$ forme une base de $Im(A)$. On veut montrer que $\{Px_1, Px_2, ..., Px_n\}$ forme une base de $Im(PA)$. Pour ce faire, nous allons commencer par montrer que ces vecteurs sont linéairement indépendant :

$$\begin{aligned}
&\lambda_1 Px_1 + \lambda_2 Px_2 + \lambda_3 Px_3 + ... + \lambda_n Px_n = 0 \\
&\Longrightarrow P(\lambda_1 x_1) + P(\lambda_2 x_2) + P(\lambda_3 x_3) + ... + P(\lambda_n x_n) = 0 \\
&\Longrightarrow P(\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + ... + \lambda_n x_n) = 0 \\
&\Longrightarrow \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + ... + \lambda_n x_n = 0 \text{ car } P \text{ est inversible} \\
&\Longrightarrow \lambda_1 = \lambda_2 = \lambda_3 = ... = \lambda_n = 0 \text{ par indépendance linéaire de } \{x_1, x_2, ..., x_n\}
\end{aligned}$$

   Les vecteurs $\{Px_1, Px_2, ..., Px_n\}$ sont donc linéairement indépendant. Nous allons maintenant montrer que ces vecteurs sont générateurs de $Im(PA)$. Pour ce faire, prenons $z \in Im(PA)$. Il existe donc un y tel que $z = PAy$ et donc comme la matrice P est inversible $P^{-1}z = Ay$. On a donc que $P^{-1}z \in Im(A)$. En utilisant notre base de $Im(A)$, on peut donc trouver des λ_i tel que :

$$P^{-1}z = \lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + ... + \lambda_n x_n$$

   Ce qui nous donne finalement :

$$z = P(\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 + ... + \lambda_n x_n) = \lambda_1(Px_1) + \lambda_2(Px_2) + \lambda_3(Px_3) + ... + \lambda_n(Px_n)$$

   On peut donc écrire z comme combinaison linéaire des vecteurs $\{Px_1, Px_2, ..., Px_n\}$, ce qui confirme qu'ils sont bien générateur. Il s'agit donc d'une base de $Im(PA)$. Comme la base de $Im(A)$ a le même nombre d'élément que la base de $Im(PA)$, on obtient donc $rang(A) = rang(PA)$.

2. Dans le chapitre précédent, nous avons vu deux méthodes nous permettant de trouver une base du span d'un ensemble de vecteurs. Dans un premier temps, il s'agissait de placer les vecteurs à l'horizontal, et dans la seconde à la verticale. Dans les deux cas on appliquait la méthode de Gauss pour obtenir une matrice sous forme échellon, le nombre de vecteurs dans une base était alors donné par le nombre de pivot. Remarquez maintenant que le rang de A est la dimension de l'espace colonne, alors que le rang de A^T est la dimension de l'espace ligne. En applicant la méthode de Gauss directement sur la matrice A, on peut alors obtenir en même temps une base pour l'espace ligne et pour l'espace colonne (il ne s'agit pas de la même base). Dans les deux cas, le nombre de pivot correspond à la dimension, ce qui nous permet de déduire que

$$rang(A) = rang(A^T)$$

3. Premièrement, par la partie précédente, nous avons $rang(A) = rang(A^T)$. De plus, si P est une matrice inversible, alors P^T est aussi inversible, car elles ont le même déterminant. On a donc que $rang(A^T) = rang(P^T A^T)$ par la première partie du théorème. Maintenant, par les propriétés de la transposé, on a $rang(P^T A^T) = rang((AP)^T)$. Puis finalement, en utilisant à nouveau la seconde partie du théorème on obtient finalement $rang((AP))^T = rang(AP)$. On peut donc résumé le tout par les égalités suivantes :

$$rang(A) = rang(A^T) = rang(P^T A^T) = rang((AP)^T) = rang(AP)$$

$\square$

Corollaire 2.5.1. Si A est une matrice de dimension $m \times n$, alors $rang(A) \leq \min\{m, n\}$.

Démonstration. Il s'agit d'une conséquence directe du théorème. Si A est une matrice de dimension $m \times n$, alors $rang(A) \leq n$ car la dimension de l'espace colonne de la matrice ne peut pas excéder le nombre de colonne (il s'agit d'un ensemble générateur). De plus, comme A^T est une matrice de dimension $n \times m$, la même idée s'applique ce qui nous donne $rang(A^T) \leq m$. Finalement, par le théorème, comme $rang(A) = rang(A^T)$ on obtient $rang(A) \leq \min\{m, n\}$. $\square$

Remarquez que le fait que $rang(A) = rang(A^T)$ signifie que l'espace ligne et l'espace colonne de la matrice A ont la même dimension. Il est cependant important de notez que bien que ces deux espaces ont la même dimension, il sont en général deux espaces très différent. En particulier, l'espace ligne et l'espace colonne n'ont pas besoin d'être des sous-espace du même espace vectoriel.

Théorème 2.5.2. **(Théorème du rang)** Si A est une matrice de dimension $m \times n$, alors :

$$\text{rang}(A) + \text{nullité}(A) = n$$

Démonstration. Supposons que $\mathcal{S}_1 = \{u_1, u_2, ..., u_k\}$ est une base de $\ker(A)$, alors $\mathcal{S}_1$ peut être complété en une base $\mathcal{S} = \{u_1, u_2, ..., u_k, w_{k+1}, ...w_n\}$ de $\mathbb{R}^n$. On veut montrer que $\mathcal{S}' = \{Aw_{k+1}, Aw_{k+2}, ..., Aw_n\}$ est une base de $Im(A)$. Premièrement, il est facile de voir que $\mathcal{S}'$ est générateur de $Im(A)$. Supposons que $x \in Im(A)$, alors il existe $y \in \mathbb{R}^n$ tel que $x = Ay$. Maintenant, comme $\mathcal{S}$ est une base de $\mathbb{R}^n$, alors

$$y = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_k u_k + \lambda_{k+1} w_{k+1} + ... + \lambda_n w_n$$

On a donc :

$$\begin{aligned} x &= Ay \\ &= A(\lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_k u_k + \lambda_{k+1} w_{k+1} + ... + \lambda_n w_n) \\ &= 0 + 0 + ... + 0 + \lambda_{k+1} A w_{k+1} + ... + \lambda_n A w_n \\ &= \lambda_{k+1} A w_{k+1} + ... + \lambda_n A w_n \end{aligned}$$

On doit donc maintenant montrer que $\mathcal{S}'$ est linéairement indépendant. On doit donc trouver les solutions de l'équations :

$$\lambda_{k+1} A w_{k+1} + \lambda_{k+2} A w_{k+2} + ... + \lambda_n A w_n = A\Big(\lambda_{k+1} w_{k+1} + \lambda_{k+2} w_{k+2} + ... + \lambda_n w_n\Big) = 0$$

ce qui nous donne que

$$\lambda_{k+1} w_{k+1} + \lambda_{k+2} w_{k+2} + ... + \lambda_n w_n \in \ker(A)$$

ce qui est possible seulement si tout les coefficients sont nul étant donné que $\mathcal{S}$ est une base (donc linéairement indépendant). On peut donc conclure que $\mathcal{S}'$ est une base de $Im(A)$, et donc $\text{rang}(A) = n - \text{nullité}(A)$. $\square$

Exemple 2.5.2. Existe-t-il une matrice A de dimension 4×6 tel que $nullite(A) = 1$? Premièrement, par le théorème du rang, on a

$$rang(A) + nullite(A) = n \quad \Rightarrow \quad rang(A) + 1 = 6 \quad \Rightarrow \quad rang(A) = 5$$

ce qui est impossible car

$$rang(A) \leq min\{m, n\} = min\{4, 6\} = 4$$

Une telle matrice ne peut donc pas exister.

Théorème 2.5.3. Si A et B sont deux matrices semblables, alors :

1. $det(A) = det(B)$
2. $trace(A) = trace(B)$
3. $rang(A) = rang(B)$
4. $nullite(A) = nullite(B)$

Démonstration.

1. Cette partie est facile et vous est laissé en exercice
2. Le cas général est plutôt technique et ne sera pas nécessaire pour la suite du cours. Nous ne le feront donc pas ici. Vous devriez cependant essayer en exercice de le démontrer pour des matrices A et B de dimension 2×2.
3. Supposons que A et B sont semblable, alors il existe une matrice P inversible tel que $B = PAP^{-1}$. Par le théorème 2.5, comme P est inversible, alors on a $rang(A) = rang(PA)$. De plus, comme la matrice P^{-1} est elle aussi inversible, alors on a :

$$rang(A) = rang(PA) = rang((PA)P^{-1}) = rang(PAP^{-1}) = rang(B)$$

4. Il s'agit d'une conséquence du théorème du rang et de la partie précédente. Premièrement, on remarque que si A et B sont des matrices semblable, alors elle doivent avoir la même dimension, et donc le même nombre de colonne, disons n. Par le théorème du rang appliqué à chacune des deux matrices, on obtient donc :

$$rang(A) + nullite(A) = n = rang(B) + nullite(B)$$

par la partie précédente, comme $rang(A) = rang(B)$, en simplifiant l'équation ci-dessus, on obtient donc :

$$nullite(A) = nullite(B)$$

□

Nous sommes maintenant en mesure d'ajouter quelques éléments à notre théorème de la matrice inverse. C'est ce que nous allons faire pour compléter le chapitre.

Théorème 2.5.4. **(Théorème de la matrice inverse)** Si A est une matrice de dimension $n \times n$, alors les énoncés suivant sont équivalents :

1. A est une matrice inversible
2. $det(A) \neq 0$
3. Le système d'équations linéaires $Ax = 0$ a une unique solution
4. Le système d'équations linéaires $Ax = b$ a une unique solution pour tout $b \in \mathbb{R}^n$.
5. Les colonnes de A forment une base de $\mathbb{R}^n$
6. Les lignes de A forment une base de $\mathbb{R}^n$.
7. $rang(A) = n$
8. $nullite(A) = 0$

Démonstration. Comme les 6 premiers énoncés ont déjà été montré équivalent, ils ne nous reste plus qu'à démontrer que les deux derniers sont équivalent aux précédents. On a donc : **(1) $\Rightarrow$ (8) :** Si A est une matrice inversible, alors le système d'équations $Ax = 0$ a une unique solution (par le théorème de la matrice inverse). Cette solution doit donc être $x = 0$. On peut donc déduire que $ker(A) = \{0\}$. On peut donc déduire que $nullite(A) = 0$.
(8) $\Rightarrow$ (1) : Si la nullité de A est 0, alors $ker(A) = \{0\}$ et donc le système d'équation $Ax = 0$ a une unique solution. Par le théorème de la matrice inverse, la matrice est donc inversible.
(7) $\iff$ (8) : Il s'agit d'une application du théorème du rang. On a donc $rang(A) + nullite(A) = n$. En remplaçant on obtient directement l'équivalence. □

Exemple 2.5.3. Complétez le tableau ci dessous :

Dimension de A	$rang(A)$	$nullite(A)$	$rang(A^T)$	$nullite(A^T)$	inversible ?
$4 \times a$	b	c	d	e	Oui
$f \times g$	5	1	h	2	i
7×8	4	j	k	l	m

a : Comme la matrice est inversible, on déduit que la matrice doit être carré, et donc $a = 4$
b : Comme la matrice est inversible, $b = rang(A) = a = 4$
c : Comme la matrice est inversible, $c = 0$
d : Par l'équation $rang(A) = rang(A^T)$, on obtient $d = b = 4$
e : Comme A est inversible, alors A^T est aussi inversible, et donc $e = 0$
g : Par le théorème du rang, on a $g = rang(A) + nullite(A) = 5 + 1 = 6$
h : Par l'équation $rang(A) = rang(A^T)$, on a $h = 5$
i : Comme $nullite(A) \neq 0$, alors la matrice n'est pas inversible
f : La dimension de A^T est $g \times f$ donc $6 \times f$. Maintenant on applique le théorème du rang : $rang(A^T) + nullite(A^T) = f$ et donc : $5 + 2 = 7$
j : Par le théorème du rang on a : $rang(A) + nullite(A) = 8$ on obtient donc $j = 8 - 4 = 4$
k : Par l'équation $rang(A) = rang(A^T)$ on obtient : $k = 4$
l : Comme la dimension de A^T est 8×7, on applique le théorème du rang ce qui nous donne $rang(A^T) + nullite(A^T) = 7$ et donc $l = 7 - rang(A^T) = 7 - 4 = 3$
m : Comme il ne s'agit pas d'une matrice carré, alors la matrice ne peut pas être inversible

On obtient donc le tableau complété suivant :

Dimension de A	$rang(A)$	$nullite(A)$	$rang(A^T)$	$nullite(A^T)$	inversible ?
4×4	4	0	4	0	Oui
7×6	5	1	5	2	Non
7×8	4	4	4	3	Non

Exemple 2.5.4. Considérez la matrice $A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \end{pmatrix}$. On veut premièrement trouver une base de $row(A)$ et $col(A)$. Pour ce faire, on commence par échelonner notre matrice :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \end{pmatrix} \sim^{6L_1 - L_2 \to L_2} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 5 & 10 & 15 & 20 \end{pmatrix}$$

On obtient donc les bases suivantes :

$$\mathcal{B}_{col(A)} = \left\{ \begin{pmatrix} 1 \\ 6 \end{pmatrix}, \begin{pmatrix} 2 \\ 7 \end{pmatrix} \right\} \quad \text{et} \quad \mathcal{B}_{row(A)} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 10 \\ 15 \\ 20 \end{pmatrix} \right\}$$

En particulier, ceci confirme que $rang(A) = rang(A^T) = 2$. Remarquons aussi que $col(A) \subseteq \mathbb{R}^2$ et $row(A) \subseteq \mathbb{R}^5$. Cherchons maintenant une base de $null(A)$. On doit donc résoudre le système d'équations linéaires $Ax = 0$:

$$\left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 0 \\ 6 & 7 & 8 & 9 & 10 & 0 \end{array}\right) \sim^{6L_1 - L_2 \to L_2} \left(\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 0 \\ 0 & 5 & 10 & 15 & 20 & 0 \end{array}\right)$$

Ce qui nous donne l'ensemble solution suivant :

$$Sol = \left\{ \begin{pmatrix} r + 2s + 3t \\ -2r - 3s - 4t \\ r \\ s \\ t \end{pmatrix} : r, s, t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} r + \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix} s + \begin{pmatrix} 3 \\ -4 \\ 0 \\ 0 \\ 1 \end{pmatrix} t : r, s, t \in \mathbb{R} \right\}$$

On obtient donc la base suivante :

$$\mathcal{B}_{Null(A)} = \left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -4 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Ce qui nous permet d'obtenir nullité$(A) = 3$. On peut donc vérifier que le théorème du rang est bel et bien satisfait dans ce cas :

$$\text{rang}(A) + \text{nullité}(A) = 2 + 3 = 5 = \text{nombre de colonne de } A$$

On veut maintenant vérifier que l'ensemble $\mathcal{B}'$ ci-dessous est aussi une base de $Null(A)$.

$$\left\{ \begin{pmatrix} 11 \\ -17 \\ 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 18 \\ -26 \\ 3 \\ 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 5 \\ -7 \\ 1 \\ -1 \\ 2 \end{pmatrix} \right\}$$

Comme nous avons le bon nombre de vecteur, nous savons que les vecteurs sont linéairement indépendant si et seulement si ils sont générateurs. Le problème étant que pour être une base, nous devons aussi nous assurer que ces vecteurs font bel et bien parti de $Null(A)$. Plusieurs approche sont possible, mais l'une des plus simple est de s'assurer que tout les vecteurs de la base $\mathcal{B}_{Null(A)}$ peuvent s'écrire comme combinaison linéaire des vecteurs de $\mathcal{B}'$. Plutôt que de résoudre 3 systèmes d'équations linéaires essentiellement identique à l'exception des constantes, nous allons résoudre les trois simultanément :

$$\left(\begin{array}{ccc|ccc} 11 & 18 & 5 & 1 & 2 & 3 \\ -17 & -26 & -7 & -2 & -3 & -4 \\ 2 & 3 & 1 & 1 & 0 & 0 \\ 3 & 0 & -1 & 0 & 1 & 0 \\ 1 & 5 & 2 & 0 & 0 & 1 \end{array}\right) \underset{\substack{3L_1-11L_4\to L_4 \\ L_1-11L_5\to L_5}}{\overset{\substack{17L_1+11L_2\to L_2 \\ 2L_1-11L_3\to L_3}}{\sim}} \left(\begin{array}{ccc|ccc} 11 & 18 & 5 & 1 & 2 & 3 \\ 0 & 20 & 8 & -5 & 1 & 7 \\ 0 & 3 & -1 & -9 & 4 & 6 \\ 0 & 54 & 26 & 3 & -5 & 9 \\ 0 & -37 & -17 & 1 & 2 & -8 \end{array}\right)$$

$$\underset{37L_2+20L_5\to L_5}{\overset{\substack{3L_2-20L_3\to L_3 \\ 27L_2-10L_4\to L_4}}{\sim}} \left(\begin{array}{ccc|ccc} 11 & 18 & 5 & 1 & 2 & 3 \\ 0 & 20 & 8 & -5 & 1 & 7 \\ 0 & 0 & 44 & 165 & -77 & -99 \\ 0 & 0 & -44 & -165 & 77 & 99 \\ 0 & 0 & -44 & -165 & 77 & 99 \end{array}\right) \underset{L_3+L_5\to L_5}{\overset{L_3+L_4\to L_4}{\sim}} \left(\begin{array}{ccc|ccc} 11 & 18 & 5 & 1 & 2 & 3 \\ 0 & 20 & 8 & -5 & 1 & 7 \\ 0 & 0 & 44 & 165 & -77 & -99 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Comme chacun des 3 systèmes d'équations linéaires possède bien une solution, nous pouvons donc affirmer que les 3 vecteurs de $\mathcal{B}'$ font partie de $Null(A)$ et qu'ils sont générateurs. De plus, comme nous savions déjà que nullité$(A) = 3$, nous avons le bon nombre de vecteur, et donc ils doivent aussi être linéairement indépendant. Il s'agit donc bien d'une base de $Null(A)$.

Chapitre 3

Valeurs et vecteurs propres

Dans le chapitre précédent, nous avons étudié les applications linéaires, et nous avons remarqué qu'il est parfois plus simple de travailler dans certaines bases plutôt qu'une autre. La question est maintenant de savoir dans quel base l'application est la plus simple ? En fait, ici plusieurs réponse sont possible. Une première option est de chercher une base dans laquelle la matrice correspondant à l'application linéaire sera diagonale. C'est ce que nous ferons dans ce chapitre. Une autre option est de chercher une base dans laquelle la matrice de passage sera facilement inversible, ce qui est le cas lorsque la matrice P est orthogonale. C'est ce que nous ferons au chapitre suivant.

3.1 Valeurs et vecteurs propres

Definition 3.1.1. Si A est une matrice, λ un nombre complexe et v un vecteur non nul, alors on dit que λ est une valeur propre et v est un vecteur propre associé à la valeur propre λ si l'équation ci dessous est satisfaite :

$$Av = \lambda v$$

Théorème 3.1.1. Si A est une matrice carré, alors :

1. l'ensemble des valeurs propres λ de la matrice A sont donné par les racines du polynôme

$$det(A - \lambda I) = 0$$

   Ce polynôme porte le nom de polynôme caractéristique.

2. l'ensemble des vecteurs propres associés à une valeur propre λ est l'ensemble des vecteurs non nul de l'espace vectoriel

$$ker(A - \lambda I)$$

Démonstration.

1. Supposons que λ est un nombre réel et v un vecteur non nul tel que $Av = \lambda v$, alors on peut écrire :

$$Av = \lambda v \quad \Longrightarrow \quad Av - \lambda I v = 0 \quad \Longrightarrow \quad (A - \lambda I)v = 0$$

   Comme $(A - \lambda I)$ est une matrice, l'équation $(A - \lambda I)v = 0$ est en fait un système d'équations linéaires homogène. De plus, comme v est une solution non nul de ce système d'équation, on peut donc conclure que le système d'équations linéaires possède au moins deux solutions, c'est à dire qu'il doit en avoir une infinité. Par le théorème de la matrice inverse, on peut donc conclure que la matrice $(A - \lambda I)$ n'est pas inversible, et donc $det(A - \lambda I) = 0$.

2. En utilisant la même notation que dans la partie précédente, nous avons vu que $Av = \lambda v \Longleftrightarrow (A-\lambda I)v = 0$. Cette dernière équation n'est en fait rien d'autre que la définition du noyau de la matrice $(A-\lambda I)$. On peut donc conclure que l'ensemble des vecteurs propres de la matrice est en fait l'ensemble des vecteurs non nul de l'espace vectoriel $ker(A-\lambda I)$.

□

Nous somme maintenant en mesure d'ajouter un élément de plus à notre théorème de la matrice inverse. C'est ce que nous allons faire immédiatement.

Théorème 3.1.2. **(Théorème de la matrice inverse)** Si A est une matrice de dimension $n \times n$, alors les énoncés suivants sont équivalents :

1. A est une matrice inversible
2. $\det(A) \neq 0$
3. Le système d'équations linéaires $Ax = 0$ a une unique solution
4. Le système d'équations linéaires $Ax = b$ a une unique solution pour tout $b \in \mathbb{R}^n$.
5. Les colonnes de A forment une base de $\mathbb{R}^n$
6. Les lignes de A forment une base de $\mathbb{R}^n$.
7. $\text{rang}(A) = n$
8. $\text{nullite}(A) = 0$
9. 0 n'est pas une valeur propre de la matrice A

Démonstration. La seule partie qui n'a pas été démontrer est la dernière. Nous allons donc démontrer seulement que (2) $\Leftrightarrow$ (9). Par le théorème précédent, l'ensemble des valeurs propres de la matrice sont donné par les racines du polynôme $det(A-\lambda I) = 0$. Si $\lambda = 0$, ceci cet ramène à l'équation $det(A) = 0$. Donc $\lambda = 0$ est une valeur propre de la matrice si et seulement si $det(A) = 0$. □

Nous allons maintenant compléter cette section avec un théorème nous permettant de relier les valeurs propres et vecteurs propres d'une matrice A, avec les valeurs propres / vecteurs propres de la matrice A^k.

Théorème 3.1.3. Si A est une matrice carré et v un vecteur propre de A associé à une valeur propre λ alors on a :

1. Pour chaque entier positif k, v est un vecteur propre de la matrice A^k associé à la valeur propre λ^k.
2. Si A est inversible, alors v est un vecteur propre de la matrice A^{-1} associé à la valeur propre $\dfrac{1}{\lambda}$.

Démonstration.

1. Par définition des valeurs / vecteurs propres, on se rappelle que $Av = \lambda v$, ce qui nous permet d'obtenir :

$$\begin{aligned} A^k v &= A^{k-1}(Av) = A^{k-1}\lambda v = \lambda A^{k-1} v \\ &= \lambda A^{k-2}(Av) = \lambda A^{k-2}(\lambda v) = \lambda^2 A^{k-2} v \\ &= .... \\ &= \lambda^k v \end{aligned}$$

Ce qui confirme que v est un vecteur propre de A^k associé à la valeur propre λ^k.

2. Par définition des valeurs / vecteurs propres, on a :

$$Av = \lambda v$$

Comme la matrice A est inversible, on sait que $\lambda \neq 0$:

$$\begin{aligned} \frac{1}{\lambda}Av &= v \\ A\Big(\frac{1}{\lambda}u\Big) &= v \\ A^{-1}A\Big(\frac{1}{\lambda}v\Big) &= A^{-1}v \\ \frac{1}{\lambda}v &= A^{-1}v \end{aligned}$$

Et donc, v est un vecteur propre de A^{-1} associé à la valeur propre $\frac{1}{\lambda}$.

□

3.2 Diagonalisation : Le cas simple

Nous allons maintenant nous intéresser à savoir comment utiliser les valeurs propres et vecteurs propres d'une matrice pour trouver une factorisation de la forme $A = PDP^{-1}$ où D est une matrice diagonale et P est une matrice inversible. Le plus grand obstacle dans notre étude vient du fait que cela n'est pas toujours possible en général. [1] Nous allons commencer par regarder le cas où toutes les valeurs propres de la matrice sont distinctes, c'est à dire le cas où une matrice A de dimension $n \times n$ possède n valeurs propres distinctes. Dans ce cas, comme l'énonce le théorème ci-dessous, une telle factorisation est toujours possible, et certaine subtilité de la méthode peuvent être complètement ignoré.

Théorème 3.2.1. Supposons que A est une matrice carré $n \times n$ contenant exactement n valeurs propres distinctes, alors il existe une base dans laquelle la matrice est diagonale. C'est à dire qu'il existe une matrice inversible P et une matrice diagonale D tel que :

$$A = PDP^{-1}$$

La matrice D est une matrice contenant les valeurs propres sur la diagonale, et la matrice P est une matrice ayant pour colonne un vecteur propre pour chacune des valeurs propres, et placé dans le même ordre.

Démonstration. Le problème reviens à démontrer que la matrice P formé des vecteurs propres est une matrice inversible, ce qui est équivalent à montrer que les vecteurs propres forme une base de $\mathbb{R}^n$, ou bien, comme nous avons n vecteurs dans un espace vectoriel de dimension n, il est donc suffisant de démontrer que les vecteurs propres sont linéairement indépendant. Pour ce faire, supposons que $\{v_1, v_2, ..., v_n\}$ sont des vecteurs propres de la matrice associé aux valeurs propres distinctes $\{\lambda_1, \lambda_2, ..., \lambda_n\}$. Supposons que l'ensemble des vecteurs propres n'est pas linéairement indépendant, donc l'un des vecteurs propres peut s'écrire comme combinaison linéaire des précédents. Supposons que v_{r+1}, $r < n$ est le premier vecteur propre qui peut s'écrire comme combinaison linéaire des précédents, on a donc :

$$v_{r+1} = c_1v_1 + c_2v_2 + ... + c_rv_r \tag{3.1}$$

1. Une autre factorisation plus générale appelé décomposition de Jordan permet de résoudre ce problème. La décomposition de Jordan d'une matrice A existe toujours et consiste à trouver non pas une matrice D qui est diagonale, mais plutôt une matrice J qui est presque diagonale. Nous n'étudierons cependant pas cette factorisation dans ce cours.

Maintenant, en multipliant cette équation par A, nous obtenons :

$$Av_{r+1} = c_1 Av_1 + c_2 Av_2 + ... + c_r Av_r$$

et donc par définition de la valeur propre :

$$\lambda_{r+1} v_{r+1} = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + ... + c_r \lambda_r v_r$$

Maintenant, si nous reprenons l'équation 3.1 et nous la multiplions par λ_{r+1} on obtient :

$$\lambda_{r+1} v_{r+1} = c_1 \lambda_{r+1} v_1 + c_2 \lambda_{r+1} v_2 + ... + c_r \lambda_{r+1} v_r$$

Et finalement, on soustrait c'est deux dernière équation, ce qui nous permet d'obtenir :

$$0 = c_1(\lambda_1 - \lambda_{r+1}) v_1 + c_2(\lambda_2 - \lambda_{r+1}) v_2 + ... c_r(\lambda_r - \lambda_{r+1}) v_r$$

Par indépendance linéaire des vecteurs $\{v_1, v_2, ... v_r\}$, tous les $c_i(\lambda_i - \lambda_{r+1})$ doivent être nul. De plus, comme toutes les valeurs propres sont distinctes, ce sont tous les c_i qui doivent être nul. Par l'équation 3.1, on conclut donc que $v_{r+1} = 0$, ce qui contredit la définition d'un vecteur propre. On conclut donc que les vecteurs propres sont linéairement indépendant, et donc que la matrice P est inversible. Maintenant, par simple calcul, on peut vérifier que $PA = PD$ en utilisant la définition de la valeur propre, ce qui nous permet d'obtenir

$$A = PDP^{-1}$$

□

Exemple 3.2.1. On veut diagonaliser la matrice suivante :

$$A = \begin{pmatrix} 39 & 54 \\ -15 & -18 \end{pmatrix}$$

On commence par trouver les valeurs propres :

$$\begin{vmatrix} 39 - \lambda & 54 \\ -15 & -18 - \lambda \end{vmatrix} = (39 - \lambda)(-18 - \lambda) + 810 = \lambda^2 - 21\lambda + 108 = (\lambda - 9)(\lambda - 12) = 0$$

Les valeurs propres sont donc 9 et 12. On doit maintenant trouver les vecteurs propres. Pour la valeur propre 9, on a donc :

$$\begin{pmatrix} 30 & 54 \\ -15 & -27 \end{pmatrix} \sim \begin{pmatrix} 30 & 54 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 5 & 9 \\ 0 & 0 \end{pmatrix}$$

Donc en posant $y = s$, on a $x = \frac{-9s}{5}$. Si on prend $s = 5$, on obtient donc le vecteur propre suivant :

$$v_1 = \begin{pmatrix} -9 \\ 5 \end{pmatrix}$$

On refait maintenant la même chose avec la valeur propre 12.

$$\begin{pmatrix} 27 & 54 \\ -15 & -30 \end{pmatrix} \sim \begin{pmatrix} 27 & 54 \\ 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

En posant $y = s$, on a $x = -2y = -2s$, et donc si on prend $s = 1$, on obtient le vecteur propre suivant :

$$v_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

On peut donc maintenant écrire la factorisation de la matrice :

$$\begin{pmatrix} 39 & 54 \\ -15 & -18 \end{pmatrix} = \begin{pmatrix} -9 & -2 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 \\ 0 & 12 \end{pmatrix} \begin{pmatrix} -9 & -2 \\ 5 & 1 \end{pmatrix}^{-1}$$

Exemple 3.2.2. On veut utiliser l'exemple précédent pour calculer A^{100} si

$$A = \begin{pmatrix} 39 & 54 \\ -15 & -18 \end{pmatrix}$$

Pour ce faire, remarquons que

$$\begin{aligned} A^{100} &= (PDP^{-1})^{100} = \underbrace{(PDP^{-1})(PDP^{-1})...(PDP^{-1})}_{100 \text{ fois}} \\ &= PD(P^{-1}P)D(P^{-1}P)...D(P^{-1}P) \\ &= PD^{100}P^{-1} \end{aligned}$$

On obtient donc :

$$\begin{aligned} \begin{pmatrix} 39 & 54 \\ -15 & -18 \end{pmatrix}^{100} &= \begin{pmatrix} -9 & -2 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 9 & 0 \\ 0 & 12 \end{pmatrix}^{100} \begin{pmatrix} -9 & -2 \\ 5 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} -9 & -2 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} 9^{100} & 0 \\ 0 & 12^{100} \end{pmatrix} \begin{pmatrix} -9 & -2 \\ 5 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} -9^{101} & -2 \cdot 12^{100} \\ 5 \cdot 9^{100} & 12^{100} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -5 & -9 \end{pmatrix} \\ &= \begin{pmatrix} -9^{101} + 10 \cdot 12^{100} & -2 \cdot 9^{101} + 18 \cdot 12^{100} \\ 5 \cdot 9^{100} - 5 \cdot 12^{100} & 10 \cdot 9^{100} - 9 \cdot 12^{100} \end{pmatrix} \end{aligned}$$

Exemple 3.2.3. Diagonaliser la matrice suivante :

$$A = \begin{pmatrix} 2 & -2 & 4 \\ -6 & -4 & 18 \\ -3 & -2 & 9 \end{pmatrix}$$

On commence par calculer les valeurs propres :

$$\begin{aligned} & \begin{vmatrix} 2-\lambda & -2 & 4 \\ -6 & -4-\lambda & 18 \\ -3 & -2 & 9-\lambda \end{vmatrix} \\ = \; & (2-\lambda) \begin{vmatrix} -4-\lambda & 18 \\ -2 & 9-\lambda \end{vmatrix} + 2 \begin{vmatrix} -6 & 18 \\ -3 & 9-\lambda \end{vmatrix} + 4 \begin{vmatrix} -6 & -4-\lambda \\ -3 & -2 \end{vmatrix} \\ = \; & (2-\lambda)\Big[(-4-\lambda)(9-\lambda) + 36\Big] + 2\Big[-6(9-\lambda) + 54\Big] + 4\Big[12 + 3(-4-\lambda)\Big] \\ = \; & (2-\lambda)\Big[\lambda^2 - 5\lambda\Big] + 2\Big[6\lambda\Big] + 4\Big[-3\lambda\Big] \\ = \; & (-\lambda^3 + 7\lambda^2 - 10\lambda) + (12\lambda) + (-12\lambda) \\ = \; & -\lambda^3 + 7\lambda^2 - 10\lambda \\ = \; & -\lambda(\lambda-2)(\lambda-5) = 0 \end{aligned}$$

On en déduit donc que les valeurs propres sont $0, 2$ et 5. On va maintenant chercher un vecteur propres pour chacune des valeurs propres, en commençant par 0, on doit donc trouver une base de $ker(A - 0I)$:

$$\begin{pmatrix} 2-0 & -2 & 4 \\ -6 & -4-0 & 18 \\ -3 & -2 & 9-0 \end{pmatrix} = \begin{pmatrix} 2 & -2 & 4 \\ -6 & -4 & 18 \\ -3 & -2 & 9 \end{pmatrix}$$

Et en appliquant la méthode de Gauss, on obtient :

$$\begin{pmatrix} 2 & -2 & 4 \\ -6 & -4 & 18 \\ -3 & -2 & 9 \end{pmatrix} \sim^{\substack{3L_1+L_2\to L_2 \\ 3L_1+2L_3\to L_3}} \begin{pmatrix} 2 & -2 & 4 \\ 0 & -10 & 30 \\ 0 & -10 & 30 \end{pmatrix}$$

$$\sim^{L_2-L_3\to L_3} \begin{pmatrix} 2 & -2 & 4 \\ 0 & -10 & 30 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim^{\substack{\frac{1}{2}L_1\to L_1 \\ \frac{-1}{10}L_2\to L_2}} \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix}$$

En posant $z = s$, on obtient : $y = 3z = 3s$ et $x = y - 2z = 3s - 2s = s$, donc en prenant $s = 1$, on obtient notre premier vecteur propre :

$$v_1 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

Nous allons maintenant refaire la même chose pour la valeur propre 2 :

$$\begin{pmatrix} 2-2 & -2 & 4 \\ -6 & -4-2 & 18 \\ -3 & -2 & 9-2 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 4 \\ -6 & -6 & 18 \\ -3 & -2 & 7 \end{pmatrix}$$

Et en appliquant la méthode de Gauss, on obtient :

$$\begin{pmatrix} 0 & -2 & 4 \\ -6 & -6 & 18 \\ -3 & -2 & 7 \end{pmatrix} \sim^{L_1\leftrightarrow L_3} \begin{pmatrix} -3 & -2 & 7 \\ -6 & -6 & 18 \\ 0 & -2 & 4 \end{pmatrix}$$

$$\sim^{2L_1-L_2\to L_2} \begin{pmatrix} -3 & -2 & 7 \\ 0 & 2 & -4 \\ 0 & -2 & 4 \end{pmatrix}$$

$$\sim^{L_2+L_3\to L_3} \begin{pmatrix} -3 & -2 & 7 \\ 0 & 2 & -4 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sim^{\frac{1}{2}L_2\to L_2} \begin{pmatrix} -3 & -2 & 7 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

En posant $z = s$, on obtient : $y = 2z = 2s$ et $-3x = 2y - 7z = 4s - 7s = -3s$ et donc $x = s$, et donc en prenant $s = 1$, on obtient le vecteur propre :

$$v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

Et finalement, on refait encore une fois la même chose, mais cette fois avec la valeur propre 5 :

$$\begin{pmatrix} 2-5 & -2 & 4 \\ -6 & -4-5 & 18 \\ -3 & -2 & 9-5 \end{pmatrix} = \begin{pmatrix} -3 & -2 & 4 \\ -6 & -9 & 18 \\ -3 & -2 & 4 \end{pmatrix}$$

Et en appliquant la méthode de Gauss, on obtient :

$$\begin{pmatrix} -3 & -2 & 4 \\ -6 & -9 & 18 \\ -3 & -2 & 4 \end{pmatrix} \sim^{\substack{2L_1-L_2\to L_2 \\ L_1-L_3\to L_3}} \begin{pmatrix} -3 & -2 & 4 \\ 0 & 5 & -10 \\ 0 & 0 & 0 \end{pmatrix}$$

En posant $z = s$, on obtient donc : $y = 2z = 2s$ et $-3x = 2y - 4z = 4s - 4s = 0$ ce qui nous donne $x = 0$. En prenant $s = 1$, on obtient donc finalement le vecteur propre :

$$v_3 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

On peut maintenant écrire la factorisation de la matrice :

$$\begin{pmatrix} 2 & -2 & 4 \\ -6 & -4 & 18 \\ -3 & -2 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 3 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 3 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix}^{-1}$$

3.3 Diagonalisation : Le cas général

Nous allons maintenant regarder le cas général. Comme nous l'avons déjà mentionné, il n'est pas toujours possible de diagonaliser une matrice. Une méthode (simple) pour vérifier si une matrice est diagonalisable consiste à étudier la multiplicité de chacune des valeurs propres. Nous allons donc commencer cette section en définissant les deux notions de multiplicité d'une valeur propre.

Definition 3.3.1. Si A est une matrice de dimension $n \times n$, alors le polynômes caractéristique de A peut s'écrire sous la forme :

$$p(\lambda) = a(\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2}(\lambda - \lambda_3)^{\alpha_3} ... (\lambda - \lambda_k)^{\alpha_k}$$

où les λ_i sont les valeurs propres (distinctes) de la matrice. Dans ce cas on définit :

1. la multiplicité algébrique de la valeur propre λ_i comme étant α_i.
2. la multiplicité géométrique de la valeur propre λ_i comme étant la dimension de l'espace vectoriel $ker(A - \lambda_i I)$.

Ces définitions nous permettent maintenant dénoncer notre premier théorème permettant de caractériser les matrices qui sont diagonalisable.

Théorème 3.3.1. Si A est une matrice carré, alors les énoncés suivant sont équivalent :

1. La matrice A est diagonalisable. C'est à dire qu'il existe une matrice D qui est diagonale et une matrice P qui est inversible tel que $A = PDP^{-1}$.
2. Pour toute valeur propre λ, la multiplicité algébrique est égale à la multiplicité géométrique
3. Il existe une base de $\mathbb{R}^n$ (ou $\mathbb{C}^n$) consistant de vecteur propre de la matrice A.

Dans le cas où la matrice A est diagonalisable, la matrice D sera obtenu en plaçant les valeurs propres sur la diagonale en tenant compte de leur multiplicité, et la matrice P sera obtenu en plaçant à la verticale une base de $ker(A - \lambda I)$ pour chacune des valeurs propres λ et dans le même ordre qu'on les a placé dans la matrice D.

Exemple 3.3.1. On veut diagonaliser (si cela est possible) la matrice

$$A = \begin{pmatrix} -1 & 4 & -6 \\ -8 & 11 & -12 \\ -4 & 4 & -3 \end{pmatrix}$$

On commence par trouver le polynôme caractéristique :

$$p(\lambda) = \begin{vmatrix} -1-\lambda & 4 & -6 \\ -8 & 11-\lambda & -12 \\ -4 & 4 & -3-\lambda \end{vmatrix} = (-1-\lambda)\begin{vmatrix} 11-\lambda & -12 \\ 4 & -3-\lambda \end{vmatrix} - 4\begin{vmatrix} -8 & -12 \\ -4 & -3-\lambda \end{vmatrix} - 6\begin{vmatrix} -8 & 11-\lambda \\ -4 & 4 \end{vmatrix}$$

$$
\begin{aligned}
&= (-1-\lambda)[(11-\lambda)(-3-\lambda)+48] - 4[-8(-3-\lambda)-48] - 6[-32+4(11-\lambda)] \\
&= (-1-\lambda)(\lambda^2-8\lambda+15) - 4(8\lambda-24) - 6(-4\lambda+12) \\
&= (-\lambda^3+7\lambda^2-7\lambda-15) + (-32\lambda+96) + (24\lambda-72) \\
&= -\lambda^3+7\lambda^2-15\lambda+9 = -(\lambda-3)^2(\lambda-1) = 0
\end{aligned}
$$

Les valeurs propres sont donc 3 avec une multiplicité algébrique de 2 et 1 avec une multiplicité algébrique de 1. On va maintenant essayer de trouver une base de vecteurs propres.

$$
\begin{pmatrix} -1-3 & 4 & -6 \\ -8 & 11-3 & -12 \\ -4 & 4 & -3-3 \end{pmatrix} = \begin{pmatrix} -4 & 4 & -6 \\ -8 & 8 & -12 \\ -4 & 4 & -6 \end{pmatrix} \sim \begin{pmatrix} -2 & 2 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}
$$

Donc les vecteurs propres associé à la valeur propre 3 doivent respecter l'équation

$$
-2x + 2y - 3z = 0
$$

on peut donc prendre les deux vecteurs propres suivants :

$$
v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix}
$$

Donc la multiplicité géométrique de la valeur propre 3 est 2. On cherche maintenant un vecteur propre associé à la valeur propre 1 :

$$
\begin{pmatrix} -1-1 & 4 & -6 \\ -8 & 11-1 & -12 \\ -4 & 4 & -3-1 \end{pmatrix} = \begin{pmatrix} -2 & 4 & -6 \\ -8 & 10 & -12 \\ -4 & 4 & -4 \end{pmatrix} \sim \begin{pmatrix} -1 & 2 & -3 \\ -4 & 5 & -6 \\ -1 & 1 & -1 \end{pmatrix}
$$

$$
\sim_{\substack{-4L_1+L_2\to L_2 \\ L_1-L_3\to L_3}} \begin{pmatrix} -1 & 2 & -3 \\ 0 & -3 & 6 \\ 0 & 1 & -2 \end{pmatrix} \sim^{L_2+3L_3\to L_3} \begin{pmatrix} -1 & 2 & -3 \\ 0 & -3 & 6 \\ 0 & 0 & 0 \end{pmatrix}
$$

En posant $z = 1$, on obtient $y = 2$ et $x = 1$ ce qui nous donne le vecteur :

$$
v_3 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}
$$

Donc la multiplicité géométrique de la valeur propre 1 est 1. La matrice est donc bien diagonalisable. Ce qui nous donne finalement la factorisation suivante :

$$
\begin{pmatrix} -1 & 4 & -6 \\ -8 & 11 & -12 \\ -4 & 4 & -3 \end{pmatrix} = \begin{pmatrix} 1 & -3 & 1 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}^{-1}
$$

Nous avons déjà vu au chapitre précédent plusieurs invariant de similitude. C'est à dire des propriétés que deux matrices semblables ont en commun. Nous allons maintenant ajouter les valeurs propres à notre liste.

Théorème 3.3.2. Si A et B sont deux matrices semblables, alors :

1. $det(A) = det(B)$
2. $trace(A) = trace(B)$
3. $rang(A) = rang(B)$
4. $nullite(A) = nullite(B)$
5. Les matrices A et B ont les mêmes valeurs propres

Démonstration. Nous allons seulement démontrer la dernière propriété. Les autres ayant déjà été démontré dans le chapitre précédent. Supposons que A et B sont des matrices carrés de même dimension pour lesquels ils existent une matrice P inversible tel que $A = PBP^{-1}$. alors on a :

$$\begin{aligned}
\det(A - \lambda I) &= \det(PBP^{-1} - \lambda I) \\
&= \det(PBP^{-1} - P\lambda IP^{-1}) \\
&= \det(P(B - \lambda I)P^{-1}) \\
&= \det(P)\det(B - \lambda I)\det(P^{-1}) \\
&= \det(B - \lambda I)
\end{aligned}$$

□

Les propriétés du théorème précédent ne garantissent cependant pas que les deux matrices en question sont semblable. Elle ne nous permettent seulement d'affirmer que si l'une de ces propriétés n'est pas respecté, alors les matrices ne sont pas semblable. Si les matrices A et B sont diagonalisable, on peut cependant aller un peu plus loin.

Théorème 3.3.3. Si A et B sont des matrices diagonalisable ayant exactement les mêmes valeurs propres (en tenant compte de leur multiplicité algébrique), alors les deux matrices sont semblables.

Démonstration. Si A et B sont des matrices diagonalisable, alors il existe des matrices diagonales D_A et D_B, ainsi que des matrices inversible P_A et P_B tel que $A = P_A D_A P_A^{-1}$ et $B = P_B D_B P_B^{-1}$. En isolant la matrice diagonalisable dans chacun des cas, on obtient donc :

$$D_A = P_A^{-1} A P_A \qquad \text{et} \qquad D_B = P_B^{-1} B P_B$$

Maintenant, comme les matrices A et B ont par hypothèse les mêmes valeurs propres, incluant leur multiplicité, on peut donc supposer que $D_A = D_B$ (Pour ce faire, les valeurs propres doivent être placé dans le même ordre). On obtient donc :

$$P_A^{-1} A P_A = P_B^{-1} B P_B \qquad \Rightarrow \qquad A = P_A P_B^{-1} B P_B P_A^{-1} = (P_A P_B^{-1}) B (P_A P_B^{-1})^{-1}$$

Donc si on pose $P = P_A P_B^{-1}$, on obtient finalement $A = PBP^{-1}$. Les matrices A et B sont donc semblables. □

Notez que le théorème est faux si on omet l'hypothèse que les matrices doivent être diagonalisable. En effet, les deux matrices ci-dessous ne sont pas semblable bien qu'elle ait les mêmes valeurs propres. Le problème étant que la première n'est pas diagonalisable.

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

3.4 La théorie des polynômes annulateurs

Nous avons déjà vu qu'une matrice est diagonalisable si et seulement si il existe une base de vecteurs propres dans $\mathbb{K}^n$. Nous allons maintenant voir une autre technique pour savoir si une matrice est diagonalisable, cette fois sans avoir besoin de trouver des bases pour chaque espace propre. Il s'agit de la théorie des polynômes annulateurs.

Definition 3.4.1. Si A est une matrice carré de dimension $n \times n$, et $p(x)$ est un polynôme, alors on dit que $p(x)$ est un polynôme annulateur si $p(A) = 0$.

Théorème 3.4.1. **(Cayley-Hamilton)** Une matrice carrée A satisfait toujours son équation caractéristique. C'est à dire que si $p(\lambda)$ est le polynôme caractéristique de la matrice A (C'est à dire $p(\lambda) = \det(A - \lambda I)$), alors $p(A) = 0$.

Le théorème de Cayley-Hamilton nous garantie donc l'existence de polynôme annulateur pour n'importe quelle matrice carré, et nous donne en même temps une méthode alternative pour calculer l'inverse d'une matrice. Cette méthode est surtout utile dans le cas où le polynôme caractéristique de la matrice nous est connu. Voici un exemple qui illustre la méthode.

Exemple 3.4.1. On veut calculer l'inverse de la matrice $A = \begin{pmatrix} 2 & 6 \\ 3 & 8 \end{pmatrix}$ en utilisant le théorème de Cayley-Hamilton. Pour ce faire, commencer par trouver son polynôme caractéristique :

$$p(\lambda) = det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 6 \\ 3 & 8-\lambda \end{vmatrix} = (2-\lambda)(8-\lambda) - 18 = \lambda^2 - 10\lambda - 2$$

Par le théorème d'Hamilton-Cayley, on peut donc affirmer que :

$$A^2 - 10A - 2I = 0$$

En factorisant le A à gauche et en envoyant le $2I$ à droite, on obtient donc :

$$A(A - 10I) = 2I \qquad \Longrightarrow \qquad A\left(\frac{1}{2}(A-10)I\right) = I$$

Par définition de l'inverse d'une matrice, on obtient donc :

$$A^{-1} = \frac{1}{2}(A - 10I) = \frac{1}{2}\begin{pmatrix} -8 & 6 \\ 3 & -2 \end{pmatrix} = \begin{pmatrix} -4 & 3 \\ 3/2 & -1 \end{pmatrix}$$

Definition 3.4.2. Si A est une matrice carré de dimension $n \times n$ alors on appelle polynôme minimal de A un polynôme annulateur non trivial de A ayant le plus petit degré possible et ayant pour coefficient de la plus grande puissance de x la valeur 1.

Théorème 3.4.2.

1. Si A est une matrice carré, $p(x)$ sont polynôme minimal, et $q(x)$ un autre polynôme annulateur de A, alors $p(x)$ divise $q(x)$.
2. Soit A une matrice carré et $p(x)$ son polynôme minimal. Si λ est une valeur propre de la matrice A, alors $p(\lambda) = 0$.

Démonstration.

1. Supposons que $p(x)$ et $q(x)$ sont des polynômes annulateurs de A avec $deg(q(x)) \geq deg(p(x))$. En effectuant la division de $q(x)$ par $p(x)$, on obtient donc des polynômes $r(x)$ et $m(x)$ tel que :

$$\frac{q(x)}{p(x)} = m(x) + \frac{r(x)}{p(x)}$$

avec $deg(r(x)) < deg(p(x))$. Dans ce cas, le $r(x)$ est appelé le reste de la division. En multipliant des deux côtés par $p(x)$, on obtient :

$$q(x) = m(x)p(x) + r(x)$$

Maintenant, comme par hypothèse $p(x)$ et $q(x)$ sont des polynômes annulateurs, en remplaçant le x par la matrice A, on obtient :

$$q(A) = m(A)p(A) + r(A) \qquad \Rightarrow \qquad r(A) = 0$$

On a donc trouver un polynôme de degré strictement inférieur à $p(x)$ qui est aussi un polynôme annulateur de la matrice A. Si on suppose maintenant que $p(x)$ est le polynôme minimal de la matrice A, on devra donc avoir $r(x) = 0$. C'est à dire que le polynôme $p(x)$ divise le polynôme $q(x)$.

2. Supposons que $m(x) = x^r + a_{r-1}x^{r-1} + \ldots + a_1x + a_0$ et supposons que que v est un vecteur propre associé à la valeur propre λ (donc $v \neq 0$). On a donc :

$$p(A) = A^r + a_{r-1}A^{r-1} + \ldots + a_1A + a_0I = 0$$

Donc en multipliant par v, on obtient :

$$\begin{aligned}
(A^r + a_{r-1}A^{r-1} + \ldots + a_1A + a_0I)v &= 0v \\
A^rv + a_{r-1}A^{r-1}v + \ldots + a_1Av + a_0Iv &= 0 \\
\lambda^rv + a_{r-1}\lambda^{r-1}v + \ldots + a_1\lambda v + a_0v &= 0 \\
(\lambda^r + a_{r-1}\lambda^{r-1} + \ldots + a_1\lambda + a_0)v &= 0
\end{aligned}$$

Comme $v \neq 0$, alors on a :

$$p(\lambda) = \lambda^r + a_{r-1}\lambda^{r-1} + \ldots + a_1\lambda + a_0 = 0$$

λ est donc une racine du polynôme minimal.

□

Ces deux derniers théorèmes nous donne en fait une bonne indication sur comment trouver le polynôme minimal d'une matrice. En particulier, toutes les racines du polynôme minimal sont des valeurs propres de la matrice A et inversement. De plus, le polynôme minimal doit diviser le polynôme caractéristique, car ce dernier est un polynôme annulateur. L'intérêt du polynôme minimal d'une matrice est qu'il nous donne une façon de vérifier si une matrice A est diagonalisable ou non. Nous allons donc pouvoir ajouter un élément à notre théorème de diagonalisibilité de la section précédente.

Théorème 3.4.3. Si A est une matrice carrée, alors les énoncés suivants sont équivalents :

1. La matrice A est diagonalisable. C'est à dire qu'il existe une matrice D qui est diagonale et une matrice P qui est inversible tel que $A = PDP^{-1}$.
2. Pour toute valeur propre λ, la multiplicité algébrique est égale à la multiplicité géométrique
3. Il existe une base de $\mathbb{R}^n$ (ou $\mathbb{C}^n$) composée de vecteurs propres de la matrice A.
4. Le polynôme minimal de A se factorise complètement dans $\mathbb{R}^n$ (ou $\mathbb{C}^n$) et n'a que des racines simples.

Il est relativement long de chercher le polynôme minimal d'une matrice. Dans la plupart des cas, ça ne sera donc pas la méthode la plus efficace. La dernière propriété du théorème peut s'utiliser à la place de la manière suivante. On commence par trouver le polynôme caractéristique. Par le théorème de Cayley-Hamilton, ce dernier est un polynôme annulateur et possède exactement les mêmes racines que le polynôme minimal. Donc le polynôme minimal sera essentiellement identique au polynôme caractéristique, sauf peut être pour les exposants qui changeront. Si la matrice est diagonalisable, tout les exposants du polynôme minimal doivent être des 1. Donc plutôt que de chercher le polynôme minimal, il suffit de vérifier si le polynôme caractéristique, auquel on remplace tous les exposants par des 1, est un polynôme annulateur. Si la réponse est positive, la matrice est diagonalisable (et il s'agit du polynôme minimal). Dans le cas contraire, la matrice n'est pas diagonalisable. Remarquez finalement que comme corollaire du théorème précédent, si une matrice $n \times n$ possède n valeurs propres distinctes, alors elle est diagonalisable car son polynôme minimal sera égal à un multiple contant près à son polynôme caractéristique.

Exemple 3.4.2. Est-ce que la matrice
$$A = \begin{pmatrix} -1 & 1 & 1 \\ -3 & 2 & 2 \\ -3 & 0 & 4 \end{pmatrix}$$
est diagonalisable ? Pour le vérifier, nous allons commencer par trouver son polynôme caractéristique :

$$p(\lambda) = \begin{vmatrix} -1-\lambda & 1 & 1 \\ -3 & 2-\lambda & 2 \\ -3 & 0 & 4-\lambda \end{vmatrix} = (-1-\lambda)\begin{vmatrix} 2-\lambda & 2 \\ 0 & 4-\lambda \end{vmatrix} - \begin{vmatrix} -3 & 2 \\ -3 & 4-\lambda \end{vmatrix} + \begin{vmatrix} -3 & 2-\lambda \\ -3 & 0 \end{vmatrix}$$

$$\begin{aligned}
&= (-1-\lambda)(2-\lambda)(4-\lambda) - (-12+3\lambda+6) + (6-3\lambda) \\
&= (-1-\lambda)(\lambda^2-6\lambda+8) + (-3\lambda+6) + (-3\lambda+6) \\
&= (-\lambda^3+5\lambda^2-2\lambda-8) + (-6\lambda+12) \\
&= -\lambda^3+5\lambda^2-8\lambda+4 \\
&= -(\lambda-2)^2(\lambda-1)
\end{aligned}$$

Donc par les théorèmes précédent, cette matrice est diagonalisable si et seulement si le polynôme

$$q(\lambda) = (\lambda-2)(\lambda-1)$$

est un polynôme annulateur.

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

Comme il ne s'agit pas d'un polynôme annulateur, la matrice n'est donc pas diagonalisable.

Exemple 3.4.3. Nous allons maintenant utiliser la théorie des polynômes annulateurs pour vérifier que la matrice de la section précédente est diagonalisable, mais cette fois sans calculer les vecteurs propres ni la factorisation. Posons :
$$A = \begin{pmatrix} -1 & 4 & -6 \\ -8 & 11 & -12 \\ -4 & 4 & -3 \end{pmatrix}$$
On commence par trouver le polynôme caractéristique :

$$p(\lambda) = \begin{vmatrix} -1-\lambda & 4 & -6 \\ -8 & 11-\lambda & -12 \\ -4 & 4 & -3-\lambda \end{vmatrix} = (-1-\lambda)\begin{vmatrix} 11-\lambda & -12 \\ 4 & -3-\lambda \end{vmatrix} - 4\begin{vmatrix} -8 & -12 \\ -4 & -3-\lambda \end{vmatrix} - 6\begin{vmatrix} -8 & 11-\lambda \\ -4 & 4 \end{vmatrix}$$

$$\begin{aligned}
&= (-1-\lambda)[(11-\lambda)(-3-\lambda)+48] - 4[-8(-3-\lambda)-48] - 6[-32+4(11-\lambda)] \\
&= (-1-\lambda)(\lambda^2-8\lambda+15) - 4(8\lambda-24) - 6(-4\lambda+12) \\
&= (-\lambda^3+7\lambda^2-7\lambda-15) + (-32\lambda+96) + (24\lambda-72) \\
&= -\lambda^3+7\lambda^2-15\lambda+9 = -(\lambda-3)^2(\lambda-1) = 0
\end{aligned}$$

Les valeurs propres sont donc 3 avec une multiplicité algébrique de 2 et 1 avec une multiplicité algébrique de 1. On va maintenant vérifier en utilisant la théorie des polynômes annulateur que la matrice est bien diagonalisable. Si elle l'est, alors le polynôme $q(x) = (x-3)(x-1)$ est un polynôme annulateur :

$$q(A) = \begin{pmatrix} -4 & 4 & -6 \\ -8 & 8 & -12 \\ -4 & 4 & -6 \end{pmatrix}\begin{pmatrix} -2 & 4 & -6 \\ -8 & 10 & -12 \\ -4 & 4 & -4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Comme il s'agit bien d'un polynôme annulateur, la matrice est diagonalisable.

3.5 Application : Les équations définies par récurrence

Exemple 3.5.1. On veut résoudre le système d'équations définies par récurrence suivant :

$$\begin{cases} x_n = -2x_{n-1} + 2y_{n-1}, & \forall n \geq 1 \\ y_n = -10x_{n-1} + 7y_{n-1}, & \forall n \geq 1 \\ x_0 = 1 \\ y_0 = 2 \end{cases}$$

On peut donc réécrire le tout sous forme matricielle :

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -10 & 7 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -10 & 7 \end{pmatrix}^2 \begin{pmatrix} x_{n-2} \\ y_{n-2} \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -10 & 7 \end{pmatrix}^n \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -10 & 7 \end{pmatrix}^n \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Pour trouver une formule explicite simple, on va donc devoir diagonaliser la matrice. Commençons par trouver les valeurs propres :

$$\begin{vmatrix} -2-\lambda & 2 \\ -10 & 7-\lambda \end{vmatrix} = (-2-\lambda)(7-\lambda) + 20 = \lambda^2 - 5\lambda + 6 = (\lambda-2)(\lambda-3) = 0$$

Les valeurs propres sont donc : $\lambda_1 = 2$ et $\lambda_2 = 3$. Maintenant, on va chercher une base de vecteurs propres :

$$\left(\begin{array}{cc|c} -2-2 & 2 & 0 \\ -10 & 7-2 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -4 & 2 & 0 \\ -10 & 5 & 0 \end{array}\right)$$

Donc $-4x + 2y = 0$ ce qui nous donne le vecteur propre

$$v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} -2-3 & 2 & 0 \\ -10 & 7-3 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -5 & 2 & 0 \\ -10 & 4 & 0 \end{array}\right)$$

Donc $-5x + 2y = 0$ ce qui nous donne le vecteur propre :

$$v_2 = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

On obtient donc la diagonalisation suivante :

$$\begin{pmatrix} -2 & 2 \\ -10 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}^{-1}$$

La solution de l'équation de récurrence est donc :

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ -10 & 7 \end{pmatrix}^n \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} 2^n & 0 \\ 0 & 3^n \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2^n & 2\cdot 3^n \\ 2^{n+1} & 5\cdot 3^n \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 2^n & 2\cdot 3^n \\ 2^{n+1} & 5\cdot 3^n \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2^n \\ 2^{n+1} \end{pmatrix}$$

Donc $x_n = 2^n$ et $y_n = 2^{n+1}$ pour tout $n \geq 0$.

Exemple 3.5.2. On veut résoudre l'équation définie par récurrence suivante :

$$\begin{cases} x_n = -2x_{n-1} + 15x_{n-2}, & \forall n \geq 2 \\ x_0 = 3 \\ x_1 = 1 \end{cases}$$

On va commencer par poser $y_n = x_{n-1}$ ce qui nous donne le système suivant :

$$\begin{cases} x_n = -2x_{n-1} + 15y_{n-1}, \quad \forall n \geq 2 \\ y_n = x_{n-1}, \quad \forall n \geq 2 \\ x_1 = 1 \\ y_1 = 3 \end{cases}$$

On peut donc résoudre ce système de la même façon que nous avions résolut le système de l'exemple précédent. On commence par réécrire le tout sous forme matricielle :

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} -2 & 15 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} -2 & 15 \\ 1 & 0 \end{pmatrix}^2 \begin{pmatrix} x_{n-2} \\ y_{n-2} \end{pmatrix} = \begin{pmatrix} -2 & 15 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} -2 & 15 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

On va donc devoir diagonaliser la matrice. On va commencer par trouver les valeurs propres :

$$\begin{vmatrix} -2-\lambda & 15 \\ 1 & -\lambda \end{vmatrix} = (-2-\lambda)(-\lambda) - 15 = \lambda^2 + 2\lambda - 15 = (\lambda - 3)(\lambda + 5) = 0$$

Les valeurs propres sont donc $\lambda_1 = 3$ et $\lambda_2 = -5$. On veut maintenant trouver une base de vecteurs propres :

$$\left(\begin{array}{cc|c} -2-3 & 15 & 0 \\ 1 & -3 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -5 & 15 & 0 \\ 1 & -3 & 0 \end{array}\right)$$

Donc un premier vecteur propre est donné par $x - 3y = 0$ ce qui nous donne le vecteur :

$$v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

Maintenant, pour trouver un second vecteur propre, on a :

$$\left(\begin{array}{cc|c} -2+5 & 15 & 0 \\ 1 & +5 & 0 \end{array}\right) = \left(\begin{array}{cc|c} 3 & 15 & 0 \\ 1 & 5 & 0 \end{array}\right)$$

Donc un second vecteur propre est donné par $x + 5y = 0$ ce qui nous donne :

$$v_2 = \begin{pmatrix} -5 \\ 1 \end{pmatrix}$$

On peut donc maintenant résoudre notre équation définie par récurrence :

$$\begin{aligned} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} -2 & 15 \\ 1 & 0 \end{pmatrix}^{n-1} \begin{pmatrix} 1 \\ 3 \end{pmatrix} &= \begin{pmatrix} 3 & -5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3^{n-1} & 0 \\ 0 & (-5)^{n-1} \end{pmatrix} \begin{pmatrix} 3 & -5 \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3^{n-1} & 0 \\ 0 & (-5)^{n-1} \end{pmatrix} \begin{pmatrix} 1/8 & 5/8 \\ -1/8 & 3/8 \end{pmatrix} \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 3^n & (-5)^n \\ 3^{n-1} & (-5)^{n-1} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3^n + (-5)^n \\ 2 \cdot 3^{n-1} + (-5)^{n-1} \end{pmatrix} \end{aligned}$$

La solution du système est donc :

$$x_n = 2 \cdot 3^n + (-5)^n, \quad \forall n \geq 0$$

3.6 Application : Les équations différentielles linéaires

Pour ceux qui ont suivi le cours MATH-2731 - Suites et séries, vous avez pris que la plupart des fonctions usuelles tel que la fonction exponentiel, les fonctions sinus et cosinus, etc, peuvent s'écrire sous forme de séries de puissance (un polynôme de degré infini). Dans ce cours, nous allons maintenant utiliser la série de e^x pour définir l'exponentiel d'une matrice.

Sachez qu'il n'est absolument pas nécessaire d'avoir suivi le cours MATH-2731 pour comprendre cette section, par contre, il est malheureusement impossible dans ce cours de justifier pourquoi nous définissons l'exponentiel d'une matrice de cette façon.

Definition 3.6.1. Si A est une matrice carré, alors on définit l'exponentiel de la matrice, dénoté e^A comme étant :

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

Théorème 3.6.1. (Calcul de l'exponentielle d'une matrice)

1. Si D est une matrice diagonale, alors on calcul son exponentielle en calculant l'exponentielle de chacun des éléments sur sa diagonale. C'est à dire :

$$D = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix} \implies e^D = \begin{pmatrix} e^{a_1} & 0 & 0 & \dots & 0 \\ 0 & e^{a_2} & 0 & \dots & 0 \\ 0 & 0 & e^{a_3} & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & e^{a_n} \end{pmatrix}$$

2. Si A est une matrice carré diagonalisable tel que $A = PDP^{-1}$, alors son exponentielle est donné par :

$$e^A = Pe^D P^{-1}$$

Definition 3.6.2. Un système d'équations différentielles linéaires est un système de la forme :

$$y'(x) = Ay(x)$$

où $y'(x)$ et $y(x)$ sont des vecteurs, et A est une matrice.

Théorème 3.6.2. Si $y'(x) = Ay(x)$ est un système d'équations différentielles, alors la solution de ce système est donné par :

$$y(x) = e^{Ax}y(0)$$

Démonstration. Il s'agit de vérifier que $y(x) = e^{Ax}y(0)$ satisfait bien l'équation différentielle et la condition initiale. Ceci démontre qu'il s'agit bien d'une solution. On peut ensuite vérifier qu'il s'agit de la seule solution en supposant qu'il en existe une deuxième $y_2(x)$ et en vérifiant que la dérivé de $y(x) - y_2(x)$ est zéro, ce qui signifit que $y(x) - y_2(x)$ est constant. Comme ils ont un point en commun (la condition initiale) il doit donc s'agit de la même fonction. □

Exemple 3.6.1. On veut résoudre le système d'équations différentielles linéaires suivant :

$$\begin{cases} y' = -58y + 252z \\ z' = -14y + 61z \\ y(0) = 1 \\ z(0) = -1 \end{cases}$$

Ce qui peut être réécrit sous la forme matricielle suivante :

$$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} -58 & 252 \\ -14 & 61 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

Par le théorème précédent, la solution nous est donc donné par :

$$\begin{pmatrix} y \\ z \end{pmatrix} = e^{\begin{pmatrix} -58x & 252x \\ -14x & 61x \end{pmatrix}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

On va donc devoir diagonaliser la matrice. Pour ce faire, commençons par trouver les valeurs propres :

$$\begin{vmatrix} -58-\lambda & 252 \\ -14 & 61-\lambda \end{vmatrix} = (-58-\lambda)(61-\lambda) + 3528 = \lambda^2 - 3\lambda - 10 = (\lambda-5)(\lambda+2) = 0$$

Les valeurs propres sont donc $\lambda = 5$ et $\lambda = -2$ On veut maintenant trouver un vecteur propre associé à la valeur propre 5.

$$\left(\begin{array}{cc|c} -58-5 & 252 & 0 \\ -14 & 61-5 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -63 & 252 & 0 \\ -14 & 56 & 0 \end{array}\right)$$

Donc un vecteur propre doit satisfaire l'équation $-14x + 56y = 0$. On peut donc prendre le vecteur

$$v_1 = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$$

On cherche maintenant un vecteur propre associé à la valeur propre -2 :

$$\left(\begin{array}{cc|c} -58+2 & 252 & 0 \\ -14 & 61+2 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -56 & 252 & 0 \\ -14 & 63 & 0 \end{array}\right)$$

Donc un vecteur propre doit satisfaire l'équation $-14x + 63y = 0$. On peut donc prendre par exemple :

$$v_2 = \begin{pmatrix} 9 \\ 2 \end{pmatrix}$$

Donc la solution du système d'équations différentielles est :

$$\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 4 & 9 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} e^{5x} & 0 \\ 0 & e^{-2x} \end{pmatrix} \begin{pmatrix} 4 & 9 \\ 1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$= \begin{pmatrix} 4e^{5x} & 9e^{-2x} \\ e^{5x} & 2e^{-2x} \end{pmatrix} \begin{pmatrix} -2 & 9 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4e^{5x} & 9e^{-2x} \\ e^{5x} & 2e^{-2x} \end{pmatrix} \begin{pmatrix} -11 \\ 5 \end{pmatrix} = \begin{pmatrix} -44e^{5x} + 45e^{-2x} \\ -11e^{5x} + 10e^{-2x} \end{pmatrix}$$

Exemple 3.6.2. Résoudre l'équation différentielle suivante :

$$\begin{cases} y'' = 9y' - 18y \\ y(0) = 1 \\ y'(0) = 2 \end{cases}$$

Commençons par poser $z = y'$ et donc l'équation peut être réécrite sous forme de système d'équations différentielles :

$$\begin{cases} z' = 9z - 18y \\ y' = z \\ y(0) = 1 \\ y'(0) = 2 \end{cases}$$

Ce qui nous donne sous forme matricielle :

$$\begin{pmatrix} y' \\ z' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -18 & 9 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix}$$

La solution nous est donc donné par :

$$\begin{pmatrix} y \\ z \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ -18 & 9 \end{pmatrix} x} \begin{pmatrix} y(0) \\ z(0) \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ -18 & 9 \end{pmatrix} x} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Pour calculer l'exponentiel, on va devoir diagonaliser la matrice. Commençons par calculer les valeurs propres :

$$\begin{vmatrix} -\lambda & 1 \\ -18 & 9-\lambda \end{vmatrix} = \lambda(\lambda-9) + 18 = \lambda^2 - 9\lambda + 18 = (\lambda-3)(\lambda-6) = 0$$

Donc les valeurs propres sont 3 et 6. On veut maintenant trouver une base de vecteurs propres :

$$\left(\begin{array}{cc|c} -3 & 1 & 0 \\ -18 & 6 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} -3 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

On vecteur propre doit donc satisfaire l'équation $-3x + y = 0$. On peut donc prendre le vecteur :

$$v_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Il ne nous reste plus qu'à trouver un vecteur propre associé à la valeur propre -6 :

$$\left(\begin{array}{cc|c} -6 & 1 & 0 \\ -18 & 3 & 0 \end{array}\right) \sim \left(\begin{array}{cc|c} -6 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

On vecteur propre doit donc satisfaire l'équation $-6x + y = 0$. On peut donc prendre le vecteur :

$$v_2 = \begin{pmatrix} 1 \\ 6 \end{pmatrix}$$

La solution de l'équation est donc :

$$\begin{pmatrix} y \\ z \end{pmatrix} = e^{\begin{pmatrix} 0 & 1 \\ -18 & 9 \end{pmatrix} x} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} e^{3x} & 0 \\ 0 & e^{6x} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 6 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} e^{3x} & e^{6x} \\ 3e^{3x} & 6e^{6x} \end{pmatrix} \begin{pmatrix} 2 & -1/3 \\ -1 & 1/3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} e^{3x} & e^{6x} \\ 3e^{3x} & 6e^{6x} \end{pmatrix} \begin{pmatrix} 4/3 \\ -1/3 \end{pmatrix} = \begin{pmatrix} \frac{4}{3}e^{3x} - \frac{1}{3}e^{6x} \\ 4e^{3x} - 2e^{6x} \end{pmatrix}$$

Donc la solution du système est

$$y(x) = \frac{4}{3}e^{3x} - \frac{1}{3}e^{6x}$$

3.7 La méthode de la puissance et de la puissance inverse

Nous avons vu que si A est une matrice carré de dimension $n \times n$, alors les valeurs propres sont les solutions de l'équation

$$\det(A - \lambda I) = 0$$

le problème étant que $\det(A - \lambda I)$ est un polynôme de degré n, et qu'il est difficile de trouver les racines d'un polynôme lorsque le degré devient grand. En particulier, si $n \geq 5$, il n'existe aucune formule permettant de trouver les racines du polynôme de manière exacte. On a donc besoin de développer des technique nous permettant d'approximer numériquement les valeurs propres de la matrice A. Comme première méthode, nous allons étudier la méthode de la puissance qui nous permet d'approximer la valeur propre la plus grande en valeur absolue.

Definition 3.7.1. Si A est une matrice carré, alors on appelle valeur propre dominante la valeur propre de la matrice qui est la plus grande en valeur absolue.

Théorème 3.7.1. **(Méthode de la puissance)** Si A est une matrice carré de dimension $n \times n$, $x \in \mathbb{R}^n$ et si

1. La matrice A est diagonalisable.
2. la matrice A a une seule valeur propre dominante de multiplicité algébrique 1.
3. v est un vecteur propre associé à la valeur propre dominante, alors x et v ne sont pas linéairement indépendant.

alors la suite des λ_k définie par :

$$\begin{cases} u_0 = v_0 = x \\ u_k = Av_{k-1}, \quad k \geq 1 \\ \lambda_k = \text{ valeur dominante de } u_k \\ v_k = \dfrac{u_k}{\lambda_k}, \quad k \geq 1 \end{cases}$$

converge vers une valeur propre de A. De plus, si x n'est pas un vecteur propre de A, alors la suite des λ_k converge vers la valeur propre dominante de A.

Démonstration. Comme la matrice A est diagonalisable, alors il existe une base $\{u_1, u_2, ..., u_n\}$ formé de vecteur propre de A associé aux valeurs propres $\{\mu_1, \mu_2, ..., \mu_n\}$. Donc le vecteur x peut s'écrire sous la forme :

$$x = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + ... + \alpha_n u_n$$

On obtient donc :

$$\begin{aligned} Ax &= A(\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 + ... + \alpha_n u_n) \\ &= \alpha_1 Au_1 + \alpha_2 Au_2 + \alpha_3 Au_3 + ... + \alpha_n Au_n \\ &= \alpha_1 \mu_1 u_1 + \alpha_2 \mu_2 u_2 + ... + \alpha_n \mu_n u_n \end{aligned}$$

Maintenant, sans perte de généralité, on va supposer que μ_1 est la valeur propre dominante. Alors on a :

$$\frac{1}{\mu_1} Ax = \alpha_1 u_1 + \left(\alpha_2 \left(\frac{\mu_2}{\mu_1}\right) u_2 + ... + \alpha_n \left(\frac{\mu_n}{\mu_1}\right) u_n\right)$$

On obtient donc :

$$\frac{1}{\mu_1^k} A^k x = \alpha_1 u_1 + \left(\alpha_2 \left(\frac{\mu_2}{\mu_1}\right)^k u_2 + ... + \alpha_n \left(\frac{\mu_n}{\mu_1}\right)^k u_n\right)$$

Maintenant, en prenant la limite, on obtient finalement :

$$\lim_{k \to \infty} \frac{1}{\mu_1^k} A^k x = \lim_{k \to \infty} \left[\alpha_1 u_1 + \left(\alpha_2 \left(\frac{\mu_2}{\mu_1}\right)^k u_2 + ... + \alpha_n \left(\frac{\mu_n}{\mu_1}\right)^k u_n\right)\right] = \alpha_1 u_1$$

Donc la limite

$$\lim_{k \to \infty} \frac{1}{\mu_1^k} A^k x$$

est un vecteur propre associé à la valeur propre dominante de A. Maintenant, pour trouver la valeur propre dominante, il s'agit de trouver μ tel que

$$\mu(\alpha_1 u_1) = A(\alpha_1 u_1)$$

ce qui peut être fait en comparant n'importe quel des coordonnées des vecteurs, en particulier la valeur dominante du vecteur. $\square$

Exemple 3.7.1. On veut trouver la valeur propre dominante de la matrice

$$A = \begin{pmatrix} -68 & 6 & 80 \\ -10 & 9 & 10 \\ -75 & 6 & 87 \end{pmatrix}$$

Pour commencer la méthode, on doit choisir un vecteur de départ. Pour cet exemple, on va prendre le vecteur :

$$u_0 = v_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

On obtient donc :

$$u_1 = \begin{pmatrix} -68 & 6 & 80 \\ -10 & 9 & 10 \\ -75 & 6 & 87 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 18 \\ 9 \\ 18 \end{pmatrix} \qquad \lambda_1 = 18 \qquad v_1 = \begin{pmatrix} 1 \\ 0,5 \\ 1 \end{pmatrix}$$

$$u_2 = \begin{pmatrix} -68 & 6 & 80 \\ -10 & 9 & 10 \\ -75 & 6 & 87 \end{pmatrix}\begin{pmatrix} 1 \\ 0,5 \\ 1 \end{pmatrix} = \begin{pmatrix} 15 \\ 4,5 \\ 15 \end{pmatrix} \qquad \lambda_2 = 15 \qquad v_2 = \begin{pmatrix} 1 \\ 0,3 \\ 1 \end{pmatrix}$$

$$u_3 = \begin{pmatrix} -68 & 6 & 80 \\ -10 & 9 & 10 \\ -75 & 6 & 87 \end{pmatrix}\begin{pmatrix} 1 \\ 0,3 \\ 1 \end{pmatrix} = \begin{pmatrix} 13,8 \\ 2,7 \\ 13,8 \end{pmatrix} \qquad \lambda_3 = 13,8 \qquad v_3 = \begin{pmatrix} 1 \\ 0,1957 \\ 1 \end{pmatrix}$$

$$u_4 = \begin{pmatrix} -68 & 6 & 80 \\ -10 & 9 & 10 \\ -75 & 6 & 87 \end{pmatrix}\begin{pmatrix} 1 \\ 0,1957 \\ 1 \end{pmatrix} = \begin{pmatrix} 13,1742 \\ 1,7613 \\ 13,1742 \end{pmatrix} \qquad \lambda_4 = 13,1742 \qquad v_4 = \begin{pmatrix} 1 \\ 0,1337 \\ 1 \end{pmatrix}$$

$$u_5 = \begin{pmatrix} -68 & 6 & 80 \\ -10 & 9 & 10 \\ -75 & 6 & 87 \end{pmatrix}\begin{pmatrix} 1 \\ 0,1337 \\ 1 \end{pmatrix} = \begin{pmatrix} 12,8022 \\ 1,2033 \\ 12,8022 \end{pmatrix} \qquad \lambda_5 = 12,8022 \qquad v_5 = \begin{pmatrix} 1 \\ 0,0940 \\ 1 \end{pmatrix}$$

Donc après 5 itérations, on obtient une approximation de $12,8022$ pour la valeur propre dominante. On peut comparer cette réponse avec la valeur propre exacte de 12.

Maintenant, en se rappelant que si λ est un valeur propre d'une matrice inversible A, alors $\frac{1}{\lambda}$ est une valeur propre de la matrice A^{-1}, ceci nous permet de modifier la méthode de la puissance pour nous permettre d'approximer une seconde valeur propre. Cette fois ci il s'agira de la plus proche de 0.

Théorème 3.7.2. **(Méthode de la puissance inverse)** Si A est une matrice carré inversible, alors on peut appliquer la méthode de la puissance à la matrice A^{-1} pour approximer la valeur propre dominante λ de A^{-1}. On a donc que $\frac{1}{\lambda}$ est la valeur propre de A la plus proche de 0.

Exemple 3.7.2. On veut trouver approximer toutes les valeurs propres de la matrice

$$A = \begin{pmatrix} 130 & -264 \\ 60 & -122 \end{pmatrix}$$

Commençons par utiliser la méthode de la puissance pour approximer la valeur propre dominante. Pour ce faire, nous allons commencer avec le vecteur $u_0 = v_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

$$u_1 = Av_0 = \begin{pmatrix} 130 & -264 \\ 60 & -122 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -134 \\ -62 \end{pmatrix} \qquad \lambda_1 = -134 \qquad v_1 = \begin{pmatrix} 1 \\ 0,4627 \end{pmatrix}$$

$$u_2 = Av_1 = \begin{pmatrix} 130 & -264 \\ 60 & -122 \end{pmatrix} \begin{pmatrix} 1 \\ 0,4627 \end{pmatrix} = \begin{pmatrix} 7,8507 \\ 3,5522 \end{pmatrix} \qquad \lambda_2 = 7,8507 \qquad v_2 = \begin{pmatrix} 1 \\ 0,4525 \end{pmatrix}$$

$$u_3 = Av_2 = \begin{pmatrix} 130 & -264 \\ 60 & -122 \end{pmatrix} \begin{pmatrix} 1 \\ 0,4525 \end{pmatrix} = \begin{pmatrix} 10,5475 \\ 4,7985 \end{pmatrix} \qquad \lambda_3 = 10,5475 \qquad v_3 = \begin{pmatrix} 1 \\ 0,4549 \end{pmatrix}$$

$$u_4 = Av_3 = \begin{pmatrix} 130 & -264 \\ 60 & -122 \end{pmatrix} \begin{pmatrix} 1 \\ 0,4549 \end{pmatrix} = \begin{pmatrix} 9,8962 \\ 4,4975 \end{pmatrix} \qquad \lambda_4 = 9,8962 \qquad v_3 = \begin{pmatrix} 1 \\ 0,4545 \end{pmatrix}$$

Donc une approximation de la valeur propre dominante est $9,8962$. Ceci est relativement proche de la vrai valeur qui est 10. Nous allons maintenant utiliser la méthode de la matrice inverse pour approximer l'autre valeur propre de la matrice. On va donc refaire la même chose, mais en utilisant la matrice inverse. On va commencer à nouveau avec le vecteur $u_0 = v_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

$$A^{-1} = \begin{pmatrix} 6,1 & -13,2 \\ 3 & -6,5 \end{pmatrix}$$

$$u_1 = A^{-1}v_0 = \begin{pmatrix} 6,1 & -13,2 \\ 3 & -6,5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -7,1 \\ -3,5 \end{pmatrix} \qquad \lambda_1 = -7,1 \qquad v_1 = \begin{pmatrix} 1 \\ 0,493 \end{pmatrix}$$

$$u_2 = A^{-1}v_1 = \begin{pmatrix} 6,1 & -13,2 \\ 3 & -6,5 \end{pmatrix} \begin{pmatrix} 1 \\ 0,493 \end{pmatrix} = \begin{pmatrix} -0,4070 \\ -0,2042 \end{pmatrix} \qquad \lambda_2 = -0,407 \qquad v_2 = \begin{pmatrix} 1 \\ 0,5017 \end{pmatrix}$$

$$u_3 = A^{-1}v_2 = \begin{pmatrix} 6,1 & -13,2 \\ 3 & -6,5 \end{pmatrix} \begin{pmatrix} 1 \\ 0,5017 \end{pmatrix} = \begin{pmatrix} -0,5228 \\ -0,2612 \end{pmatrix} \qquad \lambda_3 = -0,5228 \qquad v_3 = \begin{pmatrix} 1 \\ 0,4997 \end{pmatrix}$$

$$u_4 = A^{-1}v_2 = \begin{pmatrix} 6,1 & -13,2 \\ 3 & -6,5 \end{pmatrix} \begin{pmatrix} 1 \\ 0,4997 \end{pmatrix} = \begin{pmatrix} -0,4956 \\ -0,2478 \end{pmatrix} \qquad \lambda_4 = -0,4956 \qquad v_4 = \begin{pmatrix} 1 \\ 0,5 \end{pmatrix}$$

Donc une approximation de la valeur propre dominante de la matrice inverse est $-0,4956$, donc une approximation de la valeur propre de la matrice A qui est la plus proche de 0 est donnée par :

$$\frac{1}{-0,4956} = -2,0178$$

Ce qui est très proche de la vrai valeur propre qui est -2.

Chapitre 4

Les espaces euclidiens et l'orthogonalité

Dans ce chapitre, nous allons ajouter un autre opération à nos espaces vectoriels : Un produit scalaire. L'avantage d'avoir un produit scalaire est qu'il nous sera maintenant possible de parler de la longueur d'un vecteur, ainsi que de l'angle entre deux vecteurs. Nous allons ensuite voir que pour plusieurs problèmes, avoir une base qui est orthogonale ou orthonormale peut grandement simplifier le problème. C'est le cas entre autre des problèmes de projection. Nous allons ensuite voir qu'il est particulièrement facile de calculer l'inverse d'une matrice si ses colonnes forme une base orthonormale. Il s'agit tout simplement de sa transposé.

4.1 Produit scalaire et norme

Definition 4.1.1. Si V est un espace vectoriel, alors un produit scalaire $\langle\cdot,\cdot\rangle$, est une fonction $\langle\cdot,\cdot\rangle : V\times V \to \mathbb{R}$ ayant les 5 propriétés suivantes :

1. $\langle u, v + w\rangle = \langle u, v\rangle + \langle u, w\rangle$ pour tout $u, v, w \in V$
2. $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$ pour tout $u, v \in V$ et $\lambda \in \mathbb{R}$
3. $\langle u, v\rangle = \langle v, u\rangle$
4. $\langle u, u\rangle \geq 0$ pour tout $u \in V$
5. $\langle u, u\rangle = 0$ si et seulement si $u = 0$

On appelle espace Euclidien un espace vectoriel muni d'un produit scalaire.

Definition 4.1.2. Si V est un espace Euclidien, alors on définit la norme d'un vecteur $u \in V$ comme étant :

$$||u|| = \sqrt{\langle u, u\rangle}$$

Techniquement, bien que nous ayons définie $||u||$ comme étant le norme du vecteur u, techniquement nous n'avons pas encore démontrer qu'il s'agit d'une norme. Pour démontrer que notre fonction satisfait bien les propriété d'une norme, nous allons cependant avoir besoin en premier d'une inégalité importante que nous allons démontrer immédiatement.

Théorème 4.1.1. **(Inégalité de Cauchy-Schwartz)** Si V est un espace euclidien, $u, w \in V$, alors on a :

$$\langle u, w\rangle \leq ||u||\ ||w||$$

Démonstration. Supposons que V est un espace euclidien, $u, w \in V$ et $x \in \mathbb{R}$. En utilisant la définition de la norme et ainsi que les propriétés du produit scalaire, on obtient donc :

$$\begin{aligned} 0 &\leq ||xu + w||^2 = \langle xu + w, xu + w\rangle \\ &= \langle xu, xu\rangle + \langle xu, w\rangle + \langle w, xu\rangle + \langle w, w\rangle \\ &= \langle u, u\rangle x^2 + 2\langle u, w\rangle x + \langle w, w\rangle \end{aligned}$$

En considérant x comme une variable, et le reste comme des constantes, on remarque donc que nous avons obtenue un polynôme du second degré (une parabole). Comme notre fonction est toujours plus grande ou égale à 0, on peut donc conclure que la parabole possède 0 ou 1 zéro (il ne peut pas y en avoir 2). Ceci nous permet d'affirmer que le discriminant doit être plus petit ou égal à 0. On a donc :

$$(2\langle u, w\rangle)^2 - 4\langle u, u\rangle\langle w, w\rangle \leq 0$$

$$4\langle u, w\rangle^2 - 4||u||^2\ ||w||^2 \leq 0$$

$$\langle u, w\rangle \leq ||u||\ ||w||$$

Ce qui est exactement ce que nous voulions démontrer. □

Théorème 4.1.2. **(Propriétés de la norme)** Si V est un espace Euclidien, alors la norme d'un vecteur a les quatre propriétés suivantes :

1. $||u|| \geq 0$ pour tout $u \in V$
2. $||u|| = 0$ si et seulement si $u = 0$
3. $||\lambda u|| = |\lambda| \cdot ||u||$ pour tout $u \in V$ et $\lambda \in \mathbb{R}$
4. $||u + v|| \leq ||u|| + ||v||$ pour tout $u, v \in V$

Démonstration.

1. Il s'agit d'une conséquence directe du fait que $\langle u, u\rangle \geq 0$ pour tout $u \in V$.
2. Il s'agit d'une conséquence directe du fait que $\langle u, u\rangle = 0$ si et seulement si $u = 0$.
3. On remarque que
$$||\lambda u||^2 = \langle \lambda u, \lambda u\rangle = \lambda^2\langle u, u\rangle = \lambda^2||u||^2$$
En prenant la racine carré de chaque côté, on obtient donc :
$$||\lambda u|| = |\lambda|\ ||u||$$
4. Il s'agit d'une application de l'inégalité de Cauchy-Schwartz. Si $u, v \in V$, alors on a :
$$\begin{aligned} ||u + v||^2 &= \langle u + v, u + v >= \langle u, u\rangle + 2\langle u, w\rangle + \langle w, w\rangle \\ &\leq ||u||^2 + 2||u||\ ||v|| + ||v||^2 = (||u|| + ||v||)^2 \end{aligned}$$
Ce qui nous donne : $||u + v|| \leq ||u|| + ||v||$.

□

Definition 4.1.3. Si V est un espace Euclidien, alors on définit l'angle θ entre des vecteurs $u, v \in V$ comme étant :

$$\cos(\theta) = \frac{\langle u, v\rangle}{||u|| \cdot ||v||}$$

La plupart des espaces vectoriels que nous avons rencontrer jusqu'à présent peuvent être transformé en espace Euclidien en y ajoutant un produit scalaire. Voici les produits scalaires standard pour certain des espaces vectoriel que nous avons rencontré jusqu'à présent :

1. Dans $\mathbb{R}^n$ on a le produit scalaire suivant :

$$\left\langle \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \right\rangle = a_1b_1 + a_2b_2 + a_3b_3 + ... + a_nb_n$$

2. Dans l'espace vectoriel $\mathbb{P}_n(\mathbb{R})$ on a le produit scalaire suivant :

$$\langle a_nx^n + a_{n-1}x^{n-1} + ... + a_1x + a_0, b_nx^n + b_{n-1}x^{n-1} + ... + b_1x + b_0\rangle = a_nb_n + a_{n-1}b_{b-1} + ... + a_1b_1 + a_0b_0$$

3. Dans l'espace vectoriel $M_{n\times n}(\mathbb{R})$ on a le produit scalaire suivant :

$$\langle A, B\rangle = tr(A^TB)$$

4. Dans l'espace vectoriel $C[a,b]$ des fonctions continue sur l'intervalle $[a,b]$ à valeurs réelles, on a le produit scalaire suivant :

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx$$

Exemple 4.1.1. Dans l'espace vectoriel $M_{2\times 2}$, on veut calculer la norme de la matrice $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

$$\left\| \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \right\| = \sqrt{tr\left[\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right]} = \sqrt{tr\left[\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right]} = \sqrt{tr\begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}} = \sqrt{30}$$

Exemple 4.1.2. Dans l'espace vectoriel $C[0,1]$, on veut trouver l'angle entre les fonctions $f(x) = (x^2+1)$ et $g(x) = (x^3+x)$. Pour ce faire, en utilisant la formule, on a :

$$\begin{aligned} \cos(\theta) &= \frac{\langle f, g\rangle}{\sqrt{\langle f, f\rangle}\sqrt{\langle g, g\rangle}} = \frac{\int_0^1 (x^2+1)(x^3+x)dx}{\sqrt{\int_0^1 (x^2+1)^2dx}\sqrt{\int_0^1 (x^3+x)^2dx}} = \frac{\int_0^1 (x^5+2x^3+x)dx}{\sqrt{\int_0^1 (x^2+1)^2dx}\sqrt{\int_0^1 (x^3+x)^2dx}} \\ &= \frac{\frac{7}{6}}{\sqrt{\frac{28}{15}}\sqrt{\frac{92}{105}}} = \frac{35\sqrt{23}}{184} \end{aligned}$$

On pourrait ensuite trouver la valeur de θ, mais ce n'est pas vraiment nécessaire. Connaitre la valeur du cosinus est suffisant dans ce cas.

4.2 Bases orthogonales et méthode de Gram-Schmidt

Definition 4.2.1. Si V est un espace Euclidien, et $\{u_1, u_2, ..., u_n\}$ sont des vecteurs de V, alors on dit que :

1. les vecteurs sont orthogonaux si :

$$\langle u_i, u_j\rangle = 0 \quad \text{pour tout } i \neq j$$

2. les vecteurs sont orthonormaux si :

$$\langle u_i, u_j\rangle = \begin{cases} 0 & \text{pour tout } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Remarquez que la définition signifie que des vecteurs sont orthonormaux si ils sont orthogonaux et unitaire.

Exemple 4.2.1. Dans l'espace vectoriel $C[0,1]$ muni du produit scalaire $\langle f,g\rangle = \int_0^1 f(x)g(x)dx$, vérifier que les vecteurs suivant sont orthonormaux :

$$S = \left\{\sqrt{3}x, 4\sqrt{5}x^2 - 3\sqrt{5}x\right\}$$

On a donc :

$$\langle\sqrt{3}x, \sqrt{3}x\rangle = \int_0^1 3x^2 dx = 1$$

$$\langle 4\sqrt{5}x^2 - 3\sqrt{5}x, 4\sqrt{5}x^2 - 3\sqrt{5}x\rangle = \int_0^1 (80x^4 - 120x^3 + 45x^2)dx = 1$$

$$\langle\sqrt{3}x, 4\sqrt{5}x^2 - 3\sqrt{5}x\rangle = \int_0^1 (4\sqrt{15}x^3 - 3\sqrt{15}x^2)dx = 0$$

Il s'agit donc bien d'un ensemble orthonormal.

Definition 4.2.2. Si V est un espace Euclidien, et W est un sous-espace vectoriel de V, alors on définit le sous-espace orthogonal $W^\perp$ par :

$$W^\perp = \{u : \langle u,w\rangle = 0 \text{ pour tout } w \in W\}$$

Théorème 4.2.1.

1. Si V est un espace Euclidien, et W est un sous-espace vectoriel de V, alors $W^\perp$ est aussi un sous-espace vectoriel.
2. Si V est un espace vectoriel et W un sous-espace vectoriel de V, alors :

$$W \cap W^\perp = \{0\}$$

Démonstration.

1. Prenons $u, v \in W^\perp$ et $x \in W$, alors

$$\langle(u+v), x\rangle = \langle u,x\rangle + \langle v,x\rangle = 0 + 0 = 0.$$

Donc $u, v \in W^\perp$. Maintenant, si $u \in W^\perp$, $x \in W$ et $\lambda \in \mathbb{R}$, alors

$$\langle(\lambda u), x\rangle = \lambda\langle u,x\rangle = \lambda(0) = 0$$

Donc $\lambda u \in W^\perp$. Il s'agit donc bien d'un sous-espace vectoriel.

2. Supposons que $x \in W \cap W^\perp$. Alors $x \in W$ et $x \in W^\perp$. Par définition de $W^\perp$, on doit donc avoir que $\langle x,x\rangle = 0$. Maintenant, par la définition d'un produit scalaire, on peut donc conclure que $x = 0$. Donc $W \cap W^\perp = \{0\}$.

□

Nous avons vu au chapitre 2 que tous les espaces vectoriels admettent au moins une base. La question est maintenant de savoir si tout les espaces euclidiens admettent au moins une base orthonormale. L'idée étant qu'il est souvent plus facile de travailler avec une base orthonormale qu'avec une base quelconque. Un premier pas dans cette direction est donné par la méthode de Gram-Schmidt. Cette dernière nous permet de trouver une base orthogonale à un espace vectoriel de la forme $W = span\{u_1, u_2, ..., u_n\}$.

Théorème 4.2.2. **(Méthode de Gram-Schmidt)** Supposons que V est un espace vectoriel, et $W = span\{u_1, u_2, ..., u_n\}$ est un sous espace vectoriel de V, alors les vecteurs non nuls de $\{w_1, w_2, ..., w_n\}$ forment une base orthogonale de W, où les w_i sont définies comme l'algorithme suivant :

$$\begin{cases} w_1 = u_1 \\ w_2 = u_2 - \dfrac{\langle w_1, u_2\rangle}{\langle w_1, w_1\rangle} w_1 \\ w_3 = u_3 - \dfrac{\langle w_1, u_3\rangle}{\langle w_1, w_1\rangle} w_1 - \dfrac{\langle w_2, u_3\rangle}{\langle w_2, w_2\rangle} w_2 \\ w_4 = u_4 - \dfrac{\langle w_1, u_4\rangle}{\langle w_1, w_1\rangle} w_1 - \dfrac{\langle w_2, u_4\rangle}{\langle w_2, w_2\rangle} w_2 - \dfrac{\langle w_3, u_4\rangle}{\langle w_3, w_3\rangle} w_3 \\ \ldots.. \\ w_n = u_n - \displaystyle\sum_{i=1}^{n-1} \dfrac{\langle w_i, u_n\rangle}{\langle w_i, w_i\rangle} w_i \end{cases}$$

Démonstration. La démonstration se fait par induction. Si on a un seul vecteur, il n'y a rien à démontrer. Supposons que $\mathcal{B} = \{w_1, w_2, ..., w_m\}$ est une base orthogonale de $W = span\{x_1, x_2, ..., x_m\}$, alors on veut trouver une base orthogonale de $W' = span\{x_1, x_2, ..., x_m, x_{m+1}\}$. On a donc :

$$\begin{aligned} x_{m+1} &= z_1 + z_2, \qquad z_1 \in W, z_2 \in W^\perp \\ &= (\alpha_1 w_1 + \alpha_2 w_2 + ... + \alpha_m w_m) + z_2, \qquad \alpha_i \in \mathbb{R} \end{aligned}$$

Maintenant, en calculant le produit scalaire, on obtient :

$$\begin{aligned} \langle x_{m+1}, w_i\rangle &= \alpha_1\langle w_1, w_i\rangle + \alpha_2\langle w_2, w_i\rangle + ... + \alpha_m\langle w_m, w_i\rangle + \langle z_2, w_i\rangle, \qquad i \leq m \\ &= \alpha_i\langle w_i, w_i\rangle \end{aligned}$$

Donc :

$$\alpha_i = \frac{\langle x_{m+1}, w_i\rangle}{\langle w_i, w_i\rangle}$$

On obtient donc que :

$$z_2 = x_{m+1} - \frac{\langle x_{m+1}, w_1\rangle}{\langle w_1, w_1\rangle} - \frac{\langle x_{m+1}, w_2\rangle}{\langle w_2, w_2\rangle} - ... - \frac{\langle x_{m+1}, w_m\rangle}{\langle w_m, w_m\rangle}$$

On va maintenant poser $w_{m+1} = z_2$. Le vecteur w_{m+1} est donc orthogonal à tout les vecteur de $\mathcal{B}$, ce qui nous donne la base orthogonale de W' suivante :

$$\mathcal{B}' = \{w_1, w_2, w_3, ..., w_{m+1}\}$$

□

Théorème 4.2.3. **(Existence de bases orthonormales)** Si V est un espace vectoriel de dimension finie, alors il existe une base $\mathcal{B}$ de V qui est orthonormale.

Démonstration. Au chapitre 1, nous avons vu que tout les espace vectoriel de dimension finie admettent au moins une base. En application la méthode de Gram-Schmidt à cette base, nous obtenons donc une base orthogonale de V, disons $\mathcal{B}' = \{w_1, w_2, w_3, ..., w_n\}$. Maintenant, pour trouver une base orthonormale, il s'agit de normaliser chacun de ces vecteurs. On obtient donc la base orthonormale suivante :

$$\mathcal{B} = \left\{\frac{w_1}{||w_1||}, \frac{w_2}{||w_2||}, ..., \frac{w_n}{||w_n||}\right\}$$

□

Nous allons maintenant conclure cette section en démontrant que dans une base orthonormale, tout les produit scalaire ressemble au produit scalaire standard de $\mathbb{R}^n$.

Théorème 4.2.4. Si V est un espace euclidien de dimension finie, et $\mathcal{B} = \{e_1, e_2, e_3, ..., e_n\}$ une base orthonormale de V, alors $\langle x, y\rangle = x^T y$ pour tout $x, y \in V$, où le produit de droite est interpréter comme un produit de matrices avec les vecteurs x, y écrit dans la base $\mathcal{B}$ et représenté comme des matrices colonnes.

Démonstration. Premièrement, remarquons que le résultat est évident pour les vecteurs de notre base $\mathcal{B}$ (où les vecteurs sont écrit dans cette base) est évident. En effet, nous avons :

$$\langle e_i, e_j\rangle = e_i^T e_j = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{autrement} \end{cases}$$

Maintenant, prenons $x, y \in V$, et écrivons ces vecteurs dans notre base $\mathcal{B}$. On a donc :

$$\begin{aligned} x &= \lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 + ... + \lambda_n e_n = \sum_{i=1}^{n} \lambda_i e_i \\ y &= \mu_1 e_1 + \mu_2 e_2 + \mu_3 e_3 + ... + \mu_n e_n = \sum_{i=1}^{n} \mu_i e_i \end{aligned}$$

Ce qui nous permet d'obtenir :

$$\langle x, y\rangle = \left\langle \sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{n} \mu_j e_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \mu_j \langle e_i, e_j\rangle = \sum_{i=1}^{n} \lambda_i \mu_i = \begin{pmatrix} \lambda_1 & \lambda_2 & \lambda_3 & ... & \lambda_n \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_n \end{pmatrix} = [x]_{\mathcal{B}}^T [y]_{\mathcal{B}}$$

□

4.3 Les projections orthogonales

Dans le premier cours d'algèbre linéaire, vous avez appris comment définir la projection orthogonale d'un vecteur u sur un vecteur w, ce que l'on a dénoté par u_w ou bien $proj_w(u)$. Par des méthodes géométriques, nous avons obtenu la formule suivante :

$$u_w = \frac{u \cdot w}{\|w\|^2} w$$

Nous voulons maintenant regarder le même problème, mais de manière un peu différente, ce qui va nous permettre en même temps de généraliser cette idée. Lorsque nous avons regarder la projection du vecteur u sur le vecteur w, nous avons en fait trouver un vecteur de l'espace vectoriel

$$W = span\{w\}$$

de plus, nous aurions pu remplacer le vecteur w par n'importe quel autre vecteur de W et nous aurions obtenu la même solution (et la même formule). Donc en réalité, il aurait été plus juste de parler de la projection du vecteur u sur l'espace vectoriel W. De plus, nous avions remarquez que $u - u_w$ est un vecteur qui est perpendiculaire à w (et donc perpendiculaire à tout les vecteurs de W), on peut donc décomposer le vecteur u sous la forme :

$$u = \underbrace{u_w}_{\in W} + \underbrace{(u - u_w)}_{\in W^\perp}$$

En prenant cette décomposition comme étant la définition de la projection orthogonale, on peut alors remplacer le W par n'importe quel sous-espace vectoriel d'un espace euclidien, et obtenir une notion de projection orthogonale qui aura toujours du sens, et qui sera en fait rien d'autre qu'une généralisation de la projection orthogonale que nous avons vu dans le premier cours.

Théorème 4.3.1. Si V est un espace Euclidien, et W est un sous-espace vectoriel de V. Si $u \in V$, alors il existe $w_1 \in W$ et $w_2 \in W^{\perp}$ tel que :

$$u = w_1 + w_2$$

De plus, les vecteurs w_1 et w_2 sont unique.

Démonstration. Supposons que $\{w_1, w_2, ..., w_n\}$ est une base orthonormale de W, et $x \in V$. Alors, le vecteur

$$w = \langle x, w_1\rangle w_1 + \langle x, w_2\rangle w_2 + ... + \langle x, w_n\rangle w_n$$

est dans W. On veut montrer que $x - w \in W^{\perp}$.

$$\begin{array}{rcl} \langle (x-w), w_i\rangle & = & \langle x, w_i\rangle - \langle w, w_i\rangle \\ & = & \langle x, w_i\rangle - \Big\langle \Big(\langle x, w_1\rangle w_1 + \langle x, w_2\rangle w_2 + ... + \langle x, w_n\rangle w_n\Big), w_i\Big\rangle \\ & = & \langle x, w_i\rangle - \langle x, w_i\rangle\langle w_1, w_i\rangle - \langle x, w_2\rangle\langle w_2, w_i\rangle - ... - \langle x, w_n\rangle\langle w_n, w_i\rangle \\ & = & \langle x, w_i\rangle - \langle x, w_i\rangle \\ & = & 0 \end{array}$$

Comme $(x-w)$ est orthogonal à tout les w_i, on peut donc conclure que $(x-w) \in W^{\perp}$. On a donc écrit le vecteur x comme la somme d'un vecteur de W et de $W^{\perp}$. On veut maintenant montrer que cette décomposition est unique. Pour ce faire, supposons que

$$x = w_1 + z_1 = w_2 + z_2, \qquad \text{avec } w_i \in W \text{ et } z_i \in W^{\perp}$$

On a donc :

$$\underbrace{w_1 - w_2}_{\in W} = \underbrace{z_2 - z_1}_{\in W^{\perp}}$$

Maintenant, comme $W \cap W^{\perp} = \{0\}$, on peut donc conclure que $w_1 = w_2$ et $z_1 = z_2$. La décomposition est donc unique. ☐

Definition 4.3.1. Si V est un espace Euclidien, et W est un sous-espace vectoriel de V. Si $u \in V$, alors on définit

$$proj_W(u) = w_1$$

comme étant l'unique vecteur $w_1 \in W$ tel que $u = w_1 + w_2$, où $w_2 \in W^{\perp}$.

Dans le premier cours, nous avons utiliser la projection dans plusieurs contexte différent, entre autre pour calculer la plus courte distance en un point et une droite. Nous pouvons maintenant généraliser cette idée pour trouver la plus courte distance entre un point et un sous-espace vectoriel. Mais avant, nous avons besoin d'un théorème qui devrait vous être familier depuis le secondaire, mais que nous allons réécrire de manière vectoriel.

Théorème 4.3.2. **(Théorème de Pythagore)** Si V est un espace euclidien et $x, y \in V$ tel que $\langle x, y\rangle = 0$, alors on a :

$$||x||^2 + ||y||^2 = ||x+y||^2$$

Démonstration. Exercice. ☐

Théorème 4.3.3. Si V est un espace euclidien et W un sous-espace vectoriel de V. Prenons $x \in V$, alors $w = proj_W(x)$ est le point de W le plus proche de x.

Démonstration. Prenons z un élément quelconque de W, alors $(w - z) \in W$ et $(x - w) \in W^\perp$, ce qui nous donne par le théorème de Pythagore :

$$||w - z||^2 + ||x - w||^2 = ||x - z||^2$$

On peut donc déduire que

$$||x - w||^2 \leq ||x - z||^2 \quad \Rightarrow \quad ||x - w|| \leq ||x - z||$$

On peut donc conclure que w est le point de W le plus proche de x. □

Nous pouvons maintenant utiliser la méthode de Gram-Schmidt en combinaison avec le dernier théorème pour calculer la projection d'un vecteur sur un sous-espace vectoriel.

Théorème 4.3.4. Si V est un espace euclidien, W est un sous-espace vectoriel de V, et $x \in V$. Alors on a :

$$proj_W(x) = \langle x, w_1 \rangle w_1 + \langle x, w_2 \rangle w_2 + \ldots + \langle x, w_n \rangle w_n$$

où $\mathcal{B} = \{w_1, w_2, \ldots, w_n\}$ est une base orthonormale de W.

Démonstration. Comme $x \in V$, alors x peut s'écrire d'une unique manière sous la forme

$$x = y_1 + y_2, \qquad y_1 \in W, \quad y_2 \in W^\perp$$

Maintenant, comme w_1 est dans W, alors on a :

$$x = \alpha_1 w_1 + \alpha_2 w_2 + \ldots + \alpha_n w_n + y_2, \qquad \alpha_i \in \mathbb{R}$$

Et maintenant en calculant le produit scalaire :

$$\langle x, w_i \rangle = \alpha_i \langle w_i, w_i \rangle = \alpha_i, \qquad 1 \leq i \leq n$$

ce qui nous donne finalement :

$$x = \langle x, w_1 \rangle w_1 + \langle x, w_2 \rangle w_2 + \ldots + \langle x, w_n \rangle w_n + y_2$$

la projection nous est donc donné par :

$$proj_W(x) = \langle x, w_1 \rangle w_1 + \langle x, w_2 \rangle w_2 + \ldots + \langle x, w_n \rangle w_n$$

□

Exemple 4.3.1. Considérez l'espace euclidien $V = \mathbb{R}^3$. On veut trouver la projection orthogonale du vecteur

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

sur le sous-espace vectoriel

$$W = span\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$$

Pour ce faire, on commence par trouver une base orthogonale de W en appliquant la méthode de Gram-Schmidt :

$$\begin{cases} w_1 = \begin{pmatrix}1\\1\\0\end{pmatrix} \\ w_2 = \begin{pmatrix}1\\0\\1\end{pmatrix} - \dfrac{\left\langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix} \right\rangle} \begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\1\end{pmatrix} - \dfrac{1}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}1/2\\-1/2\\1\end{pmatrix} \end{cases}$$

On a donc la base orthogonale

$$\mathcal{B} = \left\{ \begin{pmatrix}1\\1\\0\end{pmatrix} \begin{pmatrix}1\\-1\\2\end{pmatrix} \right\}$$

Remarquez que pour simplifier les calculs, nous avons remplacer le second vecteur par un autre vecteur dans la même direction. Maintenant, avant de calculer la projection, nous devons trouver une base orthonormale en normalisant les vecteurs, ce qui nous donne :

$$\mathcal{B}' = \left\{ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-1\\2\end{pmatrix} \right\} = \left\{ \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}, \begin{pmatrix}1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix} \right\}$$

Et maintenant on peut calculer la projection :

$$\begin{aligned} proj_W(x) &= \left\langle \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix}, \begin{pmatrix}1\\2\\3\end{pmatrix} \right\rangle \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix} + \left\langle \begin{pmatrix}1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix}, \begin{pmatrix}1\\2\\3\end{pmatrix} \right\rangle \begin{pmatrix}1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix} \\ &= \frac{3}{\sqrt{2}}\begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\\0\end{pmatrix} + \frac{5}{\sqrt{6}}\begin{pmatrix}1/\sqrt{6}\\-1/\sqrt{6}\\2/\sqrt{6}\end{pmatrix} \\ &= \begin{pmatrix}3/2\\3/2\\0\end{pmatrix} + \begin{pmatrix}5/6\\-5/6\\10/6\end{pmatrix} \\ &= \begin{pmatrix}7/3\\2/3\\5/3\end{pmatrix} \end{aligned}$$

Théorème 4.3.5. Si V est un espace Euclidien de dimension finie et W est un sous-espace vectoriel, alors

$$(W^{\perp})^{\perp} = W$$

De plus, si A est une matrice, alors on a les relations d'orthogonalités suivantes :

1. $(\mathrm{Row}(A))^{\perp} = \ker(A)$
2. $(\ker(A))^{\perp} = \mathrm{Row}(A)$
3. $(\mathrm{Col}(A))^{\perp} = \ker(A^T)$
4. $(\ker(A^T))^{\perp} = \mathrm{Col}(A)$

Démonstration. Pour démontrer la première partie du théorème, supposons que $\mathcal{B}_W = \{v_1, v_2, ..., v_m\}$ est une base orthonormale de W. Alors cette base peut être complété en une base orthonormale $\mathcal{B} = \{v_1, v_2, ..., v_m, v_{m+1}, ...v_n\}$ de V. Maintenant, comme $W^\perp$ est aussi un sous-espace vectoriel de V, alors tout les vecteurs $w \in W^\perp$ peut s'écrire comme :

$$w = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + ... + \alpha_n v_n$$

Mais comme w est orthogonal au vecteur de W, on a donc que $\alpha_i = 0$ pour tout $i \in [1, m]$. On a donc montrer que

$$\mathcal{B}_{W^\perp} = \{v_{m+1}, v_{m+2}, ..., v_n\}$$

est une base orthonormale de $W^\perp$. En appliquant le même principe, on obtient que $\mathcal{B}$ doit donc aussi être une base de $(W^\perp)^\perp$, et donc :

$$W = (W^\perp)^\perp$$

Maintenant, pour les relations d'orthogonalités, on a :

1. On remarque que si A est une matrice de dimension $m \times n$ et $x \in \mathbb{R}^n$, alors le produit Ax reviens à calculer le produit scalaire entre les lignes de A et le vecteur x. Donc si $x \in (Row(A))^\perp$, cela reviens à dire que le produit scalaire de chacune des lignes de A avec le vecteur x doit être 0. On peut donc conclure que $Ax = 0$, c'est à dire que $x \in \ker(A)$.
2. Par la première relation d'orthogonalité on sait que $(Row(A))^\perp = \ker(A)$, maintenant, en prenant le complément orthogonal de chaque côté, on obtient :
$$((Row(A))^\perp)^\perp = (\ker(A))^\perp \qquad \Rightarrow \qquad Row(A) = (\ker(A))^\perp$$
3. Pour démontrer la troisième relation d'orthogonalité, on commence par se rappeler que $Row(A^T) = Col(A)$. On va donc appliquer la première relation d'orthogonalité, mais cette fois à la matrice A^T :
$$(Row(A^T))^\perp = \ker(A^T) \qquad \Rightarrow \qquad (Col(A))^\perp = \ker(A^T)$$
4. Pour cette dernière partie, on va tout simplement prendre le complément orthogonale de la partie précédente :
$$((Col(A))^\perp)^\perp = (\ker(A^T))^\perp \qquad \Rightarrow \qquad Col(A) = (\ker(A^T))^\perp$$

□

4.4 Les matrices orthogonales et transformations orthogonales

Nous allons maintenant dans cette section étudier les transformations linéaires sur un espace euclidien qui préservent le produit scalaire. Ce type de transformation linéaire est particulièrement important dans la théorie des espaces euclidiens. Nous allons premièrement montrer que dans une base orthonormale, la représentation matricielle d'une telle transformation linéaire a la propriété que son inverse correspond à sa transposer, ce qui peut particulièrement simplifier certain calcul. Nous allons ensuite étudier certaines propriétés de ces matrices.

Definition 4.4.1. Une matrice A de dimension $m \times n$ est dite orthogonale si $A^T A = I$, où I est la matrice identité de dimension $m \times m$.

Remarquez que bien que nous ayons définie une matrice orthogonale comme étant potentiellement non carré, dans le reste du texte nous seront presque exclusivement intéressé par le cas des orthogonales qui sont carrés.

Definition 4.4.2. Si V est un espace euclidien, et $T : V \to V$ une transformations linéaires, alors on dit que T est une transformation orhtogonale si $\langle T(x), T(y)\rangle = \langle x, y\rangle$, pour tout $x, y \in V$.

Nous allons maintenant démontrer que lorsque l'on fixe une base orthonormale, les notions de matrices orthogonales et de transformations orthogonales sont essentiellement équivalentes. Le point clé de la démonstration est de se rappeler que dans une base orthonormale, le produit scalaire peut s'écrire sous la forme :

$$\langle x, y \rangle = x^T y$$

Théorème 4.4.1. Si V est un espace euclidien de dimension finie, $\mathcal{B} = \{e_1, e_2, ..., e_n\}$ une base orthonormale, et $T : V \to V$ une transformation linéaire pouvant être représenté par la matrice A dans la base $\mathcal{B}$, c'est à dire $T(x) = Ax$ pour tout $x \in V$. Alors, les énoncés suivant sont équivalent :

1. A est une matrice orthogonale
2. T est une transformation orthogonale.

Démonstration.
(1) $\Rightarrow$ **(2)** Supposons que A est une matrice orthogonale, alors :

$$\langle T(x), T(y) \rangle = \langle Ax, Ay \rangle = (Ax)^T (Ay) = x^T A^T Ay = x^T y = \langle x, y \rangle$$

C'est à dire que T est une transformation orthogonale.
(2) $\Rightarrow$ **(1)** Supposons que $\langle T(x), T(y) \rangle = \langle x, y \rangle$ pour tout $x, y \in V$, et prenons $e_i, e_j \in \mathcal{B}$. Alors on a :

$$[A^T A]_{ij} = e_i^T A^T A e_j = (Ae_i)^T (Ae_j) = \langle Ae_i, Ae_j \rangle = \langle T(e_i), T(e_j) \rangle = \langle e_i, e_j \rangle = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

On obtient donc $A^T A = I$, c'est à dire que la matrice A est orthogonale. □

Remarquez qu'en suivant suivant notre habitude d'associer une transformation linéaire avec sa représentation matricielle, on obtient donc que dans une base orthogonale, une matrice est orthogonale si et seulement si $\langle Ax, Ay \rangle = \langle x, y \rangle$ pour tout $x, y \in V$. Donc les matrices orthogonales sont précisément les matrices pour lesquelles le produit scalaire est préservé. C'est à dire que les matrices orthogonale représentent les applications linéaires préservant les angles et les longueurs.

Théorème 4.4.2.

1. Si A est une matrice orthogonal carré, alors A^T est aussi une matrice orthogonale et $A^{-1} = A^T$.
2. Si A est une matrice carré, alors A est orthogonale si et seulement si les colonnes de A forme une base orthonormale.
3. Si A est une matrice carré orthogonale, alors $det(A) = \pm 1$.

Démonstration.

1. Si A est une matrice carré tel que $A^T A = I$, alors par définition de l'inverse d'une matrice $A^{-1} = A^T$. Le fait que A^T est aussi une matrice orthogonale se vérifie alors directement.
2. Exercice
3. Exercice

□

4.5 La factorisation QR

Dans cette section, nous voulons factoriser une matrice A sous la forme QR, où Q est une matrice orthogonale, et R est une matrice triangulaire supérieure. Comme application, nous aurons entre autre la résolution des systèmes d'équations linéaires et l'approximation des valeurs propres.

Théorème 4.5.1. Si A est une matrice carré ayant des colonnes linéairement indépendantes, alors la matrice A peut être factoriser sous la forme $A = QR$, où Q est une matrice orthogonale (carré) et R est une matrice triangulaire supérieure. De plus, on obtient la matrice Q en appliquant la méthode de Gram-Schmidt aux vecteurs colonnes de la matrice A, et en normalisant ensuite ces vecteurs. On peut ensuite calculer $R = Q^T A$.

Exemple 4.5.1. Trouver la factorisation QR de la matrice

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

On commence donc par appliquer la méthode de Gram-Schmidt aux vecteurs de l'ensemble

$$S = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\}$$

On obtient donc :

$$\begin{cases} w_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \\ w_2 = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \dfrac{\left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\rangle} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \dfrac{14}{10} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3/5 \\ -1/5 \end{pmatrix} \end{cases}$$

Maintenant on doit normaliser chacun des vecteurs que nous avons obtenu :

$$\begin{cases} w_1' = \dfrac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix} \\ w_2' = \dfrac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{pmatrix} \end{cases}$$

On obtient donc la matrice Q suivante :

$$Q = \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix}$$

Il ne nous reste plus qu'à trouver la matrice R, pour ce faire, on a :

$$R = Q^T A = \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} \sqrt{10} & 14/\sqrt{10} \\ 0 & 2/\sqrt{10} \end{pmatrix}$$

On obtient donc la factorisation suivante :

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{10} & 14/\sqrt{10} \\ 0 & 2/\sqrt{10} \end{pmatrix}$$

Exemple 4.5.2. Utiliser la factorisation QR que nous avons trouvé précédemment pour résoudre le système d'équations linéaires suivant :

$$\begin{cases} x + 2y = 7 \\ 3x + 4y = 19 \end{cases}$$

On commence par réécrire le système d'équations sous forme matricielle :

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 19 \end{pmatrix}$$

Donc en utilisant la factorisation QR, on obtient :

$$\begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{10} & 14/\sqrt{10} \\ 0 & 2/\sqrt{10} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 19 \end{pmatrix}$$

Puis en utilisant le fait que la matrice Q est orthogonale, on obtient :

$$\begin{pmatrix} \sqrt{10} & 14/\sqrt{10} \\ 0 & 2/\sqrt{10} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} 7 \\ 19 \end{pmatrix} = \begin{pmatrix} 64/\sqrt{10} \\ 2/\sqrt{10} \end{pmatrix}$$

On peut donc maintenant trouver les valeurs de x et y :

$$\frac{2}{\sqrt{10}} y = \frac{2}{\sqrt{10}} \Longrightarrow y = 1$$

$$\sqrt{10}x + \frac{14}{\sqrt{10}}(1) = \frac{64}{\sqrt{10}} \Longrightarrow \sqrt{10}x = \frac{50}{\sqrt{10}} \Longrightarrow y = 5$$

Exemple 4.5.3. On veut trouver la factorisation QR de la matrice

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

Pour ce faire, nous allons à nouveau commencer par appliquer la méthode de Gram-Schmidt aux vecteurs de l'ensemble :

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \right\}$$

Ce qui nous donne :

$$\begin{cases} w_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \\ w_2 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \dfrac{\left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} - \dfrac{3}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1 \\ -1/2 \end{pmatrix} \to \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \\ w_3 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} - \dfrac{\left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\rangle} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \dfrac{\left\langle \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \right\rangle}{\left\langle \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\rangle} \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -1/3 \\ -1/3 \end{pmatrix} \to \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \end{cases}$$

Puis on normalise chacun des vecteurs ci dessus :

$$\begin{cases} w_1' = \dfrac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{pmatrix} \\ w_2' = \dfrac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\-1\end{pmatrix} = \begin{pmatrix}1/\sqrt{6}\\2/\sqrt{6}\\-1/\sqrt{6}\end{pmatrix} \\ w_3' = \dfrac{1}{\sqrt{3}}\begin{pmatrix}1\\-1\\-1\end{pmatrix} = \begin{pmatrix}1/\sqrt{3}\\-1/\sqrt{3}\\-1/\sqrt{3}\end{pmatrix} \end{cases}$$

On obtient donc la matrice Q suivante :

$$Q = \begin{pmatrix}1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3}\\ 0 & 2/\sqrt{6} & -1/\sqrt{3}\\ 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3}\end{pmatrix}$$

Puis on calcul la matrice R :

$$\begin{pmatrix}1/\sqrt{2} & 0 & 1/\sqrt{2}\\ 1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6}\\ 1/\sqrt{3} & -1/\sqrt{3} & -1/\sqrt{3}\end{pmatrix}\begin{pmatrix}1 & 2 & 3\\ 0 & 1 & 2\\ 1 & 1 & 0\end{pmatrix} = \begin{pmatrix}\sqrt{2} & 3/\sqrt{2} & 3/\sqrt{2}\\ 0 & \sqrt{6}/2 & 7/\sqrt{6}\\ 0 & 0 & 1/\sqrt{3}\end{pmatrix}$$

Ce qui nous donne finalement la factorisation suivante :

$$\begin{pmatrix}1 & 2 & 3\\ 0 & 1 & 2\\ 1 & 1 & 0\end{pmatrix} = \begin{pmatrix}1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3}\\ 0 & 2/\sqrt{6} & -1/\sqrt{3}\\ 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{3}\end{pmatrix}\begin{pmatrix}\sqrt{2} & 3/\sqrt{2} & 3/\sqrt{2}\\ 0 & \sqrt{6}/2 & 7/\sqrt{6}\\ 0 & 0 & 1/\sqrt{3}\end{pmatrix}$$

4.6 Application : La méthode des moindres carrés

Dans le cours d'algèbre linéaire I, beaucoup de temps et d'effort on été mis sur la résolution d'un système d'équation linéaire de la forme $Ax = b$, où x est le vecteur contenant les inconnus. Nous avons vu que 3 cas peuvent se produire :

— Le système n'a aucune solution

— Le système a exactement une solution

— Le système a une infinité de solution

Le problème est que dans plusieurs cas, bien qu'un système admette aucune solution, on a besoin d'avoir la meilleure solution possible. C'est ce qu'on appelle le problème des moindres carrés.

Théorème 4.6.1. Si A est une matrice de dimension $m \times n$ ayant des colonnes linéairement indépendante, et $b \in \mathbb{R}^n$, alors il existe un unique vecteur $y \in \mathbb{R}^m$ tel que :

$$||Ay - b|| \leq ||Ax - b|| \quad \text{pour tout } x \in \mathbb{R}^m$$

Démonstration. Supposons que V est l'espace vectoriel $\mathbb{R}^n$ et W est un sous-espace vectoriel de V, alors nous avons vu que pour tout $b \in V$, il existe un unique vecteur $z \in W$ qui est le point de W le plus proche de b. Si on suppose que l'espace $W = Im(A)$, alors ce point peut s'écrire sous la forme $z = Ay$. □

Definition 4.6.1. Si A est une matrice de dimension $m \times n$ ayant des colonnes linéairement indépendante, et $b \in \mathbb{R}^n$, alors $y \in \mathbb{R}^m$ est une solution au sens des moindres carrés, si y est l'unique vecteur tel que :

$$||Ay - b|| \leq ||Ax - b|| \quad \text{pour tout } x \in \mathbb{R}^m$$

Théorème 4.6.2. Si A est une matrice de dimension $m \times n$ ayant des colonnes linéairement indépendante, et $b \in \mathbb{R}^n$, alors $y \in \mathbb{R}^m$ est une solution au sens des moindres carrés si et seulement si

$$Ay = \text{proj}_{\text{Im}(A)} b$$

Démonstration. Il s'agit d'une conséquence directe du fait que $proj_{Im(A)}b$ est le point de $Im(A)$ le plus proche de b. □

Théorème 4.6.3. Si A est une matrice de dimension $m \times n$ ayant des colonnes linéairement indépendante, et $b \in \mathbb{R}^n$, alors $y \in \mathbb{R}^m$ est une solution au sens des moindres carrés si et seulement si

$$A^T Ay = A^T b$$

Démonstration. Premièrement, posons

$$b' = proj_{Im(A)} b$$

donc par définition de la projection, on a que

$$(b - b') \in [Im(A)]^{\perp}$$

Comme b' est dans l'image de la matrice, il existe donc un $y \in \mathbb{R}^m$ tel que $b' = Ay$ et donc $b' - Ay$ est orthogonal à chaque colonne de A, car l'image de A est l'espace généré (span) par les colonnes de la matrice A. On en déduit donc que $A^T(b - Ay) = 0$ et donc en réarrangeant les termes :

$$A^T Ay = A^T b$$

□

Exemple 4.6.1. On veut résoudre le système d'équation suivant au sens des moindres carrés :

$$\begin{cases} x + 2x = 1 \\ 3x + 4y = 2 \\ 5x + 6y = 4 \end{cases}$$

Bien que ce ne soit pas vraiment nécessaire, nous allons commencer par essayer de résoudre le système de la manière habituelle. C'est à dire qu'on va essayer de trouver la solution du système :

$$\left(\begin{array}{cc|c} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 5 & 6 & 4 \end{array}\right) \sim^{3L_1 - L_2 \to L_2}_{5L_1 - L_3 \to L_3} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 4 & 1 \end{array}\right) \sim^{2L_2 - L_3 \to L_3} \left(\begin{array}{cc|c} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{array}\right)$$

Donc le système n'a aucune solution. Nous allons donc chercher la meilleur solution possible (C'est à dire les valeurs de x et y qui sont les plus proche d'être des solutions) en utilisant les deux méthodes que nous avons vu, en commençant par la méthode de la projection. On commence par réécrire le système d'équation sous forme matricielle :

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

On commence donc par trouver une base orthonormale de

$$W = span\left\{\begin{pmatrix}1\\3\\5\end{pmatrix}, \begin{pmatrix}2\\4\\6\end{pmatrix}\right\}$$

Pour ce faire on commence par appliquer la méthode de Gram-Schmidt :

$$\begin{cases} w_1 = \begin{pmatrix}1\\3\\5\end{pmatrix} \\ w_2 = \begin{pmatrix}2\\4\\6\end{pmatrix} - \dfrac{\left\langle \begin{pmatrix}1\\3\\5\end{pmatrix}, \begin{pmatrix}2\\4\\6\end{pmatrix}\right\rangle}{\left\langle \begin{pmatrix}1\\3\\5\end{pmatrix}, \begin{pmatrix}1\\3\\5\end{pmatrix}\right\rangle} \begin{pmatrix}1\\3\\5\end{pmatrix} = \begin{pmatrix}2\\4\\6\end{pmatrix} - \dfrac{44}{35}\begin{pmatrix}1\\3\\5\end{pmatrix} = \begin{pmatrix}26/35\\8/35\\-2/7\end{pmatrix} \longrightarrow \begin{pmatrix}26\\8\\-10\end{pmatrix} \end{cases}$$

Puis on normalise les vecteurs :

$$\begin{cases} w_1' = \dfrac{1}{\sqrt{35}}\begin{pmatrix}1\\3\\5\end{pmatrix} = \begin{pmatrix}1/\sqrt{35}\\3/\sqrt{35}\\5/\sqrt{35}\end{pmatrix} \\ w_2' = \dfrac{1}{\sqrt{840}}\begin{pmatrix}26\\8\\-10\end{pmatrix} = \begin{pmatrix}13/\sqrt{210}\\4/\sqrt{210}\\-5/\sqrt{210}\end{pmatrix} \end{cases}$$

On veut maintenant calculer la projection :

$$proj_W\begin{pmatrix}1\\2\\4\end{pmatrix} = \left\langle\begin{pmatrix}1\\2\\4\end{pmatrix}, \begin{pmatrix}1/\sqrt{35}\\3/\sqrt{35}\\5/\sqrt{35}\end{pmatrix}\right\rangle \begin{pmatrix}1/\sqrt{35}\\3/\sqrt{35}\\5/\sqrt{35}\end{pmatrix} + \left\langle\begin{pmatrix}1\\2\\4\end{pmatrix}, \begin{pmatrix}13/\sqrt{210}\\4/\sqrt{210}\\-5/\sqrt{210}\end{pmatrix}\right\rangle \begin{pmatrix}13/\sqrt{210}\\4/\sqrt{210}\\-5/\sqrt{210}\end{pmatrix} = \begin{pmatrix}5/6\\7/3\\23/6\end{pmatrix}$$

Puis on doit résoudre le système d'équations linéaires suivant :

$$\begin{cases} x + 2y = 5/6 \\ 3x + 4y = 7/3 \\ 5x + 6y = 23/6 \end{cases}$$

On applique donc la méthode de Gauss :

$$\left(\begin{array}{cc|c}1 & 2 & 5/6\\3 & 4 & 7/3\\5 & 6 & 23/6\end{array}\right) \sim_{\substack{3L_1-L_2\to L_2\\5L_1-L_3\to L_3}} \left(\begin{array}{cc|c}1 & 2 & 5/6\\0 & 2 & 1/6\\0 & 4 & 1/3\end{array}\right) \sim^{2L_2-L_3\to L_3} \left(\begin{array}{cc|c}1 & 2 & 5/6\\0 & 2 & 1/6\\0 & 0 & 0\end{array}\right)$$

On obtient donc $y = \dfrac{1}{12}$ et $x = \dfrac{2}{3}$ qui est une solution au sens des moindres carrés.

On va maintenant refaire le calcul, mais cette fois en utilisant la seconde méthode. Nous allons voir que cette seconde méthode est en fait beaucoup plus rapide. Comme nous l'avons déjà fait, nous allons écrire le système d'équations sous forme matricielle :

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

Puis nous multiplions de chaque côté par la transposé de la matrice des coefficients :

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

Ce qui nous donne le système d'équations suivant :

$$\begin{pmatrix} 35 & 44 \\ 44 & 56 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 27 \\ 34 \end{pmatrix}$$

que l'on résout à l'aide de la méthode de Gauss :

$$\left(\begin{array}{cc|c} 35 & 44 & 27 \\ 44 & 56 & 34 \end{array}\right) \sim^{44L_1-35L_2\to L_2} \left(\begin{array}{cc|c} 35 & 44 & 27 \\ 0 & -24 & -2 \end{array}\right)$$

Ce qui nous donne comme solution : $x = \dfrac{2}{3}$ et $y = \dfrac{1}{12}$. On voit donc que les deux méthodes nous donne bien la même solution. La seconde méthode est cependant beaucoup plus simple.

Exemple 4.6.2. On veut trouver l'équation de la parabole qui est le plus proche (au sens des moindres carrés) de passer par les points suivants :

x	y
-3	8
-2	3
-1	4
0	3
1	8
2	11
3	20

Comme l'équation d'une parabole est de la forme

$$y = ax^2 + bx + c,$$

on va remplacer les valeurs de x et y par chacun des points ci dessus, ce qui nous donne le système d'équations linéaires suivant :

$$\begin{cases} 9a - 3b + c = 8 \\ 4a - 2b + c = 3 \\ a - b + c = 4 \\ c = 3 \\ a + b + c = 8 \\ 4a + 2b + c = 11 \\ 9a + 3b + c = 20 \end{cases}$$

Comme ce système n'a aucune solution, on va appliquer la méthode des moindres carrés. Commençons par écrire le système d'équations sous forme matricielle :

$$\begin{pmatrix} 9 & -3 & 1 \\ 4 & -2 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 8 \\ 3 \\ 4 \\ 3 \\ 8 \\ 11 \\ 29 \end{pmatrix}$$

On doit donc résoudre le système suivant :

$$\begin{pmatrix} 9 & 4 & 1 & 0 & 1 & 4 & 9 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 9 & -3 & 1 \\ 4 & -2 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 9 & 4 & 1 & 0 & 1 & 4 & 9 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 8 \\ 3 \\ 4 \\ 3 \\ 8 \\ 11 \\ 29 \end{pmatrix}$$

Ce qui nous donne :

$$\begin{pmatrix} 196 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 7 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 401 \\ 83 \\ 66 \end{pmatrix}$$

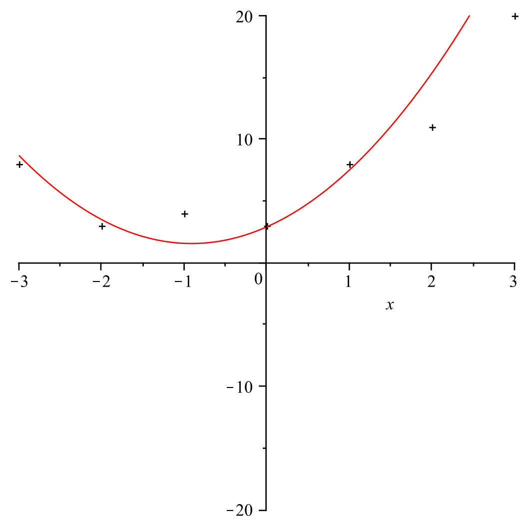
On applique donc la méthode de Gausse, ce qui nous donne :

$$\left(\begin{array}{ccc|c} 196 & 0 & 28 & 401 \\ 0 & 28 & 0 & 83 \\ 28 & 0 & 7 & 66 \end{array}\right) \sim^{L_1 - 7L_3 \to L_3} \left(\begin{array}{ccc|c} 196 & 0 & 28 & 401 \\ 0 & 28 & 0 & 83 \\ 0 & 0 & -21 & -61 \end{array}\right)$$

Donc la solution est $c = \frac{61}{21}$, $b = \frac{83}{28}$ et $a = \frac{137}{84}$ ce qui nous donne l'équation suivante pour la parabole :

$$y = \left(\frac{137}{84}\right) x^2 + \left(\frac{83}{28}\right) x + \left(\frac{61}{21}\right)$$

Voici un graphique représentant le problème :



On peut donc voir que la parabole que nous avons trouvez semble bien représenter le problème.

Théorème 4.6.4. **(Théorème des moindres carrés)** Si A est une matrice de dimension $m \times n$, alors les énoncés suivant sont équivalent :

1. Les colonnes de A sont linéairement indépendante
2. Le système d'équation $Ax = 0$ a une unique solution
3. Le système d'équation $A^T Ax = 0$ a une unique solution
4. La matrice $A^T A$ est inversible

Démonstration.
(1) $\Leftrightarrow$ **(2)** Il s'agit de remarquer que si

$$A = [v_1 | v_2 | v_3 | \ldots | v_n]$$

où les v_i représente les colonnes de la matrice A, alors on a que A a des colonnes linéairement indépendantes si et seulement si

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \ldots + \lambda_n v_n = 0 \text{ a une seule solution}$$

Ce qui est le cas si et seulement si

$$[v_1|v_2|\ldots|v_n]\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}=\begin{pmatrix}0\\0\\ \vdots\\0\end{pmatrix} \text{ a une seule solution}$$

Ce qui est le cas si et seulement si :

$$Ax = 0 \text{ a une seule solution}$$

(2) ⇔ (3) On commence par remarquer que si x est une solution de $Ax = 0$, alors x est aussi une solution de $A^T Ax = 0$ (Il s'agit de multiplier par A^T de chaque côté. Maintenant, supposons que x est une solution de $A^T Ax = 0$, alors x est aussi une solution de $x^T A^T Ax = 0$ (Il s'agit de multiplier par x^T de chaque côté). On obtient donc :

$$0 = x^T A^T Ax = (Ax)^T Ax = \langle Ax, Ax\rangle$$

et donc par les propriétés du produit scalaire, on peut conclure que $Ax = 0$. Donc x est une solution de $Ax = 0$ si et seulement si x est une solution de $A^T Ax = 0$. Donc si on suppose que la solution est unique dans l'une des deux équations, alors elle sera unique dans l'autre.
(3) ⇔ (4) Comme $A^T A$ est toujours une matrice carré, alors il s'agit d'une conséquence directe du théorème de la matrice inverse. □

4.7 Application : Algorithme QR pour l'approximation des valeurs propres

Pour terminer se chapitre, nous allons voir une application particulièrement utile de la factorisation QR. Nous avons déjà vu les méthodes de la puissance et de la puissance inverse pour calculer la plus grande et la plus petite valeur propre. La question demeure cependant ouverte à savoir comment trouver les autres valeurs propres dans le cas d'une matrice A de grande dimension. Dans cette section, nous allons voir une méthode permettant la plupart des cas d'approximer toutes les valeurs propres d'une matrice.

Théorème 4.7.1. Si A est une matrice inversible, et $A = QR$ où Q est une matrice orthogonale et R est triangulaire supérieure, alors A et RQ ont les mêmes valeurs propres.

Démonstration. Si $A = QR$ alors $Q^T A = R$ ce qui nous donne :

$$RQ = Q^T AQ = Q^{-1} AQ$$

Donc A et RQ sont deux matrices semblables, et donc elles ont les mêmes valeurs propres. □

On peut donc utiliser ce théorème pour créer un algorithme (l'algorithme QR) pour approximer les valeurs propres d'une matrice

$$\begin{cases} A_0 = A \\ A_k = Q_k R_k, \quad \forall k \geq 0 & \text{la factorisation QR de la matrice} A_k \\ A_{k+1} = R_k Q_k, \quad \forall k \geq 1 \end{cases}$$

Pour chaque k (itérations) les valeurs sur la diagonale de A_k sont des approximations des valeurs propres de la matrice A.

Exemple 4.7.1. On veut utiliser l'algorithme QR pour approximer toutes les valeurs propres de la matrice

$$A = \begin{pmatrix}-70 & 150\\ -45 & 95\end{pmatrix}$$

On commence par calculer une première factorisation QR :

$$A_0 = Q_0 R_0 \begin{pmatrix} -0,8412 & 0,5408 \\ -0,5408 & -0,8412 \end{pmatrix} \begin{pmatrix} 83,2166 & -177,5487 \\ 0 & 1,2017 \end{pmatrix}$$

Puis on inverse les deux matrices et on calcul le produit :

$$A_1 = R_0 Q_0 = \begin{pmatrix} 83,2166 & -177,5487 \\ 0 & 1,2017 \end{pmatrix} \begin{pmatrix} -0,8412 & 0,5408 \\ -0,5408 & -0,8412 \end{pmatrix} = \begin{pmatrix} 26,0108 & 194,3502 \\ -0,6498 & -1,0108 \end{pmatrix}$$

Puis on calcule la factorisation QR de cette nouvelle matrice :

$$A_1 = Q_1 R_1 = \begin{pmatrix} 0,9997 & 0,025 \\ -0,025 & 0,9997 \end{pmatrix} \begin{pmatrix} 26,0189 & 194,3148 \\ 0 & 3,8434 \end{pmatrix}$$

Puis on inverse à nouveau la matrice, ce qui nous donne :

$$A_2 = R_1 Q_1 = \begin{pmatrix} 26,0189 & 194,3148 \\ 0 & 3,8434 \end{pmatrix} \begin{pmatrix} 0,9997 & 0,025 \\ -0,025 & 0,9997 \end{pmatrix} = \begin{pmatrix} 21,1578 & 194,9040 \\ -0,096 & 3,8422 \end{pmatrix}$$

Et on continue de la même manière en calculant la factorisation QR de cette dernière matrice :

$$A_2 = Q_2 R_2 = \begin{pmatrix} 1 & 0,0045 \\ -0,0045 & 1 \end{pmatrix} \begin{pmatrix} 21,1581 & 194,8846 \\ 0 & 4,7263 \end{pmatrix}$$

Puis on inverse l'ordre des deux matrices et on les remultiplies :

$$A_3 = \begin{pmatrix} 21,1581 & 194,8846 \\ 0 & 4,7263 \end{pmatrix} \begin{pmatrix} 1 & 0,0045 \\ -0,0045 & 1 \end{pmatrix} = \begin{pmatrix} 20,2737 & 194,9786 \\ -0,0214 & 4,7263 \end{pmatrix}$$

Donc après 3 itérations, nous avons comme approximation des valeurs propres les valeurs : 20,2737 et 4,7263, que nous pouvons comparer avec les vrais valeurs propres qui sont 20 et 5.

4.8 Retour sur les 4 espaces fondamentaux

Nous allons maintenant compléter ce chapitre en revenant sur l'étude des 4 espaces fondamentaux que nous avons commencé au chapitre 2. Dans un premier temps, nous allons regarder le lien entre les projections, l'orthogonalité et l'application linéaire $f(x) = Ax$. Pour ce faire, illustrons le problème avec un schéma comme nous l'avons fait au chapitre 2. [1]

Dans le diagramme, V et W représentent des espaces vectoriels, et $f : V \to W$ est une application linéaire représenté par la matrice A. Dans un premier temps, comme $Row(A)$ et $null(A)$ sont orthogonaux, ainsi que $Col(A)$ et $null(A^T)$ sont orthogonaux, on le représente sur notre schéma à l'aide d'un symbol d'angle droit. Ensuite, si x est un vecteur quelconque de V, alors le concept de projection orthogonal nous permet de décomposer x sous la forme $x = x_1 + x_2$ avec $\langle x_1, x_2 \rangle = 0$, $x_1 \in Row(A)$ et $x_2 \in null(A)$. Finalement, comme $x_2 \in null(A)$, on a $Ax_2 = 0$, ce qui nous permet de trouver :

$$Ax = A(x_1 + x_2) = Ax_1 + Ax_2 = Ax_1 + 0 = Ax_1$$

Ce qui signifie que la composante de x qui se trouve dans la direction du sous-espace vectoriel $Row(A)$ est en fait la seule qui est importante pour la transformation linéaire f, le reste étant tout simplement envoyer sur 0. Regardons maintenant le lien entre les 4 espaces fondamentaux et le problème des moindres carrés. Pour ce faire, nous allons à nouveau faire un diagramme :

Le problème est de résoudre le système d'équations linéaires $Ax = b$. Si $b \in Col(A)$, le système possède au moins une solution. Il n'y a donc pas lieu d'utiliser la méthode des moindres carrés. Si au contraire $b \notin Col(A)$, alors le système possède aucune solution. Dans ce cas, on peut décomposer le vecteur b sous la

1. Notez que le diagramme, bien que légèrement différent, est fortement inspiré de l'article de Strang [6]

FIGURE 4.1 – Les 4 espaces fondamentaux et système d'équations linéaires

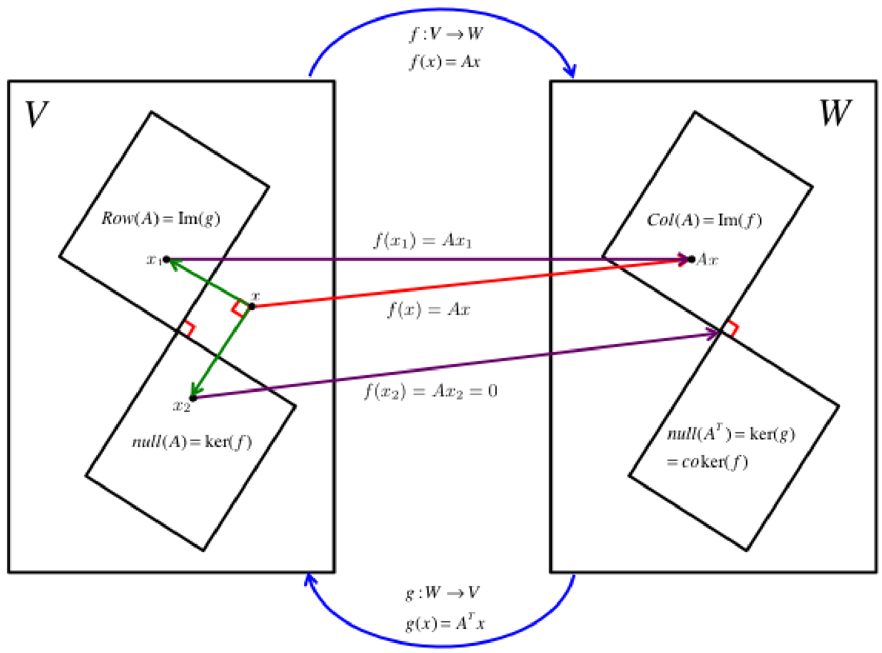
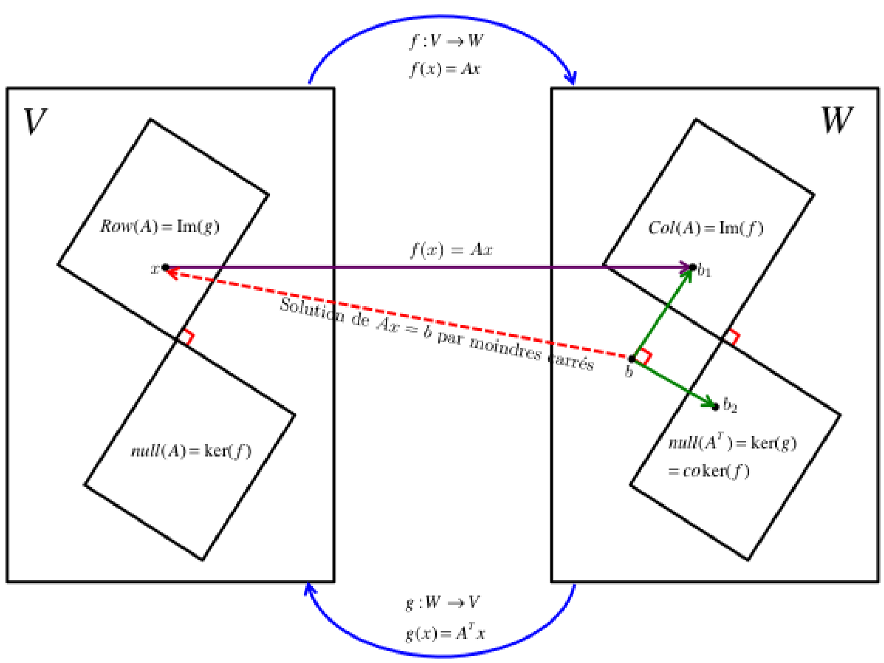


FIGURE 4.2 – Les 4 espaces fondamentaux et moindres carrés



forme $b = b_1 + b_2$ où $\langle b_1, b_2 \rangle = 0$, $b_1 \in Col(A)$ et $b_2 \in null(A^T)$. Dans ce cas, b_1 est en fait le vecteur le plus proche de b qui va permettre de trouver une solution. Donc la solution de $Ax = b$ au sens des moindres carrés est en fait la solution du système d'équation linéaire $Ax = b_1$.

Finalement, remarquez que le diagramme que nous venons de produire ressemble étrangement à une réflexion du diagramme précédent. Ceci donne l'impression que nous venons de trouver une méthode pour calculer l'inverse de n'importe quelle matrice. Ceci est complètement faux, mais n'est pas non plus très loin de la vérité. Il existe une notion appelé pseudo-inverse qui permet de produire une matrice qui ressemble à un inverse pour n'importe quelle matrice. Bien que ce concept est plusieurs applications intéressantes, nous ne l'étudierons pas ici. La plupart des idées importantes reliés à la notion de pseudo-inverse se trouve cependant dans le diagramme.

Chapitre 5

Les matrices symétriques et les formes bilinéaires

5.1 Les matrices symétriques et les applications symétriques

Dans ce chapitre, nous allons étudier l'un des types de matrices qui se retrouve le plus couramment dans les applications : Les matrices symétriques. L'avantage de travailler avec des matrices qui sont symétriques est qu'elles sont toujours diagonalisable. Mieux encore, il est toujours possible de diagonaliser la matrice en utilisant une base orthonormale.

Comme pour les matrices / applications orthogonales, nous allons commencer par définir une notion de matrice symétrique et une notion d'application symétrique, puis démontrer que dans le cas des bases orthogonales, c'est deux notions sont en fait équivalentes.

Definition 5.1.1. Si A est une matrice, alors on dit que A est symétrique si $A^T = A$.

Théorème 5.1.1.

1. Si A est une matrice symétrique, alors A est une matrice carré.
2. Si A est une matrice quelconque, alors $A^T A$ est une matrice symétrique.

Démonstration.

1. Il s'agit de comparer la dimension de A avec celle de A^T.
2. Exercice

□

Definition 5.1.2. Si V est un espace euclidien et $T : V \to V$ une application linéaire, alors on dit que T est symétrique si $\langle T(x), y\rangle = \langle x, T(y)\rangle$ pour tout $x, y \in V$.

Théorème 5.1.2. Si V est un espace euclidien muni d'une base orthonormale $\mathcal{B} = \{e_1, e_2, e_3, ..., e_n\}$, et $T : V \to V$ une application linéaire représenté dans la base $\mathcal{B}$ par la matrice A. C'est à dire $T(x) = Ax$ pour tout $x \in V$. Alors, les énoncés suivants sont équivalent :

1. A est une matrice symétrique.
2. T est une application symétrique.

Démonstration. Comme dans le cas des matrices / applications orthogonales, le point clé de notre démonstration est de se rappeler que dans une base orthogonale, le produit scalaire prend la forme $\langle x, y\rangle = x^T y$ pour tout $x, y \in V$. On obtient donc :
(1) $\Rightarrow$ (2) : Supposons que A est une matrice symétrique, alors on a :

$$\langle T(x), y\rangle = \langle Ax, y\rangle = (Ax)^T y = x^T A^T y = x^T Ay = \langle x, Ay\rangle = \langle x, T(y)\rangle$$

L'application linéaire T est donc symétrique.
(2) $\Rightarrow$ (1) : Supposons maintenant que l'application linéaire T est symétrique. On a donc :

$$A_{ij} = e_i^T Ae_j = \langle e_i, Ae_j\rangle = \langle e_i, T(e_j)\rangle = \langle T(e_i), e_j\rangle = \langle Ae_i, e_j\rangle = (Ae_i)^T e_j = e_i^T A^T e_j = A_{ij}^T$$

Étant donné que les différentes composante de la matrice A et A^T sont égales, on obtient donc que $A = A^T$, c'est à dire que la matrice A est symétrique. □

En conservant notre habitude d'associer une application linéaire avec la matrice qui la représente, on obtient donc que dans une base orthonormale, une matrice A est symétrique si et seulement si $\langle Ax, y\rangle = \langle x, Ay\rangle$ pour tout $x, y \in V$.

Remarquer finalement la similarité avec les matrices orthogonale que nous avons étudié au chapitre précédent. On se rappel qu'une matrice A est orthogonale si et seulement si $\langle Ax, Ay\rangle = \langle x, y\rangle$ pour tout $x, y \in \mathbb{R}^n$. Les connexions entre ces deux types de matrices ne s'arrête pas là, mais nous y reviendrons un peu plus tard.

5.2 Les formes bilinéaires et quadratiques

Nous avons déjà vu qu'une application linéaire peut être représenté sous forme de matrice. Dans cette section, nous allons voir qu'une forme quadratique peut être représenté sous forme de matrice symétrique.

Definition 5.2.1. Si A est une matrice symétrique de dimension $n \times n$, alors on définit la forme bilinéaire $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ associé à A comme étant :

$$b(x, y) = x^T Ay$$

Nous allons maintenant montrer que la forme bilinéaire que nous venons de définir est bien une forme bilinéaire au sens classique du terme, c'est à dire une fonction qui est linéaire selon chacun de ses paramètre.

Théorème 5.2.1. Si A est une matrice symétrique, alors la forme bilinéaire b associé à la matrice A a les propriétés suivantes :

1. $b(x + y, z) = b(x, z) + b(y, z), \quad \forall x, y, z \in \mathbb{R}^n$
2. $b(x, y + z) + b(x, z) + b(y, z), \quad \forall x, y, z \in \mathbb{R}^n$
3. $b(\lambda x, y) = \lambda b(x, y) = b(x, \lambda y), \quad \forall x, y \in \mathbb{R}^n, \quad \forall \lambda \in \mathbb{R}$

Démonstration. Exercice □

Théorème 5.2.2. Tout les produits scalaire $\langle \cdot, \cdot\rangle$ sur $\mathbb{R}^n$ sont des formes bilinéaire associé à une matrice symétrique A. En d'autre mot, pour tout produit scalaire $\langle \cdot, \cdot\rangle$, il existe une matrice symétrique A tel que :

$$\langle x, y\rangle = x^T Ay, \qquad \forall x, y \in \mathbb{R}^n$$

Démonstration. Considérons la base habituelle de $\mathbb{R}^n$ suivante $\{e_1, e_2, ..., e_n\}$. Si $x, y \in \mathbb{R}^n$, alors on peut écrire :

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ v_n \end{pmatrix}$$

On obtient donc :

$$\begin{aligned}
\langle x, y\rangle &= \langle x_1e_1 + x_2e_2 + ... + x_ne_n, y_1e_1 + y_2e_2 + ... + y_ne_n\rangle \\
&= \sum_{i,j=1}^{n} x_i\langle e_i, e_j\rangle y_j \\
&= \begin{pmatrix} x_1 & x_2 & ... & x_n \end{pmatrix} \begin{pmatrix} \langle e_1, e_1\rangle & \langle e_1, e_2\rangle & ... & \langle e_1, e_n\rangle \\ \langle e_2, e_1\rangle & \langle e_2, e_2\rangle & ... & \langle e_2, e_n\rangle \\ \vdots & \vdots & & \vdots \\ \langle e_n, e_1\rangle & \langle e_n, e_2\rangle & ... & \langle e_n, e_n\rangle \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\
&= x^TAy
\end{aligned}$$

où la matrice A est définit par $A_{ij} = \langle e_i, e_j\rangle$. □

Exemple 5.2.1. Pour l'espace euclidien $V = \mathbb{R}^3$ muni du produit scalaire habituel, alors :

$$\langle x, y\rangle = x^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} y$$

Definition 5.2.2. Si A est une matrice symétrique de dimension $n \times n$, alors on définit la forme quadratique $q : \mathbb{R}^n \to \mathbb{R}$ comme étant :

$$q(x) = x^TAx$$

5.3 Le théorème spectral

Le théorème que nous allons voir dans cette section est l'un des théorèmes les plus important de l'algèbre linéaire, au même titre que sont les théorèmes du rang, de la matrice inverse et des moindres carrés. Le théorème spectrale voit aussi plusieurs généralisation dans les nombres complexes, puis dans le cadre de l'analyse fonctionnelle.

Definition 5.3.1. Si A est une matrice, alors on dit que A est diagonalisable orthogonalement s'il existe une matrice orthogonale P et une matrice diagonale D tel que :

$$A = P^TDP = P^{-1}DP$$

Théorème 5.3.1. **(Théorème spectral)**

1. Toutes les valeurs propres d'une matrice symétrique sont réelles
2. Les espaces propres d'une matrice symétrique sont mutuellement orthogonaux
3. Une matrice A est symétrique si et seulement si A peut être diagonalisé orthogonalement

Démonstration.

1. Commençons par définir $q(x) = \overline{x}^T Ax$ où $x \in \mathbb{C}^n$. Alors on a :

$$\overline{q(x)} = \overline{\overline{x}^T Ax} = x^T \overline{Ax} = x^T A\overline{x} = (x^T A\overline{x})^T = \overline{x}^T A^T x = q(x), \quad \forall x \in \mathbb{C}^n$$

Donc pour n'importe quelle vecteur $x \in \mathbb{C}^n$, on a que $q(x)$ est un nombre réel. Maintenant, supposons que x est un vecteur propre (complexe) associé à une valeur propre (complexe) λ de la matrice A. alors on a :

$$q(x) = \overline{x}^T Ax = \overline{x}^T \lambda x = \lambda(\overline{x}^T x)$$

Comme $\overline{x}^T x$ est toujours un nombre réel, et que $q(x)$ est aussi réel, on en déduit que :

$$\lambda = \frac{q(x)}{\overline{x}^T x} \in \mathbb{R}$$

Donc toutes les valeurs propre de la matrice A sont réels.

2. Supposons que λ_1 et λ_2 sont des valeurs propres distinctes de la matrice A, et v_1, v_2 sont des vecteurs propres associé à λ_1 et λ_2 respectivement. Alors on a :

$$\begin{aligned}\lambda_1\langle v_1, v_2\rangle &= (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = (v_1^T A^T)v_2 = v_1^T(A^T v_2)\\ &= v_1^T(Av_2) = v_1^T(\lambda_2 v_2) = \lambda_2\langle v_1, v_2\rangle\end{aligned}$$

Donc $\lambda_1 = \lambda_2$ ou $\langle v_1, v_2\rangle = 0$. Comme nous avons exclut par hypothèse ce premier cas, on peut donc conclure que les vecteurs v_1, v_2 sont orthogonaux.

3. Supposons que la matrice A est diagonalisable orthogonalement. Alors il existe une matrice orthogonal P et une matrice diagonal D tel que $A = PDP^{-1}$. On a donc :

$$A^T = (PDP^{-1})^T = (P^{-1})^T D^T P^T = (P^T)^T D^T P^T = PDP^{-1} = A$$

La matrice est donc symétrique. Maintenant pour l'autre direction, supposons que A est une matrice symétrique de dimension $n \times n$, prenons λ_1 ayant comme valeur propre $\lambda_1, \lambda_2, ..., \lambda_n$ que nous savons toutes réelles. Prenons v_1 un vecteur propre unitaire associé à la valeur propre λ_1. On complète ensuite l'ensemble $\{v_1\}$ en une base orthonormale de $\mathbb{R}^n$, et on place ces vecteurs dans une matrice Q_1 qui sera alors orthogonale. On a donc :

$$(Q_1^T AQ_1)^T = Q_1^T AQ_1$$

et donc la matrice $Q_1^T AQ_1$ est aussi une matrice symétrique, et :

$$Q_1^T AQ_1 = \left(\begin{array}{c|cccc} \lambda_1 & 0 & 0 & \dots & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ \vdots & & A_2 & & \\ 0 & & & & \end{array}\right)$$

où A_2 est une matrice symétrique de dimension $(n-1)\times(n-1)$ ayant comme valeur propres $\lambda_2, \lambda_3, ..., \lambda_n$. Maintenant, considérons un vecteur propre unitaire v_2 associé à la matrice A_2. On complète ensuite l'ensemble $\{v_2\}$ en une base orthonormale de $\mathbb{R}^{n-1}$, puis on place ces vecteurs dans une matrice Q_2'. On pose ensuite :

$$Q_2 = \left(\begin{array}{c|cccc} 1 & 0 & 0 & \dots & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ \vdots & & Q_2' & & \\ 0 & & & & \end{array}\right)$$

La matrice Q_2 est donc orthogonale. La matrice $Q_2^T Q_1^T A Q_1 Q_2$ est symétrique et on a :

$$Q_2 = \left(\begin{array}{cc|cccc} \lambda_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & 0 & \dots & 0 \\ \hline 0 & 0 & & & & \\ 0 & 0 & & & & \\ \vdots & \vdots & & Q_3' & & \\ 0 & 0 & & & & \end{array}\right)$$

En continuant de la même manière, on obtient un ensemble de matrice $Q_1, Q_2, Q_3, ..., Q_n$ qui sont toutes orthogonales tel que :

$$Q_n^T Q_{n-1}^T ... Q_2^T Q_1^T A Q_1 Q_2 ... Q_{n-1} Q_n = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} = D$$

En posant $Q = Q_1 Q_2 ... Q_n$, on obtient donc :

$$A = QDQ^T$$

où Q est une matrice orthogonale, car Q est un produit de matrices orthogonales. On a donc montré que si A est une matrice symétrique, alors A peut être diagonalisé orthogonalement.

□

Théorème 5.3.2. **(Théorème des axes principaux)** Si $q(\vec{x})$ est une forme quadratique, alors il existes un changement de variable $\vec{y} = P\vec{x}$, où P est une matrice orthogonale, pour lequel la forme quadratique peut s'écrire sans termes croisés. C'est à dire que :

$$q(\vec{y}) = a_1 y_1^2 + a_2 y_2^2 + ... + a_n y_n^2$$

Démonstration. Par définition d'une forme quadratique, il existe une matrice symétrique A telle que $q(\vec{x}) = \vec{x}^T A x$. Comme la matrice A est symétrique, le théorème spectral s'applique. On peut donc trouver une matrice orthogonale P et une matrice diagonale D telle que $A = PDP^T$. Maintenant, posons $\vec{y} = P\vec{x}$, donc $\vec{x} = P^T\vec{y}$. En remplaçant dans l'équation, on obtient donc :

$$q(\vec{y}) = \vec{x}^T A \vec{x} = (P^T\vec{y})^T A (P^T\vec{y}) = \vec{y}^T PAP^T \vec{y} = \vec{y}^T D \vec{y}$$

Maintenant, si on pose :

$$D = \begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_3 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_n \end{pmatrix} \qquad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}$$

Ce qui nous permet d'obtenir $q(\vec{y}) = a_1 y_1^2 + a_2 y_2^2 + ... + a_n y_n^2$. □

Théorème 5.3.3. **(Loi d'inertie de Sylvester)** Si $q(\vec{x})$ est une forme quadratique, alors il existe une base $\{u_1, u_2, ..., u_n\}$ pour laquelle la forme quadratique peut s'écrire sous la forme :

$$q(\vec{x}) = x_1^2 + x_2^2 + ... + x_k^2 - x_{k+1}^2 - x_{k+2}^2 - ... - x_r^2$$

De plus, les valeurs de k et r sont des invariants. C'est à dire que dans toutes les bases où $q(\vec{x})$ aura cette forme, on aura les mêmes valeurs de k et r.

5.4 La diagonalisation orthogonale

Nous avons vu dans la section précédente que toute les matrices symétriques A peuvent être diagonalisé orthogonalement. La question est maintenant de savoir comment peut on accomplir ceci. La méthode est relativement simple. Il s'agit de commencer par diagonaliser la matrice A comme nous l'avons fait au chapitre 3. Ensuite, pour chacune des valeurs propres (distinctes), on applique la méthode de Gram-Schmidt aux vecteurs propres qui y sont associés. Finalement, on normalise tout les vecteurs propres, ce qui nous donne notre matrice P qui diagonalise orthogonalement la matrice A.

Exemple 5.4.1. On veut diagonaliser orthogonalement la matrice suivante :

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

On commence par trouver les valeurs propres de la matrice A :

$$\begin{aligned}
\begin{vmatrix} 4-\lambda & 2 & 2 \\ 2 & 4-\lambda & 2 \\ 2 & 2 & 4-\lambda \end{vmatrix} &= (4-\lambda)\begin{vmatrix} 4-\lambda & 2 \\ 2 & 4-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & 2 \\ 2 & 4-\lambda \end{vmatrix} + 2\begin{vmatrix} 2 & 4-\lambda \\ 2 & 2 \end{vmatrix} \\
&= (4-\lambda)[(4-\lambda)(4-\lambda)-4] - 2[2(4-\lambda)-4] + 2[4-2(4-\lambda)] \\
&= (4-\lambda)(\lambda^2-8\lambda+12) - 2(-2\lambda+4) + 2(2\lambda-4) \\
&= (-\lambda^3+12\lambda^2-44\lambda+48) + (4\lambda-8) + (4\lambda-8) \\
&= -\lambda^3+12\lambda^2-36\lambda+32 \\
&= -(\lambda-8)(\lambda-2)^2
\end{aligned}$$

On doit maintenant trouver une base de vecteurs propres associé à ces valeurs propres. Pour la valeur propre 8, on a donc :

$$\left(\begin{array}{ccc|c} -4 & 2 & 2 & 0 \\ 2 & -4 & 2 & 0 \\ 2 & 2 & -4 & 0 \end{array}\right) \sim^{L_1+2L_2\to L_2}_{L_1+2L_3\to L_3} \left(\begin{array}{ccc|c} -4 & 2 & 2 & 0 \\ 0 & -6 & 6 & 0 \\ 0 & 6 & -6 & 0 \end{array}\right) \sim^{L_2+L_3\to L_3} \left(\begin{array}{ccc|c} -4 & 2 & 2 & 0 \\ 0 & -6 & 6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Donc si $z = s$, alors $y = z = s$ et $x = \dfrac{2y+2z}{4} = s$. On peux donc prendre le vecteur propre

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Puis on normalise le vecteur pour obtenir une base orthonormale de l'espace propre associé à la valeur propre 8 :

$$w_1' = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1\sqrt{3} \end{pmatrix}$$

On fait maintenant de même pour trouver des vecteurs propres associé à la valeur propre 2 :

$$\left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 0 \end{array}\right) \sim^{L_1-L_2\to L_2}_{L_1-L_3\to L_3} \left(\begin{array}{ccc|c} 2 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Donc si $z = s$ et $y = t$, on obtient $x = -s-t$. Donc on peut prendre les deux vecteurs propres suivant :

$$v_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \qquad v_3 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

On doit maintenant trouver une base orthonormale de l'espace propre associé à la valeur propre 2. On commence donc par appliquer l'algorithme de Gram-Schmidt :

$$w_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

$$w_3 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1 \\ -1/2 \end{pmatrix} \to \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

Puis on normalise les deux vecteurs qu'on a obtenu :

$$w_2' = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}$$

$$w_3' = \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{pmatrix}$$

On peut donc maintenant écrire la diagonalisation orthogonale de la matrice A :

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{pmatrix} \begin{pmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{pmatrix}^T$$

Nous allons maintenant regarder un exemple d'application du théorème des axes principaux.

Exemple 5.4.2. On veut trouver un changement de variable préservant les angles et les longueurs, et permettant de réécrire la forme quadratique ci-dessous sans terme croisé.

$$q(\vec{x}) = 3x_1^2 + 12x_1x_2 + 8x_2^2$$

Pour ce faire, rappelons qu'un changement de variable qui préserve les angles et les longueurs, doit être donné par une matrice orthogonale. On cherche donc une matrice orthogonale telle que $\vec{x} = P\vec{y}$. Commençons par réécrire la forme quadratique sous forme matricielle. On a donc :

$$q(\vec{x}) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \underbrace{\begin{pmatrix} 3 & 6 \\ 6 & 8 \end{pmatrix}}_{A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Nous allons donc diagonaliser orthogonalement la matrice A. Cherchons les valeurs propres :

$$\begin{vmatrix} 3-\lambda & 6 \\ 6 & 8-\lambda \end{vmatrix} = (3-\lambda)(8-\lambda) - 36 = \lambda^2 - 11\lambda - 12 = (\lambda - 12)(\lambda + 1) = 0$$

Les valeurs propres sont donc 12 et −1. Cherchons un vecteur propre associé à la valeur propre $\lambda_1 = 12$:

$$\left(\begin{array}{cc|c} 3-12 & 6 & 0 \\ 6 & 8-12 & 0 \end{array}\right) \overset{L_1+2L_2 \to L_2}{\sim} \left(\begin{array}{cc|c} -9 & 6 & 0 \\ 0 & 0 & 0 \end{array}\right) \overset{\frac{1}{3}L_1 \to L_1}{\sim} \left(\begin{array}{cc|c} -3 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right) \quad \Rightarrow \quad \vec{v_1} = \frac{1}{2^2+3^2}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{13} \\ 3/\sqrt{13} \end{pmatrix}$$

Puis on cherche un vecteur propre associé à $\lambda_2 = -1$:

$$\left(\begin{array}{cc|c} 3+1 & 6 & 0 \\ 6 & 8+1 & 0 \end{array}\right) \overset{3L_1-2L_2 \to L_2}{\sim} \left(\begin{array}{cc|c} 4 & 6 & 0 \\ 0 & 0 & 0 \end{array}\right) \overset{\frac{1}{2}L_1 \to L_1}{\sim} \left(\begin{array}{cc|c} 2 & 3 & 0 \\ 0 & 0 & 0 \end{array}\right) \quad \Rightarrow \quad \vec{v_2} = \frac{1}{\sqrt{3^2+2^2}}\begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -3/\sqrt{13} \\ 2/\sqrt{13} \end{pmatrix}$$

Donc la matrice P est donnée par $P = \begin{pmatrix} 2/\sqrt{13} & -3/\sqrt{13} \\ 3/\sqrt{13} & 2/\sqrt{13} \end{pmatrix}$. En posant $\vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, on obtient donc le changement de variable suivant :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{13} & -3/\sqrt{13} \\ 3/\sqrt{13} & 2/\sqrt{13} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{13}} y_1 - \frac{3}{\sqrt{13}} y_2 \\ \frac{3}{\sqrt{13}} y_1 + \frac{2}{\sqrt{13}} y_2 \end{pmatrix}$$

Remplaçons maintenant dans la forme quadratique :

$$\begin{aligned}
q(\vec{y}) &= 3x_1^2 + 12x_1x_2 + 8x_2^2 \\
&= 3\left(\frac{2}{\sqrt{13}} y_1 - \frac{3}{\sqrt{13}} y_2\right)^2 + 12\left(\frac{2}{\sqrt{13}} y_1 - \frac{3}{\sqrt{13}} y_2\right)\left(\frac{3}{\sqrt{13}} y_1 + \frac{2}{\sqrt{13}} y_2\right) + 8\left(\frac{3}{\sqrt{13}} y_1 + \frac{2}{\sqrt{13}} y_2\right)^2 \\
&= \frac{3}{13}(2y_1 - 3y_2)^2 + \frac{12}{13}(2y_1 - 3y_2)(3y_1 + 2y_2) + \frac{8}{13}(3y_1 + 2y_2)^2 \\
&= \frac{3}{13}(4y_1^2 - 12y_1y_2 + 9y_2^2) + \frac{12}{13}(6y_1^2 - 5y_1y_2 - 6y_2^2) + \frac{8}{13}(9y_1^2 + 12y_1y_2 + 4y_2^2) \\
&= \left(\frac{12}{13} + \frac{72}{13} + \frac{72}{13}\right) y_1^2 + \left(\frac{-36}{13} - \frac{60}{13} + \frac{96}{13}\right) y_1y_2 + \left(\frac{27}{13} - \frac{72}{13} + \frac{32}{13}\right) y_2^2 \\
&= 12y_1^2 + y_2^2
\end{aligned}$$

Remarquez qu'il n'était pas nécessaire de faire ce dernier calcul. La solution étant directement relié aux valeurs propres selon la théorie. L'intérêt d'une telle transformation deviendra évident dans notre étude des coniques et des quadriques un peu plus tard dans ce chapitre.

5.5 Les formes définies positives

Nous avons déjà vu que tout les produits scalaire sont des formes bilinéaires associés à une matrice symétrique. La question est maintenant de savoir si le contraire est aussi vrai ? En fait le contraire est complètement faux. Il existe des formes bilinéaires qui ne définissent pas des produits scalaires. Nous allons donc essayer de caractériser les matrices symétriques pour lesquels leur forme bilinéaire qui leur est associé est un produit scalaire.

Definition 5.5.1. Si A est une matrice symétrique, alors la forme quadratique $q(x) = x^T Ax$ est dites définie positive si $q(x) > 0$ pour tout $x \in \mathbb{R}^n \setminus \{0\}$. Dans ce cas, on dit que la matrice A est définie positive.

Théorème 5.5.1. Si A est une matrice symétrique définie positive, alors les éléments sur la diagonale de A sont tous positif.

Démonstration. Définissons la forme quadratique $q(x) = x^T Ax$. Pour démontrer le théorème, il s'agit de remarquer que

$$A_{ii} = e_i^T Ae_i = q(e_i) > 0$$

Donc tous les éléments sur la diagonale de A doivent être positif. $\square$

Remarquez que le théorème précédent n'affirme absolument pas que tout les éléments de A sont positif. Il affirme seulement que ceux sur la diagonale le sont.

Théorème 5.5.2. Si A est une matrice symétrique, alors les énoncés suivants sont équivalents :

1. A est une matrice définie positive
2. Toutes les valeurs propres de la matrice A sont positive (i.e. > 0).
3. La forme bilinéaire $\langle x, y\rangle = x^T Ay$ est un produit scalaire
4. A peut s'écrire sous la forme LL^T où L est une matrice triangulaire inférieure inversible. (Factorisation de Choleski)

Démonstration.
(1) ⇒ (2) : Supposons que A est une matrice définie positive, $q(x) = x^T Ax, \ \forall x \in \mathbb{R}^n$, et prenons v un vecteur propre associé à la valeur propre λ. Alors on a :

$$0 < q(v) = v^T Av = v^T \lambda v = \lambda v^T v = \lambda ||v||^2 \quad \Rightarrow \quad \lambda = \frac{q(v)}{||v||^2} > 0$$

Donc toutes les valeurs propres de la matrice A sont positive.
(2) ⇒ (1) : Supposons que toutes les valeurs propres de A sont positive. Comme la matrice A est symétrique, par le théorème spectral il existe une base orthonormal $\{v_1, v_2, ..., v_n\}$ de vecteurs propres associés aux valeurs propres $\lambda_1, \lambda_2, ..., \lambda_n$. Posons

$$q(x) = x^T Ax, \quad \forall x \in \mathbb{R}^n$$

Prenons $x \in \mathbb{R}^n$, alors il existe des scalaires $\alpha_1, \alpha_2, ..., \alpha_n$ tel que :

$$x = \alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n$$

ce qui nous donne :

$$\begin{array}{rcl} q(x) & = & (\alpha_1 v_1^T + \alpha_2 v_2^T + ... + \alpha_n v_n^T)A(\alpha_1 v_1 + \alpha_2 v_2 + ... + \alpha_n v_n) \\ & = & (\alpha_1 v_1^T + \alpha_2 v_2^T + ... + \alpha_n v_n^T)(\alpha_1 Av_1 + \alpha_2 Av_2 + ... + \alpha_n Av_n) \\ & = & (\alpha_1 v_1^T + \alpha_2 v_2^T + ... + \alpha_n v_n^T)(\alpha_1 \lambda_1 v_1 + \alpha_2 \lambda_2 v_2 + ... + \alpha_n \lambda_n v_n) \\ & = & \alpha_1^2 \lambda_1 ||v_1||^2 + \alpha_2^2 \lambda_2 ||v_2||^2 + ... + \alpha_n^2 \lambda_n ||v_n||^2 \\ & > & 0 \end{array}$$

(1) ⇒ (3) : Supposons que A est une matrice définie positive, et définissons

$$\langle x, y\rangle = x^T Ay, \quad \forall x, y \in \mathbb{R}^n$$

On va vérifier les 5 propriétés d'un produit scalaire.

1. Si $x, y, z \in \mathbb{R}^n$, alors on a :
$$\langle x, y + z\rangle = x^T A(y + z) = x^T Ay + x^T Az = \langle x, y\rangle + \langle x, z\rangle$$
2. Si $x, y \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$ alors on a :
$$\langle \lambda x, y\rangle = (\lambda x)^T Ay = \lambda x^T Ay = \lambda\langle x, y\rangle$$
3. Si $x, y \in \mathbb{R}^n$, alors :
$$\langle x, y\rangle = x^T Ay = (y^T A^T x)^T = y^T Ax = \langle y, x\rangle$$
4. Si $x \in \mathbb{R}^n$, alors on a :
$$\langle x, x\rangle = x^T Ax > 0 \text{ car la matrice est définie positive}$$
5. Il est facile de vérifier que $\langle 0, 0\rangle = 0$. Si on suppose que $x \in \mathbb{R}^n$ tel que $\langle x, x\rangle = 0$, alors on a $x^T Ax = 0$, ce qui est possible que si $x = 0$ par définie d'une matrice définie positive.

(3) ⇒ (1) : Supposons que la forme bilinéaire $\langle x,y\rangle = x^T Ay$ est un produit scalaire, alors pour tout $x \in \mathbb{R}^n \backslash \{0\}$ on a :

$$q(x) = x^T Ax = \langle x,x\rangle > 0$$

par définition d'un produit scalaire. Donc la matrice A est définie positive.
(2) ⇒ (4) : La démonstration de cette partie est plus technique et ne sera pas faites en détail ici. L'idée est cependant la suivante. On commence par démontrer qu'une matrice définie positive possède toujours une factorisation LU, c'est à dire que $A = LU$, où L est une matrice triangulaire inférieure, et U une matrice triangulaire supérieure. Ceci est fait en remarquant que l'algorithme de Gauss appliquer à la matrice A n'a jamais de 0 comme pivot. Comme toutes les valeurs propres de A sont positive, la matrice A est inversible, ce qui implique que L et U le sont aussi. On peut donc écrire $U = L^{-1}A$. Maintenant, si on pose $D = L^{-1}A(L^T)^{-1} = U(L^T)^{-1}$, on peut montrer que la matrice D doit être triangulaire supérieure (Produit de deux matrice triangulaire supérieure) et symétrique (par les propriétés de la transposé). On peut donc conclure que D est une matrice diagonale, et une simple substitution nous montre que $A = LDL^T$. Maintenant, si on démontre que la matrice D est définie positive, on obtiendra que la matrice D, en plus d'être diagonale, possède uniquement des entrés positives sur la diagonale, ce qui nous permet de calculer facilement une racines carrés (En prenant la racine carré de chaque élément). Finalement, en posant $L_2 = L\sqrt{D}$, on obtient :

$$A = L_2 L_2^T$$

où la matrice L_2 est triangulaire inférieure et ne possède aucun zéro sur la diagonale. Elle est donc inversible.
(4) ⇒ (1) : Supposons que $A = LL^T$ avec L une matrice triangulaire inférieure et inversible, on définissons la forme quadratique $q(x) = x^T Ax$. On obtient donc :

$$q(x) = x^T Ax = x^T LL^T x = (L^T x)^T (L^T x) = \langle L^T x, L^T x\rangle > 0, \quad \text{si } x \neq 0$$

La matrice A est donc définie positive. □

Dans le théorème précédent, nous avons vu qu'une matrice définie positive A a toujours une factorisation de Choleski. C'est à dire qu'on peut écrire $A = LL^T$ où L est une matrice triangulaire inférieure. Les exemples ci dessous illustre comment trouver une telle factorisation :

Exemple 5.5.1. Trouvez la factorisation de Choleski de la matrice suivante :

$$A = \begin{pmatrix} 4 & 10 \\ 10 & 26 \end{pmatrix}$$

Posons une matrice

$$L = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix}$$

on a donc :

$$\begin{pmatrix} 4 & 10 \\ 10 & 26 \end{pmatrix} = \begin{pmatrix} a & 0 \\ b & c \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} a^2 & ab \\ ab & b^2 + c^2 \end{pmatrix}$$

Ce qui nous donne :

$$\begin{cases} a^2 = 4 \Rightarrow a = 2 \\ ab = 10 \Rightarrow b = 5 \\ b^2 + c^2 = 26 \Rightarrow 25 + c^2 = 26 \Rightarrow c^2 = 1 \Rightarrow c = 1 \end{cases}$$

On obtient donc la matrice L suivante :

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

Exemple 5.5.2. Trouvez la factorisation de Choleski de la matrice suivante :

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 13 & 23 \\ 4 & 23 & 77 \end{pmatrix}$$

Posons une matrice

$$L = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}$$

on a donc :

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 13 & 23 \\ 4 & 23 & 77 \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix} \begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix} = \begin{pmatrix} a^2 & ab & ad \\ ab & b^2 + c^2 & bd + ce \\ ad & bd + ce & d^2 + e^2 + f^2 \end{pmatrix}$$

Ce qui nous donne :

$$\begin{cases} a^2 = 1 \Rightarrow a = 1 \\ ab = 2 \Rightarrow b = 2 \\ b^2 + c^2 = 13 \Rightarrow 4 + c^2 = 13 \Rightarrow c^2 = 9 \Rightarrow c = 3 \\ ad = 4 \Rightarrow d = 4 \\ bd + ce = 23 \Rightarrow 8 + 3e = 23 \Rightarrow 3e = 15 \Rightarrow e = 5 \\ d^2 + e^2 + f^2 = 77 \Rightarrow 16 + 25 + f^2 = 77 \Rightarrow f^2 = 36 \Rightarrow f = 6 \end{cases}$$

On obtient donc la matrice L suivante :

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

Exemple 5.5.3. On veut démontrer que pour tout $x, y \in \mathbb{R} \backslash \{0\}$ on a $5x^2 + 4xy + 5y^2 > 0$. Pour ce faire, on commence par réécrire cette équation sous forme matricielle :

$$q = 5x^2 + 4xy + 5y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Il s'agit donc de trouver les valeurs propres de la matrice symétrique

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

On a donc :

$$\begin{vmatrix} 5 - \lambda & 2 \\ 2 & 5 - \lambda \end{vmatrix} = \lambda^2 - 10\lambda + 21 = (\lambda - 7)(\lambda - 3)$$

Les valeurs propres sont donc $\lambda = 3$ et $\lambda = 7$. Comme ces deux valeurs propres sont positives, on peut donc affirmer que $5x^2 + 4xy + 5y^2 > 0$ pour tout $(x, y) \neq (0, 0)$.

5.6 Application : Maximum et minimum sous contrainte

Nous allons maintenant regarder dans cette section une autre application des valeurs propres d'une matrice. Cette fois dans le contexte des maximum et minimum d'une forme quadratique selon certaine contrainte. Ce type de problème peut bien sur être résolut à l'aide des techniques du calcul différentiel et intégral, mais pour ce type de problème précis, l'algèbre linéaire offre une solution beaucoup plus simple.

Théorème 5.6.1. Si A est une matrice symétrique, et $q(x) = x^T Ax$ la forme quadratique qui lui est associé. Alors :

1. Le maximum de la forme quadratique $q(x)$ sous la contrainte $||x|| = 1$ est la plus grande valeur propre de la matrice A.
2. Le minimum de la forme quadratique $q(x)$ sous la contrainte $||x|| = 1$ est la plus petite valeur propre de la matrice A.

Démonstration. 1. Supposons que $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ sont les valeurs propres de la matrice A. Alors il existe une matrice P orthogonale tel que

$$A = PDP^T = P \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & ... & \lambda n \end{pmatrix} P^T$$

Maintenant, prenons $x \in \mathbb{R}^n$ tel que $||x|| = 1$, et posons $y = P^T x$. alors

$$||y|| = \sqrt{\langle y, y\rangle} = \sqrt{\langle P^T x, P^T x\rangle} = \sqrt{x^T P P^T x} = \sqrt{x^T x} = \sqrt{\langle x, x\rangle} = ||x|| = 1$$

On obtient donc :

$$\begin{aligned} q(x) &= x^T A x = x^T P D P^T x = (P^T x)^T D (P^T x) = y^T D y \\ &= \begin{pmatrix} y_1 & y_2 & ... & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & ... & \lambda n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + ... + \lambda_n y_n^2 \\ &\leq \lambda_1 y_1^2 + \lambda_1 y_2^2 + ... + \lambda_1 y_n^2 \\ &= \lambda_1 (y_1^2 + y_2^2 + ... + y_n^2) \\ &= \lambda_1 ||y||^2 \\ &= \lambda_1 \end{aligned}$$

Donc le maximum de $q(x)$ sous la contrainte $||x|| = 1$ est λ_1, la plus grande valeur propre de la matrice A.

2. La démonstration est pratiquement identique à la précédente.

□

Exemple 5.6.1. Trouver le maximum et le minimum de la fonction $q(x) = x^2 + 4xy + y^2$ sous la contrainte $x^2 + y^2 = 1$. Pour ce faire, on commence par écrire la fonction $q(x)$ sous forme matricielle :

$$q(x) = x^2 + 4xy + y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

La fonction $q(x)$ est donc une forme quadratique. Maintenant, on remarque que la contrainte peut s'écrire comme étant la norme d'un vecteur :

$$\left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\| = \sqrt{x^2 + y^2} = 1$$

Donc le maximum et minimum et la fonction sous cette contrainte sont donc donné par les valeurs propres de la matrice

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

On a donc :

$$\begin{vmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 - 4 = \lambda^2 - 2\lambda - 3 = (\lambda + 1)(\lambda - 3) = 0$$

Le maximum est donc 3 et le minimum est -1.

5.7 Application : Étude des coniques et quadriques

Dans toute sa généralité, une conique est une équation de la forme

$$ax^2 + bxy + cy^2 + dx + ey + f = 0, \qquad \text{où } a, b, c, d, e, f \in \mathbb{R}$$

Une quadrique est une équation similaire avec trois variables. Le graphique d'une conique peut prendre l'une des forme suivantes dans le plan cartésien :

1. Un cercle
2. Une parabole
3. Une ellipse
4. Une hyperbole
5. Deux droites sécantes
6. Deux droites parallèles
7. Une droite
8. Un point
9. Aucun point

Ce sont les 4 premiers cas que nous appellerons des coniques non-dégénérés, et ce sont celles qui nous intéressent plus particulièrement dans cette section. La question revient donc à savoir comment étudier une telle équation. Remarquez premièrement qu'il est possible de réécrire cette équation sous forme matricielle :

$$\vec{x}^T \underbrace{\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}}_{A} \vec{x} + \underbrace{\begin{pmatrix} d & e \end{pmatrix}}_{B} \vec{x} + f = 0, \qquad \vec{x} = \begin{pmatrix} x \\ y \end{pmatrix}$$

Dans un premier temps, nous allons faire un changement de base qui va nous permettre de diagonaliser orthogonalement la matrice A. Ceci va nous permettre d'identifier les rotations qui sont présentent dans notre conique. Comme A est symétrique, il existe donc une matrice orthogonale P telle que $A = PDP^T$. En posant $\vec{z} = P\vec{x} = \begin{pmatrix} z \\ w \end{pmatrix}$, on obtient donc :

$$\begin{aligned} \vec{x}^T A\vec{x} + \begin{pmatrix} d & e \end{pmatrix} \vec{x} + f = 0 \quad &\Rightarrow \quad (P^T\vec{z})^T A(P^T\vec{z}) + B(P^T\vec{z}) + f = 0 \\ &\Rightarrow \quad \vec{z}^T PAP^T\vec{z} + BP^T\vec{z} + f = 0 \\ &\Rightarrow \quad \vec{z}^T D\vec{z} + BP^T\vec{z} + f = 0 \end{aligned}$$

Notez que les colonnes de la matrice P représente la direction de nos nouveaux axes. En developpant l'équation que nous venons d'obtenir, l'équation de notre conique peut donc être écrite sous la forme :

$$a'z^2 + c'w^2 + dz + ew + f = 0$$

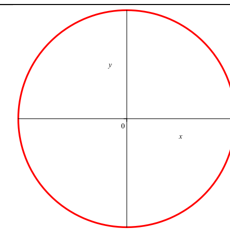
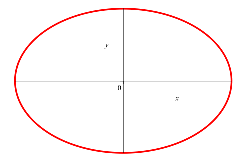
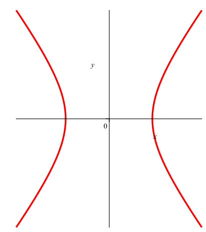
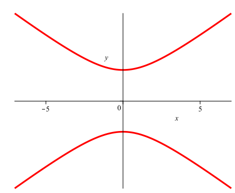
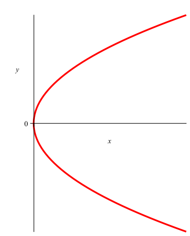
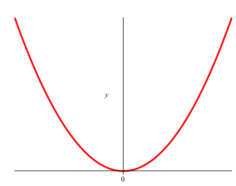
où z et w représente nos nouveaux axes, c'est à dire des axes dans la direction de nos valeurs propres, ou de manière équivalente dans la direction des colonnes de la matrice P. Maintenant, il nous faut éliminer les termes en z et en w. Ceci peut être facilement accomplie en complétant les carrés, ce qui va nous permettre d'identifier les translations qui sont présente. On obtient donc :

$$\begin{aligned} a'z^2 + c'w^2 + dz + ew + f = 0 \quad &\Rightarrow \quad a'\left(z^2 + \frac{d'}{a'}z\right) + c'\left(w^2 + \frac{e'}{c'}w\right) + f = 0 \\ &\Rightarrow \quad a'\left(z + \frac{d'}{2a'}\right)^2 + c'\left(w + \frac{e'}{2c'}\right)^2 + \left(f - \left(\frac{d'}{2a'}\right)^2 - \left(\frac{e'}{2c'}\right)^2\right) = 0 \end{aligned}$$

En posant $h = \frac{-d'}{2a'}, k = \frac{-e'}{2c'}$ et $g = f - \left(\frac{d'}{2a'}\right)^2 - \left(\frac{e'}{2c'}\right)^2$, on obtient donc l'équation suivante pour notre conique :

$$a'(z-h)^2 + c'(w-k)^2 + g = 0$$

dans ce cas, le paramètre h représente une translation dans la direction de l'axe z, et le k représente une translation dans la direction de l'axe w. Rendu à cette étape, il est relativement facile d'identifier notre conique en réécrivant notre équation sous l'une des formes standard d'une conique illustré dans le tableau suivant.

cercle L'équation générale d'un cercle a la forme $$(x-h)^2+(y-k)^2=r^2$$ où r représente le rayon du cercle et (h,k) représente le centre du cercle.	
Ellipse L'équation générale d'une ellipse a la forme $$\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1$$ où (h,k) représente le centre de l'ellipse, et les paramètres a et b représente la longueur des demi-axes dans la direction de l'axe de x et y respectivement.	
Hyperbole horizontale L'équation générale d'une hyperbole horizontale a la forme $$\frac{(x-h)^2}{a^2}-\frac{(y-k)^2}{b^2}=1$$ où (h,k) représente le centre de l'hyperbole, et les paramètres a et b représente la longueur des demi-axes dans la direction de l'axe de x et y respectivement.	
Hyperbole verticale L'équation générale d'une ellipse a la forme $$\frac{(x-h)^2}{a^2}-\frac{(y-k)^2}{b^2}=-1$$ où (h,k) représente le centre de l'hyperbole, et les paramètres a et b représente la longueur des demi-axes dans la direction de l'axe de x et y respectivement.	
Parabole horizontale L'équation générale d'une parabole horizontale a la forme $$x=a(y-k)^2+h$$ où (h,k) représente le sommet de la parabole.	
Parabole verticale L'équation générale d'une parabole $$y=a(x-h)^2+k$$ où (h,k) représente le sommet de la parabole.	

Exemple 5.7.1. On veut trouver quel est la figure géométrique représenté par la relation

$$5x^2-4xy+5y^2=48$$

Pour ce faire, on commence par réécrire l'équation sous forme matricielle :

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 48$$

Puis on diagonalise orthogonalement la matrice

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

On commence par trouver les valeurs propres :

$$\begin{vmatrix} 5-\lambda & -2 \\ -2 & 5-\lambda \end{vmatrix} = (5-\lambda)^2 - 4 = \lambda^2 - 10\lambda + 21 = (\lambda-7)(\lambda-3) = 0$$

Les valeurs propres sont donc 3 et 7. On trouve ensuite un vecteur propre associé à la valeur propre 3 :

$$\left(\begin{array}{cc|c} 2 & -2 & 0 \\ -2 & 2 & 0 \end{array}\right) \sim^{L_1+L_2\to L_2} \left(\begin{array}{cc|c} 2 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

On peut donc prendre le vecteur propre

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Pour la valeur propre 7 on a ensuite :

$$\left(\begin{array}{cc|c} -2 & -2 & 0 \\ -2 & -2 & 0 \end{array}\right) \sim^{L_1-L_2\to L_2} \left(\begin{array}{cc|c} -2 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

On peut donc prendre le vecteur propre

$$v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

On normalise ensuite les deux vecteurs :

$$v_1' = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \qquad v_2' = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

Ce qui nous donne l'équation suivante :

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1\sqrt{2} \\ 1\sqrt{2} & -1\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix} \begin{pmatrix} 1\sqrt{2} & 1\sqrt{2} \\ 1\sqrt{2} & -1\sqrt{2} \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} = 48$$

$$\begin{pmatrix} (x+y)/\sqrt{2} & (x-y)/\sqrt{2} \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 7 \end{pmatrix} \begin{pmatrix} (x+y)/\sqrt{2} \\ (x-y)/\sqrt{2} \end{pmatrix} = 48$$

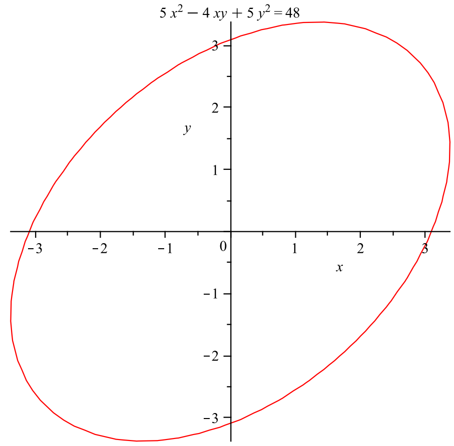
On va donc poser

$$u = \frac{x+y}{\sqrt{2}} \qquad \text{et} \qquad v = \frac{x-y}{\sqrt{2}}$$

Ce qui nous donne l'équation :

$$3u^2 + 7v^2 = 48 \Rightarrow \frac{u^2}{4^2} + \frac{v^2}{(\sqrt{48/7})^2} = 1$$

Il s'agit donc d'un éllipse ayant un grand demi-axe de longueur 4 dans la direction du vecteur $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et un petit demi-axe de longueur $\sqrt{\frac{48}{7}}$ dans la direction du vecteur $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Le graphique ci dessous illustre l'ellipse.



Exemple 5.7.2. On veut trouver quelle conique est représenté par l'équation ci-dessous, et en trouver une représentation graphique.

$$33x^2 - 20xy + 54y^2 + 18x - 12y - 100 = 0$$

On commence par écrire l'équation sous forme matricielle :

$$\begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 33 & -10 \\ -10 & 54 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\vec{x}} + \underbrace{\begin{pmatrix} 18 & -12 \end{pmatrix}}_{B} \begin{pmatrix} x \\ y \end{pmatrix} - \underbrace{100}_{f} = 0$$

C'est à dire qu'on peut écrire l'équation de la conique sous la forme $\vec{x}^T A\vec{x} + B\vec{x} + f = 0$. On commence par diagonaliser orthogonalement la matrice A. Pour ce faire, on cherche premièrement les valeurs propres :

$$\begin{vmatrix} 33-\lambda & -10 \\ -10 & 54-\lambda \end{vmatrix} = (33-\lambda)(54-\lambda) - 100 = \lambda^2 - 87\lambda + 1682$$

Ce qui nous donne :

$$\lambda = \frac{87 \pm \sqrt{87^2 - 4(1682)}}{2} = \frac{87 \pm 29}{2} = 58 \text{ et } 29$$

Les valeurs propres sont donc 58 et 29. Cherchons maintenant un vecteur propre associé à 29. On a donc :

$$\left(\begin{array}{cc|c} 4 & -10 & 0 \\ -10 & 25 & 0 \end{array}\right) \sim^{5L_1+2L_2 \to L_2} \left(\begin{array}{cc|c} 4 & -10 & 0 \\ 0 & 0 & 0 \end{array}\right) \sim^{\frac{1}{2}L_1 \to L_1} \left(\begin{array}{cc|c} 2 & -5 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

On a donc le vecteur propre $v_1 = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$. Comme la matrice est symétrique, le vecteur propre associé à la valeur propre 58 doit être orthogonal, ce qui nous donne le vecteur propre $v_2 = \begin{pmatrix} -2 \\ 5 \end{pmatrix}$. En normalisant ces vecteurs, on obtient la factorisation suivante de la matrice A :

$$\begin{pmatrix} 33 & -10 \\ -10 & 54 \end{pmatrix} = \underbrace{\begin{pmatrix} 5/\sqrt{29} & -2/\sqrt{29} \\ 2/\sqrt{29} & 5/\sqrt{29} \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} 29 & 0 \\ 0 & 58 \end{pmatrix}}_{D} \underbrace{\begin{pmatrix} 5/\sqrt{29} & -2/\sqrt{29} \\ 2/\sqrt{29} & 5/\sqrt{29} \end{pmatrix}^T}_{P^T}$$

On peut donc réécrire l'équation de la conique sous la forme : $\vec{x}^T PDP^T\vec{x} + B\vec{x} + f = 0$. Maintenant, si on fait le changement de variable $\vec{z} = P^T\vec{x}$, l'équation devient : $\vec{z}^T D\vec{z} + BP\vec{z} + f = 0$. Maintenant, si on suppose que $\vec{z} = \begin{pmatrix} z \\ w \end{pmatrix}$, on obtient donc :

$$\begin{pmatrix} z & w \end{pmatrix} \begin{pmatrix} 29 & 0 \\ 0 & 58 \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} + \begin{pmatrix} 18 & -12 \end{pmatrix} \begin{pmatrix} 5/\sqrt{29} & -2/\sqrt{29} \\ 2/\sqrt{29} & 5/\sqrt{29} \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} - 100 = 0$$

$$29z^2 + 58w^2 + \frac{66}{\sqrt{29}}z - \frac{-96}{\sqrt{29}}w - 100 = 0$$

Puis, on complète les carrés pour nous permettre de trouver le centre de la conique.

$$\begin{aligned}
29\left(z + \frac{33}{29\sqrt{29}}\right)^2 + 58\left(w - \frac{24}{29\sqrt{29}}\right)^2 &= 100 - 29\left(\frac{33}{29\sqrt{29}}\right)^2 - 58\left(\frac{24}{29\sqrt{29}}\right)^2 \\
&= 100 - \frac{1089}{841} - \frac{1152}{841} \\
&= \frac{81859}{841}
\end{aligned}$$

Maintenat, pour réécrire la conique sous forme standard, et ainsi pouvoir l'identifier, il nous faut diviser par la constante de droite, ce qui nous donne :

$$\frac{24389}{81859}\left(z + \frac{33}{29\sqrt{29}}\right)^2 + \frac{48778}{81859}\left(w - \frac{24}{29\sqrt{29}}\right)^2 = 1$$

$$\frac{\left(z + \frac{33}{29\sqrt{29}}\right)^2}{\left(\sqrt{81859/24389}\right)^2} + \frac{\left(w - \frac{24}{29\sqrt{29}}\right)^2}{\left(\sqrt{81859/48778}\right)^2} = 1$$

On remarque donc qu'il s'agit de l'équation d'une éllipse. La longueur des demi-axes est donné par :

$$\text{Dans la direction de l'axe des } z: \quad \sqrt{\frac{81859}{24389}} \approx 1,83205$$

$$\text{Dans la direction de l'axe des } w: \quad \sqrt{\frac{81859}{48778}} \approx 1,29545$$

Notez que la longueur des demi-axes est rester identique car notre changement de variable a été fait à l'aide d'une matrice orthogonale (une matrice orthogonale préserve les angles et les longueurs). Maintenant, pour les coordonnés du centre de la conique, d'après l'équation nous avons :

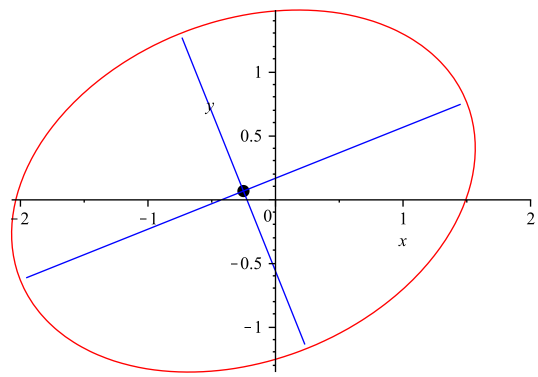
$$\text{Centre} = \begin{pmatrix} \frac{-33}{29\sqrt{29}} \\ \frac{24}{29\sqrt{29}} \end{pmatrix} \quad \text{Dans les coordonées } z, w$$

Il nous faut maintenant trouver les coordonnées du centre dans la base originale. Pour ce faire, il suffit de réutiliser notre matrice P. On a donc :

$$\begin{pmatrix} \frac{5}{\sqrt{29}} & \frac{-2}{\sqrt{29}} \\ \frac{2}{\sqrt{29}} & \frac{5}{\sqrt{29}} \end{pmatrix} \begin{pmatrix} \frac{-33}{29\sqrt{29}} \\ \frac{24}{29\sqrt{29}} \end{pmatrix} = \begin{pmatrix} \frac{-213}{841} \\ \frac{54}{841} \end{pmatrix} \quad \text{Dans les coordonées } x, y$$

Maintenant que nous savons qu'il s'agit d'une élipse, que nous connaissons la direction des axes, que nous connaissons le centre ainsi que la longueur des demi-axes, il nous est facile de la dessiner. On a donc :

$$33x^2 - 20xy + 54y^2 + 18x - 12y - 100 = 0$$



Exemple 5.7.3. On veut trouver quelle conique est représenté par l'équation ci-dessous, et en trouver une représentation graphique.

$$-5x^2 + 20xy + 10y^2 + 20x - 50y - 10 = 0$$

On commence par écrire l'équation sous forme matricielle :

$$\begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} -5 & 10 \\ 10 & 10 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\vec{x}} + \underbrace{\begin{pmatrix} 20 & -50 \end{pmatrix}}_{B} \begin{pmatrix} x \\ y \end{pmatrix} = 10$$

Ou de manière équivalence on peut écrire $\vec{x}^T A\vec{x} + B\vec{x} = 10$. On commence par diagonaliser orthogonalement la matrice A, ce qui est possible car la matrice A est symétrique. On obtient donc :

$$\begin{pmatrix} -5 & 10 \\ 10 & 10 \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} 15 & 0 \\ 0 & -10 \end{pmatrix}}_{D} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^T$$

Ce qui nous permet d'écrire $\vec{x}^T PDP^T\vec{x} + B\vec{x} = 10$. Nous allons donc faire le changement de variable $\vec{z} = P^T\vec{x}$. L'équation se réduit donc à $\vec{z}^T D\vec{z} + BP\vec{z} = 10$. Maintenant si on suppose que $\vec{z} = \begin{pmatrix} z \\ w \end{pmatrix}$, l'équation devient donc :

$$\begin{pmatrix} z & w \end{pmatrix} \begin{pmatrix} 15 & 0 \\ 0 & -10 \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} + \begin{pmatrix} 20 & -50 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} = 10$$

$$15z^2 - 10w^2 + \frac{-80}{\sqrt{5}} z + \frac{-90}{\sqrt{5}} w = 10$$

Il nous faut maintenant compléter les carrés :

$$15\left(z - \frac{8}{3\sqrt{5}}\right)^2 - 10\left(w + \frac{9}{2\sqrt{5}}\right)^2 = 10 + 15\left(\frac{8}{3\sqrt{5}}\right)^2 - 10\left(\frac{9}{2\sqrt{5}}\right)^2$$

$$15\left(z - \frac{8}{3\sqrt{5}}\right)^2 - 10\left(w + \frac{9}{2\sqrt{5}}\right)^2 = \frac{-55}{6}$$

$$\frac{-90}{55}\left(z - \frac{8}{3\sqrt{5}}\right)^2 + \frac{60}{55}\left(w + \frac{9}{2\sqrt{5}}\right)^2 = 1$$

$$\frac{-18}{11}\left(z - \frac{8}{3\sqrt{5}}\right)^2 + \frac{12}{11}\left(w + \frac{9}{2\sqrt{5}}\right)^2 = 1$$

$$-\frac{\left(z-\frac{8}{3\sqrt{5}}\right)^2}{\left(\sqrt{\frac{11}{18}}\right)^2}+\frac{\left(w+\frac{9}{2\sqrt{5}}\right)^2}{\left(\sqrt{\frac{11}{12}}\right)^2}=1$$

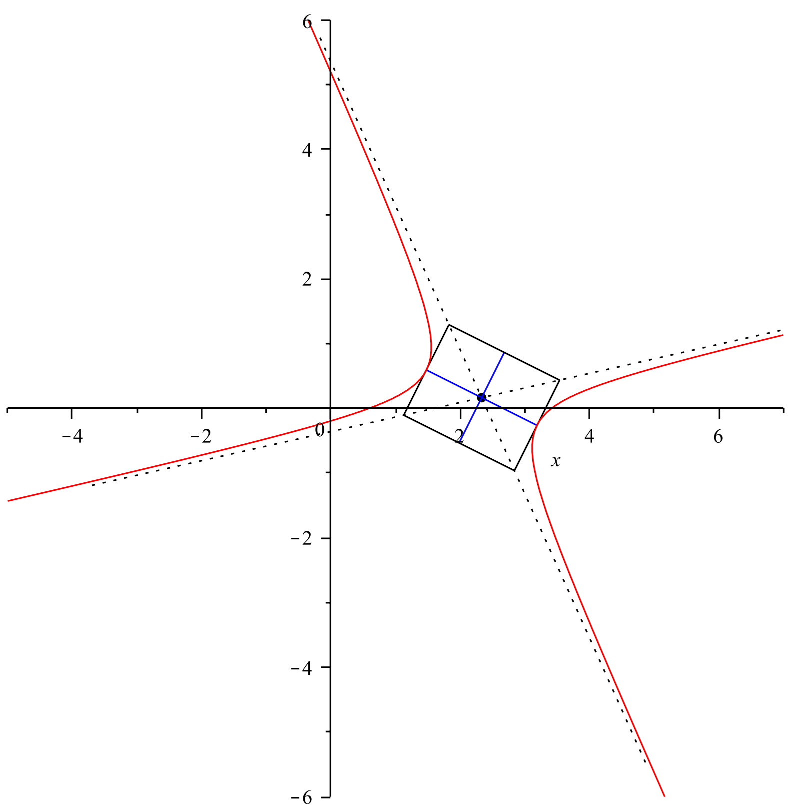
On obtient donc qu'il s'agit d'une hyperbole, et que la longueur des demi-axes est respectivement $\sqrt{\frac{11}{18}}$ et $\sqrt{\frac{11}{12}}$. De plus, nous savons que les coordonnées du centre selon les axes z et w est donné par :

$$\begin{pmatrix} \frac{8}{3\sqrt{5}} \\ \frac{-9}{2\sqrt{5}} \end{pmatrix}$$

Il nous faut maintenant réécrire le centre dans la base habituelle. Pour ce faire, on doit utiliser la matrice de passage P. On a donc :

$$\text{Centre} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{8}{3\sqrt{5}} \\ \frac{-9}{2\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{7}{3} \\ \frac{1}{6} \end{pmatrix}$$

Maintenant, nous avons toute les informations nécessaire pour dessiner la conique. Comme il s'agit d'une hyperbole, on commence par place le centre sur le graphique, puis on dessiner les demi-axes (en bleu sur le graphique). Les demi-axes sont obtenu en dessiner des segments commençant au centre, dans la direction de nos axes z et w (Les colonnes de P), et de la longueur que nous avons calculer. Maintenant que nous avons la croix bleu qui représente nos demi-axes, on dessine le rectangle en noir. Ceci nous permet d'obtenir les asymptotes de l'hyperbole en reliant les coins opposés. Une fois que ceci est fait, il ne nous reste plus qu'à dessiner l'hyperbole. Notez que l'hyperbole est ouverte dans la direction de l'axe des z, car c'est selon cette axe que nous avons obtenu le signe négatif. Voici le graphique représentant notre conique.



Exemple 5.7.4. On veut trouver quelle conique est représenté par l'équation ci-dessous, et en trouver une représentation graphique.

$$-2x^2 - 12xy - 18y^2 + 15x + 20y - 50 = 0$$

On commence par écrire l'équation sous forme matricielle :

$$\begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} -2 & -6 \\ -6 & -18 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\vec{x}} + \underbrace{\begin{pmatrix} 15 & 20 \end{pmatrix}}_{B} \begin{pmatrix} x \\ y \end{pmatrix} = 50$$

On a donc : $\vec{x}^T A\vec{x} + B\vec{x} = 50$. On commence donc par diagonaliser orthogonalement la matrice A :

$$\begin{pmatrix} -2 & -6 \\ -6 & -18 \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & -20 \end{pmatrix}}_{D} \begin{pmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix}^T$$

L'équation de notre conique peut donc être réécrite sous la forme : $\vec{x}^T PDP^T\vec{x} + B\vec{x} = 50$. En faisant le changement de variable $\vec{z} = P^T\vec{x}$, on obtient donc : $\vec{z}^T D\vec{z} + BP\vec{z} = 50$. Donc si on pose $\vec{z} = \begin{pmatrix} z \\ w \end{pmatrix}$, on obtient donc :

$$\begin{pmatrix} z & w \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -20 \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} + \begin{pmatrix} 15 & 20 \end{pmatrix} \begin{pmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} = 50$$

$$-20w^2 - \frac{25}{\sqrt{10}}z + \frac{75}{\sqrt{10}}w = 50$$

Le fait d'avoir un terme en w^2, mais aucun terme en z^2 signie qu'il doit s'agir d'une parabole. On doit donc isoler le z et compléter le carré.

$$-\frac{25}{\sqrt{10}}z = 20w^2 - \frac{75}{\sqrt{10}}w + 50$$

$$-25z = 20\sqrt{10}w^2 - 75w + 50\sqrt{10}$$

$$z = \frac{-4\sqrt{10}}{5}w^2 + 3w - 2\sqrt{10}$$

$$z = \frac{-4\sqrt{10}}{5}\left(w^2 - \frac{15w}{4\sqrt{10}}\right) - 2\sqrt{10}$$

$$z = \frac{-4\sqrt{10}}{5}\left(w - \frac{15}{8\sqrt{10}}\right)^2 - 2\sqrt{10} + \frac{4\sqrt{10}}{5}\left(\frac{15}{8\sqrt{10}}\right)^2$$

$$z = \frac{-4\sqrt{10}}{5}\left(w - \frac{15}{8\sqrt{10}}\right)^2 - \frac{55\sqrt{10}}{32}$$

Donc le centre de la parabole dans les coordonnées z, w est donné par :

$$\begin{pmatrix} \frac{-55\sqrt{10}}{32} \\ \frac{15}{8\sqrt{10}} \end{pmatrix}$$

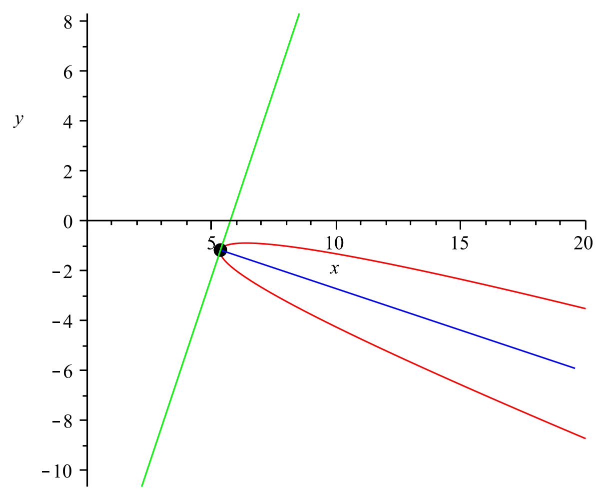
Il nous faut maintenant trouver l'équation du centre dans les coordonnées habituelle x, y. Pour ce faire, on utilise la matrice de passage P.

$$\text{Centre} = \begin{pmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} \begin{pmatrix} \frac{-55\sqrt{10}}{32} \\ \frac{15}{8\sqrt{10}} \end{pmatrix} = \begin{pmatrix} \frac{171}{32} \\ \frac{-37}{32} \end{pmatrix}$$

Maintenant, pour la direction dans laquelle la parabole est ouverte, il suffit de remarquer que la forme de l'équation nous dit que la parabole est ouverte dans la direction négative de l'axe des z. C'est à dire dans la direction opposé au vecteur :

$$\begin{pmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix}$$

Ces informations nous permet donc d'obtenir le graphique approximatif de notre parabole.



Exemple 5.7.5. Comme dernier exemple, nous vous étudier la fonction

$$f(x) = \frac{8x-5}{2x+4}$$

En particulier, nous somme intéressé à trouver les asymptotes de cette fonction (horizontale, verticale et possiblement oblique), sans utiliser les techniques vu dans les cours de calcul. Pour ce faire, nous allons traiter la fonction comme étant l'équation d'une conique, ce qui peut être fait facilement en remplaçant $f(x)$ par y. On a donc :

$$y = \frac{8x-5}{2x+4} \qquad \Rightarrow \qquad (2x+4)y = 8x-5 \qquad \Rightarrow \qquad 2xy - 8x + 4y = -5$$

Suivant notre habitude, nous réécrivons l'équation sous forme matricielle. Nous utilisons ici la même notation et le même changement de variable que dans les exemples précédent.

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -8 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -5$$

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix}^T \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -8 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -5$$

$$\begin{pmatrix} z & w \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} + \begin{pmatrix} -8 & 4 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix} = -5$$

$$z^2 - w^2 - \frac{4}{\sqrt{2}} z - \frac{12}{\sqrt{2}} w = -5$$

$$\left(z - \frac{2}{\sqrt{2}} \right)^2 - \left(w + \frac{6}{\sqrt{2}} \right)^2 = -21$$

$$\frac{-\left(z - \frac{2}{\sqrt{2}} \right)^2}{\left(\sqrt{21}\right)^2} + \frac{\left(w + \frac{6}{\sqrt{2}} \right)^2}{\left(\sqrt{21}\right)^2} = 1$$

On remarque donc que notre conique est une hyperbole. Le centre de l'hyperbole, dans les coordonnées habituelles est donc :

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{2}{\sqrt{2}} \\ \frac{-6}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \end{pmatrix}$$

Il nous reste maintenant à trouver la direction des asymptotes. Pour ce faire, nous devons trouver des vecteurs représentant les demi-axes. Ces vecteurs sont :

$$\sqrt{21} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{21}}{\sqrt{2}} \\ \frac{\sqrt{21}}{\sqrt{2}} \end{pmatrix} \qquad \sqrt{21} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{21}}{\sqrt{2}} \\ \frac{-\sqrt{21}}{\sqrt{2}} \end{pmatrix}$$

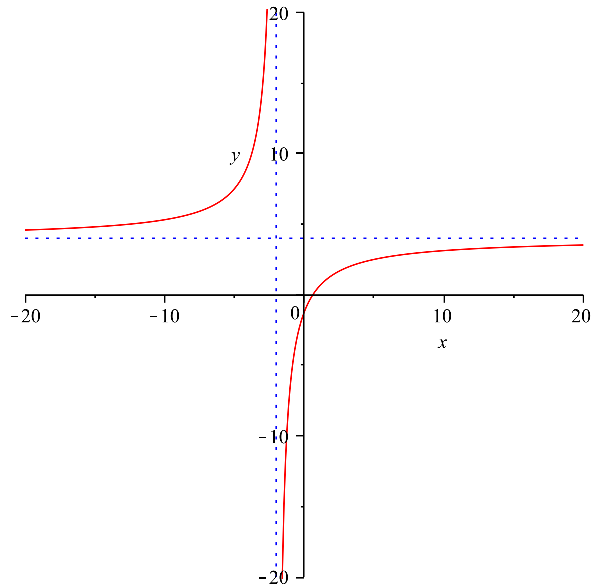
La somme et la différence de ces vecteurs nous donne la direction des asymptotes. Les asymptotes sont donc dans la direction des vecteurs :

$$\begin{pmatrix} \frac{2\sqrt{21}}{\sqrt{2}} \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ \frac{2\sqrt{21}}{\sqrt{2}} \end{pmatrix}$$

Les asymptotes sont donc horizontal et vertical respectivement. Comme les asymptotes doivent passer par le centre de l'hyperbole, on obtient donc les deux asymptotes suivante :

$$x = -2 \qquad \text{et} \qquad y = 4$$

Voici une graphique représentant notre hyperbole et en conséquence notre fonction de départ.


20
y
10
-20
-10
0
10
20
x
-10
-20

Appendice 1 : Quelques rappels

6.1 Systèmes d'équations linéaires

Dans cette section, nous allons explorer différentes méthodes pour résoudre le système d'équations suivant :

$$\begin{cases} 2x + 4y + 7z = 1 \\ 2x + 5y + 9z = 2 \\ 3x + 7y + 13z = 3 \end{cases}$$

Méthode de Gauss :

Rappel : Les opérations suivantes ne change pas la (les) solution d'un système d'équations :

1. $L_i + cL_j \to L_i$ avec $i \neq j$
2. $cL_i \to L_i$ si $c \neq 0$
3. $L_i \leftrightarrow L_j$
4. $aL_i + bL_j \to L_i$ avec $i \neq j$ et $a \neq 0$

On écrit le système sous forme de matrice augmenté puis on applique des opérations élémentaires jusqu'à obtenir une matrice échelon.

$$\left(\begin{array}{ccc|c} 2 & 4 & 7 & 1 \\ 2 & 5 & 9 & 2 \\ 3 & 7 & 13 & 3 \end{array}\right) \sim^{L_2 - L_1 \to L_2}_{L_3 - \frac{3}{2}L_1 \to L_3} \left(\begin{array}{ccc|c} 2 & 4 & 7 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 1 & 5/2 & 3/2 \end{array}\right) \sim^{L_3 - L_2 \to L_3} \left(\begin{array}{ccc|c} 2 & 4 & 7 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1/2 & 1/2 \end{array}\right)$$

Donc la solution du système est :

$$\begin{cases} z = 1 \\ y = 1 - 2z = 1 - 2 = -1 \\ x = \dfrac{1 - 4y - 7z}{2} = \dfrac{1 + 4 - 7}{2} = -1 \end{cases}$$

Méthode de Gauss-Jordan :

On applique des opérations élémentaires jusqu'à obtenir une matrice échelon réduite. Il s'agit donc de continuer à réduire la matrice obtenu dans la méthode de Gauss.

$$\left(\begin{array}{ccc|c} 2 & 4 & 7 & 1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1/2 & 1/2 \end{array}\right) \sim^{\frac{1}{2}L_1 \to L_1}_{2L_3 \to L_3} \left(\begin{array}{ccc|c} 1 & 2 & 7/2 & 1/2 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right) \sim^{L_2 - 2L_3 \to L_2}_{L_1 - \frac{7}{2}L_3 \to L_1} \left(\begin{array}{ccc|c} 1 & 2 & 0 & -3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

$$\sim^{L_1 - 2L_2 \to L_1} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right)$$

Donc la solution du système est :

$$\begin{cases} z = 1 \\ y = -1 \\ x = -1 \end{cases}$$

Factorisation LU

On réécrit les opérations élémentaires de la méthode de Gauss sous forme de matrices pour obtenir la matrice L. La matrice U étant la matrice obtenu dans la méthode de Gauss.

$$\begin{pmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{3}{2} & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 7 \\ 0 & 1 & 2 \\ 0 & 0 & 1/2 \end{pmatrix} = LU$$

Pour résoudre le système d'équations, on va donc devoir résoudre les deux systèmes d'équations suivants :

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3/2 & 1 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$\Longrightarrow \begin{cases} x' = 1 \\ y' = 2 - x' = 1 \\ z' = 3 - y' - \frac{3}{2}x' = \frac{1}{2} \end{cases}$$

$$\begin{pmatrix} 2 & 4 & 7 \\ 0 & 1 & 2 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1/2 \end{pmatrix}$$

$$\Longrightarrow \begin{cases} z = 1 \\ y = 1 - 2z = -1 \\ x = \frac{1}{2}(1 - 4y - 7z) = -1 \end{cases}$$

Méthode de la matrice inverse (Partie 1)

On commence par calculer l'inverse de la matrice des coefficients (ici avec la méthode de Gauss-Jordan) :

$$\left(\begin{array}{ccc|ccc} 2 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 9 & 0 & 1 & 0 \\ 3 & 7 & 13 & 0 & 0 & 1 \end{array}\right) \sim^{L_2 - L_1 \to L_2}_{L_3 - \frac{3}{2}L_1 \to L_3} \left(\begin{array}{ccc|ccc} 2 & 4 & 7 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 1 & 5/2 & -3/2 & 0 & 1 \end{array}\right) \sim^{L_3 - L_2 \to L_3} \left(\begin{array}{ccc|ccc} 2 & 4 & 7 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1/2 & -1/2 & -1 & 1 \end{array}\right)$$

$$\sim^{\frac{1}{2}L_1 \to L_1}_{2L_3 \to L_3} \left(\begin{array}{ccc|ccc} 1 & 2 & 7/2 & 1/2 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -2 & 2 \end{array}\right) \sim^{L_2 - 2L_3 \to L_2}_{L_1 - \frac{7}{2}L_3 \to L_1} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 4 & 7 & -7 \\ 0 & 1 & 0 & 1 & 5 & -4 \\ 0 & 0 & 1 & -1 & -2 & 2 \end{array}\right)$$

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

La matrice inverse est donc :

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

Donc la solution du système est donnée par :

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 & -3 & 1 \\ 1 & 5 & -4 \\ -1 & -2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$$

Méthode de la matrice inverse (Partie 2)

À la place d'utiliser la méthode de Gauss-Jordan pour trouver la matrice inverse, on peut aussi utiliser la formule suivante :

$$A^{-1} = \frac{1}{det(A)}[cof(A)]^t$$

On commence donc par calculer le déterminant :

$$\begin{vmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix} = 2\begin{vmatrix} 5 & 9 \\ 7 & 13 \end{vmatrix} - 4\begin{vmatrix} 2 & 9 \\ 3 & 13 \end{vmatrix} + 7\begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix} = 2(65-63) - 4(26-27) + 7(14-15) = 4 + 4 - 7 = 1$$

Puis on calcul la matrice des cofacteurs :

$$cof(A) = \begin{pmatrix} \begin{vmatrix} 5 & 9 \\ 7 & 13 \end{vmatrix} & -\begin{vmatrix} 2 & 9 \\ 3 & 13 \end{vmatrix} & \begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix} \\ \\ -\begin{vmatrix} 4 & 7 \\ 7 & 13 \end{vmatrix} & \begin{vmatrix} 2 & 7 \\ 3 & 13 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 3 & 7 \end{vmatrix} \\ \\ \begin{vmatrix} 4 & 7 \\ 5 & 9 \end{vmatrix} & -\begin{vmatrix} 2 & 7 \\ 2 & 9 \end{vmatrix} & \begin{vmatrix} 2 & 4 \\ 2 & 5 \end{vmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ -3 & 5 & -2 \\ 1 & -4 & 2 \end{pmatrix}$$

Donc la matrice inverse est :

$$A^{-1} = \frac{1}{1}\begin{pmatrix} 2 & 1 & -1 \\ -3 & 5 & -2 \\ 1 & -4 & 2 \end{pmatrix}^T = \begin{pmatrix} 2 & -3 & 1 \\ 1 & 5 & -4 \\ -1 & -2 & 2 \end{pmatrix}$$

Méthode de Cramer

La méthode de Cramer utilise les déterminants pour résoudre le système d'équations linéaires. Elle fonctionne seulement dans le cas ou nous avons exactement une solution. Comme exemple nous allons trouver la valeur de x.

$$x = \frac{\begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix}}{\begin{vmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix}} = \frac{\begin{vmatrix} 5 & 9 \\ 7 & 13 \end{vmatrix} - 4\begin{vmatrix} 2 & 9 \\ 3 & 13 \end{vmatrix} + 7\begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix}}{2\begin{vmatrix} 5 & 9 \\ 7 & 13 \end{vmatrix} - 4\begin{vmatrix} 2 & 9 \\ 3 & 13 \end{vmatrix} + 7\begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix}} = \frac{2+4-7}{4+4-7} = \frac{-1}{1} = -1$$

On peut faire de manière similaire pour trouver y et z.

6.2 Le déterminant

Definition 6.2.1. On définit le déterminant d'une matrice 2×2 comme étant :

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

Definition 6.2.2. Si A est une matrice $n \times n$, alors on définit le mineur i,j dénoté M_{ij} comme étant le déterminant de la matrice A à laquelle on a enlever la i-eme ligne et la j-eme colonne. De plus, on définit le cofactor i,j dénoté C_{ij} comme étant

$$C_{ij} = (-1)^{i+j} M_{ij}$$

Definition 6.2.3. Si A est une matrice de dimension $n \times n$ comme étant :

$$det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} = \sum_{j=1}^{n} a_{ij} C_{ij}$$

Les opérations élémentaires de la section précédente peuvent être traduite en terme d'opération sur le déterminant de la façon suivante :

1. $L_i + cL_j \to L_i$ ne change pas le déterminant si $i \neq j$

2. $cL_i \to L_i$ multiplie le déterminant par c
3. $L_i \leftrightarrow L_j$ change le signe du déterminant

On veut calculer le déterminant suivant :

$$\begin{vmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix}$$

Comme nous avons déjà calculer ce déterminant à l'aide de la définition dans la méthode de Cramer, nous allons le faire ici en utilisant des opérations sur les lignes. En appliquant une méthode de Gauss, on obtient :

$$\begin{pmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{pmatrix} \sim^{\frac{1}{2}L_1 \to L_1} \begin{pmatrix} 1 & 2 & 7/2 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{pmatrix} \sim^{\substack{L_2-2L_1\to L_2 \\ L_3-3L_1\to L_3}} \begin{pmatrix} 1 & 2 & 7/2 \\ 0 & 1 & 2 \\ 0 & 1 & 5/2 \end{pmatrix}$$

$$\sim^{L_3-L_2\to L_3} \begin{pmatrix} 1 & 2 & 7/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1/2 \end{pmatrix} \sim^{2L_3\to L_3} \begin{pmatrix} 1 & 2 & 7/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

En terme de déterminant, on a donc :

$$\begin{vmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix} = 2\begin{vmatrix} 1 & 2 & 7/2 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 7/2 \\ 0 & 1 & 2 \\ 0 & 1 & 5/2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 7/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1/2 \end{vmatrix} = \frac{1}{2}\cdot 2\begin{vmatrix} 1 & 2 & 7/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = \frac{1}{2}\cdot 2\cdot 1 = 1$$

Théorème 6.2.1. *Si A et B sont des matrices carrés pour lesquels les opérations si dessous sont permise. Alors on a :*

1. $\det(AB) = \det(A)\det(B)$
2. $\det(A^T) = \det(A)$
3. $\det(A^{-1}) = \dfrac{1}{\det(A)}$

Remarquez qu'en combinant le théorème ci dessus avec la factorisation LU, on peut donc à nouveau calculer le déterminant de la matrice. On a donc :

$$\begin{vmatrix} 2 & 4 & 7 \\ 2 & 5 & 9 \\ 3 & 7 & 13 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{3}{2} & 1 & 1 \end{vmatrix}\begin{vmatrix} 2 & 4 & 7 \\ 0 & 1 & 2 \\ 0 & 0 & 1/2 \end{vmatrix} = 1\cdot 1 = 1$$

6.3 Inversion de matrice et théorème de la matrice inverse

Nous avons déjà vu dans la section précédente que nous pouvons calculer l'inverse d'une matrice carré A (s'il existe) par la formule suivante :

$$A^{-1} = \frac{1}{det(A)}[cof(A)]^T$$

Dans le cas d'une matrice 2×2, on obtient donc la formule suivante qui nous sera très utile :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{pmatrix}$$

Théorème 6.3.1 (Théorème de la matrice inverse)**.** *Si A est une matrice carré de dimension $n\times n$, alors les énoncés suivants sont équivalent :*

1. *A est une matrice inversible*
2. $det(A) \neq 0$
3. *Le système d'équations linéaires* $Ax = 0$ *a une unique solution*
4. *Le système d'équations linéaires* $Ax = b$ *a une unique solution pour tout* $b \in \mathbb{R}^n$.

Appendice 2 : Les nombres complexes

7.1 Introduction : Les ensembles de nombres

Au primaire et secondaire, vous avez appris à travailler avec différent ensemble de nombre. Vous avez premièrement définie l'ensemble des nombres naturels :

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, ...\}$$

Remarquez que certain auteurs préfère exclure le zéro de l'ensemble des nombres naturels. Il s'agit d'une question de préférence qui n'aura pas vraiment d'incidence dans ce cours. L'ensemble des nombres naturels vous a permis d'apprendre à compter, et faire des additions simple. Par contre, lorsqu'est venu le temps d'apprendre la soustraction, vous avez vite réalisé que cette opération n'est pas toujours possible. Par exemple, il n'existe aucun nombre dans $\mathbb{N}$ qui est égal à $1 - 2$. Vous avez donc été amené à élargir votre concept de nombre à l'ensemble des nombres entiers :

$$\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}$$

Il était donc alors possible d'additionner et soustraire sans problème, et même d'effectuer des multiplications. Par contre, lorsqu'il est venu le temps d'apprendre à diviser, vous avez appris que ce n'était pas toujours possible. Par exemple, il n'existe aucun nombre de $\mathbb{Z}$ qui est égal à $1 \div 2$. Il a donc été encore une fois nécessaire d'élargir votre concept de nombre, ce qui vous a amené à définir l'ensemble des nombres rationnelles (les fractions) :

$$\mathbb{Q} = \left\{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\right\}$$

Dans cette ensemble, les choses vont particulièrement bien. Nous pouvons maintenant effectuer les 4 opérations de base : addition, soustraction, multiplication et division. Par contre, certain problème se posait toujours. Quelle est la longueur de l'hypoténuse d'un triangle rectangle pour lequel les deux côté adjacent à l'angle droit mesure 1 cm ? Quel est la circonférence et l'aire d'un cercle de rayon 1 ? Comme c'est problème n'avait toujours pas de solution, vous avez encore une fois étendu votre concept de nombre. Vous avez donc commencer à parler des nombres réels :

$$\mathbb{R} = \text{complétion de } \mathbb{Q}$$

La complétude est de procédé relativement complexe qui sera étudié en plus de détail dans le cours MATH-2721 : Suites et séries. Nous ne le décrirons donc pas en détail dans ce cours. Il s'agit essentiellement de boucher les trous qui était toujours présent dans $\mathbb{Q}$, de sorte que toute les longueurs soit maintenant des nombres. La plupart des opérations fonctionne très bien, et vous avez terminé vos études secondaire (et même première année universitaire) en utilisant seulement les nombres réels. Il restait cependant encore un problème. Certaines opérations qui semblent très naturelles restaient incorrectes. Nous avons donc :

$$(\sqrt{1})^2 = 1$$

$$(\sqrt{2})^2 = 2$$

$$(\sqrt{3})^2 = 3$$

$$(\sqrt{4})^2 = 4$$

Ce qui nous amène à supposer que $(\sqrt{x})^2 = x$ pour tout x. Par contre, ceci est complètement faux. En effet, si $x = -1$, nous avons :

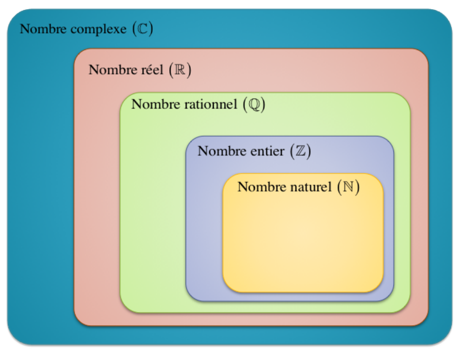
$$(\sqrt{-1})^2 = -1$$

qui semble juste à première vue, par contre $\sqrt{-1}$ n'existe pas dans les nombres réels. Ceci nous amène donc à définir un nouvel ensemble de nombre encore plus grand. Nous parlerons donc de l'ensemble des nombres complexes :

$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}, i^2 = -1\}$$

L'ensemble des nombres complexes est donc l'ensemble des nombres réels auquel nous ajoutons un nouveau nombre correspondant à $\sqrt{-1}$. Nous appelons ce nouveau nombre i. Nous allons donc voir dans ce chapitre comment travailler avec ces nouveaux nombres, et en particulier comment effectuer les opérations arithmétiques de base dans ce nouvel ensemble.

FIGURE 7.1 – Les ensembles de nombres



7.2 Les nombres complexes

L'ensemble des nombres complexes $\mathbb{C}$ est définie comme étant l'ensemble des nombres réels auxquels on ajoute un nombre dénoté i correspondant à $\sqrt{-1}$. Formellement, on définit l'ensemble des nombres complexes comme étant :

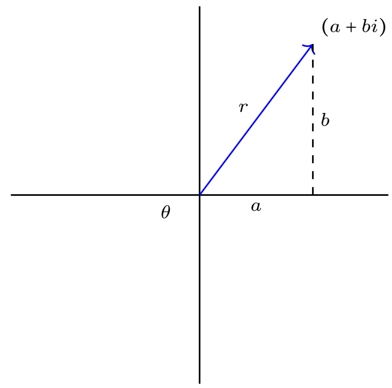
$$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}, i^2 = -1\}$$

L'avantage de travailler avec les nombres complexes plutôt que les nombres réels, c'est essentiellement que dans les nombres complexes, tous les polynômes qui ne sont pas constant admettent au moins une racine. En fait, dans tous les cas, ils se factorise complètement.

Un nombre complexe peut donc être représenté par un couple de deux nombres réels :

$$(a, b) \longleftrightarrow a + bi$$

Ce qui nous permet de représenter un nombre complexe dans le plan (complexe).



Ceci nous permet d'introduire nos 4 premières opérations sur les nombres complexes. Si z est un nombre complexe, alors on peut écrire z sous la forme cartésienne $z = a + bi$, où $a, b \in \mathbb{R}$. On obtient donc :

1. **La partie réelle :** $\Re(z) = a$
2. **La partie imaginaire :** $\Im(z) = b$
3. **Le module :** $|z| = r = \sqrt{a^2 + b^2}$
4. **L'argument :** $arg(z) = \theta = \arctan\left(\frac{b}{a}\right)$

Maintenant, un peu de trigonométrie nous permet d'obtenir les relations suivantes :

$$a = r\cos(\theta) \qquad \text{et} \qquad b = r\sin(\theta)$$

Ce qui nous permet d'obtenir la forme polaire d'un nombre complexe. Si z est un nombre complexe, alors on peut écrire z sous la forme :

$$z = r\Big(\cos(\theta) + i\sin(\theta)\Big)$$

où $r = ||z||$ et $\theta = arg(z)$.

Exemple 7.2.1. On veut trouver la forme polaire du nombre complexe $4 + 4i$. On a donc :

$$r = \sqrt{4^2 + 4^2} = \sqrt{32} \qquad \text{et} \qquad \theta = \arctan\left(\frac{4}{4}\right) = \frac{\pi}{4}$$

On obtient donc :

$$4 + 4i = \sqrt{32}\left(\cos\left(\frac{\pi}{4}\right) + i\sin\left(\frac{\pi}{4}\right)\right)$$

7.3 Opérations sur les nombres complexes

Notre but dans cette section est maintenant d'étendre les opérations habituelles d'addition, de soustraction, de multiplication et de division à l'ensemble des nombres complexes. Les opérations d'addition et de soustraction se font relativement facilement. Il suffit de penser au i comme s'il s'agissait d'une simple variable :
Si $(a + bi)$ et $(c + di)$ alors on a :

1. **Addition :** $(a + bi) + (c + di) = (a + c) + (b + d)i$
2. **Soustraction :** $(a + bi) - (c + di) = (a - c) + (b - d)i$

Maintenant, pour ce qui est de la multiplication, la méthode est relativement semblable à ce que nous avons fait pour l'addition et la soustraction. Il faut cependant se rappeler que $i^2 = -1$.
Si $(a + bi)$ et $(c + di)$ alors on a :

1. **Multiplication :** $(a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i$

La seule opération qui nous reste à définir est la division. Ici l'idée se complique un peu et nous aurons besoin d'une opération que l'on appelle conjugué pour arriver à nos fins.
Si $(a + bi)$ et $(c + di)$ alors on a :

1. **Conjugué :** $\overline{a+bi} = a - bi$
2. **Division :** $\dfrac{a+bi}{c+di} = \dfrac{a+bi}{c+di} \cdot \dfrac{c-di}{c-di} = \dfrac{(ac+bd)+(bc-ad)i}{c^2+d^2} = \dfrac{ac+bd}{c^2+d^2} + \dfrac{bc-ad}{c^2+d^2}i$

Exercice 7.3.1.

1. Démontrez que si z est un nombres complexes tel que $\overline{z} = z$, alors z est aussi un nombre réel.
2. Démontrez que si z est un nombre complexe, alors $z\overline{z}$ est un nombre réel.

Exemple 7.3.1. On veut effectuer l'opération suivante dans les nombres complexes :

$$\frac{7i-11}{3i+5}$$

On a donc :

$$\frac{7i-11}{3i+5} = \frac{7i-11}{3i+5} \cdot \frac{3i-5}{3i-5} = \frac{21i^2-35i-33i+55}{9i^2-25} = \frac{-68i+34}{-34} = 2i+1$$

7.4 Forme exponentielle et racines

Nous avons déjà vu dans les sections précédente qu'un nombres complexes peut s'écrire sous forme cartésienne et sous forme polaire. Dans cette section, nous allons voir qu'il est aussi possible d'écrire un nombre complexe sous forme exponentielle. L'idée provient des définitions des fonctions exponentielle, sinus et cosinus à l'aire des séries de Taylor.

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad \sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} \qquad \cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$$

Ces définitions ont aussi du sens dans les nombres complexes, et ce sont celles qui sont habituellement utilisées. En manipulant les sommes, on obtient la relation d'Euler suivante qui est particulièrement utile :

Identité d'Euler : $e^{ix} = \cos(x) + i\sin(x)$

Ce qui nous permet entre autre d'en déduire la formule de De Moivre qui est souvent très utile :

Formule de De Moivre : $(\cos(x) + i\sin(x))^n = \cos(nx) + i\sin(nx)$

Cette identité nous permet donc d'écrire un nombre complexe z sous les trois formes suivantes :

1. **Forme cartésienne :** $z = x + iy$ où $x = \Re(z)$ et $y = \Im(z)$
2. **Forme polaire :** $z = r(\cos(\theta) + i\sin(\theta))$ où $r = |z|$ et $\theta = arg(z)$
3. **Forme exponentielle :** $z = re^{i\theta}$ où $r = |z|$ et $\theta = arg(z)$

À partir de l'identité d'Euler, on peut aussi retrouver les expressions suivantes pour les fonctions sinus et cosinus :

$$\sin(z) = \frac{e^{iz} - e^{-iz}}{2i} \qquad \cos(z) = \frac{e^{ix} + e^{-iz}}{2}$$

Au secondaire, vous avez appris que les fonctions sinus et cosinus sont périodiques avec une période 2π. C'est à dire :

$$\sin(z + 2k\pi) = \sin(z), \quad \forall k \in \mathbb{Z}$$
$$\cos(z + 2k\pi) = \cos(z), \quad \forall k \in \mathbb{Z}$$

L'identité d'Euler nous permet maintenant de remarquer que la fonction exponentielle aussi est périodique, mais cette dernière a une période de $2ki\pi$. C'est à dire :

$$e^{iz+2ki\pi} = e^{iz}, \quad \forall k \in \mathbb{Z}$$

Cette dernière propriété va maintenant nous permettre de calculer toutes les racines d'un nombre. Par exemple, si on souhaite calculer toutes les valeurs de $\sqrt{4}$, on aura :

$$\sqrt{4} = \sqrt{4e^{0i+2ki\pi}} = \left(4e^{2ki\pi}\right)^{1/2} = 2e^{ki\pi} = 2e^0 \text{ ou } 2e^{i\pi} = 2 \text{ ou } -2$$

En générale, si z est un nombre complexe différent de 0, alors $\sqrt[n]{z}$ aura n solutions. Il sera possible de trouver toutes ces solutions en utilisant la forme exponentielle d'un nombre complexe.

Exemple 7.4.1. On veut calculer la valeur de $\sqrt{-16}$. On a donc :

$$\sqrt{-16} = \sqrt{(-1)(16)} = \sqrt{-1}\sqrt{16} = \pm 4i$$

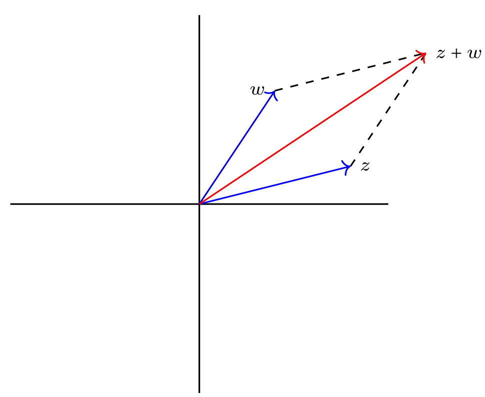
Exemple 7.4.2. Trouvez toutes les racines (dans les nombres complexes) du polynôme $x^2 + 2x + 6$. En appliquant la formule quadratique, on obtient :

$$\begin{aligned}
\frac{-b \pm \sqrt{b^2 - 4ac}}{2a} &= \frac{-2 \pm \sqrt{4-24}}{2} \\
&= \frac{-2 \pm \sqrt{-20}}{2} \\
&= -1 \pm \sqrt{-5} \\
&= -1 \pm \sqrt{5}\sqrt{-1} \\
&= -1 \pm i\sqrt{5}
\end{aligned}$$

7.5 Interprétation géométrique des opérations élémentaires

Nous allons maintenant regarder ce que signifie géométriquement les opérations élémentaires que nous avons vu sur les nombres complexes. Commençons par l'addition de deux nombres complexes. Si $z = a + bi$ et $w = c + di$ sont deux nombres complexes, alors on peut représenter leur somme en commençant par dessiner graphiquement les nombres z et w (en bleu), puis en complétant le parallélogramme (les lignes pointillés). La somme $z + w$, sera donc le nombre complexe se trouvant à l'extrémité du parallélogramme (représenté en rouge). Il s'agit en fait d'un addition de vecteurs.

FIGURE 7.2 – Addition de nombres complexes



Pour soustraire d'un nombre complexe, le principe est semblable. Il s'agit essentiellement de refaire les opérations que nous venons de faire à l'envers. On commence par dessiner les nombres complexes z et w (en bleu), puis on relie les deux points ainsi obtenu. On complete ensuite le parallélogramme en tenant compte que la ligne allant de l'origine jusqu'à z représente maintenant la diagonale du parallélogramme. Le nombre complexe $z - w$ sera alors l'autre sommet du parallélogramme.

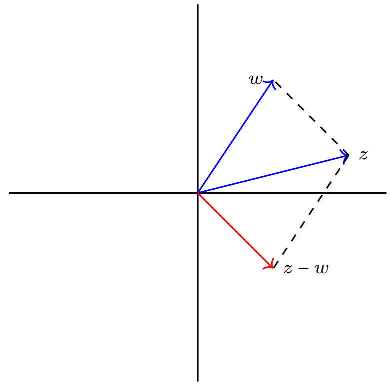
La multiplication et la division de deux nombres complexes est un peu plus étrange. Supposons que z et w sont deux nombres complexes, alors nous avons vu qu'on peut les écrire sous forme exponentielle :

$$z = r_1 e^{i\theta_1}, \qquad w = r_2 e^{i\theta_2}$$

En multipliant, en obtient donc :

$$zw = r_1 r_2 e^{i(\theta_1 + \theta_2)}$$

Figure 7.3 – Soustraction de nombres complexes



Donc un obtient le produit de deux nombres complexes en multipliant leur module, et en addition leur argument. De manière similaire, en divisant, on obtient :

$$\frac{z}{w} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)}$$

Donc la division est obtenu en divisant les modules, et en soustrayant les arguments.

7.6 Théorème fondamental de l'algèbre

Dans cette section, nous allons voir un théorème particulièrement important en mathématiques, et pour lequel la démonstration dépasse largement la niveau du cours. Pourtant, ce théorème est particulièrement simple à énoncé et est souvent appris dès le secondaire. Nous allons cependant inclure une version pour les nombres complexes qui sera sans doute nouvelle pour vous. Il s'agit du théorème fondamental de l'algèbre.

Théorème 7.6.1. **(Théorème fondamental de l'algèbre dans les nombres réels)** Si $p(x)$ est un polynôme de degré n à coefficients réels, alors l'équation $p(x) = 0$ possède au plus n solutions dans les nombres réels. De plus, $p(x)$ peut se factoriser sous forme d'un produit de facteur linéaire et / ou de facteur quadratique irréductible.

Par exemple, bien que nous ne soyons pas capable de calculer facilement les racines du polynôme $p(x) = 3x^5 + 4x^3 + 2x^2 - 5$, nous pouvons affirmer qu'il aura au plus 5 racines dans les nombres réels. Nous ne pouvons cependant pas garantir qu'il en a exactement 5. Par exemple, le théorème fondamental de l'algèbre nous garantie que le polynôme $q(x) = x^2 + 1$ possède au plus 2 racines dans les nombres réels, mais un calcul relativement simple nous permet de nous convaincre qu'il en a en fait aucune. Nous allons maintenant voir une version du théorème pour les nombres complexes qui va nous permettre d'aller beaucoup plus loin.

Théorème 7.6.2. **(Théorème fondamental de l'algèbre)** Si $p(x)$ est un polynôme de degré n à coefficients complexes tel que $p(x) \neq 0$, alors $p(x)$ possède exactement n racines dans les nombres complexes lorsque ces dernières sont comptés avec leur multiplicité.

Pour pouvoir bien comprendre le théorème, il nous est nécessaire d'expliquer ce que l'on entend par multiplicité d'une racine. L'existence d'une racine est équivalente à dire que le polynôme peut se factoriser. Par exemple, si a est une racine du polynôme $p(x)$, alors il existe un polynôme $q(x)$ tel que $p(x) = (x-a)q(x)$.

Le théorème fondamental de l'algèbre revient donc à dire qu'il nous est possible de factoriser complètement n'importe quel polynôme. Par exemple, on aura :

$$x^4 - 17x^3 + 105x^2 - 275x + 250 = (x-2)(x-5)^3$$

On a donc que 2 et 5 sont les racines du polynôme. On appelle la multiplicité d'une racine, l'exposant qui accompagne cette dernière. On dira donc que 2 est une racine de multiplicité 1 (c'est à dire une racine simple), et on dira que 5 est une racines de multiplicité 3. On peut donc vérifier que le théorème fondamental de l'algèbre fonctionne bien dans ce cas. En effet, comme il s'agit d'un polynôme non nul de degré 4, le théorème fondamental nous affirme que nous devons avoir exactement 4 racines, ce qui est bien le cas lorsqu'on les comptes avec leur multiplicité : $1 + 3 = 4$.

Remarquez qu'il s'agit uniquement d'un théorème d'existence. Le théorème ne nous donne aucune indication sur comment trouver les racines du polynôme. Dans les deux prochaines sections, nous allons voir comment trouver toutes les racines de certain polynômes relativement simple, et dans le chapitre suivant, nous allons voir comment trouver toutes les racines rationnelles d'un polynôme à coefficient entier. Par contre, en général il est impossible de trouver toutes les racines d'un polynôme de manière exacte, comme l'affirme le théorème ci-dessous. Il sera donc nécessaire d'utiliser des méthodes numérique pour les approximer, ce que nous ne ferons pas dans ce cours.

Théorème 7.6.3. **(Abel / Galois)** Il n'existe aucune formule permettant de trouver (en calcul exacte) toutes les racines d'un polynôme général de degré ≥ 5.

Faites attention cependant à ce que signifie ce théorème. Dans certain cas il est possible de calculer toutes les racines d'un polynôme de degré plus grand ou égal à 5. Ce que le théorème nous affirme c'est que ce n'est pas toujours possible.

7.7 Résolution de l'équation $z^n = a$

Nous allons maintenant nous intéresser à trouver toutes les solutions d'une équation de la forme $z^n = a$. Pour ce faire, il suffit d'utiliser la forme exponentiel comme nous l'avons déjà fait avec les racines carrés.

Exemple 7.7.1. On veut trouver toutes les solutions de l'équation $z^3 = 27$. Pour ce faire, on se rappel que $27 = 27e^{0+2\pi ik}$ avec $k \in \mathbb{Z}$. En calculant la racine cubique, on obtient donc :

$$z = (27e^{2\pi ik})^{1/3} = 27^{\frac{1}{3}} e^{\frac{2}{3}\pi ik} = 3e^{\frac{2}{3}\pi ik}, \quad k \in \mathbb{Z}$$

On doit maintenant se rappeler que la fonction exponentielle est périodique, et par le théorème fondamental de l'algèbre nous savons que nous cherchons exactement 3 solutions. Nous allons donc considérer uniquement les valeurs $k = 0, 1, 2$. On obtient donc les trois solutions suivantes :

$$z_0 = 3e^{\frac{2}{3}\pi i(0)} = 3e^0 = 3$$

$$z_1 = 3e^{\frac{2}{3}\pi i(1)} = 3e^{\frac{2}{3}\pi i} = 3\left(\cos\left(\frac{2\pi}{3}\right) + i\sin\left(\frac{2\pi}{3}\right)\right) = \frac{-3}{2} + i\frac{3\sqrt{3}}{2}$$

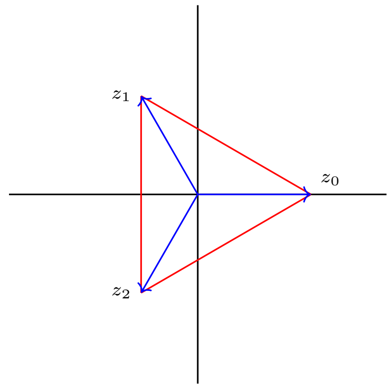
$$z_2 = 3e^{\frac{2}{3}\pi i(2)} = 3e^{\frac{4}{3}\pi i} = 3\left(\cos\left(\frac{4\pi}{3}\right) + i\sin\left(\frac{4\pi}{3}\right)\right) = \frac{-3}{2} + i\frac{-3\sqrt{3}}{2}$$

Les solutions de l'équations $z^n = a$ est particulièrement curieuse et possède un lien important avec la géométrie comme le montre le théorème suivant :

Théorème 7.7.1. En reliant ensemble les solutions de l'équations $z^n = a$ où $a \in \mathbb{R}\backslash\{0\}$, on obtient toujours un polygone régulier à n côtés.

Bien qu'en apparence le théorème soit plutôt une curiosité mathématiques, il peut aussi être utile pour vérifier qu'aucune erreur de calcul n'a été commise, ou bien pour nous permettre de trouver géométriquement les solutions de l'équation lorsqu'une première solution est connue. Pour illustrer le théorème, regardons géométriquement ou se trouve les solutions de l'équation $z^3 = 27$ que nous avons trouvé précédemment.

FIGURE 7.4 – Solutions de l'équations $z^3 = 27$



On obtient donc que les solutions de l'équations $z^3 = 27$ représentent bien un triangle équilatéral comme le théorème l'avait prédis.

7.8 Racines d'une fonction quadratique

Nous avons déjà vu à l'aide d'un exemple comment trouver les racines d'une fonction quadratique (i.e. un polynôme de degré 2). Nous allons maintenant le refaire de manière un peu plus précise. Au secondaire, vous avez appris qu'un polynôme de degré 2 (à coefficients réels) possède soit aucune, une seule, ou exactement deux racines. Ceci correspond au théorème fondamental de l'algèbre dans le cas réel. Vous avez aussi appris que les racines (lorsqu'il y en a) du polynôme $p(x) = ax^2 + bx + c$ sont donné par la formule :

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

et que le nombre de solution est obtenu en étudiant le signe du discriminent $\Delta = b^2 - 4ac$ qui se trouve sous la racine. Le cas où $\Delta < 0$ correspondait alors à aucune solution, car il n'était pas possible de calculer la racine carré d'un nombre négatif. Maintenant que vous connaissez les nombres complexes, il nous ait maintenant possible de calculer la racine carré d'un nombre négatif. Le cas où nous avions aucune solution dans les nombres réels, correspondra donc à deux solutions dans les nombres complexes.

Théorème 7.8.1. Si $p(z) = az^2 + bz + c$ est un polynôme à coefficient complexes, alors les racines du polynôme sont donné par la formule :

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Exemple 7.8.1. On veut trouver toutes les racines du polynôme $p(z) = z^2 - 4z + 29$ dans les nombres complexes. Pour ce faire, on utilise la formule quadratique, ce qui nous donne :

$$z = \frac{4 \pm \sqrt{(-4)^2 - 4(1)(29)}}{2(1)} = \frac{4 \pm \sqrt{16 - 116}}{2} = \frac{4 \pm \sqrt{-100}}{2} = \frac{4 \pm 10i}{2} = 2 \pm 5i$$

Théorème 7.8.2. Si $p(x) = ax^2 + bx + c$ est un polynôme à coefficient réel pour lequel $z \in \mathbb{C} \backslash \mathbb{R}$ est l'une des racines, alors l'autre racine sera $\overline{z}$. C'est à dire que les racines non-réel apparaissent toujours par paire qui sont conjugué l'une à l'autre.

Exemple 7.8.2. Sachant que $\frac{5}{2} - i\frac{\sqrt{11}}{2}$ est l'une des racines du polynôme $x^2 - 5x + 9$, on veut trouver l'autre sans faire de calcul. Comme il s'agit d'un polynôme de degré 2 à coefficient réel, le théorème précédent nous garantie que l'autre racine sera donné par le conjugué. On a donc que l'autre racine est :

$$\frac{5}{2} + i\frac{\sqrt{11}}{2}$$

Exemple 7.8.3. Sachant que $2i$ et 5 sont des racines du polynôme $p(x) = x^3 - 5x^2 + 4x - 20$, on veut trouver la troisième. Pour ce faire, remarquons que comme 5 est une racine du polynôme $p(x)$, alors on peut trouver un polynôme $q(x)$ qui sera aussi à coefficient réels tel que $p(x) = (x - 5)q(x)$. On aura donc que $q(x)$ est un polynôme de degré 2 pour lequel $2i$ est une racine. La troisième racine devra donc être le conjugué de $2i$, c'est à dire que la troisième racine est $-2i$.

7.9 Application à la trigonométrie

Nous allons maintenant voir dans cette section une application des nombres complexes à la trigonométrie. En effet, il est possible de retrouver la plupart des identités trigonométriques en comparant les formes polaires et exponentielles d'un nombre complexe.

Exemple 7.9.1. À l'aide de la formule d'Euler et des nombres complexes, nous pouvons retrouver les identités trigonométriques pour $\cos(2x)$ et $\sin(2x)$.

$$\begin{aligned} \cos(2x) + i\sin(2x) &= e^{2xi} \\ &= (e^{xi})^2 \\ &= (\cos(x) + i\sin(x))^2 \\ &= \cos^2(x) + 2i\cos(x)\sin(x) - \sin^2(x) \end{aligned}$$

Ce qui nous donne :

$$\begin{aligned} \cos(2x) &= \Re(\cos(2x) + i\sin(2x)) \\ &= \Re(\cos^2(x) + 2i\cos(x)\sin(x) - \sin^2(x)) \\ &= \cos^2(x) - \sin^2(x) \end{aligned}$$

$$\begin{aligned} \sin(2x) &= \Im(\cos(2x) + i\sin(2x)) \\ &= \Im(\cos^2(x) + 2i\cos(x)\sin(x) - \sin^2(x)) \\ &= 2\cos(x)\sin(x) \end{aligned}$$

Exercice 7.9.1. Utiliser la formule d'Euler et les nombres complexes pour obtenir des identités trigonométrique pour $\cos(a + b)$ et $\sin(a + b)$.

Bibliographie

[1] Howard Anton and Chris Rorres. *Elementary linear algebra.* Wiley-Interscience [John Wiley & Sons], tenth edition edition, 2010.

[2] Sheldon Axler. *Linear algebra done right.* Undergraduate Texts in Mathematics. Springer, Cham, third edition, 2015.

[3] Joseph Grifone. *Algèbre linéaire.* Cépaduès Éditions, Toulouse, 1990.

[4] Peter D. Lax. *Linear algebra and its applications.* Pure and Applied Mathematics (Hoboken). Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2007.

[5] David C. Lay. *Algèbre Linéaire : Théorie, Exercices et Applications.* De Boeck, 2009.

[6] Gilbert Strang. The fundamental theorem of linear algebra. *Amer. Math. Monthly*, 100(9) :848–855, 1993.

Index