L2 SVT : ANALYSE DE DONNÉES ET CALCUL MATRICIEL

KEVIN TANGUY

Ce cours est destiné à un enseignement de 5 créneaux de $2h40$ dans la filière SVT de l'université d'Angers. Le public étant composé de biologistes et de chimistes, la plupart des résultats théorique sont omis (et nous renvoyons le lecteur curieux vers l'ouvrage de son choix pour combler ces manques, par exemple « Toute l'algèbre de la licence » de J.P. Escoffier dont nous nous inspirons) afin de mettre l'accent sur le côté pratique du calcul matriciel.

TABLE DES MATIÈRES

1. Introduction 2
1.1. Systèmes linéaires 2
1.2. Un peu de génétique 2
2. Eléments de réponses 3
2.1. Systèmes linéaires 3
2.2. Génétique 4
3. Systèmes linéaires 4
4. Calcul matriciel 6
4.1. Opérations sur les matrices 7
5. Inverse de matrice et déterminant 10
5.1. Calcul d'inverse de matrice 10
5.2. Notions de déterminant 14
6. Diagonalisation de matrice 16

1. Introduction

Voici quelques exemples permettant de motiver l'utilisation de matrices (que nous étudierons plus en détail dans la suite de ce cours).

1.1. **Systèmes linéaires.** Soit $a \in \mathbb{R}_*$ et $b \in \mathbb{R}$ des réels, il est bien simple de résoudre l'équation suivante :

$$ax = b \iff x = a^{-1}b \tag{1.1}$$

Il n'est pas non plus très difficile de résoudre le système suivant :

$$\begin{cases} 2x + y = 3 \\ -x + 6y = -1 \end{cases} \tag{1.2}$$

Bien que de nature similaire, le système suivant semble encore plus complexe à résoudre :

$$\begin{cases} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \ldots + a_{2n}x_n = b_2 \\ \qquad\qquad \vdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = b_n \end{cases}$$

pour des réels données $(a_{ij})_{1\leq i\leq j\leq n}$ et $(b_i)_{i=1,\ldots,n}$.

1.2. **Un peu de génétique.** Au 19-ème siècle, un moine du nom de Gregor Mendel procéde à des expériences de croisement de variétés de pois pour étudier certaines traits héréditaires (la couleur du pois par exemple). Après quelques expériences, il observe le résultat surprenant suivant : en croisant deux pois jaunes (a priori identique d'une point de vu extérieur) il a obtenu un pois vert. Il propose alors en 1866 la théorie suivante : le génotype et le phénotype (d'un trait particulier) sont deux choses distinctes.

Il suggère alors un modèle permettant de mieux saisir ce qui se produit pour les pois. D'après Mendel, la couleur d'un pois est déterminé par deux gênes se présentant sous deux types différents (allèles) :

(1) un type dominant que l'on notera (d)
(2) un type récessif que l'on notera (r).

Ainsi, un individu peut avoir un génotype de la forme suivante :

(1) D = dd (dominant pur)

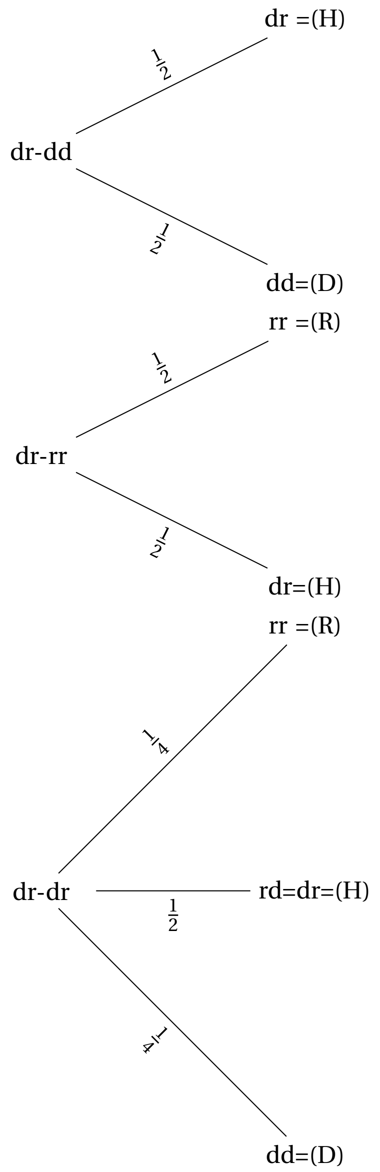
(2) H = dr (hybride)

(3) R = rr (récessif pur).

Si l'on croise deux individus, la descendance obtient un gêne (au hasard) de chacun de ses parents (reproduction diploïd). De plus, Mendel propose le postulat suivant : lorsque le gêne d est présent il détermine le phénotype de l'individu (ici la couleur du poids). Le gêne récessif r determine le phénotype uniquement lorsque le génotype est rr.

Remarque. Notons donc que les génotypes dd et dr donnent lieu au même phénotype (ici la couleur).

Voici l'expérience que l'on peu mener avec ce trait héréditaire : on prend un individu quelconque (un pois) et on le croise un hybride. On prend un descendant d'une telle opération et on le croise à nouveau avec un hybride. Il est facile de décrire les différents cas de figures lors d'une premier croisement. En effet, nous pouvons résumer ceci par les diagrammes suivants :



Que se passe-t-il si l'on itère indéfiniment cette expérience ? Quelle notation (mathématique) pourrait-on adopter pour que cela plus simple de comprendre ce qui se produit ?

2. Eléments de réponses

2.1. **Systèmes linéaires.** Peut-être que l'on pourrait collecter les coefficients du systéme (1.2) sous la forme d'un tableau ?

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix},$$

Il serait alors tentant de proposer l'écriture suivante de (1.2) :

$$AX = B$$

afin d'obtenir une solution similaire à (1.1)

$$X = A^{-1}B$$

Seulement quel sens peut-on donner à A^{-1} ou au produit $A^{-1}B$? D'ailleurs a-t-on l'égalité suivante ? $A^{-1}B = BA^{-1}$?

2.2. **Génétique.** Pourrait-on représenter l'expérience de croisement par un tableau similaire ?

$$\begin{array}{cc} \begin{array}{ccc} D & H & R \end{array} & \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} & \begin{array}{c} D \\ H \\ R \end{array} \end{array}$$

Ce tableau se lit comme suit : supposons qu'on le parte d'un individu de phénotype D alors, après croissement avec un hybride H, nous avons 50% de chance d'obtenir un individu de phénotype D, 50% de chance d'obtenir un individu de phénotype H et 0% de chance d'avoir un individu de phénotype R. Ceci est encodé, dans notre tableau, par la première ligne. Les lignes suivantes de lisent de la même manière (toujours en effectuant un croisement avec un hybride).

3. Systèmes linéaires

Certains problèmes mathématiques donnent lieu à des systèmes d'équations impliquant plusieurs inconnues. Voici un exemple assez ancien. Liu Hui (Chine 250 – 300 avant J.C.) était confronté à problème de céréales. Il connaisant les informations concernant la qualité des bottes de céréales et combien de quantité (quantifier avec l'unité de mesure Chinoise, le dou) de grain cela permettait de produire :

(1) 3 bottes de qualités supérieur combinées à 2 bottes de qualités moyennes et 1 botte de qualité médiocre donnent 39 dou (de grains).

(2) 2 bottes de qualités supérieur combinées à 3 bottes de qualités moyennes et 3 bottes de qualités médiocres donnent 34 dou (de grains).

(3) 1 botte de qualité supérieur combinée à 2 bottes de qualités moyennes et 3 bottes de qualités médiocres donnent 26 dou (de grains).

Il s'interrogea alors sur la quantité de grains obtenue pour chaque type de botte de céréales. Si l'on désigne par (x, y, z) les bottes de qualités supérieurs, moyennes et médiocre, il n'est pas difficile d'obtenir le système d'équation suivant :

$$\left\{ \begin{array}{ll} 3x + 2y + z = 39 & L_1 \\ 2x + 3y + z = 34 & L_2 \\ x + 2y + 3z = 26 & L_3 \end{array} \right. \tag{3.1}$$

Pour résoudre ce système nous allons mettre en place l'algorithme du pivot de Gauss. Bien que celui fu connu de manière pragmatique par Liu Hui, il fallu attendre les années 1800 pour que Gauss démontre de manière rigoureuse toutes les propriétés nécessaires à sa mise en place. Voyons par l'exemple comment celui-ci fonctionne.

(1) Tout d'abord choisissons une variable qui nous servira de pivot. Par exemple x, le but est alors d'effectuer des opérations sur les lignes permettant de faire disparaitre la variables x des lignes L_2 et L_3 (et de ne conserver la variable x uniquement dans la ligne L_1). Bien entendu, tout ceci est arbitraire et l'on aurait pu choisir la variables y ou z pour procéder identiquement. On aurait également pu échanger les lignes L_1 et L_3 (par exemple) si l'on préfère travailler avec un coefficient 1 devant la variable de pivot.

Toutes les opérations précédemment citées ne modifie pas les système (ceci ayant été démontrer par Gauss) et permettent d'aboutir au même ensemble de solutions. Nous n'entrerons pas dans ces détails théorique par la suite. Voyons plutôt comme éliminer x de la ligne L_2.

(2) Nous sommes donc tenter de multiplier la deuxième ligne par 3 et de lui soustraire la première ligne aprés l'avoir multiplier par 2. Si l'on procède ainsi, il ne restera plus de variable x dans la ligne L_2. Mathématiquement cela se note :

$$L_2 \leftarrow 3L_2 - 2L_1.$$

Ces opérations faites, nous obtenons le système équivalent suivant :

$$\begin{cases} 3x+2y+z=39 \qquad L_1 \\ 0+5y+z=24 \qquad L_2 \\ x+2y+3z=26 \qquad L_3 \end{cases}$$

Il est alors naturel de procèder de manière similaire pour éliminer la variable x du la ligne L_3. Autrement dit, nous effectuons l'opération suivante :

$$L_3 \leftarrow 5L_3 - 4L_2$$

pour obtenir

$$\begin{cases} 3x+2y+z=39 \qquad L_1 \\ 5y+z=24 \qquad L_2 \\ 36z=99 \qquad L_3 \end{cases}$$

(3) Il est maintenant assez simple de résoudre le système précédent. On obtient aisément que $z = \frac{99}{36}$ (d'après L_3), on peut donc substituer sa valeur dans la L_2 pour en déduire la valeur de y. Une fois y déterminer on peut substituer à nouveau dans la ligne L_1 pour trouver la valeur de la variable x.

Remarque. Il se trouve qu'un heureux hasard fait que la dernière opération ($L_3 \leftarrow 5L_3 - 4L_2$) supprime également la variable y de la ligne 3. Si jamais cela n'avait pas été le cas, nous aurions procéder au même raisonnement que precédemment sur les lignes L_2 et L_3. Plus précisement, les lignes L_2 et L_3 ne contiendraient que les variables y et z. Il aura alors fallu choisir une nouvelle variable de pivot dans la ligne L_2 (y par exemple) et d'effectuer des opérations sur les lignes pour qu'il n'y ait plus de variables y dans les ligne L_3.

Nous allons à présent mettre en place cette méthode sur différents exemples (exercice 1 et 2) pour observer ce qu'il se produit. Nous allons notamment voir qu'il n'y a pas toujours une unique solution. Parfois, il peut exister une infinité de solutions ou, au contraire, aucune solutions. Nous verrons également comment résoudre une système d'équation dépendant d'un paramètre.

4. Calcul matriciel

Ces objets, dont nous avons esquissé la forme durant l'introduction, ont été inventé par les mathématiciens Sylvester (1814 – 1897) et Cayley (1821 – 1895). Nous allons voir quelles opérations classiques, entre de tels objets, sont licites et de quelle manière celle-ci peuvent servir à representer, de manière plus compacte, un système linéaire.

Définition 4.1. *Une matrice* A *à coefficients réels de* $n \geq 1$ *lignes et* $p \geq 1$ *colonnes est un tableau rectangulaire de* $n \times p$ *nombres réels* $(a_{ij})_{1\leq i\leq n, 1\leq j\leq p}$ *:*

$$\mathrm{A} = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ a_{12} & \dots & a_{2p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$$

L'ensemble de telles matrices est noté $\mathrm{M}_{n\times p}(\mathbb{R})$. *Lorsque* $n = p$ *on parlera de matrice carrées de taille* n.

Exemple 4.1. Voici quelques exemples :

$$\mathrm{A} = \begin{pmatrix} 3 & -1 & 5 \\ -1 & 2 & 0 \end{pmatrix} \in \mathrm{M}_{2\times 3}(\mathbb{R}), \quad \mathrm{B} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} \in \mathrm{M}_{3\times 1}(\mathbb{R}), \quad \mathrm{C} = \begin{pmatrix} 2 & 0 & 7 \\ 1 & -1 & 8 \\ 25 & -12 & 0 \end{pmatrix} \in \mathrm{M}_{3\times 3}(\mathbb{R})$$

Remarque. La matrice B est aussi appelée vecteur colonne.

Certaines matrices joueront un rôle particulier par la suite, il s'agit de la matrice nulle, la matrice identité et des matrices diagonales.

Définition 4.2. *Soit $n \in \mathbb{N}_*$.*

(1) La matrice nulle $0_{M_n(\mathbb{R})}$ (que l'on notera la plupart du temps 0) est définie comme suit :

$$0_{M_n(\mathbb{R})} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

(2) La matrice identité $\mathrm{I}d_{\mathbb{R}^n}$, que nous notera la plupart du temps par I_d, est définie comme suit :

$$\mathrm{I}d_{\mathbb{R}^n} = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

où les constantes 1 sont uniquement placcées sur les termes diagonaux de la matrices.

(3) La matrice identité est un cas particulier des matrices diagonales. Un matrice diagonale D est de la forme suivante :

$$\mathrm{D} = \begin{pmatrix} a_{11} & \dots & \dots & \dots & 0 \\ 0 & a_{22} & \dots & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \dots & \dots & 0 & a_{nn} \end{pmatrix}$$

avec $a_{11}, \dots, a_{nn}$ des réels.

4.1. **Opérations sur les matrices.** Soient $p \geq 1, n \geq 1$, considérons deux matrices de même taille : $\mathrm{A}, \mathrm{B} \in M_{n\times p}(\mathbb{R})$. Quel sens donné à $\mathrm{A}+\mathrm{B}$? Si de plus, nous considérons un réel (scalaire) λ, comment définir $\lambda\mathrm{A}$?

Définition 4.3. *La matrice $\mathrm{C} = (c_{ij})_{1\leq i\leq n, 1\leq j\leq p}$, correspondant à la somme $\mathrm{A}+\mathrm{B}$, est définie comme suit :*

$$\mathrm{C} = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1p}+b_{1p} \\ a_{12}+b_{12} & \dots & a_{2p}+b_{2p} \\ \vdots & & \vdots \\ a_{n1}+b_{n1} & \dots & a_{np}+b_{np} \end{pmatrix}$$

Il suffit simplement d'additionner termes à termes les coefficients des matrices.

Remarque. **Attention** : il n'est pas possible d'additionner des matrices de tailles différentes.

Exemple 4.2. Considérons les matrices suivantes :

(1)

$$\mathrm{A} = \begin{pmatrix} 1 & -2 & 4 \\ 0 & 16 & -1 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 3 & 6 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

alors la matrice $\mathrm{C} = \mathrm{A}+\mathrm{B}$ vaut

$$\mathrm{C} = \begin{pmatrix} 4 & 4 & 4 \\ -1 & 16 & 0 \end{pmatrix}$$

(2)

$$A = \begin{pmatrix} -7 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & -2 & 4 \end{pmatrix}$$

alors

$$C = \begin{pmatrix} -6 & -2 & 5 \end{pmatrix}$$

Définition 4.4. *Le produit d'un scalaire* $\lambda \in \mathbb{R}$ *avec une matrice* $A = (a_{ij})_{1\le i\le n, 1\le j\le p}$ *est défini comme suit :*

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1p} \\ \lambda a_{12} & \dots & \lambda a_{2p} \\ \vdots & & \vdots \\ \lambda a_{n1} & \dots & \lambda a_{np} \end{pmatrix}$$

en mulipliant tout les termes de la matrices par le réel λ.

Exemple 4.3. Reprenons la matrice A de l'exemple précédent et multiplions la par $\lambda = 2$, nous obtenons alors la matrice

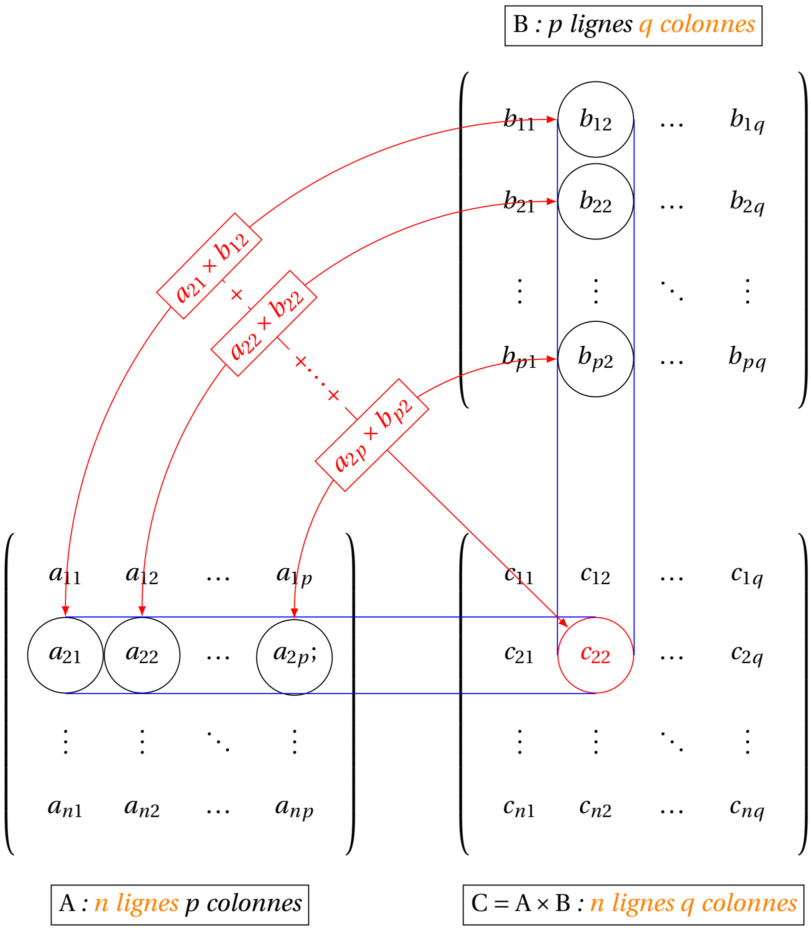
$$2A = \begin{pmatrix} 2 & -4 & 8 \\ 0 & 32 & -2 \end{pmatrix}$$

Etant donné deux matrices A et B comment définir le produit AB ? Tout les produits sont-ils envisageables ou bien faut-il imposer certaines restriction aux tailles des matrices A et B ?

Définition 4.5. *Soient* $n \ge 1$, $p \ge 1$, $q \ge 1$, *considérons alors des matrices* $A = (a_{ij})_{1\le i\le n, 1\le j\le p} \in M_{n\times p}(\mathbb{R})$ *et* $B = (b_{ij})_{1\le i\le p, 1\le j\le q} \in M_{p\times q}(\mathbb{R})$. *Le produit* C *de telles matrices est défini comme suit :*

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \quad 1 \le i \le n, 1 \le j \le q$$

De manière plus visuelle, cela correspond au schéma suivant :



Remarque. **Attention** : il est essentiel que le nombre de lignes de la matrices A coïncide avec le nombre de colonnes de la matrice B pour effectuer le produit AB. Lorsque les matrices sont carrées les deux produits AB et BA ont alors du sens. Cependant, en général, $AB \neq BA$ contrairement au produit de nombre réels que l'on a plus l'habitude de rencontrer.

Exemple 4.4. Soient A et B deux matrices définies comme suit :

$$A = \begin{pmatrix} 0 & 2 & -1 \\ 3 & 4 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 12 & -1 \\ 1 & -1 \\ 0 & 5 \end{pmatrix}$$

alors

$$AB = \begin{pmatrix} 0 \times 12 + 2 \times 1 + (-1) \times 0 & 0 \times (-1) + 2 \times (-1) - 1 \times 5 \\ 3 \times 12 + 4 \times 1 + (-3) \times 0 & 3 \times (-1) + 4 \times (-1) - 3 \times 5 \end{pmatrix} = \begin{pmatrix} 2 & -7 \\ 40 & -22 \end{pmatrix}$$

Comme nous avons pu le découvrir dans l'introduction, les matrices s'avèrent un outil commode pour représenter un système linéaire d'équation à plusieurs inconnues. Observons le dans l'exemple suivant :

Exemple 4.5. Soient

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \end{pmatrix},$$

Vérifiez que le produit AX = B correspond bien à l'écriture matricielle du système linéaire (1.2) de l'introduction :

$$\begin{cases} 2x + y = 3 \\ -x + 6y = -1 \end{cases}$$

Exemple 4.6. En statistiques, un modèle de regression linéaire est un modèle qui cherche à exprimer une relation linéaire entre une variable (dite variable expliquée) et une ou plusieurs variables (dites explicatives).Par exemple, on pourrait chercher à expliquer la concentration d'ozone sur différents lieux $(y_1, \dots, y_n)$ par différents paramètres comme la température, l'altitude, etc... (ce qui correspondrait aux variables $x_{11}, \dots, x_{np}$). Un tel modèle pourrait se représenter sous la forme suivante :

$$\mathrm{Y} = \mathrm{X}\beta + \epsilon,$$

avec

$$\mathrm{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathrm{X} = \begin{pmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix},$$

avec ϵ_i, $i = 1, \dots, n$ des variables aléatoires représentant les erreurs de mesures. Le problème du statisticien, pour analyser ces données, correspondrait à déterminer la valeur des coefficients $\beta_0, \dots, \beta_p$ associés aux variables explicatives $x_{11}, \dots, x_{np}$.

5. Inverse de matrice et déterminant

5.1. **Calcul d'inverse de matrice.**

Revenons à l'un de nos premiers exemples :

$$\begin{cases} 2x + y = 3 \\ -x + 6y = -1 \end{cases}$$

Nous venons de voir qu'il était possible d'exprimer ceci en termes matricielles : AX = B avec

$$\mathrm{A} = \begin{pmatrix} 2 & 1 \\ -1 & 6 \end{pmatrix}, \quad \mathrm{B} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad \mathrm{X} = \begin{pmatrix} x \\ y \end{pmatrix},$$

Il est donc tentant de procéder comme avec des réels : $ax = b \iff x = a^{-1}b$ (pour $a \in \mathbb{R}_*$ et $b \in \mathbb{R}$) en proposant

$$\mathrm{X} = \mathrm{A}^{-1}\mathrm{B}$$

Toutefois, il faut donner un sens à A^{-1} ! A quelle condition un tel objet existe-t-il ? Dans $\mathbb{R}$ c'est assez simple, il faut et il suffit que a soit non nul pour que a^{-1} existe. Lorsqu'il s'agit de matrice on parle d'inversibilité. Dans un premier temps nous allons donner une définition de ceci, puis voir de quelle manière peut-on calculer l'inverse d'une matrice. La notion sous-jacente de rang d'une matrice sera également, briévement, abordée.

Définition 5.1. *Soient* $\mathrm{A}, \mathrm{B} \in \mathrm{M}_n(\mathbb{R})$. *On dira que la matrice* B *est l'inverse de la matrice* A *si la relation suivante est satisfaite :*

$$\mathrm{AB} = \mathrm{BA} = \mathrm{I}_d \tag{5.1}$$

Si elle existe, une telle matrice est unique. Nous la désignerons par A^{-1}.

Remarque. La relation (5.1) est similaire à celle des réels : $a \times a^{-1} = 1$.

Exemple 5.1.
Plaçons nous dans l'espace des matrices de tailles 2×2 et considérons la matrice suivante :

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

Vérifiez que la matrice $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ est bien l'inverse de la matrice A.

Montrons que la matrice suivante n'admet pas d'inverse :

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

En effet, si une telle matrice existait elle serait de la forme $D = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et satisfera la relation $CD = DC = I_d$. Autrement dit, la matrice

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \times \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$$

doit être égale à la matrice $I_d = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Ce qui est absurde car $0 \neq 1$. Donc la matrice C n'est pas inversible.

Comment obtenir de manière algorithmique l'inverse (si celui-ci existe) d'une matrice donnée ? Tout ceci s'obtient par le biais de l'algorithme de Gauss que nous avons utilisé auparavant. Nous verrons dans la section suivante que la notion de déterminant permet d'obtenir une condition nécéssaire et suffisante à l'inversibilité d'une matrice. En attendant, voici un exemple

Exemple 5.2. Considérons la matrice A suivante :

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

et cherchons à l'inverser. Pour cela nous allons adopter la notation suivante

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

qui corresponf à concatener les matrices A et I_d. Maintenant que ceci est fait, nous allons mettre en place l'algorithme de Gauss sur la partie de gauche du tableau précédent. Plus précisement, pour commencer, nous allons faire des opérations sur les lignes pour faire apparaitre des 0 en position a_{21} et a_{31}. Il est important que les opérations que nous procédions aux mêmes opérations dans la partie de droite du tableau. Voyons plutôt :

L'opération $L_2 \leftarrow L_2 - 3L_1$ permet de rendre nul le terme a_{21} tandis que l'opération $L_3 \leftarrow L_3 + 2L_1$ annule le terme a_{31}. Effectuons ces opérations, nous obtenons alors le tableau suivant :

$$\left(\begin{array}{ccc|ccc} 1 & 3 & -1 & 1 & 0 & 0 \\ 0 & -10 & 4 & -3 & 1 & 0 \\ 0 & 7 & -5 & 2 & 0 & 1 \end{array}\right)$$

Continuons l'agorithme de Gauss pour, cette fois-ci, rendre nul le terme a_{32}. Cela correspond à l'opération

$$L_3 \leftarrow 10L_3 + 7L_2.$$

Nous obtenons donc

$$\left(\begin{array}{ccc|ccc} 1 & 3 & -1 & 1 & 0 & 0 \\ 0 & -10 & 4 & -3 & 1 & 0 \\ 0 & 0 & -22 & -1 & 7 & 1 \end{array}\right)$$

L'étape suivante est alors de renormalisée L_3 de telle sorte que $a_{33} = 1$. C'est-à-dire

$$L_3 \leftarrow -\frac{1}{22} \times L_3.$$

$$\left(\begin{array}{ccc|ccc} 1 & 3 & -1 & 1 & 0 & 0 \\ 0 & -10 & 4 & -3 & 1 & 0 \\ 0 & 0 & 1 & -\frac{1}{22} & -\frac{7}{22} & -\frac{1}{22} \end{array}\right)$$

On peut maintenant, en quelque sorte, remonter l'algorithme de Gauss pour annuler les coefficients a_{23} et a_{13}. Cela correspond aux opérations suivantes :

$$L_2 \leftarrow L_2 - 4L_3 \qquad L_1 \leftarrow L_1 + L_3.$$

Nous obtenons donc

$$\left(\begin{array}{ccc|ccc} 1 & 3 & 0 & \frac{23}{22} & -\frac{7}{22} & -\frac{1}{22} \\ 0 & -10 & 0 & -\frac{70}{22} & \frac{50}{22} & \frac{4}{22} \\ 0 & 0 & 1 & -\frac{1}{22} & -\frac{7}{22} & -\frac{1}{22} \end{array}\right)$$

Il ne reste plus qu'à renormaliser la deuxième ligne pour avoir $a_{22} = 1$ puis de répéter les opérations pour annuler le terme a_{12} et renormaliser la première ligne pour avoir $a_{11} = 1$. En poursuivant les calculs, nous obtenons quelque chose de la forme

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & a & b & c \\ 0 & 1 & 0 & d & e & f \\ 0 & 0 & 1 & i & j & k \end{array}\right)$$

où $a, b, c, d, e, f, i, j, k$ sont des réels. La matrice composée des ces nombres correspond alors à l'inverse de la matrice A. Ici, nous devrions obtenir :

$$A^{-1} = \frac{1}{22}\begin{pmatrix} 2 & 8 & 2 \\ 7 & -5 & -4 \\ 1 & -7 & -10 \end{pmatrix}$$

Remarque. Dans l'exemple précédent nous avons pu mener à bien l'algorithme de Gauss et obtenir l'inverse de la matrice A. On dit alors que la rang de la matrice A vaut 3 :

$$\mathrm{rg}(A) = 3$$

Si $A \in M_{n\times p}(\mathbb{R})$ alors $0 \le \mathrm{rg}(A) \le n$ (dans l'exemple précédent $n = 3$). Bien entendu, il se peut que $\mathrm{rg}(A) < n$, ce phénomène s'observe lorsque l'on met en place l'algorithme du pivot de Gauss et qu'une des opérations sur les lignes aboutie à une ligne entièrement nulle. Le rang de la matrice correspond alors aux nombres de lignes non-nulles.

Etudions quelques exemples pour éclaircir ceci :

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

Tout d'abord $A \in M_{2\times 2}(R)$, donc $0 \leq rg(A) \leq 2$. Mettons en place l'algorithme du pivot de Gauss pour échelonner la matrice :

$$L_2 \leftarrow L_2 + L_1 \qquad \text{fournit} \qquad \begin{pmatrix} 1 & 3 \\ 0 & 3 \end{pmatrix}$$

Il est alors évident que l'on peut poursuivre l'algorithme du pivot de Gauss sans obtenir de ligne complètement nulle, alors $rg(A) = 2$.

(2) Etudions la matrice
$$B = \begin{pmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{pmatrix}$$

Puisque $B \in M_{3\times 3}(\mathbb{R})$ nous avons que $0 \leq rg(B) \leq 3$. Procédons à l'algorithme de Gauss :

$$L_2 \leftarrow L_2 - L1 \qquad L_3 \leftarrow L_3 - L2$$

fournissent

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

Nous ne pouvons poursuivre l'algorithme jusqu'au bout. Donc, d'une part la matrice B est non inversible et d'autres part, puisqu'il reste deux lignes non nulles (après avoir effectuer le pivot de Gauss), on en déduit que $rg(B) = 2$.

(3) Et enfin, un dernier exemple.
$$C = \begin{pmatrix} 1 & 3 & 0 \\ -1 & 0 & 3 \\ 2 & 6 & 0 \\ 0 & 3 & 3 \end{pmatrix}$$

Remarquons que $C \in M_{4,3}(\mathbb{R})$. Ainsi, puisqu'il ne s'agit pas d'une matrice carrée, la matrice C ne peut-être inversible. De plus, nous savons que $0 \leq rg(C) \leq 4$. Bien qu'elle ne soit pas inversible, il est possible de mettre en place l'algorithme de Gauss pour déterminer son rang : les opérations suivantes

$$L_2 \leftarrow L_2 + L1, \qquad L_3 \leftarrow L_3 - 2L_1, \qquad L_4 \leftarrow L_4 - L_2$$

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

C'est pourquoi rg(C) = 2. Bien que nous n'explorerons pas ceci plus avant, la notion de rang d'une matrice est très importante et apparait naturellement dans la résolution de certains problèmes statistiques.

Nous verrons en exercice de quelle manière résoudre un système linéaire en inversant une matrice.

5.2. **Notions de déterminant.** Considérons deux vecteurs du plan :

$$u = (1,2) \quad v = (1,0)$$

quel est l'aire du parallélogramme engendré par les vecteurs u et v ? Une manière de répondre à cette question est la notion de déterminant. Ce concept mathématique étant assez difficile, nous allons nous concentrer sur l'aspect calculatoire de celui-ci afin d'obtenir une condition nécéssaire et suffisante à l'inversibilité d'une matrice. La notion de déterminant permettra aussi d'instaurer les bases permettant d'aborder la diagonalisation d'une matrice.

Définition 5.2. *Soit* $\mathrm{A} \in \mathrm{M}_{2\times 2}(\mathbb{R})$*, autrement dit*

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

alors $\det(\mathrm{A}) = ad - bc$.

Remarque. (1) Si $\det(\mathrm{A}) = 0$ cela signifie que les vecteurs $u = (a,c)$ et $v = (b,d)$ sont colinéaires (en fait il s'agit d'une équivalence) et donc l'aire du parallélogramme engendré est nulle.

(2) Si $\det(\mathrm{A}) \neq 0$, nous avons une expression simple de A^{-1} faisant intervenir $\det(\mathrm{A})$:

$$\mathrm{A}^{-1} = \frac{1}{\det(\mathrm{A})} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Cette formule se généralise (et se complexifie) pour des matrices $\mathrm{A} \in \mathrm{M}_{n\times n}(\mathrm{R})$, nous ne l'évoquerons pas. Cependant l'assertion suivante reste vraie :

$$\mathrm{A} \quad \text{est inversible} \quad \Longleftrightarrow \quad \det(\mathrm{A}) \neq 0 \quad \Longleftrightarrow \quad \mathrm{rg}(\mathrm{A}) = n$$

La notion de déterminant étant assez complexe, nous n'entrerons pas dans les détails théorique. Il sera plus simple d'illustrer les propriétés au travers plusieurs exemples en dimension 3. La généralisation aux dimensions supérieur est assez simple à imaginer.

Exemple 5.3. La façon de repreésenter un determinant est assez similaire à la représentation matricielle :

Si A est la matrice suivante :

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

alors on notera son determinant comme suit :

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

Remarque. A nouveau, le réel obtenu en calculant det(A) correspondra au volume du parallélépipède engendré par les vecteurs $v_1 = (1,1,0)$, $v_2 = (2,4,2)$, $v_3 = (-3,-5,-2)$.

Exemple 5.4. Une propriété remarquable du déterminant est qu'il est linéaire en chacune de ses colonnes et chacunes de ses lignes. Autrement dit, si l'on procède aux opérations suivantes :

$$C_2 \leftarrow C_2 - 2C_1 \qquad \text{ou} \qquad C_2 \leftarrow \frac{1}{2}C_2 \qquad \text{ou} \qquad L1 \leftarrow L_1 - L_2$$

on ne modifie pas la valeur du déterminant. Dans les opérations précédentes, nous aurions obtenu (respectivement)

$$\det(A) = \begin{vmatrix} 1 & 0 & -3 \\ 1 & 2 & -5 \\ 0 & 2 & -2 \end{vmatrix} \qquad \text{ou} \qquad \det(A) = \begin{vmatrix} 1 & 1 & -3 \\ 1 & 2 & -5 \\ 0 & 1 & -2 \end{vmatrix} \qquad \text{ou} \qquad \det(A) = \begin{vmatrix} 0 & -2 & 2 \\ 1 & 2 & -5 \\ 0 & 2 & -2 \end{vmatrix}$$

Exemple 5.5. Un des moyen le plus pratique pour calculer un determinant est de développer celui-ci par rapport à une ligne ou une colonne pour se ramener à des determinants de taille 2×2 que l'on sait calculer (cf. définition 5.2) . Pour faire ceci, il est important d'avoir en tête les signes attribués aux élements contenus dans le déterminant :

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

Supposons qu'on souhaite développer le det(A) par rapport à la première colonne, nous obtiendrons alors

$$\begin{aligned} \det(A) = \begin{vmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{vmatrix} &= +1 \times \begin{vmatrix} 4 & -5 \\ 2 & -2 \end{vmatrix} -1 \times \begin{vmatrix} 2 & -3 \\ 2 & -2 \end{vmatrix} +0 \times \begin{vmatrix} 2 & -3 \\ 4 & -5 \end{vmatrix} \\ &= +1 \times \begin{vmatrix} 4 & -5 \\ 2 & -2 \end{vmatrix} -1 \times \begin{vmatrix} 2 & -3 \\ 2 & -2 \end{vmatrix} \\ &= \left[4 \times (-2) + 5 \times 2\right] - \left[2 \times (-2) + 3 \times 2\right] \\ &= 0 \end{aligned}$$

On pourrait également développer le déterminant de A par rapport à la première ligne. Voyons ce que nous obtenons :

$$\begin{aligned} \det(A) = \begin{vmatrix} 1 & 2 & -3 \\ 1 & 4 & -5 \\ 0 & 2 & -2 \end{vmatrix} &= +1 \times \begin{vmatrix} 4 & -5 \\ 2 & -2 \end{vmatrix} -2 \times \begin{vmatrix} 1 & -5 \\ 0 & -2 \end{vmatrix} -3 \times \begin{vmatrix} 1 & 4 \\ 0 & 2 \end{vmatrix} \\ &= [-8+10] - 2 \times [-2-0] - 3 \times [2-0] \\ &= 0 \end{aligned}$$

Remarque. Faisons quelques commentaires sur ce calcul. D'une part, il est rassurant de voir que l'on obtient le même résultat (que lorsqu'on avait développer par rapport à la première colonne). D'autre part, on comprend pourquoi les opérations sur les lignes et colonnes peuvent s'avérer utile pour annuler certains termes du déterminant et faciliter le développement de celui-ci.

Intuitivement, on comprend aussi quel déterminant (d'ordre inférieur) on doit conserver lorsque l'on procède au développement de celui-ci. En reprenant, le premier calcul nous avons obtenu le premier terme du développement de la manière suivante :

$$\begin{vmatrix} 1 & \cancel{2} & \cancel{-3} \\ \cancel{1} & 4 & -5 \\ \cancel{0} & 2 & -2 \end{vmatrix} = 1 \times \begin{vmatrix} 4 & -5 \\ 2 & -2 \end{vmatrix}$$

où l'on aurait « rayer » les coefficients se trouvant la même ligne et colonne que le terme suivant lequel on souhaite développer. Pour le second terme, nous avons obtenu

$$\begin{vmatrix} \cancel{1} & 2 & -3 \\ 1 & \cancel{4} & \cancel{-5} \\ \cancel{0} & 2 & -2 \end{vmatrix} = -1 \times \begin{vmatrix} 2 & -3 \\ 2 & -2 \end{vmatrix}$$

et enfin, pour le dernier terme de la colonne,

$$\begin{vmatrix} \cancel{1} & 2 & -3 \\ \cancel{1} & 4 & -5 \\ 0 & \cancel{2} & \cancel{-2} \end{vmatrix} = 0 \times \begin{vmatrix} 2 & -3 \\ 4 & -5 \end{vmatrix}$$

Il est important de ne pas oublier la règle des signes lorsque l'on effectue ce genre de développement. Si l'on considère une matrice de taille supérieur (4×4 par exemple), il est possible de procéder similaire pour se ramener à des déterminants de taille 3×3. Voyons plutôt, en développant par rapport à la première ligne

$$\begin{vmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & -2 & 1 \\ 4 & 0 & 1 & -1 \\ 7 & 2 & 1 & 0 \end{vmatrix} = 1 \times \begin{vmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 2 & 1 & 0 \end{vmatrix} - 2 \times \begin{vmatrix} 0 & -2 & 1 \\ 4 & 1 & -1 \\ 7 & 1 & 0 \end{vmatrix} + (-1) \times \begin{vmatrix} 0 & 1 & 1 \\ 4 & 0 & -1 \\ 7 & 2 & 0 \end{vmatrix} - 0 \times \begin{vmatrix} 0 & 1 & -2 \\ 4 & 0 & -1 \\ 7 & 2 & 1 \end{vmatrix}$$

où nous avons utiliser la règle de signe suivante

$$\begin{vmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{vmatrix}$$

Il est alors aisé de terminer le calcul en calculant les déterminant de taille 3×3.

6. Diagonalisation de matrice

Considérons la matrice suivante :

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

et imaginons un court instant que cette matrice et ses coefficient représentent de jeunes mariés accompagnés de leurs proches. Puisque nous souhaitons étudier cette matrice, nous pourrions nous assimiler au photographe présent au marriage. Il apparait donc désirable de chercher le meilleur angle de vue (s'il existe) permettant d'immortaliser au mieux un tel moment. Précisons un peu plus l'analogie précédente : est-il possible d'observer la matrice sous un angle différent afin d'observer uniquement les éléments essentiels de celle-ci (les marriés et leurs témoins par exemple, tout en évitant la présence d'objets superflus dans le champ de vision) ?

Soient P et son inverse P^{-1} les matrices suivantes :

$$\mathrm{P} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 1 \end{pmatrix}, \qquad \mathrm{P}^{-1} = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

Il se trouve que le produit $\mathrm{A}' = \mathrm{P}^{-1}\mathrm{AP}$ permet d'exprimer la matrice A sous un autre angle et la matrice A' ainsi obtenue est beaucoup plus simple :

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

Nous venons de procéder à un changement de base (selon l'analogie précédente, cela revient à trouver un meilleur point de vue pour observer la matrice A) et les nombres apparaissant sur la diagonale de la matrice A' s'appellent les valeurs propres de A (celles-ci correspondent aux informations essentielles contenues dans la matrice).

Il est alors naturel de s'interroger : comment déterminer, de manière algorithmique, les valeurs propres d'une matrice ? Est-il toujours possible de trouver des matrices $\mathrm{P}, \mathrm{P}^{-1}$ telles que $\mathrm{A}' = \mathrm{P}^{-1}\mathrm{AP}$ s'exprime aussi simplement ? Bien entendu, cette question est un peu complexe et la réponse dépend de la matrice A étudiée. Sans rentrer dans des détails trop technique, voyons comment nous pouvons procéder dans des situations favorables. Nous informons le lecteur que les définitions qui vont suivre ne sont pas toutes standards et ne sont uniquement le reflet d'une contorsion pédagogique permettant d'éviter trop de théorie.

Définition 6.1. *Soient* $\mathrm{A} \in \mathrm{M}_{n\times n}(\mathbb{R})$ *une matrice et* $\lambda \in \mathbb{R}$*. Le polynôme caractéristique* χ_{A} *associé à la matrice* A *est défini par la formule suivante :*

$$\chi_{\mathrm{A}}(\lambda) = \det(\mathrm{A} - \lambda \mathrm{I}_d).$$

Il s'agit d'un polynôme de degré n en λ *et ses racines sont appelées valeurs propres de* A*. L'ensemble des valeurs propres d'une matrice* A *est appelé spectre de la matrice, on le notera par* spec(A).

Reprenons l'exemple de l'introduction.

Exemple 6.1.

$$\det(\mathrm{A} - \lambda \mathrm{I}_d) = \begin{vmatrix} 1-\lambda & 2 & -3 \\ 1 & 4-\lambda & -5 \\ 0 & 2 & -2-\lambda \end{vmatrix}$$

Développons ce déterminant par rapport à la première colonne. C'est-à-dire

$$\begin{aligned} \begin{vmatrix} 1-\lambda & 2 & -3 \\ 1 & 4-\lambda & -5 \\ 0 & 2 & -2-\lambda \end{vmatrix} &= (1-\lambda)\begin{vmatrix} 4-\lambda & -5 \\ 2 & -2-\lambda \end{vmatrix} - \begin{vmatrix} 2 & -3 \\ 2 & -2-\lambda \end{vmatrix} \\ &= (1-\lambda)\big[(4-\lambda)(-2-\lambda)+10\big] + 2(2+\lambda) - 6 \\ &= (1-\lambda)(\lambda^2 - 2\lambda + 2) + 2(\lambda - 1) \\ &= (1-\lambda)\lambda(\lambda - 2) \end{aligned}$$

Autrement dit, nous avons montré que $\chi_{\mathrm{A}}(\lambda) = (1-\lambda)\lambda(\lambda-2)$ et retrouver le fait que $\mathrm{spec(A)} = \{0, 1, 2\}$.

Il reste à trouver un moyen de calculer la matrice P (a partir de celle-ci P^{-1} sera facilement obtenue) à partir du spectre de A : il s'agit de déterminer les espaces propres associés à une valeur propre donnée.

Définition 6.2. *Soit* $A \in M_{n\times n}(\mathbb{R})$ *et considérons une valeur propre* λ *de cette matrice. L'ensemble des solutions* E_λ *du système suivant est appelé espace propre associé à la valeur propre* λ *:*

$$(A - \lambda I_d)X = 0,$$

avec $X = (x_1, \dots, x_n)$.

Voyons plutôt ce que l'on obtient pour notre exemple introductif :

Exemple 6.2. Débutons par la valeur propre $\lambda = 1$. Il faut donc résoudre le système suivant, avec $X = (x, y, z)$,

$$(A - I_d)(X) = 0 \qquad \Longleftrightarrow \qquad S_1 : \begin{cases} 2y - 3z = 0 \\ x + 3y - 5z = 0 \\ 2y - 3z = 0 \end{cases}$$

Observons que $L_1 = L_3$, le système S_1 est donc équivalent au système suivant et il est alors possible d'exprimer x et y en fonction de z. C'est à dire :

$$S_1 \qquad \Longleftrightarrow \qquad \begin{cases} x + 3y - 5z = 0 \\ 2y - 3z = 0 \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} x = \frac{1}{2} z \\ y = \frac{3}{2} z \end{cases}$$

Autrement dit : $E_1 = \left\{(x, y, z) \in \mathbb{R}^3 \text{ solution de } S_1\right\} = \left\{\left(\frac{1}{2}z, \frac{3}{2}z, z\right), \quad z \in \mathbb{R}\right\}$. Par exemple, pour le choix $z = 2$, le vecteur $v_2 = (1, 3, 2) \in E_1$.

Il est possible de procéder de même pour $\lambda = 0$ et $\lambda = -2$, on trouve les espaces propres suivants :

$$E_0 = \left\{(z, z, z), \quad z \in \mathbb{R}\right\} \qquad \text{et} \qquad v_1 = (1, 1, 1) \in E_0$$

et

$$E_2 = \left\{(z, 2z, z), \quad z \in \mathbb{R}\right\} \qquad \text{et} \qquad v_3 = (1, 2, 1) \in E_2$$

Remarquons alors que la matrice composée des vecteurs (mis en colonne) v_1, v_2, v_3 est exactement la matrice P (qui servira à calculer P^{-1}.

Résumons ce qui précède sous la forme d'une théorème.

Théorème 1. *Soit* $A \in M_{n\times n}(\mathbb{R})$ *telle que* $\chi_A(\lambda)$ *admette n racines rélles distinctes* $\lambda_1, \dots, \lambda_n$. *Alors, il existe une matrice* P *inversible telle que*

$$A' = P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & \lambda_n \end{pmatrix}$$

Remarque. Rappelons que la matrice P est obtenue en placcant successivement en colonne un vecteur (non nul) de chaque espace propre.