UE PIN401 - Initiation à l'algèbre linéaire, Probabilités-Statistique

Alain Yger

9 mai 2012

Table des matières

Chapitre 1

Initiation à l'algèbre linéaire

1.1 Introduction heuristique : espaces vectoriels, bases, linéarité, matrices

1.1.1 L'espace $\mathbb{R}^N$

Pour introduire la notion d'espace vectoriel (et dépasser les exemples géométriques standard que sont le plan $\mathbb{R}^2$, l'espace $\mathbb{R}^3$ de la mécanique Newtonienne ou l'espace-temps $\mathbb{R}^4$ de la mécanique relativiste, modèles classiques sur lesquels bien sûr nous reviendrons), nous prendrons comme premier exemple une expérience physique répétée N fois et fournissant, à chacune de ses occurrences, un résultat x_k $(k = 1, ..., N)$ qui est un nombre réel. L'ensemble de tous les résultats possibles (de cette suite de N expériences) est l'ensemble

$$\{(x_1, ..., x_N)\,;\, x_j \in \mathbb{R}\,,\ j = 1, ..., N\}$$

et dépend donc de N "degrés de liberté". Cet ensemble est noté

$$\mathbb{R}^N = \mathbb{R} \times \cdots \times \mathbb{R} \quad (N \text{ fois})\,.$$

Bien souvent, on est amené à imposer des contraintes à ces diverses expériences : par exemple, si x_k représente la variation (positive ou négative) d'une certaine action boursière pendant l'année k, l'exigence

$$x_1 + x_2 + \cdots + x_N = 0 \tag{1.1}$$

est une exigence raisonnable (si l'on veut ne prendre aucun rique) ; mais, sous cette contrainte, on voit que l'ensemble des résultats possibles ne dépend plus que de $N - 1$ degrés de liberté car la relation (1.1) s'exprime aussi sous la forme

$$x_N = -(x_1 + x_2 + \cdots + x_{N-1})\,,$$

ce qui implique qu'une fois que les $N - 1$ degrés de liberté correspondant aux valeurs prises par $x_1, ..., x_{N-1}$ sont figés, celui correspondant au résultat de la N-ème expérience l'est automatiquement.

On peut changer notre fusil d'épaule et imaginer que $x_1, ..., x_N$ sont N paramètres réels sur lesquels repose la conception d'un robot. Supposons que l'objectif de la conception du robot soit le suivant : étant données N corps ponctuels de masses

respectives $m_1, ..., m_N$ et une poutre homogène en équilibre horizontal sur un pivot placé en son centre (l'origine, voir la figure 1.1 ci-dessous), les paramètres du robot sont les distances (algébriques) d_i auxquelles il faut poser ces corps sur la poutre pour que l'équilibre soit préservé ($x_k = d_k$, $k = 1, ..., N$), c'est-à-dire ici le centre de gravité de l'ensemble poutre + masses situé à l'origine.

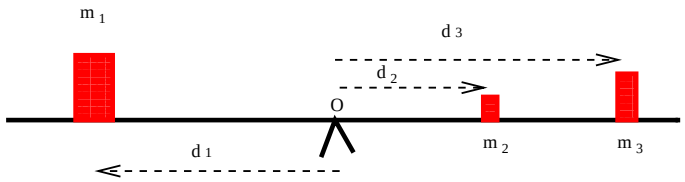

FIGURE 1.1 – Les masses en équilibre

Cette fois, la contrainte liant les paramètres x_i est

$$m_1 x_1 + \cdots + m_N x_N = 0$$

et l'on voit à nouveau que le choix des paramètres ne dépend plus que de $N - 1$ degrés de lberté car

$$x_N = -\frac{m_1}{m_N} x_1 - \cdots - \frac{m_{N-1}}{m_N} x_{N-1} .$$

Il existe plusieurs manières de "visualiser" un élément de cet ensemble $\mathbb{R}^N$; la première consiste à afficher les valeurs $x_1, ..., x_N$ (fonction de l'indice k variant de 1 à N) sur un graphe, comme indiqué sur la figure ci-dessous (ici on a pris $N = 20$) :

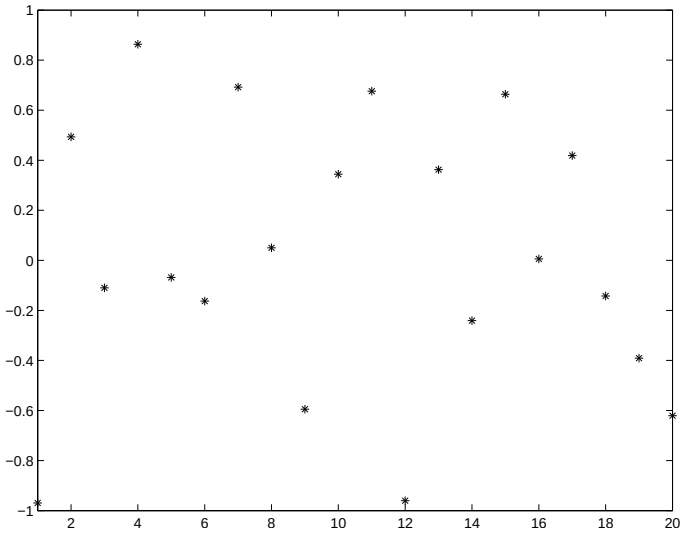

FIGURE 1.2 – La représentation graphique d'un élément de $\mathbb{R}^N$

La représentation de l'étoile au dessus de l'abscisse k correspond à la représentation de l'élément de $\mathbb{R}^N$

$$(0, ...0, x_k, 0, ..., 0) ,$$

où x_k figure en k-ème position ; cet élément peut aussi s'interpréter comme x_k fois l'élément de référence

$$e_k = (0, ..., 0, 1, 0, ..., 0)$$

où 1 figure ici en k-ème position, ce que l'on écrit encore

$$(0, ..., 0, x_k, 0, ..., 0) = x_k \cdot (0, ..., 0, 1, 0, ..., 0)\,,$$

ce qui permet au bilan final pour $(x_1, ..., x_N)$ l'écriture "éclatée" :

$$(x_1, x_2, ..., x_N) = x_1 \cdot e_1 + x_2 \cdot e_2 + \cdots + x_{N-1} \cdot e_{N-1} + x_N \cdot e_N$$

dont nous reparlerons très vite ; on vient de voir apparaître deux opérations :

– une addition "interne" (notée $+$) entre éléments de $\mathbb{R}^N$, consistant à additionner deux éléments comme suit :

$$(x_1, ..., x_k, ..., x_N) + (y_1, ..., y_k, ..., y_N) := (x_1 + y_1, ..., x_k + y_k, ..., x_N + y_N)\,;$$

– une action "externe" (notée $\cdot$) de $\mathbb{R}$ sur $\mathbb{R}^N$ définie par

$$\lambda \cdot (x_1, ..., x_k, ..., x_N) := (\lambda \times x_1, ..., \lambda \times x_k, ..., \lambda \times x_N)\,.$$

On verra que ce sont ces deux opérations (que nos modèles répercutant la notion de contrainte dans l'analyse de la liste des résultats de N expériences ou dans la conception d'un robot dépendant de N paramètres faisaient manifestement intervenir) qui président au concept de structure de $\mathbb{R}$-*espace vectoriel*.

Ce que montre un écran d'ordinateur si on lui demande d'afficher un élément donné $(x_1, ..., x_N)$ de $\mathbb{R}^N$ est de nature tout autre ; on voit sur l'écran, pour le même élément de $\mathbb{R}^20$ que précédemment) l'image reproduite sur la figure 1.3 ci-dessous.

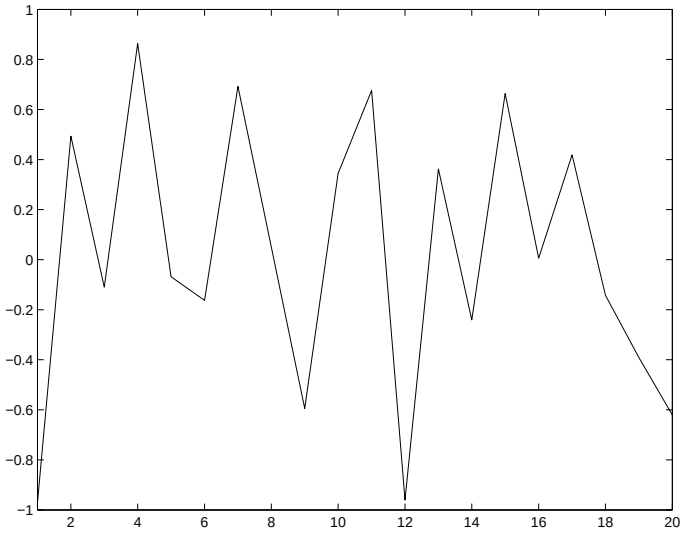

FIGURE 1.3 – Une autre visualisation graphique du même élément de $\mathbb{R}^N$

Ce que nous montre l'écran est ici non plus la suite des valeurs $(x_1, ..., x_N)$, mais en fait le graphe de l'unique fonction $f_{x_1,...,x_N} : [1, N] \longrightarrow \mathbb{R}$, linéaire par morceaux

avec noeuds aux N points $1, 2, ..., N$, qui interpole exactement la valeur x_k au point k :

$$f_{x_1,...,x_N}(k) = x_k \, .$$

Nous voyons ici apparaître un nouvel ensemble, qui n'est plus $\mathbb{R}^N$ mais qui est l'ensemble F_N des fonctions $f \ : [1, N] \longrightarrow \mathbb{R}$, affines par morceaux avec noeuds aux points $1, 2, ..., N$, ce qui signifie qu'une fois que des valeurs $f(k)$, $k = 1, ..., N$, ont été assignées aux points k, $k = 1, ..., N$, on construit le graphe en joignant par des segments les points $(k, f(k))$, $k = 1, ..., N$. Cet ensemble F_N est en correspondance bijective avec $\mathbb{R}^N$. On peut encore l'équiper de deux opérations :

- une addition "interne" (notée $+$) entre éléments de F_N, consistant à additionner deux fonctions $f \in F_N$ et $g \in F_N$ comme suit :

$$f + g \ : t \in [1, N] \longmapsto f(t) + g(t) \, ;$$

- une action "externe" (notée $\cdot$) de $\mathbb{R}$ sur F_N définie par

$$\lambda \cdot f \ : t \in [1, N] \longmapsto \lambda \times f(t) \, .$$

Parmi les éléments de F_N, on trouve les fonctions Δ_k, $k = 1, ..., N$, définies par

$$\Delta_k \ : t \in [1, N] \longmapsto \max(0, 1 - |t - k|) \, , \quad k = 1, ..., N \, ,$$

dont nous avons représenté les graphes sur la figure 1.4 pour $N = 5$ (avec différentes couleurs).

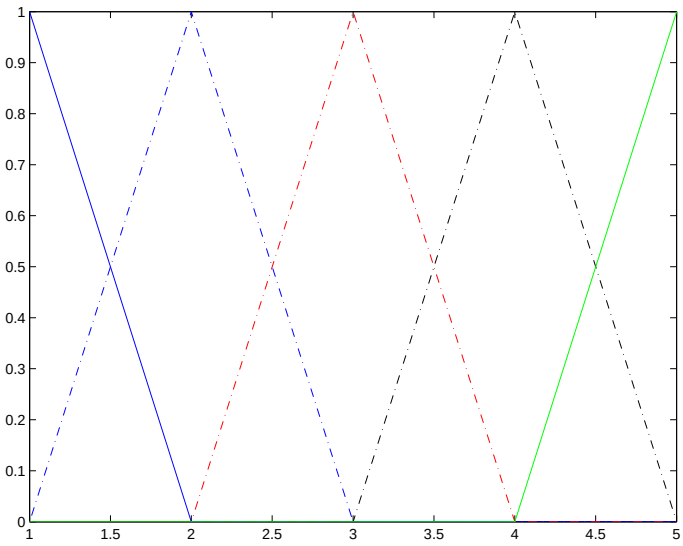

FIGURE 1.4 – Les fonctions Δ_k, $k = 1, ..., N$

Cette fois, on voit que l'on peut écrire (en faisant apparaître l'addition interne et l'action externe de $\mathbb{R}$ sur F_N) l'élément $f_{x_1,...,x_N}$ de F_N sous la forme

$$f_{x_1,...,x_N} = \sum_{k=1}^{N} x_k \cdot \Delta_k \, ;$$

en effet, pour chaque valeur $k = 1, ..., N$, on vérifie aisément que les deux fonctions ci-dessus coïncident et valent toutes les deux x_k ; comme ce sont toutes les deux des fonctions dont le graphe est obtenu en joignant par des segments les points (k, x_k), $k = 1, ..., N$, elles coïncident partout sur $[1, N]$.

1.1.2 Les espaces $\mathbb{R}^{n\times p}$

Lorsque $N = n \times p$, il existe une autre manière de représenter les éléments de $\mathbb{R}^N$, cette fois sous forme d'un tableau à n lignes et p colonnes de nombres réels. Un tel tableau est ce que l'on appellera une *matrice* et jouera pour nous un rôle crucial. On le note $[a_{i,j}]_{\substack{1\le i\le n\\1\le j\le p}}$ si le nombre réel $a_{i,j}$ représente l'entrée du tableau au carrefour de la i-ème ligne et de la j-ème colonne (c'est une convention, que l'on retrouve dans les logiciels de calcul scientifique classiques tels MATLAB ou SCILAB).

L'ensemble $\mathcal{M}_{n,p}(\mathbb{R})$ des tableaux de nombres réels à n lignes et p colonnes hérite ici encore de deux opérations clef :

- une addition (opération interne dans l'ensemble des tableaux à n lignes et p colonnes) définie par le fait que la somme de deux tableaux $A = [a_{i,j}]_{\substack{1\le i\le n\\1\le j\le p}}$ et $B = [b_{i,j}]_{\substack{1\le i\le n\\1\le j\le p}}$ à n lignes et p colonnes est le tableau

$$A + B := \Big[a_{i,j} + b_{i,j}\Big]_{\substack{1\le i\le n\\1\le j\le p}} ;$$

- une action externe de $\mathbb{R}$ sur $\mathcal{M}_{n,p}(\mathbb{R})$ définie par

$$\lambda \cdot [a_{i,j}]_{\substack{1\le i\le n\\1\le j\le p}} := [\lambda \cdot a_{i,j}]_{\substack{1\le i\le n\\1\le j\le p}}$$

  pour tout nombre réel λ et toute matrice A de $\mathcal{M}_{n,p}(\mathbb{R})$.

Si $I_{i,j}$ est le tableau à n lignes et p colonnes présentant un unique coefficient non nul (et égal à 1) au carrefour de la i-ème ligne et de la j-ème colonne, nous pouvons alors écrire

$$[a_{i,j}]_{\substack{1\le i\le n\\1\le j\le p}} = \sum_{i=1}^{n}\sum_{j=1}^{p} a_{i,j}\cdot I_{i,j}$$

et les matrices $I_{i,j}$, $1 \le i \le n$, $1 \le j \le p$ apparaîssent comme les "briques de lego" de base permettant d'exprimer (*via* les deux opérations) toute matrice à n lignes et p colonnes.

Bien sûr un tableau de nombres réels (par exemple, ici un tableau à 5 lignes et 10 colonnes) peut se "visualiser" directement sous la forme :

$$\begin{pmatrix}
0.950 & 0.762 & 0.615 & 0.405 & 0.058 & 0.203 & 0.015 & 0.419 & 0.838 & 0.503\\
0.231 & 0.456 & 0.792 & 0.935 & 0.353 & 0.199 & 0.747 & 0.846 & 0.020 & 0.709\\
0.607 & 0.018 & 0.922 & 0.917 & 0.813 & 0.604 & 0.445 & 0.525 & 0.681 & 0.429\\
0.486 & 0.821 & 0.738 & 0.410 & 0.010 & 0.272 & 0.932 & 0.203 & 0.379 & 0.305\\
0.891 & 0.445 & 0.176 & 0.894 & 0.139 & 0.199 & 0.466 & 0.672 & 0.832 & 0.190
\end{pmatrix}$$

Cette fois encore, il existe bien d'autres manières de "visualiser" une telle matrice.

Copiant ce que nous avons fait pour les suites de N nombres réels et le transposant au cadre des tableaux à n lignes et p colonnes, nous pouvons représenter un tel tableau de nombres réels

$$A = [a_{i,j}]_{\substack{i=1,\dots,n\\j=1,\dots,p}}$$

en visualisant par exemple le graphe (au dessus de $[1,n]\times[1,p]$) de l'unique fonction $f_A : [1,n]\times[1,p] \longmapsto \mathbb{R}$ telle que

$$f_A(i,j) = a_{i,j}\,, \quad 1 \le i \le n\,,\ 1 \le j \le p\,,$$

et que l'image de $[1,n]\times[1,p]$ soit l'unique surface polyédrale obtenue à partir des arêtes joignant chaque point (i,j) aux quatre points

$$(i-1,j),(i+1,j),(i,j-1),(i,j+1)$$

(on ignore parmi ces quatre points ceux qui sortent des limites du tableau $[1,n]\times[1,p]$). Voici, par exemple, une représentation sur ce modèle du tableau à 5 lignes et 10 colonnes proposé ci-dessus.

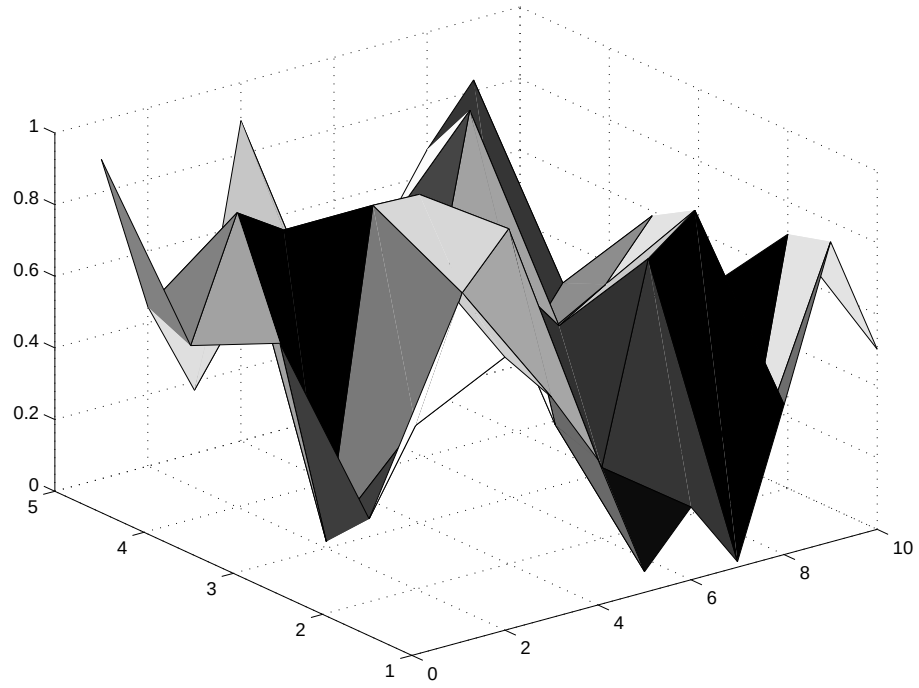

FIGURE 1.5 – Une visualisation 3D d'une matrice à 5 lignes et 10 colonnes

Il est clair que le tableau est ici identifié à une fonction F s'écrivant sous la forme

$$F\ :(x,y)\in[1,n]\times[1,p]\longmapsto\sum_{i=1}^{n}\sum_{j=1}^{p}a_{i,j}\cdot\Delta_{i,j}$$

où

$$\Delta_{i,j}(x,y)=(1-\max(|x-i|,0))\times(1-\max(|y-j|,0))\,;$$

on a ainsi représenté l'image comme élément d'un certain ensemble de fonctions : $F\ :[1,n]\times[1,p]\longrightarrow\mathbb{R}$ (ensemble noté $F_{n,p}$) constitué des fonctions de $[1,n]\times[1,p]$ dans $\mathbb{R}$ dont le graphe est polyédral par morceaux avec noeuds précisément aux points (i,j), $1\leq i\leq n$, $1\leq j\leq p$. Les "atomes" $\Delta_{i,j}$ sont maintenant les "pierres de base" d'une telle représentation et la fonction représentant le tableau $A:[a_{i,j}]_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ est alors la fonction

$$F_A\ :(x,y)\in[1,n]\times[1,p]\longmapsto\sum_{i=1}^{n}\sum_{j=1}^{p}a_{i,j}\,\Delta_{i,j}(x,y)\,.$$

ou encore, plus synthétiquement

$$F_A=\sum_{j=1}^{p}a_{i,j}\cdot\Delta_{i,j}\,;$$

notons que l'on fait encore apparaître ici une addition (opération interne) dans l'ensemble $F_{n,p}$ ainsi qu'une action externe de $\mathbb{R}$ sur cet ensemble.

Il existe d'autres modes de visualisation d'une image $[a_{i,j}]_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ conduisant à d'autres types d'ensembles équipés encore d'une addition et d'une action externe de $\mathbb{R}$ dessus. Par exemple, la représentation de la matrice envisagée ci-dessus sous la forme d'une image $U = U_A$ avec différentes nuances de gris, comme sur la figure suivante :

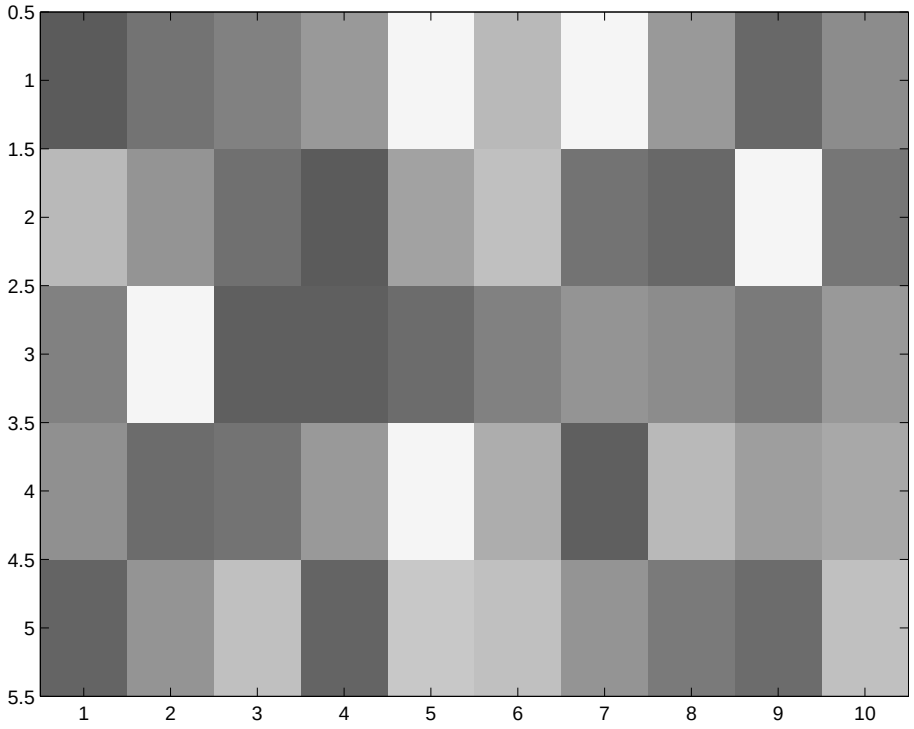

FIGURE 1.6 – Une visualisation sous forme d'image de la même matrice à 5 lignes et 10 colonnes

Ici encore, il y a un système de $n \times p$ images "type", notées $U_{i,j}$, $1 \leq i \leq n$, $1 \leq j \leq p$, où $U_{i,j}$ est la fonction de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ prenant la valeur 1 sur le carré fondamental de côté 1 centré au point étiqueté (i, j) ; la matrice $A = [a_{i,j}]_{i,j}$ se visualise alors sous la forme de l'image $U - A$

$$U_A = \sum_{i=1}^{n} \sum_{j=1}^{p} a_{i,j} \cdot U_{i,j}$$

et les nombres $a_{i,j}$ sont traduits en "intensité de brilliance" pondérant les images "de base" $U_{i,j}$.

1.1.3 Applications de $\mathbb{R}^p$ dans $\mathbb{R}^n$; notion de linéarité

Parmi les exigences naturelles que l'on peut imposer à un système d'appareils $\mathcal{S}$ consommant en entrée un élément de $\mathbb{R}^p$ (c'est-à-dire une suite ordonnée de p nombres réels) pour donner en sortie un élément de $\mathbb{R}^n$ (c'est-à-dire une suite ordonnée de n nombres réels), l'exigence de linéarité est la plus naturelle du point de vue physique.

Elle consiste à supposer que le système d'appareils réagit à la combinaison

$$\lambda \cdot (x_1, ..., x_p) + \mu \cdot (y_1, ..., y_p)$$

de deux entrées $(x_1, ..., x_p)$ et $(y_1, ..., y_p)$, λ et μ étant ici des coefficients réels "pondérateurs", en donnant en sortie la combinaison

$$\lambda \cdot (X_1, ..., X_n) + \mu \cdot (Y_1, ..., Y_n)$$

des deux sorties $(X_1, ..., X_n)$ et $(Y_1, ..., Y_n)$ correspondant aux entrées respectives $(x_1, ..., x_p)$ et $(y_1, ..., y_p)$.

On dit alors que le système $\mathcal{S}$ agit linéairement ou encore que l'application

$$(x_1, ..., x_p) \in \mathbb{R}^p \longmapsto \mathcal{S}\,(x_1, ..., x_p) = (X_1, ..., X_n) \in \mathbb{R}^n$$

est une *application linéaire*.

Il est capital de remarquer que, si $\mathcal{S}$ est un système transformant de manière linéaire les entrées dans $\mathbb{R}^p$ en sorties de $\mathbb{R}^n$, la donnée de toutes les sorties

$$\mathcal{S}(e_k) = (a_{1,k}, ..., a_{n,k})$$

pour $k = 1, ..., p$ suffit à déterminer la sortie correspondant à une entrée arbitraire $(x_1, ..., x_p)$; on s'empresse en effet d'exprimer cette entrée arbitraire $(x_1, ..., x_p)$ suivant les "pierres de base" $e_1, ..., e_p$, sous la forme

$$(x_1, ..., x_p) = x_1 \cdot e_1 + \cdots + x_p \cdot e_p\,.$$

Comme $\mathcal{S}$ agit de manière linéaire, on a

$$\begin{aligned}
\mathcal{S}((x_1, ..., x_p)) &= \mathcal{S}(x_1 \cdot e_1 + \cdots + x_p \cdot e_p) \\
&= x_1 \cdot (\mathcal{S}(e_1)) + \cdots + x_p \cdot (\mathcal{S}(e_p)) \\
&= x_1 \cdot (a_{1,1}\,, ..., a_{n,1}) + \cdots + x_p(a_{1,p}\,, ..., a_{n,p}) \\
&= \Big(\sum_{j=1}^{p} a_{1,j}\,x_j, ..., \sum_{j=1}^{p} a_{n,j}\,x_j\Big)\,.
\end{aligned}$$

Pour respecter l'idée suivant laquelle le premier indice dans la notation $a_{i,j}$ est un indice de ligne, il semble plus judicieux de représenter $\mathcal{S}(e_k)$, $k = 1, ..., p$, sous forme du tableau à n lignes et p colonnes, soit

$$\mathcal{S}(e_k) = \begin{pmatrix} a_{1,k} \\ a_{2,k} \\ \vdots \\ a_{n-1,k} \\ a_{n,k} \end{pmatrix}$$

et donc, par conséquent, réécrire ce qui précède sous la forme

$$\mathcal{S}(x_1 \cdot e_1 + \cdots + x_p \cdot e_p) = \begin{pmatrix} \sum\limits_{i=1}^{p} a_{1,j}\,x_j \\ \sum\limits_{j=1}^{p} a_{2,j}\,x_j \\ \vdots \\ \sum\limits_{j=1}^{p} a_{n-1,j}\,x_j \\ \sum\limits_{j=1}^{p} a_{n,j}\,x_j \end{pmatrix} = \sum_{j=1}^{p} x_j \cdot \begin{pmatrix} a_{1,k} \\ a_{2,k} \\ \vdots \\ a_{n-1,k} \\ a_{n,k} \end{pmatrix}\,.$$

Pour rester cohérents avec nous mêmes, nous noterons également comme des tableaux colonnes les éléments d'entrée (appartenant à $\mathbb{R}^n$) et la relation que l'on vient d'exprimer s'écrit encore

$$\mathcal{S}\left[\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{p-1} \\ x_p \end{pmatrix}\right] = \begin{pmatrix} \sum_{i=1}^{p} a_{1,j}\, x_j \\ \sum_{j=1}^{p} a_{2,j}\, x_j \\ \vdots \\ \sum_{j=1}^{p} a_{n-1,j}\, x_j \\ \sum_{j=1}^{p} a_{n,j}\, x_j \end{pmatrix}$$

Le tableau à n lignes et p colonnes

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p-1} & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p-1} & a_{2,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,p-1} & a_{n-1,p} \\ a_{n,1} & a_{n,2} & \dots & a_{n,p-1} & a_{n,p} \end{pmatrix}$$

qui résume l'action de l'application linéaire $\mathcal{S}$ est dit *matrice* de cette application linéaire relativement aux systèmes d'atomes de base $(e_1, ..., e_p)$ (pour l'ensemble $\mathbb{R}^p$ des entrées) et $(e_1, ..., e_n)$ (pour l'ensemble $\mathbb{R}^n$ des sorties). On remarque que la k-ème colonne de cette matrice correspond à la liste des "poids" $a_{1,k}, ..., a_{n,k}$ affectant $e_1, ..., e_n$ (système de briques de base pour l'ensemble $\mathbb{R}^n$ des sorties) pour réaliser la "décomposition"

$$\mathcal{S}(e_k) = \sum_{i=1}^{n} a_{i,k}\, e_i$$

suivant précisément le système de briques de base $(e_1, ..., e_n)$ dans l'ensemble $\mathbb{R}^n$ des sorties.

1.2 Notion d'espace vectoriel ; systèmes générateurs, systèmes libres, bases

1.2.1 Notion d'espace vectoriel, exemples

Nous allons définir dans cette section les diverses notions concernant les $\mathbb{R}$-espaces vectoriels ; elles se transposent mot pour mot au cas des $\mathbb{C}$-espaces vctoriels ($\mathbb{C}$ étant l'ensemble des nombres complexes), juste en remplaçant partout $\mathbb{R}$ par $\mathbb{C}$.

Pour se donner un $\mathbb{R}$-espace vectoriel, il faut se donner tout d'abord un ensemble E (on appellera ses éléments les *vecteurs*) équipé d'une addition interne (notée $+$) ; cette addition doit satisfaire les 4 propriétés suivantes :

- l'addition doit être *commutative*, c'est-à-dire : $u + v = v + u$ pour u et v éléments quelconques de E ;
- l'addition doit être *associative*, c'est-à-dire : $u + (v + w) = (u + v) + w$ pour u, v, w éléments quelconques de E ;

- il existe dans E un élément "neutre" (ou d'action nulle pour reprendre un langage physique), noté 0 tel que $u + 0 = 0 + u = u$ pour tout u dans E ;
- tout élément u de E admet un élément qui le compense (dit symétrique et noté $-u$), tel que $u + (-u) = 0$.

Un tel ensemble E équipé d'une telle addition $+$ génère une structure $(E, +)$ de *groupe commutatif*.

Ce n'est pas tout ; il nous faut aussi une action externe de $\mathbb{R}$ sur E que l'on notera

$$(\lambda, u) \in \mathbb{R} \times E \longmapsto \lambda \cdot u \,,$$

action qui se prête aux quatre exigences suivantes :

$$\begin{aligned} \lambda \cdot (u + v) &= \lambda \cdot u + \lambda \cdot v \\ (\lambda + \mu) \cdot u &= \lambda \cdot u + \mu \cdot u \\ \lambda \cdot (\mu \cdot u) &= (\lambda \times \mu) \cdot u \\ 1 \cdot u &= u \end{aligned}$$

pour tout u, v dans E, pour tout λ, μ dans $\mathbb{R}$.

Le triplet $(E, +, \cdot)$ constitué de E et de ses deux opérations (interne et externe) constitue une structure de $\mathbb{R}$-espace vectoriel. Les éléments de E sont dits *vecteurs*[1] de E.

Exemples 1.1. Voici quelques exemples importants :

- L'ensemble $\mathbb{R}^N$ est bien sûr le prototype. Un autre exemple est l'espace $\mathcal{M}_{n,p}(\mathbb{R})$ des tableaux de nombres réels à n lignes et p colonnes. On a défini l'addition et l'action externe dans la section 1.1.2
- L'ensemble F_N des fonctions de $[1, N]$ dans $\mathbb{R}$, de graphe une ligne brisée avec noeuds aux points $1, 2, ..., N$ peut être équipé d'une addition et d'une action externe pour générer une stucture d'espace vectoriel, d'ailleurs en bijection avec $\mathbb{R}^N$ (voir la section 1.1) ;il en est de même de même de l'ensemble $F_{n,p}$ (en bijection avec $\mathcal{M}_{n,p}(\mathbb{R})$ comme nous l'avons vu).
- L'ensemble des fonctions d'un ensemble quelconque X et à valeurs dans un espace vectoriel F hérite d'une structure de $\mathbb{R}$-espace vectoriel ; on définit l'addition par

  $$\forall x \in X \,, \quad (f + g)(x) := f(x) + g(x)$$

  et l'action externe de $\mathbb{R}$ par

  $$\forall x \in X \,, \quad (\lambda \cdot f)(x) := \lambda \times f(x) \,.$$

  On peut d'ailleurs, si $E = \mathbb{R}^n$ et si $X = (a, b)$ est un intervalle de $\mathbb{R}$, préciser si l'on veut "des applications continues de (a, b) dans $\mathbb{R}^n$", ce qui définit un $\mathbb{R}$-espace vectoriel inclus dans le précédent (on dit aussi un *sous-espace vectoriel* du précédent).
- L'ensemble $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ des applications linéaires de $\mathbb{R}^p$ dans $\mathbb{R}^n$ hérite aussi d'une structure de $\mathbb{R}$-espace vectoriel ; dans le cas où $n = 1$, on appelle espace dual de $\mathbb{R}^n$ et on le note $(\mathbb{R}^n)^*$.

1. On les notera en général sans les surmonter d'une flèche pour ne pas alourdir les notations ; on privilègiera les lettres u, v, w, e, f pour les vecteurs et les caractères grecs $(\alpha, \beta, \gamma, \lambda, \mu)$ ou x, y, z pour les éléments de $\mathbb{R}$ (que l'on appelle aussi les scalaires) agissant sur E *via* l'action externe.

1.2.2 Systèmes générateurs, libres ; $\mathbb{R}$-espaces vectoriels de dimension finie

Soit E un $\mathbb{R}$-espace vectoriel [2] ; on dit qu'une famille $\mathcal{G} = (g_i)_{i\in I}$ (pas nécessairement finie) est une *famille génératrice* de E si tout élément u de E s'écrit sous la forme

$$u = \sum_{k=1}^{N} x_{i_k} \cdot g_{i_k} \,,$$

où $g_{i_1}, ..., g_{i_N}$ sont N éléments de $\mathcal{G}$ (ces N éléments étant dépendants de u). On dit encore que tout élément de E est *combinaison linéaire finie* d'éléments de $\mathcal{G}$.

Bien sûr, E lui-même est toujours une famille génératrice de E ; ce qui est plus intéressant est de trouver une famille génératrice minimale, au sens où, si on lui en retire un élément, elle cesse d'être génératrice.

L'autre notion importante dans un $\mathbb{R}$-espace vectoriel est celle de *famille libre* ; une telle famille $\mathcal{L} = (l_i)_{i\in I}$ est une famille telle qu'il n'existe aucune relation du type

$$\sum_{k=1}^{N} x_{i_k} \cdot l_{i_k} = 0$$

autre que la relation triviale ($x_{i_k} = 0$, $k = 1, ..., N$).

Bien sûr, tout élément non nul de E (si E n'est pas réduit à 0) constitue une famille libre ; en revanche, E non car on a toujours $u - u = 0$ pour tout u dans E. Les familles libres intéressantes sont les familles libres maximales, au sens où, si on leur ajoute un élément, elles cessent d'être libres. On admettra que tout $\mathbb{R}$-espace vectoriel admet toujours une famille à la fois génératrice et libre ; une telle famille s'appelle une *base* de E.

Il n'y a aucune raison pour qu'un $\mathbb{R}$-espace vectoriel admette une partie génératrice finie ; les $\mathbb{R}$-espaces vectoriels ayant cette propriété sont dits $\mathbb{R}$-*espaces vectoriels de dimension finie* et c'est à cette catégorie d'espaces vectoriels que nous nous intéresserons dans la suite de ce cours. Les espaces vectoriels $\mathbb{R}^N$, F_N, $\mathcal{M}_{n,p}(\mathbb{R})$, $F_{n,p}$ sont tous des $\mathbb{R}$-espaces vectoriels de dimension finie : en effet :

- en ce qui concerne $\mathbb{R}^N$, le système $(e_1, ..., e_N)$ introduit dans la section (1.1.1) est un système générateur ;
- en ce qui concerne F_N, l'ensemble constitué des fonctions "triangle" $\Delta_1, ..., \Delta_N$ introduit dans la section 1.1.1 est un système générateur ;
- en ce qui concerne $\mathcal{M}_{n,p}(\mathbb{R})$, la famille $\{I_{i,j}\,;\, 1 \le i \le n\,,\, 1 \le j \le p\}$ (section 1.1.2) est un système générateur ;
- en ce qui concerne $F_{n,p}$, la famille $\{\Delta_{i,j}\,;\, 1 \le i \le n\,,\, 1 \le j \le p\}$ (section 1.1.2) est un système générateur .

Par contre, d'autres $\mathbb{R}$-espaces vectoriels intéressants ne sont pas, eux, de dimension finie et nous les perdons ici ! C'est le cas par exemple de l'espace $\mathcal{C}((a,b),\mathbb{R})$ des fonctions continues d'un intervalle (a,b) de $\mathbb{R}$ et à valeurs réelles ou de l'espace des fonctions continues sur $\mathbb{R}$, à valeurs réelles et T-périodiques. Concernant ce dernier exemple, on pourrait penser, au vu de la théorie de Fourier (vue en PIN301) que la collection

$$\{t \mapsto \cos(2\pi nt/T)\,,\, t \mapsto \sin(2\pi nt/T)\,;\, n \in \mathbb{N}\}$$

2. Ici encore, on peut remplacer partout $\mathbb{R}$ par $\mathbb{C}$ et tout se transpose.

est un système générateur ; ce n'est pas vrai car tout ce que l'on peut affirmer est qu'une fonction continue sur $\mathbb{R}$ à valeurs réelles et T-périodique s'approche (uniformément sur $[0,T]$) par des polynômes trigonométriques

$$a_0 + \sum_{k=1}^{N} \Big(a_k \cos(2\pi kt/T) + b_k \sin(2\pi kt/T) \Big)$$

mais n'est pas en général un polynôme trigonométrique !

Si $(e_1, ..., e_k)$ est une famille libre d'un $\mathbb{R}$-espace E de dimension finie, on peut toujours la compléter par un nombre fini d'éléments $e_{k+1}, ..., e_N$ de manière à ce que le système $(e_1, ..., e_N)$ soit libre maximal (si on rajoute un élément, il cesse d'être libre), donc libre et générateur. En dimension finie, toute famille libre peut se compléter en une base. Cette propriété importante est, comme son nom l'indique, le *théorème de la base incomplète*.

1.2.3 Bases d'un $\mathbb{R}$-espace vectoriel de dimension finie ; notion de dimension

Si E est un $\mathbb{R}$ de dimension finie, il existe un système fini de générateurs ; quitte à retirer certains d'entre eux (jusqu'à ce que cela ne soit possible qu'au prix de perdre la propriété d'être un système générateur), on voit qu'il existe toujours un système générateur minimal au sens suivant : ce n'est plus un système générateur si on lui retire un élément !

Un système générateur minimal est nécessairement libre (car sinon, on pourrait exprimer un de ses éléments comme une combinaison des autres) et c'est donc une *base* de E.

Exemple 1.2. On reprend les exemples de la section 1.1.
- le système $(e_1, ..., e_N)$ est une base de $\mathbb{R}^N$, dit *base canonique* de cet espace vectoriel ;
- le système $(\Delta_1, ..., \Delta_N)$ est une base de F_N ;
- le système $\{I_{i,j}\,;\, 1 \le i \le n\,,\, 1 \le j \le p\}$ est une base de $\mathcal{M}_{n,p}(\mathbb{R})$;
- le système $\{\Delta_{i,j}\,;\, 1 \le i \le n\,,\, 1 \le j \le p\}$ est une base de $F_{n,p}$;
- le système $\{t \mapsto \cos(2\pi nt/T)\,,\, n \in \mathbb{N}\,; t \mapsto \sin(2\pi nt/T)\,, n \in \mathbb{N}^*\}$ est une base de l'espace des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ de la forme

$$t \longmapsto a_0 + \sum_{k=1}^{N} \Big(a_k \cos(2\pi kt/T) + b_k \sin(2\pi kt/T) \Big)\,.$$

Propriété "clef" *Si $(e_1, ..., e_N)$ est une base de E, alors dès que l'on prend strictement plus de $p > N$ éléments dans E $(u_1, ..., u_p)$, ils sont liés par une relation linéaire :*

$$x_1 \cdot u_1 + \cdots + x_p \cdot u_p = 0\,.$$

On va se contenter de prouver en supposant $N = 2$ et en prenant 3 éléments

$$\begin{aligned} u_1 &= x_{11} \cdot e_1 + x_{21} \cdot e_2 \\ u_2 &= x_{12} \cdot e_1 + x_{22} \cdot e_2 \\ u_3 &= x_{13} \cdot e_1 + x_{23} \cdot e_2 \end{aligned}$$

Si

$$\Delta = x_{11}x_{22} - x_{12}x_{21} = \begin{vmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{vmatrix} = 0\,,$$

on voit que les vecteurs u_1 et u_2 sont liés par l'une des relations (forcément non triviale)

$$x_{22} \cdot u_1 - x_{21} \cdot u_2 = 0$$

ou

$$x_{12} \cdot u_1 - x_{11} \cdot u_2 = 0\,;$$

si $\Delta = 0$, alors on peut exprimer e_1 et e_2 sous forme de combinaisons linéaires de u_1 et u_2, reporter dans la relation $u_3 = x_{13} \cdot e_1 + x_{23} \cdot e_2$ et obtenir ainsi une relation non triviale de la forme

$$u_3 + X_1 \cdot u_1 + X_2 \cdot u_2 = 0\,,$$

ce qui prouve que u_1, u_2, u_3 sont liés.

La première propriété clef que nous venons de mentionner montre que si E est un $\mathbb{R}$ de dimension finie, toutes les bases ont le même cardinal. Ce cardinal est appelé *dimension* du $\mathbb{R}$-espace vectoriel ; il correspond au nombre de degrés de liberté dont dépend un élément de E, ou encore au nombre de paramètres réels nécessaires à figer pour déterminer un élément précis de E. En effet on a la caractérisation suivante de la notion de base :

Caractérisation pratique de la notion de base : *dire que $(e_1, ..., e_N)$ est une base d'un $\mathbb{R}$-espace vectoriel de dimension N, c'est dire que tout vecteur u de E s'écrit de manière unique sous la forme*

$$u = x_1 \cdot e_1 + \cdots + x_N \cdot e_N\,,$$

où $x_1, ..., x_N$ sont N nombres réels, appelés coordonnées de u dans cette base.

Bien sûr, la liste des coordonnées d'un vecteur est intrinsèquement liée à la base ; il est d'ailleurs toujours important, lorsque l'on travaille avec des phénomènes physiques modélisés par des vecteurs appartenant à un certain $\mathbb{R}$-espace vectoriel E de dimension finie, de travailler avec une base de E dans laquelle les coordonnées des vecteurs correspondant aux phénomènes physiques étudiés forment un système le mieux organisé possible (au niveau de la notion physique ou informatique d'entropie).

Exemple 1.3. Choisissons pour changer un exemple de $\mathbb{C}$-espace vectoriel, l'espace $\mathbb{C}^N$ dont une base est la base canonique (notée encore $(e_1, ..., e_N)$. Une autre base de ce même espace est la base constituée des vecteurs

$$w_k := (W_N^0, W_N^k, ..., W_N^{k(N-1)})\,,\ k = 0, ..., N-1\,,$$

où $W_N := \exp(-2i\pi/N)$. Si l'on visualise les parties réelles et imaginaires de ces vecteurs comme sur la figure 1.7 (avec $N = 64$, on a représenté sous forme de graphes polygonaux les vecteurs $\mathrm{Re}\, w_k$, $k = 5, k = 10$), on voit qu'autant la base $(w_1, ..., w_k)$ est une base judicieuse pour décomposer les phénomènes ondulatoires (d'ailleurs calculer les coordonnées de $v = (s(0), ..., s(N-1))$ revient de fait à prendre le spectre du signal digital complexe $(s(0), ..., s(N-1))$, autant ce n'est certainement pas une base judicieuse pour décomposer par exemple le vecteur $(0, ..., 0, X, 0, ..., 0)$ qui correspond juste à une impulsion ; la base $(e_1, ..., e_N)$ est bien plus judicieuse dans ce cas !

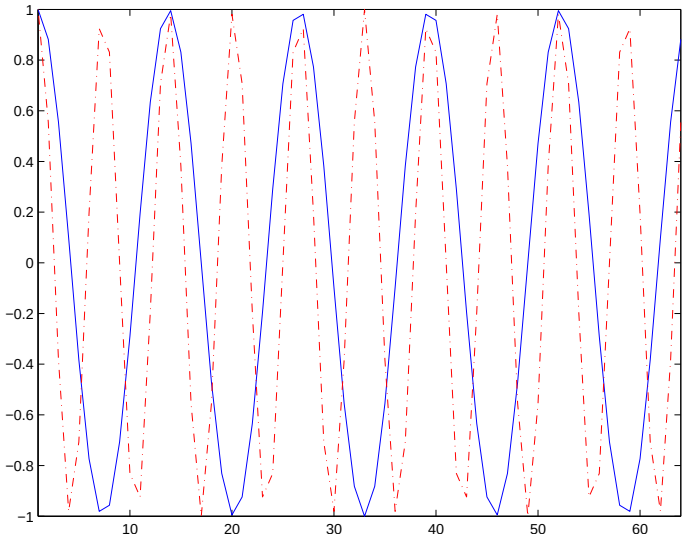

FIGURE 1.7 – La base $(w_1, ..., w_{64})$ et les vecteurs $\mathrm{Re}\, w_5$, $\mathrm{Re}\, w_{10}$

1.3 Applications linéaires entre $\mathbb{R}$-espaces de dimension finie

Dans toute cette section encore $\mathbb{R}$ peut être remplaçé par $\mathbb{C}$ et les résultats peuvent tous être transposés mot pour mot à ce nouveau cadre.

1.3.1 Matrice relativement à un choix de bases ; produit de matrices

On considère ici deux $\mathbb{R}$-espaces vectoriels, rapportés à des bases $(e_1, ..., e_p)$ pour E (qui est donc supposé de dimension p) et $(f_1, ..., f_n)$ pour F (qui est donc supposé de dimension n).

Si L est une application $\mathbb{R}$-linéaire de E dans F, c'est-à-dire une application de E dans F telle que

$$L(\lambda \cdot u + \mu \cdot v) = \lambda \cdot L(u) + \mu \cdot L(v)$$

pour u, v quelconques dans E et λ, μ quelconques dans $\mathbb{R}$ (une telle application est réalisée par un système d'appareils agissant linéairement), la connaissance de L est entièrement déterminée par la connaissance du tableau A à n lignes et p colonnes dont les entrées en $a_{i,j}$ (i indice de ligne, j indice de colonne) sont déterminées par les relations

$$L(e_j) = \sum_{i=1}^{n} a_{i,j} \cdot f_i\,,\ \ j = 1, ..., k\,.$$

En effet, connaître ce tableau, c'est connaître $L(e_1), ..., L(e_p)$ et par conséquent connaître

$$L(x_1 \cdot e_1 + \cdots + x_p \cdot e_p) = x_1 \cdot L(e_1) + \cdots + x_p \cdot L(e_p)$$

pour tout $x_1, ..., x_p$ dans $\mathbb{R}$. L'opération réalisant le calcul des coordonnées $(y_1, ..., y_n)$ dans la base $(f_1, ..., f_n)$ du vecteur $L(u)$ en fonction des coordonnées $(x_1, ..., x_p)$ de

u dans la base $(e_1, ..., e_p)$ s'obtient selon une "gymnastique opérationnelle" que nous décrirons ci dessous :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p-1} & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p-1} & a_{2,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,p-1} & a_{n-1,p} \\ a_{n,1} & a_{n,2} & \dots & a_{n,p-1} & a_{n,p} \end{pmatrix} \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{p-1} \\ x_p \end{pmatrix} \quad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}$$

Comme

$$y_k = \sum_{j=1}^{p} a_{k,j} x_j \,, \;\; k = 1, ..., n \,,$$

on voit que, si le tableau A et la liste des coordonnées $(x_1, ..., x_n)$ de u sont présentées comme ci-dessus, y_k s'obtient en multipliant terme à terme (puis en sommant) la k-ème ligne de A (exactement de longueur p) par la colonne X (également de longueur p, ce qui est cohérent).

Le tableau A est appelé matrice de L, les espaces d'entrées et de sortie étant rapportés aux bases respectives $(e_1, ..., e_p)$ pour E et $(f_1, ..., f_n)$ pour F.

Préciser le choix des bases à l'entrée et à la sortie est indispensable pour parler de matrice d'une application linéaire de E dans F ; changer de base, c'est aussi (on le verra), changer de matrice ! Ceci est très important.

Les applications linéaires peuvent être composées ; par exemple, si l'on considère trois $\mathbb{R}$-espaces vectoriels E, F, G, de dimensions finie, de dimensions respectives p pour E, n pour F, m pour G, et deux applications linéaires

$$\begin{aligned} L \; &: E \longrightarrow F \\ M \; &: F \longrightarrow G \end{aligned}$$

on peut introduire trois matrices :

- la matrice A (à n lignes et p colonnes), matrice de L lorsque les espaces E et F sont rapportés aux bases respectives $(e_1, ..., e_p)$ et $(f_1, ..., f_n)$;
- la matrice B (à m lignes et n colonnes), matrice de M lorsque les espaces F et G sont rapportés aux bases respectives $(f_1, ..., f_n)$ et $(g_1, ..., g_m)$;
- la matrice C (à m lignes et p colonnes), matrice de $M \circ L \; : u \longmapsto M(L(u))$ lorsque les espaces E et G sont rapportés aux bases respectives $(e_1, ..., e_p)$ et $(g_1, ..., g_m)$.

Il est naturel de se demander comment C se calcule à partir de A et B. Ici encore, on présente les choses comme suit :

$$\begin{pmatrix} b_{1,1} & b_{1,2} & \dots & b_{1,n-1} & b_{1,n} \\ b_{2,1} & b_{2,2} & \dots & b_{2,n-1} & b_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ b_{m-1,1} & b_{m-1,2} & \dots & b_{m-1,n-1} & b_{m-1,n} \\ b_{m,1} & b_{m,2} & \dots & b_{m,n-1} & b_{m,n} \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,p-1} & a_{1,p} \\ a_{2,1} & a_{2,2} & \dots & a_{2,p-1} & a_{2,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \dots & a_{n-1,p-1} & a_{n-1,p} \\ a_{n,1} & a_{n,2} & \dots & a_{n,p-1} & a_{n,p} \end{pmatrix} = \begin{pmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,p-1} & c_{1,p} \\ c_{2,1} & c_{2,2} & \dots & c_{2,p-1} & c_{2,p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{m-1,1} & c_{m-1,2} & \dots & c_{m-1,p-1} & c_{m-1,p} \\ c_{m,1} & c_{m,2} & \dots & c_{m,p-1} & c_{m,p} \end{pmatrix}$$

Le coefficient $c_{k,l}$ au carrefour de la k-ème ligne et de la l-ème colonne se calcule comme

$$c_{k,l} = \sum_{j=1}^{n} b_{k,j} a_{j,l} \ ;$$

on multiplie terme à terme la k-ème ligne de B par la l-ème colonne de B (les longueurs de ces deux suites de nombres valent toutes les deux n, ce qui est cohérent), puis on ajoute.

On définit ainsi le produit de matrices $B \bullet A$ défini sous un logiciel tel que MATLAB par

```
>>C=B*A;
```

Attention ! Il est essentiel pour que ce produit soit défini que le nombre de colonnes de B soit égal au nombre de lignes de A, sinon le logiciel vous renverra un message d'erreur tel celui ci :

```
>>C=B*A;
??? Error using ==> mtimes
Inner matrix dimensions must agree.
```

Il ne faut pas confondre ce produit avec le produit "terme-à-terme" (on multiplie $a_{i,j}$ par le $b_{i,j}$ correspondant) de deux matrices A et B de même dimension, opération en général formulée

```
>>C=B.*A;
```

dans un logiciel de calcul scientifique tel MATLAB et SCILAB. Il faut savoir que le produit de matrices est l'outil de base dans la conception de tels logiciels de calcul scientifique.

Règle essentielle : *Si E, F, G sont trois espaces vectoriels rapportés à des bases respectives $(e_1, ..., e_p)$, $(f_1, ..., f_n)$, $(g_1, ..., g_m)$, si $L : E \longrightarrow F$ est une application $\mathbb{R}$-linéaire de E dans F de matrice A lorsque E et F sont rapportés à ces bases, M une application $\mathbb{R}$-linéaire de F dans G de matrice B lorsque E et F sont rapportés à ces bases, la matrice de $M \circ L : E \longrightarrow G$, les espaces E et G étant rapportés à*

leurs bases $(e_1, ..., e_p)$ et $(g_1, ..., g_m)$, est égale à $C = B \bullet A$, le produit de matrices étant défini comme ci-dessus.

Deux tableaux carrés B et A de même taille se multiplient suivant la règle $B \bullet A$ et l'on a ainsi une multiplication sur l'ensemble $\mathcal{M}_{N,N}(\mathbb{R})$ des tableaux carrés à N lignes et N colonnes. Un tel tableau carré correspond (voir la section 1.1.2) à la matrice d'une application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^N$ (rapportés tous deux à la base canonique) ; si A correspond ainsi à L et B à M, $B \circ A$ correspond à l'application linéaire composée $M \circ L : \mathbb{R}^N \longrightarrow \mathbb{R}^N$. Attention, en général $B \bullet A \neq A \bullet B$ (deux applications linéaires ne commutent pas en général).

1.3.2 Noyau, image, rang

Soient E et F deux $\mathbb{R}$-espaces vectoriels (pas nécessairement de dimension finie) et $L : E \longrightarrow F$ une application linéaire correspondant à un système d'appareils agissant linéairement. L'ensemble des vecteurs u de E que *le système ne reconnaît pas*, c'est-à-dire les vecteurs $u \in E$ tels que $L(u) = 0$, est une partie de E, dite *noyau* de L (notée $\operatorname{Ker} L$, en anglais *kernel*), stable sous prise de combinaison linéaire. On a donc par définition

$$\operatorname{Ker} L = \text{Noyau de } L := \{u \in E \,;\, L(u) = 0\}\,.$$

Si l'on restreint l'addition et l'action externe de $\mathbb{R}$ à $\operatorname{Ker} L$, on équipe le noyau de L d'une structure de $\mathbb{R}$-espace vectoriel ; c'est un sous-espace vectoriel de E.

Si E est de dimension finie, les éléments d'un tel sous-espace dépendent de moins de degrés de liberté que ceux de E tout entier. Le noyau de L est donc dans ce cas un sous-espace de dimension finie, de dimension inférieure ou égale à la dimension de E.

Toujours dans un $\mathbb{R}$-espace vectoriel de dimension 1, les sous-espaces de dimension 1 sont appelés *droites vectorielles*, ceux de dimension $\dim E - 1$ *hyperplans vectoriels* (ou plans vectoriels si $\dim E = 3$). Sauf si L est l'application identiquement nulle (auquel cas $\operatorname{Ker} L = E$), la dimension du noyau de L est un nombre entre 0 et $\dim E - 1$, le cas $\dim(\operatorname{Ker} L) = 0$ correspondant au cas où L est injective (tout vecteur de F a au plus un antécédent par L).

L'*image* de E par L (même si E et F ne sont pas de dimension finie)

$$\operatorname{Im} L : \{L(u) \,;\, u \in E\}$$

est une partie de F stable sous l'addition de F et sous l'action externe de $\mathbb{R}$ sur F. C'est donc un sous-espace vectoriel de F cette fois.

Si E est de dimension finie, on a

$$\dim(\operatorname{Im} L) \leq \dim F\,.$$

Le cas où $\dim(\operatorname{Im} L) = \dim F$ correspond au cas où $\operatorname{Im} L = F$, c'est-à-dire au cas où L est surjective (tout vecteur de F a au moins un antécédent dans E). Le cas où $\dim(\operatorname{Im} L) = 0$ correspond au cas où L est l'application linéaire indentiquement nulle.

Si E et F sont deux espaces de dimension finie, il est clair que l'on ne peut que perdre des degrés de liberté en tranformant E par L. En fait, le nombre de degrés de liberté perdus correspond exactement à la dimension du noyau et on a donc la formule essentielle suivante :

$$\dim E = \dim(\operatorname{Ker} L) + \dim(\operatorname{Im} L)\,. \tag{1.2}$$

Pour s'en convaincre, on pensera au cas où $E = F = \mathbb{R}^3$ et où L est la projection sur un plan vectoriel Π parallèlement à une droite vectorielle D. Le plan Π est l'image de L, la droite D exactement son noyau (attention, droite et plan vectoriels sont à comprendre comme des droites et plans au sens usuel passant par l'origine !). Il est clair alors que les points de la droite sont les points envoyés sur l'origine par projection.

Toujours si E et F sont de dimension finie, le *rang* d'une application linéaire $L : E \longrightarrow F$ (noté aussi $\operatorname{rg}(L)$ ou $\operatorname{rk} L$) est par définition la dimension du sous-espace vectoriel $\operatorname{Im} L$ de F. On peut donc formuler (1.2) en

$$\dim E = \dim(\operatorname{Ker} L) + \operatorname{rg} L\,,$$

formule basique en algèbre linéaire et connue comme le *théorème du rang*.

Si E est un $\mathbb{R}$-espace vectoriel, une application linéaire de E dans $\mathbb{R}$ est appelée *forme linéaire* ; le noyau d'une forme linéaire l non identiquement nulle est donc d'après le théorème du rang un hyperplan vectoriel de E ; si l'on fixe une base $(e_1, ..., e_p)$, la matrice de l lorsque E est rapporté à la base $(e_1, ..., e_p)$ et $\mathbb{R}$ à la base canonique (1) est le tableau ligne :

$$(l(e_1), ..., l(e_p)) = (a_1, a_2, ..., a_p)$$

et on a

$$l(x_1 \cdot e_1 + \cdots + x_p \cdot e_p) = a_1 x_1 + \cdots + a_p x_p\,;$$

le noyau de l est donc l'hyperplan vectoriel

$$\operatorname{Ker} l = \{x_1 \cdot e_1 + \cdots + x_p \cdot e_p \in E\,;\, a_1 x_1 + \cdots a_p x_p = 0\}$$

qui est donc défini, une fois la base $(e_1, ..., e_p)$ précisée, par *l'équation linéaire*

$$a_1 x_1 + \cdots + a_p x_p = 0$$

liant entre elles les coordonnées d'un vecteur dans la base $(e_1, ..., e_p)$. On peut interpréter physiquement une telle équation linéaire comme une contrainte liant les degrés de liberté dont dépendent les éléments de l'espace vectoriel E en jeu.

1.3.3 Isomorphismes ; changement de base

Une autre conséquence du théorème du rang est très importante ; elle concerne le cas où E et F ont même dimension n.

Cas $\dim E = \dim F$. *Si L est une application $\mathbb{R}$-linéaire entre deux $\mathbb{R}$-espaces vectoriels de même dimension, le fait que L soit injective ($\operatorname{Ker} L = \{0\}$ ou encore $\dim(\operatorname{Ker} L) = 0$) équivaut au fait que L soit surjective ($\operatorname{Im} L = F$ ou encore $\dim(\operatorname{Im} L) = \dim F = \dim E$), c'est-à-dire en fait au fait que L soit bijective. L'application inverse est dans ce cas aussi $\mathbb{R}$-linéaire (de F dans E cette fois) et on dit que L réalise un isomorphisme entre E et F.*

Revenons maintenant au cas de deux $\mathbb{R}$-espaces vectoriels de dimensions respectives p et n et d'une application $\mathbb{R}$-linéaire $L : E \longrightarrow F$.

Si $(e_1, ..., e_p)$ et $(\tilde{e}_1, ..., \tilde{e}_p)$ sont deux bases de E, l'identité de E est un isomorphisme particulier de E dans lui même. La matrice de cet isomorphisme lorsque E est rapporté à la base $(e_1, ..., e_p)$ au départ et à la base $(\tilde{e}_1, ..., \tilde{e}_p)$ à l'arrivée est la matrice à p lignes et p colonnes dont la k-ème colonne $(k = 1, ..., p)$ correspond à la liste des coordonnées du vecteur e_k exprimé dans la base $(\tilde{e}_1, ..., \tilde{e}_p)$; on la note P. On appelle cette matrice matrice de passage de la base $(e_1, ..., e_p)$ dans la base $(\tilde{e}_1, ..., \tilde{e}_p)$ (attention à l'ordre dans lequel on prend les deux bases en jeu !). La matrice de passage de la base $(\tilde{e}_1, ..., \tilde{e}_p)$ dans la base $(e_1, ..., e_p)$ est, elle, une matrice notée P^{-1} qui vérifie $P^{-1} \bullet P = P \bullet P^{-1} = I_p$ où I_p est la matrice $p \times p$ constituée de 1 sur la diagonale et de zéros partout ailleurs. La matrice P^{-1} est d'ailleurs la matrice de passage de la base $(\tilde{e}_1, ..., \tilde{e}_p)$ dans la base $(e_1, ..., e_p)$.

De même, si $(f_1, ..., f_n)$ et $(\tilde{f}_1, ..., \tilde{f}_n)$ sont deux bases de F, l'identité de F est un isomorphisme particulier de F dans lui même. La matrice de cet isomorphisme lorsque F est rapporté à la base $(f_1, ..., f_n)$ au départ et à la base $(\tilde{f}_1, ..., \tilde{f}_n)$ à l'arrivée est la matrice à p lignes et p colonnes dont la k-ème colonne $(k = 1, ..., n)$ correspond à la liste des coordonnées du vecteur f_k exprimé dans la base $(\tilde{f}_1, ..., \tilde{f}_n)$; on la note Q. On appelle cette matrice matrice de passage de la base $(f_1, ..., f_n)$ dans la base $(\tilde{f}_1, ..., \tilde{f}_n)$. La matrice de passage de la base $(\tilde{f}_1, ..., \tilde{f}_n)$ dans la base $(f_1, ..., f_n)$ est une matrice notée Q^{-1} qui vérifie $Q^{-1} \bullet Q = Q \bullet Q^{-1} = I_n$ où I_n est la matrice $n \times n$ constituée de 1 sur la diagonale et de zéros partout ailleurs. La matrice Q^{-1} est d'ailleurs la matrice de passage de la base $(\tilde{f}_1, ..., \tilde{f}_n)$ dans la base $(f_1, ..., f_n)$.

Formule de changement de base. *Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimensions respectives p et n et L une application $\mathbb{R}$-linéaire de E dans F.*

- *On suppose que l'on dispose de deux bases $(e_1, ..., e_p)$ et $(\tilde{e}_1, ..., \tilde{e}_p)$ de E, P étant la matrice de passage de la première dans la seconde ;*
- *On suppose que l'on dispose de deux bases $(f_1, ..., f_n)$ et $(\tilde{f}_1, ..., \tilde{f}_n)$ de F, Q étant la matrice de passage de la première dans la seconde ;*
- *On note A la matrice de L lorsque E est rapporté à la base $(e_1, ..., e_p)$ et F à la base $(f_1, ..., f_n)$;*
- *On note $\widetilde{A}$ la matrice de L lorsque E est rapporté à la base $(\tilde{e}_1, ..., \tilde{e}_p)$ et F à la base $(\tilde{f}_1, ..., \tilde{f}_n)$.*

On a alors la formule dite de changement de base suivante

$$A = Q^{-1} \bullet \widetilde{A} \bullet P$$

compte tenu de la correspondance entre composition des actions linaires et produit de matrices dégagée dans la section 1.3.1.

Cette formule clef nous servira constamment par la suite.

1.4 La notion de déterminant

Dans toute cette section, l'ensemble des scalaires $\mathbb{K}$ désignera indifféremment $\mathbb{R}$ ou $\mathbb{C}$ car il est important pour la suite de jouer avec les deux tableaux. Bien des problèmes physiques impliquent (ne serait-ce que pour des aspects opérationnels

comme en électronique) l'entrée en jeu de données ou d'inconnues complexes ; d'autre part, on verra plus loin que, même lorsque les problèmes ne font intervenir *a priori* que des entrées scalaires réelles, il est intéressant de les traiter "en faisant comme si" ces entrées étaient complexes.

1.4.1 Déterminant d'une matrice carrée d'entrées scalaires

Une matrice à n lignes et n colonnes à entrées dans $\mathbb{K}$ peut être considérée comme la donnée de n éléments $u_1, ..., u_n$ de $\mathbb{K}^n$, en l'occurrence les n vecteurs "colonne" de la matrice. On souhaite construire une application

$$\det \ : \ \mathcal{M}_{n,n}(\mathbb{K}) \ \longrightarrow \ \mathbb{K}$$

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,n} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{pmatrix} \ \longmapsto \ \begin{vmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,n} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix}$$

telle que, pour tout choix de scalaires λ, μ dans $\mathbb{K}$, pour tout choix d'entrées scalaires $a_{i,j}$, $b_{i,j}$, $1 \leq i, j \leq n$,

$$\begin{vmatrix} a_{1,1} & \cdots & \lambda a_{1,j} + \mu b_{1,j} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \cdots & \lambda a_{i,j} + \mu b_{i,j} & \cdots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \cdots & \lambda a_{n,j} + \mu b_{n,j} & \cdots & a_{n,n} \end{vmatrix} = \lambda \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \cdots & a_{i,j} & \cdots & a_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix}$$

$$+ \ \mu \begin{vmatrix} a_{1,1} & \cdots & b_{1,j} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i,1} & \cdots & b_{i,j} & \cdots & b_{i,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & \cdots & b_{n,j} & \cdots & a_{n,n} \end{vmatrix},$$

que de plus $\det A = 0$ dès que A a deux colonnes identiques et qu'enfin $\det I_n = 1$ si I_n est la matrice constituée de 1 sur la diagonale et de zéros ailleurs. Ces diverses exigences simultanées nous guident, étant donnée une matrice $A \in \mathcal{M}_{n,n}(\mathbb{K})$, vers la construction d'un <u>unique</u> objet $\det A \in \mathbb{K}$ dont le calcul peut s'effectuer selon la règle suivante

$$\det A \ = \ \sum_{j=1}^{n} a_{i,j}(-1)^{i+j} \det B_{i,j} \tag{1.3}$$

$$= \ \sum_{i=1}^{n} a_{i,j}(-1)^{i+j} \det B_{i,j} \, , \tag{1.4}$$

$B_{i,j}$ étant la matrice à $n-1$ lignes et $n-1$ colonnes déduite de A en "rayant" la ligne d'indice i et la colonne d'indice j ; on réïtère ce calcul pour exprimer chaque $\det B_{i,j}$,

$1 \leq i, j \leq n$ suivant le même principe (la taille des matrices en jeu a diminué d'un cran).

Le calcul d'un déterminant, avant d'être conduit comme l'indique la règle ci-dessus (dite *règle de Sarrus*) consistant à le développer suivant une ligne ou une colonne, gagne en simplicité à être "préparé" suivant les deux règles suivantes :

- on transforme le déterminant de A en son opposé si l'on permute deux colonnes (ou deux lignes) ;
- on ne modifie pas le déterminant de A si l'on ajoute à une colonne C_j de A une combinaison

$$\lambda_1 \cdot C_1 + \cdots + \lambda_{j-1} \cdot C_{j-1} + \lambda_{j+1} \cdot C_{j+1} + \cdots + \lambda_n \cdot C_n$$

des autres colonnes de A ; on ne change par le déterminant de A si l'on ajoute à une ligne l_j de A une combinaison

$$\lambda_1 \cdot l_1 + \cdots + \lambda_{j-1} \cdot l_{j-1} + \lambda_{j+1} \cdot l_{j+1} + \cdots + \lambda_n \cdot l_n$$

des autres lignes de A.

Si A est une matrice carrée à n lignes et n colonnes, on appelle *cofacteur* de l'entrée $a_{i,j}$ le scalaire

$$(-1)^{i+j} \det B_{i,j} \,,$$

où $B_{i,j}$ est, rappelons le, la matrice obtenue à partir de A en "rayant" la ligne d'indice i et la colonne d'indice j.

La matrice $\operatorname{cofact} A := [\det B_{i,j}]_{1 \leq i,j \leq n}$ est appelée *matrice des cofacteurs* de A.

1.4.2 Inversion d'une matrice carrée à entrées dans $\mathbb{K}$

La *transposée* d'une matrice A (à n lignes et p colonnes) étant définie comme la matrice ${}^tA$ à p lignes et n colonnes obtenue en transformant les colonnes de A en lignes et les lignes de A en colonnes (et en en conservant l'ordre), on s'aperçoit aisément que les identités (1.3) et (1.4) impliquent que les termes diagonaux des matrices $A \bullet {}^t(\operatorname{cofact} A)$ et ${}^t(\operatorname{cofact} A) \bullet A$ valent tous $\det A$.

Si i et k sont des indices distincts pris dans $\{1, n\}$, on constate d'autre part que

$$\sum_{j=1}^{n} a_{i,j} (-1)^{k+j} \det B_{k,j}$$

correspond au calcul d'un déterminant ayant deux lignes identiques, ce qui implique

$$\sum_{j=1}^{n} a_{i,j} (-1)^{k+j} \det B_{k,j} = 0 \;\; \forall i, k \in \{1, ..., n\} \,, \;\; i \neq k \,.$$

Si l et j sont enfin des indices distincts pris dans $\{1, n\}$, on constate enfin que

$$\sum_{i=1}^{n} a_{i,j} (-1)^{i+l} \det B_{i,l}$$

correspond au calcul d'un déterminant ayant deux lignes identiques, ce qui implique

$$\sum_{i=1}^{n} a_{i,j}(-1)^{i+l} \det B_{i,l} = 0 \ \ \forall j, l \in \{1, ..., n\}\,,\ \ j \neq l\,.$$

Compte tenu de la définition du produit de matrices de même taille (opération notée $\bullet$) envisagé dans la section 1.3.1, on déduit de ce qui précède les identités matricielles

$$A \bullet {}^t(\text{cofact}\, A) = {}^t(\text{cofact}\, A) \bullet A = \det A \cdot I_n\,,$$

ce qui permet d'en déduire la proposition suivante :

Proposition 1.1 *Si A est une matrice carrée à n lignes et n colonnes de déterminant non nul, la matrice*

$$\frac{1}{\det A} \cdot {}^t(\text{cofact}\, A) = A^{-1}$$

est un inverse à gauche et à droite pour A par rapport à l'opération $\bullet$ de produit matriciel, i.e

$$A \bullet A^{-1} = A^{-1} \bullet A = I_n\,.$$

On admettra que si A et B sont deux matrices à n lignes et n colonnes

$$\det(A \bullet B) = \det A \times \det B\,.$$

Par conséquent, on peut affirmer que la matrice A admet une matrice inverse pour le produit matriciel si et seulement si $\det A = 0$; si $\det A \neq 0$, l'inverse est

$$A^{-1} = \frac{1}{\det A} \cdot {}^t(\text{cofact}\, A)\,.$$

Remarque. Concernant les matrices à entrées complexes et non réelles, il faut prendre garde au fait que la commande

```
>> B=A';
```

(sous les logiciels de calcul tels MATLAB ou SCILAB) ne correspond pas à la transposition mais à la transposition couplée avec la conjugaison des coefficients ; si A est une matrice à entrées complexes, la matrice

$$A' = {}^t\overline{A}$$

est dite *matrice adjointe* de A (on transpose après avoir conjugué les coefficients).

1.4.3 Déterminant d'une application linéaire d'un espace vectoriel de dimension n dans lui-même

Soit E un $\mathbb{K}$-espace vectoriel et $(e_1, ..., e_n)$ une base de E. Si L est une application $\mathbb{K}$-linéaire de E dans lui-même, le déterminant de L est par définition le déterminant de la matrice $A = [a_{i,j}]_{1 \leq i,j \leq n}$ dont la j-ème colonne, $j = 1, ..., n$, correspond à la liste des coordonnées du vecteur $L(e_j)$ exprimé dans la base $(e_1, ..., e_n)$:

$$L(e_j) = \sum_{i=1}^{n} a_{i,j} \cdot e_i\,.$$

En fait, du fait de la formule de changement de base, la matrice $\widetilde{A}$ de L lorsque E est rapporté à la base $(\tilde{e}_1, ..., \tilde{e}_n)$ est reliée à A par la relation

$$A = P^{-1} \bullet \widetilde{A} \bullet P \,,$$

où P est la matrice de passage de la base $(e_1, ..., e_n)$ dans la base $(\tilde{e}_1, ..., \tilde{e}_n)$.

On constate que

$$\det A = \det(P^{-1}) \times \det \widetilde{A} \times \det P = \det \widetilde{A} \,,$$

ce qui prouve que le déterminant de l'application $\mathbb{K}$-linéaire $L : E \longrightarrow E$ ne dépend pas du choix de la base de E choisie pour exprimer la matrice de L. On peut donc définir le *déterminant de l'application linéaire* L comme le déterminant de la matrice de L lorsque E est rapporté (à la source et au but) à une base $(e_1, ..., e_n)$ arbitraire. Ceci ne dépend pas de la base choisie !

1.4.4 Polynôme caractéristique d'une application linéaire d'un espace vectoriel de dimension finie dans lui-même

Dans cette sous-section, l'ensemble de scalaires $\mathbb{K}$ désignera toujours indifféremment $\mathbb{R}$ ou $\mathbb{C}$.

Si E est un $\mathbb{K}$-espace vectoriel de dimension n et L une application $\mathbb{K}$-linéaire de E dans E, le polynôme

$$P(X) := \det(X \cdot \mathrm{Id}_E - L)$$

est un polynôme unitaire de degré n,

$$P(X) = X^n - \sigma_1 X^{n-1} + \cdots + (-1)^n \sigma_n \,,$$

dont les coefficients (éléments de $\mathbb{K}$) sont des scalaires intimement liés à L.

Parmi ces coefficients, on notera que

$$\sigma_1 := \mathrm{Trace}\, L$$

est la somme des termes diagonaux de la matrice $[a_{i,j}]_{1 \leq i,j \leq n}$ de L lorsque E est rapporté à une base arbitraire $(e_1, ..., e_n)$ et que

$$\sigma_n = \det L \,.$$

On admettra le résultat suivant (théorème de Cayley-Hamilton) :

Théorème 1.1 *Soit L une application $\mathbb{K}$-linéaire d'un $\mathbb{K}$-espace vectoriel E de dimension n dans lui-même et*

$$P(X) = X^n - (\mathrm{Trace}\, L)\, X^{n-1} + \cdots + (-1)^n \det L$$

son polynôme caractéristique ; si L^k désigne L composé k fois avec lui-même ($k = 1, ..., n$), l'application $\mathbb{K}$-linéaire de E dans E définie comme :

$$L^n - (\mathrm{Trace}\, L)\, L^{n-1} + \cdots + (-1)^n (\det L)\, \mathrm{Id}_E$$

est l'application linéaire identiquement nulle.

1.4.5 Déterminant et rang d'une application linéaire

Soit E et F deux $\mathbb{K}$ espaces vectoriels de dimensions respectives p et n et L une application $\mathbb{K}$-linéaire de E. On suppose que A est la matrice de L lorsque E est rapporté à la base $(e_1, ..., e_p)$ et F à la base $(f_1, ..., f_n)$. La matrice $A = [a_{i,j}]_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ est donc une matrice à n lignes et p colonnes (voir la section 1.3.1).

Le rang de L (c'est-à-dire la dimension du sous-espace $\operatorname{Im} L \subset F$) est donnée comme la taille du plus grand déterminant non nul (obtenu en rayant des lignes et des colonnes de manière à avoir un tableau carré) extrait de la matrice A.

Une fois connu ce rang (et donc identifié un déterminant extrait de la matrice A non nul et de taille maximale), on peut construire assez facilement une base du noyau de L. Supposons (pour fixer les idées) que le rang de L vaille k et que le tableau carré extrait de la matrice A et de taille (k, k) soit le tableau

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & \vdots & \vdots \\ a_{k,1} & \cdots & a_{k,k} \end{vmatrix}.$$

Pour chercher les vecteurs $x_1 \cdot e_1 + \cdots + x_p \cdot e_p$ appartenant au noyau de L (qui est un sous-espace de E de dimension $n - k$ d'après la relation (1.2)), on doit résoudre le système linéaire

$$A \bullet X = 0$$

où X désigne le vecteur colonne des inconnues $x_1, ..., x_p$. On ne retient en fait de ce système que les k premières équations (celles qui correspondent aux k lignes isolées pour réaliser le déterminant extrait), c'est-à-dire le système

$$\begin{aligned} a_{1,1}x_1 + \cdots + a_{1,k}x_k + a_{1,k+1}x_{k+1} + \cdots + a_{1,p}x_p &= 0 \\ \vdots &= \vdots \\ a_{k,1}x_1 + \cdots + a_{k,k}x_k + a_{k,k+1}x_{k+1} + \cdots + a_{k,p}x_p &= 0\,, \end{aligned}$$

système que l'on exprime sous la forme

$$\begin{aligned} a_{1,1}\mathbf{x_1} + \cdots + a_{1,k}\mathbf{x_k} &= -a_{1,k+1}x_{k+1} - \cdots - a_{1,p}x_p \\ \vdots &= \vdots \\ a_{k,1}\mathbf{x_1} + \cdots + a_{k,k}\mathbf{x_k} &= -a_{k,k+1}x_{k+1} - \cdots - a_{k,p}x_p\,, \end{aligned}$$

et que l'on résout sous la forme

$$\begin{pmatrix} \mathbf{x_1} \\ \vdots \\ \mathbf{x_k} \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} \\ \vdots & \vdots & \vdots \\ a_{k,1} & \cdots & a_{k,k} \end{pmatrix}^{-1} \bullet \begin{pmatrix} -a_{1,k+1}x_{k+1} - \cdots - a_{1,p}x_p \\ \vdots \\ -a_{k,k+1}x_{k+1} - \cdots - a_{k,p}x_p \end{pmatrix}.$$

Les variables x_{k+1},...,x_p matérialisent les $p - k$ degrés de liberté dont dépendent les éléments du noyau de L. Si u_l, $l = k + 1, ..., p$, désigne le vecteur $x_1 \cdot e_1 + \cdots + x_p \cdot e_p$ obtenu en résolvant le système précédent lorsque les degrés de liberté sont fixés par $x_{k+1} = 0, ..., x_l = 1, ..., x_p = 0$ (et les coordonnées manquantes $x_1, ..., x_p$ calculées en conséquence), on constate que les vecteurs $u_{k+1}, ..., u_p$ forment une base du noyau de L.

Remarque. Théoriquement puissante, cette méthode ne s'avère pas la plus efficace pour trouver le noyau d'une application linéaire dont on connaît le rang (c'est-à-dire résoudre un système d'équations linéaires) car elle implique (déjà pour le calcul du rang !) des calculs de déterminants ou d'inversion de matrices qui sont en général très coûteux au niveau de la complexité de calcul (temps, capacité mémoire : un déterminant de taille 100 implique le calcul de (100)! produits de 100 termes, puis leur addition !). On préfèrera la démarche algorithmique basée sur la méthode du pivot de Gauss (voir le cours de MIS101) et c'est elle que l'on évoquera ultérieurement (en fin de section 1.5.5) pour les questions de recherche de bases de sous-espaces propres ou caractéristiques.

1.5 Valeurs propres, vecteurs propres d'une application linéaire

Dans cette section, $\mathbb{K}$ désigne un $\mathbb{R}$ ou un $\mathbb{C}$ espace vectoriel (indifféremment). Nous préciserons lorsque cela s'avèrera nécessaire.

1.5.1 Notion de vecteur propre d'une application linéaire d'un espace vectoriel dans lui-même

Si v est un vecteur non nul d'un $\mathbb{K}$-espace vectoriel E (de dimension finie ou non), la droite vectorielle

$$\mathbb{K} \cdot v := \{u \in E\,;\, u = \lambda \cdot v\,,\ \lambda \in \mathbb{K}\}$$

est un sous-$\mathbb{K}$-espace vectoriel de dimension 1 constitué des vecteurs déduits de v par simple "dilatation", donc épousant toutes les caractéristiques de v (normalisation mise à part).

Si L est une application $\mathbb{K}$-linéaire d'un $\mathbb{K}$-espace vectoriel E (non nécessairement de dimension finie) dans lui même, on dit qu'un vecteur non nul v est un *vecteur propre* de L si et seulement si

$$L(v) \in \mathbb{K} \cdot v\,,$$

c'est-à-dire si et seulement s'il existe un scalaire $\lambda \in \mathbb{K}$ tel que $L(v) = \lambda \cdot v$.

Si v est un vecteur propre de L, l'unique scalaire λ tel que $L(v) = \lambda \cdot v$ est dit *valeur propre* associée au vecteur propre v. Le vecteur v est alors dit *vecteur propre pour la valeur propre* λ.

S'il existe un vecteur propre v de L associé au scalaire $\lambda \in K$ (c'est-à-dire $v \neq 0$ et $L(v) = \lambda \cdot v$), le noyau de l'application linéaire $\lambda \cdot \mathrm{Id}_E - E$ est un $\mathbb{K}$-sous-espace vectoriel de E qui n'est pas réduit à 0 (il contient $v \neq 0$!) que l'on appelle *sous-espace propre de L associé à la valeur propre* λ. Les éléments du sous-espace propre associé à la valeur propre λ sont donc, outre le vecteur nul, tous les vecteurs propres de E pour cette valeur propre λ.

Exemple 1.4. Si E désigne l'ensemble des fonctions d'énergie finie de $\mathbb{R}$ dans $\mathbb{C}$, *i.e* des fonctions "raisonnables" $f : \mathbb{R} \longrightarrow \mathbb{C}$ telles que

$$\int_{-\infty}^{\infty} |f(t)|^2\, dt < +\infty$$

(E peut naturellement être équipée d'une structure de $\mathbb{C}$-espace vectoriel), l'opération physique de "prise de spectre" qui à f associe la fonction

$$\widehat{f} : \omega \longmapsto \int_{-\infty}^{+\infty} f(t) e^{-i\omega t}\, dt$$

est une opération linéaire ; une fonction propre particulière est la gaussienne

$$g \ : t \longmapsto \exp(-t^2/2)$$

dont on vérifie que le spectre $\widehat{g}$ est la gaussienne

$$\widehat{g} \ : \omega \longmapsto \sqrt{2\pi}\, e^{-\omega^2/2}\,;$$

la valeur propre associée à cette fonction propre particulière vaut $\sqrt{2\pi}$. Ce fait important (la gaussienne convenablement normalisée est une fonction propre pour la prise de spectre) justifie l'intérêt de la modélisation des particules par des gaussiennes en mécanique quantique (on réalise ainsi le meilleur "compromis" relatif au principe d'incertitude entre la localisation simultanée d'une particule et de son spectre).

1.5.2 Recherche des valeurs propres dans le contexte de la dimension finie

Soit E un $\mathbb{K}$-espace vectoriel de dimension n et L une application $\mathbb{K}$-linéaire de E dans lui-même. D'après le résultat important mentionné au début de la section 1.3.3, dire que λ est une valeur propre pour L (on dit aussi que λ est un élément du *spectre* de L) équivaut à dire qu'il existe un vecteur propre v pour cette valeur propre λ, c'est-à-dire que l'application linéaire

$$L_\lambda := \lambda \cdot \mathrm{Id}_E - L$$

est non injective ou, ce qui revient au même, non bijective. Ceci revient à dire que $\det L_\lambda = 0$; sinon L_λ aurait un inverse car l'on pourrait formuler le théorème de Cayley-Hamilton (pour L_λ) en disant que :

$$L_\lambda \circ \left[\frac{(-1)^{n-1}}{\det L_\lambda}\Big(L_\lambda^{n-1} - (\mathrm{Trace}\, L_\lambda)\cdot L_\lambda^{n-2} + \cdots + (-1)^{n-1}\sigma_{n-1}\cdot \mathrm{Id}_E\Big)\right] = \mathrm{Id}_E\,.$$

On a donc ainsi la règle cruciale suivante :

Proposition 1.2 *Si E est un $\mathbb{K}$-espace vectoriel de dimension n et L une application $\mathbb{K}$-linéaire de E dans lui-même, les valeurs propres de L (dans $\mathbb{K}$) sont exactement les racines (dans $\mathbb{K}$) du polynôme caractéristique de L ou encore du polynôme (unitaire de degré n)*

$$P(X) = \det(X \cdot I_n - A)$$

où A désigne la matrice de L dans une base arbitraire $(e_1, ..., e_n)$ du $\mathbb{K}$-espace vectoriel E.

1.5.3 Pourquoi le cas $\mathbb{K} = \mathbb{C}$ est plus "riche"(en termes de recherche de valeurs propres) que le cas $\mathbb{K} = \mathbb{R}$?

À cet instant, on conçoit que le cas $\mathbb{K} = \mathbb{C}$ soit plus intéressant que le cas $\mathbb{K} = \mathbb{R}$; en effet, on sait dans ce cas qu'un polynôme de degré n à coefficients complexes admet toujours n racines $\lambda_1, ..., \lambda_n$ dans $\mathbb{C}$, ces racines étant répétées avec leur ordre de multiplicité. Le polynôme caractéristique de L se factorise dans ce cas sous la forme

$$P(X) = (X - \lambda_1)^{\mu_1} \cdots (X - \lambda_p)^{\mu_p}$$

et les nombres complexes supposés distincts $\lambda_1, ..., \lambda_p$ sont exactement les valeurs propres de l'application $\mathbb{C}$-linéaire L (leurs multiplicités respectives sont ici notées $\mu_1, ..., \mu_p$, ces entiers étant tous supérieurs ou égaux à 1 et de somme n puisque P est de degré n).

Dans le cas $\mathbb{K} = \mathbb{R}$, il est parfaitement possible qu'une application $\mathbb{R}$-linéaire $L : E \longrightarrow E$ (E étant un $\mathbb{R}$-espace vectoriel) n'ait aucune valeur propre, donc qu'il n'y ait aucun vecteur propre ; par exemple la rotation vectorielle de $\mathbb{R}^2$ dans lui-même d'angle $\theta \neq 0$ (modulo π), de matrice

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

lorsque E est rapporté (à la source et au but) à la base canonique, n'a aucun vecteur propre car le polynôme caractéristique

$$X^2 - 2\cos\theta\, X + 1$$

a deux racines complexes conjuguées $\cos\theta \pm i \sin\theta$ toutes deux non réelles !

On remarque d'ailleurs, que $\mathbb{K}$ soit $\mathbb{R}$ ou $\mathbb{C}$, que les racines complexes distinctes $\lambda_1, ..., \lambda_p$ affectées de leurs multiplicités respectives $\mu_1, ..., \mu_p$ comme zéros du polynôme caractéristique de L sont telles que :

$$\begin{aligned} \operatorname{Trace} L &= \sum_{j=1}^{p} \mu_j \lambda_j \in \mathbb{K} \\ \det L &= \prod_{j=1}^{p} \lambda_j^{\mu_j} \in \mathbb{K}\,. \end{aligned}$$

Mais il faut prendre garde au fait que si $\mathbb{K} = \mathbb{R}$, λ_j n'est valeur propre de L que si c'est un nombre réel.

1.5.4 Diagonalisation d'une application $\mathbb{K}$-linéaire en dimension finie

Étant donné un $\mathbb{K}$-espace vectoriel de dimension n et une application $\mathbb{K}$-linéaire de E dans lui-même, il est souvent judicieux de rechercher une base de E dans laquelle L s'exprime *à l'économie*, c'est-à-dire de manière à ce que sa matrice dans cette base ressemble le plus à la matrice la plus simple possible, à savoir une matrice dont seuls les termes diagonaux sont non nuls.

Si tel est le cas, les termes diagonaux de la matrice de L dans une telle base sont à prendre parmi les valeurs propres de L, donc parmi les racines du polynôme caractéristique P de L. On peut d'ailleurs montrer que, si tel est le cas, toutes les p racines distinctes éventuelles $\lambda_1, ..., \lambda_p$ de P doivent figurer dans cette matrice diagonale, la racine λ_j étant répétée μ_j fois, où μ_j désigne sa multiplicité comme racine de P. Ceci est équivalent au fait que E admette une base de vecteurs propres pour L ou encore à ce que pour chaque $j = 1, ..., p$, la dimension du sous-espace propre

$$\operatorname{Ker}(\lambda_j \operatorname{Id}_E - L)$$

soit exactement égale à la multiplicité μ_j de λ_j comme racine du polynôme caractéristique de L.

Ceci impose bien sûr que les p racines distinctes (*a priori* complexes) du polynôme caractéristique P soient toutes dans $\mathbb{K}$ (ce qui est très contraignant lorsque $\mathbb{K} = \mathbb{R}$) et qu'en plus les sous-espaces propres correspondant soient assez "riches" (chacun d'eux devant avoir comme dimension la multiplicité de la valeur propre correspondante).

Définition 1.1 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et L une application $\mathbb{K}$-linéaire de E dans lui-même, telle que l'une des conditions (équivalentes) suivantes est remplie :*

- *il existe une base de E constituée de vecteurs propres ;*
- *les racines complexes distinctes $\lambda_1, ..., \lambda_p$ du polynôme caractéristique P de L sont toutes dans $\mathbb{K}$ et pour chaque $j = 1, ..., p$,*

$$\dim(\mathrm{Ker}\,(\lambda_j \cdot \mathrm{Id}_E - L) = \mu_j = \mathrm{mult}\,(\lambda_j)\,;$$

on dit que L est diagonalisable. La matrice de L dans une telle base de vecteurs propres est une matrice diagonale, dont la diagonale est constituée des scalaires distincts $\lambda_1, ..., \lambda_p$, chacun d'eux étant répété autant de fois que sa multiplicité comme zéro du polynôme caractéristique de L.

Remarque. Si $\mathbb{K} = \mathbb{R}$ et si le polynôme caractéristique de L admet (considéré comme polynôme à coefficients complexes) une racine $\lambda \in \mathbb{C}\backslash\mathbb{R}$, il n'y a aucune chance pour que l'application $\mathbb{R}$-linéaire L soit diagonalisable !

Il y a cependant une situation favorable (surtout dans le cas $\mathbb{K} = \mathbb{C}$) :

Théorème 1.2 *Soit E un $\mathbb{K}$-espace vectoriel de dimension n et L une application $\mathbb{K}$-linéaire de E dans lui-même. Si le polynôme caractéristique de P admet n racines distinctes $\lambda_1, ..., \lambda_n$ dans $\mathbb{K}$, alors chaque sous-espace propre $\mathrm{Ker}\,(\lambda_j \cdot \mathrm{Id}_E - L)$ est une droite vectorielle et L est diagonalisable.*

Ce théorème est surtout intéressant dans le cas $\mathbb{K} = \mathbb{C}$: en effet, si les entrées d'une matrice carrée A sont choisies avec une marge d'erreur dans $\mathbb{C}$ autour de valeurs précises, la probabilité pour que le polynôme caractéristique

$$\det(X \cdot I_n - A)$$

ait au moins une racine complexe multiple dans $\mathbb{C}$ est nulle. Presque sûrement, l'application $\mathbb{C}$-linéaire ayant A pour matrice dans une base $(e_1, ..., e_n)$ fixée sera donc diagonalisable !

1.5.5 Le cas $\mathbb{K} = \mathbb{C}$; décomposition spectrale d'une application $\mathbb{C}$-linéaire en dimension finie

Si E est un $\mathbb{C}$-espace vectoriel de dimension n et L une application linéaire de E dans lui-même, le polynôme caractéristique P de L admet p racines complexes distinctes $\lambda_1, ..., \lambda_p$, de multiplicités respectives $\mu_1, ..., \mu_p$ avec

$$\mu_1 + \mu_2 + \cdots + \mu_p = n\,.$$

Pour chaque $j = 1, ..., p$, le noyau de l'application $\mathbb{C}$-linéaire

$$(\lambda_j \cdot \mathrm{Id}_E - L) \circ \cdots \circ (\lambda_j \cdot \mathrm{Id}_E - L) \quad (\mu_j \text{ fois})$$

est un sous-$\mathbb{C}$-espace vectoriel de E, contenant le sous-espace propre correspondant à la valeur propre λ_j (mais en général plus gros que lui), dit *sous-espace caractéristique associé à la valeur propre* λ_j.

On admettra le théorème suivant, dit *de décomposition spectrale* :

Théorème 1.3 *Soit E un $\mathbb{C}$-espace vectoriel de dimension n, L une application $\mathbb{C}$-linéaire de E dans lui même et $E_1, ..., E_p$ les p sous-espaces caractéristiques*

$$E_j := \mathrm{Ker}\,(\lambda_j \cdot \mathrm{Id}_E - L)^{\mu_j}\,,\ j = 1, ..., p\,,$$

associés aux p valeurs propres complexes distinctes $\lambda_1, ..., \lambda_p$ (affectées de multiplicités éventuelles respectives $\mu_1, ..., \mu_p$) du polynôme caractéristique de L. On a $\dim E_j = \mu_j$ pour $j = 1, ..., p$ et de plus il existe une base de E réalisée en concaténant une base de E_1, une base de E_2,...,une base de E_p.

La matrice de L dans une telle base s'écrit comme la somme d'une matrice diagonale D (la diagonale étant constituée des racines distinctes $\lambda_1, ..., \lambda_p$ du polynôme caractéristique de L, chacune répétée autant de fois que sa multiplicité) et d'une matrice N qui a la particularité de vérifier

$$N \bullet \cdots (n \text{ fois}) \cdots \bullet N = 0\,, \tag{1.5}$$

les matrices N et D étant telles que $N \bullet D = D \bullet N$. En fait, on peut même être plus précis. La matrice de L dans cette base judicieuse s'écrit sous la forme d'une matrice "blocs"

$$\begin{pmatrix}
 & & & 0 & \cdots & 0 & & 0 & \cdots & 0 & & 0 & \cdots & 0 \\
\lambda_1 \cdot I_{\mu_1} + N_1 & & & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
 & & & 0 & \cdots & 0 & & 0 & \cdots & 0 & & 0 & \cdots & 0 \\
0 & \cdots & 0 & & & & & & & & & & & \\
\vdots & \vdots & \vdots & \lambda_2 \cdot I_{\mu_2} + N_2 & & & \cdots & & \cdots & & & & \cdots & \\
0 & \cdots & 0 & & & & & & & & & & & \\
 & \vdots & & & \vdots & & \vdots & & \vdots & & & & \vdots & \\
0 & \cdots & 0 & & & & & & & & & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \cdots & & \cdots & \lambda_{p-1} \cdot I_{\mu_{p-1}} + N_{p-1} & & & & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & & & & & & & & & 0 & \cdots & 0 \\
0 & \cdots & 0 & & & & & 0 & \cdots & 0 & & & & \\
\vdots & \vdots & \vdots & & \cdots & & \cdots & \vdots & \vdots & \vdots & & \lambda_p \cdot I_{\mu_p} + N_p & & \\
0 & \cdots & 0 & & & & & 0 & \cdots & 0 & & & &
\end{pmatrix} \tag{1.6}$$

où N_j, $j = 1, ..., p$, est une matrice carrée de taille (μ_j, μ_j) dont toutes les entrées non nulles sont strictement au dessus de la diagonale et qui vérifie

$$N_j \bullet \cdots (\mu_j \text{ fois}) \cdots \bullet N_j = 0\,; \tag{1.7}$$

on rappelle d'autre par que I_{μ_j} désigne la matrice de taille (μ_j, μ_j) dont les entrées sont des 1 sur la diagonale, des zéros partout ailleurs.

Remarque. Au prix d'un travail supplémentaire, on peut même faire en sorte que les seules entrées non nulles de N_j, $j = 1, ..., p$, soient sur la sur-diagonale et vaillent soit 0 soit 1 ; cette forme est dite *réduite de Jordan* de la matrice de L.

Une matrice carrée N de taille (k, k) et dont la puissance k-ème (au sens du produit de matrices) est nulle (comme dans (1.5) pour N avec $k = n$ et dans (1.7) pour N_j avec $k = \mu_j$) est dite *nilpotente*, l'application linéaire qu'elle représente dans une base donnée étant dite *application $\mathbb{C}$-linéaire nilpotente*.

La décomposition de L comme la somme

$$L = \mathcal{D} + \mathcal{N}$$

d'une application $\mathbb{C}$-linéaire diagonalisable $\mathcal{D}$ et d'une $\mathbb{C}$-application linéaire nilpotente $\mathcal{N}$ commutant avec $\mathcal{D}$ (les choses sont même plus précises si l'on prend en compte la forme "blocs" (1.6) de la matrice de L dans une base judicieuse) est dite *décomposition de Dunford*. Cette décomposition est unique et réalisable pour toute application $\mathbb{C}$-linéaire d'un $\mathbb{C}$-espace de dimension finie dans lui-même.

Rappel : résolution de systèmes d'équations linéaires. Qu'il s'agisse de la recherche d'une base du sous-espace propre associé à une valeur propre non nulle λ ou d'une base de l'un des sous-espaces caractéristiques, elle passe, une fois définie une base de E, par la résolution d'un certain système d'équations linéaires

$$\mathcal{A} \bullet X = 0$$

où $\mathcal{A}$ est une certaine matrice carrée à coefficients dans $\mathbb{K}$ (de déterminant nul), X désignant ici le vecteur des coordonnées du vecteur générique recherché, exprimé dans la base de E choisie. Cette résolution se fait en utilisant la méthode du pivot de Gauss (voir le cours de MIS101) et fait apparaître la dépendance de la solution en un certain nombre de degrés de lberté. Le nombre de ces degrés de liberté ν correspond à la dimension du sous-espace auquel on s'intéresse (sous-espace propre associé à une valeur propre λ ou sous-espace caractéristique correspondant à cette même valeur propre λ). La méthode du pivot doit aussi faire surgir une base $V_1, ..., V_\nu$ du sous espace étudié (les V_l étant donnés par leurs vecteurs de coordonnées (en colonne) dans la base de E choisie).

1.5.6 Retour aux matrices

Si A est une matrice carrée à n lignes et n colonnes avec entrées dans $\mathbb{K}$, on peut voir A (voir la section 1.3.1) comme la matrice d'une application $\mathbb{K}$-linéaire de $\mathbb{K}^n$ dans lui-même définie par

$$L(e_j) = \sum_{i=1}^{n} a_{i,j} \cdot e_i \,, \;\; j = 1, ..., n \,,$$

où $(e_1, ..., e_n)$ désigne la base canonique de $\mathbb{K}^n$.

Si les racines complexes du polynôme $\det(X \cdot I_n - A)$ sont toutes dans $\mathbb{K}$ et de plus simples, il résulte du théorème 1.2 que l'application $\mathbb{K}$-linéaire L est diagonalisable, ce que l'on exprime en disant qu'il existe une base $(v_1, ..., v_n)$ de $\mathbb{K}^n$ composée de vecteurs propres (v_j vecteur propre couplé avec la valeur propre λ_j, $j = 1, ..., n$) telle que la matrice de L dans cette nouvelle base soit

$$\mathrm{diag}\,(\lambda_1, ..., \lambda_n) \,,$$

matrice dont la diagonale consiste (dans cet ordre) en $\lambda_1, ..., \lambda_n$ et dont les autres entrées sont des zéros. Si P désigne la matrice de passage de la base $(v_1, ..., v_n)$ dans la base $(e_1, ..., e_n)$, c'est-à-dire la matrice (inversible) dont les colonnes sont les coordonnées des vecteurs v_j dans la base canonique, on a donc

$$D = \mathrm{diag}\,(\lambda_1, ..., \lambda_n) = P^{-1} \bullet A \bullet P\,,$$

ou encore

$$A = P \bullet \mathrm{diag}\,(\lambda_1, ..., \lambda_n) \bullet P^{-1} = P \bullet D \bullet P^{-1}\,. \tag{1.8}$$

Cette formule (1.8) permet de calculer très vite les puissances de A :

$$\begin{aligned} A^k &= P \bullet D \bullet P^{-1} \bullet P \bullet A \bullet P^{-1} \bullet \cdots \bullet P \bullet A \bullet P^{-1} \\ &= P \bullet D^k \bullet P^{-1} \\ &= P \bullet \mathrm{diag}\,(\lambda_1^k, ..., \lambda_n^k) \bullet P^{-1} \end{aligned}$$

ou encore

$$\begin{aligned} \exp(A) : &= \sum_{k=0}^{\infty} \frac{A^k}{k!} = P \bullet \Big(\sum_{k=0}^{\infty} \frac{\mathrm{diag}\,(\lambda_1^k, ..., \lambda_n^k)}{k!} \Big) \bullet P^{-1} \\ &= P \bullet \mathrm{diag}\,(e^{\lambda_1}, ..., e^{\lambda_n}) \bullet P^{-1}\,. \end{aligned}$$

Ce calcul nous sera utile lors de la résolution des systèmes différentiels linéaires à coefficients dans $\mathbb{K}$.

Si A est à entrées réelles, on dispose, pourvu que les racines complexes de

$$X \longrightarrow \det(X \cdot I_n - A)$$

soient simples (ce qui est le cas si l'on s'autorise des marges d'erreur pour la saisie des entrées de A), d'une écriture du type (1.8), les entrées de P et de $\mathrm{diag}\,(\lambda_1, ..., \lambda_n)$ étant cette fois complexes (malgré que A soit supposée à entrées réelles). Une telle formule permet encore le calcul de A^k et de $\exp A$ en remarquant qu'en prenant les parties réelles des entrées des matrices au second membre (les parties imaginaires sont nulles) :

$$\begin{aligned} A^k &= \mathrm{Re}\,[P \bullet \mathrm{diag}\,(\lambda_1^k, ..., \lambda_n^k) \bullet P^{-1}] \\ \exp A &= \mathrm{Re}\,[P \bullet \mathrm{diag}\,(e^{\lambda_1}, ..., e^{\lambda_n}) \bullet P^{-1}]\,. \end{aligned}$$

Dans le cas $\mathbb{K} = \mathbb{C}$, il résulte du théorème 1.3 que, si A est une matrice (n, n) à entrées complexes et si les racines complexes distinctes de $\det(X \cdot I_n - A)$ sont $\lambda_1, , ..., \lambda_p$ (avec les multiplicités respectives $\mu_1, ..., \mu_p$), alors A peut s'écrire sous la forme

$$A = P \bullet (D + N) \bullet P^{-1}$$

où les colonnes de P sont constituées des coordonnées de vecteurs

$$(v_{1,1}, ..., v_{1,\mu_1}, ..., v_{p,1}, ..., v_{p,\mu_p})$$

(dans cet ordre), où, pour $j = 1, ..., p$, les $v_{j,l}$, $l = 1, ..., \mu_j$, forment une base du noyau de l'application linéaire dont la matrice est

$$(\lambda_j \cdot I_n - A)^{\mu_j}$$

dans la base canonique, et $D+N$ est la matrice somme d'une matrice diagonale et d'une matrice nilpotente (qui commutent) explicitée en (1.7). La diagonale de D est constituée des entrées λ_1 (répétée μ_1 fois),..., λ_p (répétée μ_p fois) dans cet ordre. Le calcul des puissances de A se fait encore, mais en utilisant la formule du binôme :

$$\begin{aligned} A^k &= P \bullet (D+N)^k \bullet P^{-1} \\ &= P \bullet \Big(\sum_{l=0}^{\min(k,n-1)} \binom{k}{l} N^l \bullet D^{k-l} \Big) \bullet P^{-1} . \end{aligned}$$

Le calcul des puissances de D est immédiat, celui des puissances de N jusqu'à l'ordre $n-1$ se fait en respectant la forme "bloc" de

$$N = \operatorname{diag}(N_1, ..., N_p)$$

(voir la forme (1.7)). Le calcul de $\exp A$ s'en déduit de même.

Remarque. On peut remarquer que la connaissance de $A, A^2, ..., A^{n-1}$ induit celle de A^k pour tout $k \in \mathbb{N}$ en remarquant que $A^k = R_k(A)$, où le polynôme R_k (à coefficients dans $\mathbb{K}$) s'obtient comme le reste de la division euclidienne de A^k par $\det(X \cdot I_n - A)$ du fait du théorème de Cayley-Hamilton. Ceci est toujours vrai, que $\mathbb{K}$ soit $\mathbb{R}$ ou $\mathbb{C}$, que A soit diagonalisable ou non.

1.6 Résolution des systèmes différentiels linéaires à coefficients constants

1.6.1 Les modèles

Les contraintes que la physique impose à p fonctions (définies sur un même intervalle I de $\mathbb{R}$, par exemple un intervalle de l'axe des temps, où un intervalle de l'espace ou du plan matérialisé par exemple par une poutre ou une étagère en mécanique, et en général à valeurs complexes) sont souvent matérialisées par un système d'équations différentielles

$$F_j(t\,;\, y_1(t), ..., y_p(t)\,;\, y_1'(t), ..., y_p'(t)) = 0\,,\ j = 1, ..., n\,, \tag{1.9}$$

F_j désignant une fonction de $2p+1$ variables. Pour que les p fonctions $y_1, ..., y_p$ de t soient déterminées sous les n contraintes du système (sans qu'il n'y ait de contrainte "redondante" ni que le problème soit sous-déterminé), il est naturel de supposer $p = n$.

Nous supposerons ici bien plus en supposant que le système (1.9) s'écrit :

$$\begin{array}{ccccc} y_1'(t) & = & a_{1,1} \cdot y_1(t) + \cdots + a_{1,n} \cdot y_n(t) & + & b_1(t) \\ y_2'(t) & = & a_{2,1} \cdot y_1(t) + \cdots + a_{2,n} \cdot y_n(t) & + & b_2(t) \\ \vdots & = & \vdots & + & \vdots \\ y_{n-1}'(t) & = & a_{n-1,1} \cdot y_1(t) + \cdots + a_{n-1,n} \cdot y_n(t) & + & b_{n-1}(t) \\ y_n'(t) & = & a_{n,1} \cdot y_1(t) + \cdots + a_{n,n} \cdot y_n(t) & + & b_n(t) \end{array} \tag{1.10}$$

où $A = [a_{i,j}]_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ est une matrice d'entrées scalaires (disons en général dans $\mathbb{K}$, où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) et $b_1, ..., b_n$ des fonctions continues données *a priori* sur l'intervalle I et à valeurs dans $\mathbb{K}$; le tableau A correspond à un tableau de paramètres de contraintes

scalaires, le vecteur b à un vecteur de fonctions-contraintes, tandis que $y_1, ..., y_n$ sont les fonctions inconnues du problème. Un tel système (1.10) s'exprime donc sous la forme

$$Y'(t) = A \bullet Y(t) + B(t)\,, \tag{1.11}$$

où $t \longmapsto Y(t)$ désigne le vecteur de fonctions inconnues

$$t \longmapsto \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{n-1}(t) \\ y_n(t) \end{pmatrix}$$

et $t \longmapsto B(t)$ le vecteur de fonctions-contraintes

$$t \longmapsto \begin{pmatrix} b_1(t) \\ b_2(t) \\ \vdots \\ b_{n-1}(t) \\ b_n(t) \end{pmatrix}.$$

Se ramènent à ce modèle (1.11) les problèmes physiques mettant en jeu une seule fonction $y : I \longrightarrow \mathbb{K}$ assujettie à se plier sur I à une équation différentielle d'ordre n :

$$y^{(n)}(t) = \alpha_1 \cdot y^{(n-1)}(t) + \cdots + \alpha_{n-1} \cdot y'(t) + \alpha_n \cdot y(t) + \beta(t)\,, \tag{1.12}$$

où $\alpha_1, ..., \alpha_n$ désignent n paramètres de contraintes scalaires et $\beta : I \longrightarrow \mathbb{K}$ une fonction-contrainte. En effet, on peut dans ce cas associer à la fonction inconnue y le vecteur de fonctions inconnues

$$t \longmapsto Y(t) := \begin{pmatrix} y(t) \\ y'(t) \\ \vdots \\ y^{(n-1)}(t) \\ y^{(n-1)}(t) \end{pmatrix}$$

et remarquer que la recherche de y équivaut à celle du vecteur de fonctions inconnues Y tel que

$$Y'(t) \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ \alpha_n & \alpha_{n-1} & \alpha_{n-2} & \cdots & \alpha_2 & \alpha_1 \end{pmatrix} \bullet Y(t) + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \beta(t) \end{pmatrix}. \tag{1.13}$$

Le modèle (1.11) correspond à un *système différentiel linéaire du premier ordre à coefficients constants*; si $B \equiv 0$, on dit que le système est *homogène* ou encore *sans second membre*. Le modèle (1.12) (qui s'y ramène) correspond une *équation différentielle linéaire d'ordre n à coefficients constants*; si $\beta \equiv 0$, on dit que l'équation est *homogène* ou encore *sans second membre*.

1.6.2 Les résultats

On a les deux résultats majeurs suivants ; tout d'abord concernant la résolution du système (1.11) :

Théorème 1.4 *Le sous ensemble (dans le $\mathbb{K}$-espace vectoriel des vecteurs de n fonctions dérivables de I dans $\mathbb{K}$) des vecteurs Y solutions du système différentiel d'ordre 1 homogène*

$$Y'(t) = A \bullet Y(t)\,,\ t \in I\,,$$

est un $\mathbb{K}$-sous-espace vectoriel de dimension exactement n. Les n degrés de liberté correspondent aux "valeurs initiales" imposées à $y_1(t_0), ..., y_n(t_0)$ en un point arbitraire t_0 de I.

Les solutions du système différentiel d'ordre 1 (avec second membre cette fois)

$$Y'(t) = A \bullet Y(t) + B(t)\,,\ t \in I\,,$$

sont de la forme

$$t \in I \longmapsto Y(t) = Y_{\text{part}}(t) + Y_{\text{gen}}(t)\,,$$

où Y_{part} est une solution particulière du sytème (non homogène) $Y'(t) = A \bullet Y(t) + B(t)$ et Y_{gen} la solution générale du système (homogène) $Y'(t) = A \bullet Y(t)$.

Si les "degrés de liberté" $y_1(t_0), ..., y_n(t_0)$ sont imposés en un point arbitraire t_0 de I, le vecteur solution du système différentiel d'ordre 1 (avec second membre)

$$Y'(t) = A \bullet Y(t) + B(t)\,,\ t \in I\,,$$

est uniquement déterminé ; on peut donc encore dire (même s'ils ne constituent plus un $\mathbb{K}$-sous-espace vectoriel) que les vecteurs solutions de $Y'(t) = A \bullet Y(t) + B(t)$ dépendent de n degrés de libertés (gelés dès que l'on impose des valeurs "initiales" à $y_1(t_0), ..., y_n(t_0)$ en un point arbitraire t_0 de I).

On verra concrètement comment résoudre ce système lorsque la matrice A est diagonalisable sur $\mathbb{C}$.

Concernant l'equation d'ordre n (1.12), le théorème 1.4 "rebondit" en le résultat suivant :

Théorème 1.5 *Le sous ensemble (dans le $\mathbb{K}$-espace vectoriel des fonctions n-fois dérivables de I dans $\mathbb{K}$) des fonctions y solutions de l'équation différentielle d'ordre n homogène*

$$y^{(n)}(t) = \alpha_1\, y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t)\,,\ t \in I\,,$$

est un $\mathbb{K}$-sous-espace vectoriel de dimension exactement n. Les n degrés de liberté correspondent aux "valeurs initiales" imposées à $y(t_0), ..., y^{(n-1)}(t_0)$ en un point arbitraire t_0 de I.

Les solutions de l'équation différentielle d'ordre n (avec second membre cette fois)

$$y^{(n)}(t) = \alpha_1\, y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t) + \beta(t)\,,\ t \in I\,,$$

sont de la forme

$$t \in I \longmapsto y(t) = y_{\text{part}}(t) + y_{\text{gen}}(t)\,,$$

où y_{part} *est une solution particulière de l'équation différentielle d'ordre* n*(non homogène)*

$$y^{(n)}(t) = \alpha_1\, y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t) + \beta(t)\,,\ t \in I\,,$$

et y_{gen} *la solution générale de l'équation différentielle d'ordre* n *(homogène)*

$$y^{(n)}(t) = \alpha_1\, y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t)\,,\ t \in I\,.$$

Si les "degrés de liberté" $y(t_0), ..., y^{(n-1)}(t_0)$ *sont imposés en un point arbitraire* t_0 *de* I*, la fonction* $t \longmapsto y(t)$ *solution de l'équation différentielle d'ordre* n *(avec second membre)*

$$y^{(n)}(t) = \alpha_1 \cdot y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t) + \beta(t)\,,\ t \in I\,,$$

est uniquement déterminée; on peut donc encore dire (même si elles ne constituent plus un $\mathbb{K}$*-sous-espace vectoriel) que les solutions de l'équation différentielle d'ordre* n

$$y^{(n)}(t) = \alpha_1\, y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t) + \beta(t)\,,\ t \in I\,,$$

dépendent de n *degrés de libertés (gelés dès que l'on impose des valeurs "initiales" à* $y(t_0), ..., y^{n-1}(t_0)$ *en un point arbitraire* t_0 *de* I*).*

1.6.3 La méthode de résolution lorsque A est diagonalisable sur $\mathbb{C}$

Avant de détailler une méthode de résolution d'un système différentiel (n, n) d'ordre 1 linéaire et à coefficients constants (non nécessairement homogène) du type

$$Y'(t) = A \bullet Y(t) + B(t)\,,\ t \in I\,, \tag{1.14}$$

sous conditions initiales imposées ($Y(t_0)$ fixé pour un certain t_0 dans I), nous allons en rappeler la "pierre d'angle", à savoir la résolution de l'équation linéaire d'ordre 1 (avec second membre)

$$y'(t) = a \cdot y(t) + b(t)\,,\ t_0 \in I\,, \tag{1.15}$$

sous condition initiale imposée ($y(t_0) = y_0$ fixé pour un certain t_0 dans I), a désignant un scalaire complexe et b une fonction continue de I dans $\mathbb{C}$. La solution de (1.15) (cela a été vu dans le cours de MIS101) se cherche sous la forme

$$y(t) = C(t) \cdot e^{at}$$

en faisant "varier la constante" $t \longmapsto C(t)$; cela donne

$$C'(t) = b(t)e^{-at}\,,$$

soit

$$C(t) = \int_{t_0}^{t} b(\tau)e^{-a\tau}\, d\tau + y_0 e^{-at_0}\,;$$

finalement, la solution $y : I \longmapsto \mathbb{C}$ de (1.15) telle que $y(t_0) = y_0$ est la fonction

$$t \in I \longmapsto y(t) := \left(y_0 e^{-at_0} + \int_{t_0}^{t} b(\tau)e^{-a\tau}\, d\tau\right) e^{at}\,.$$

Supposons maintenant que la matrice A soit diagonalisable (comme matrice à coefficients complexes) et s'écrive donc

$$A = P \bullet \mathrm{diag}\,(\lambda_1, ..., \lambda_n) \bullet P^{-1}\,,$$

où les nombres complexes $\lambda_1, ..., \lambda_n$ (répétés éventuellement avec leurs multiplicités respectives) sont les valeurs propres de la matrice A supposée diagonalisable [3]. Les colonnes de P correspondent, rappelons le, aux coordonnées des vecteurs propres $v_1, ..., v_n$ (correspondant respectivement aux valeurs propres $\lambda_1, ..., \lambda_n$) exprimés dans la base canonique de $\mathbb{C}^n$; les colonnes de P^{-1} correspondent, elles, aux coordonnées des vecteurs $e_1, ..., e_n$ de la base canonique de $\mathbb{C}^n$, exprimés dans la base de vecteurs propres $v_1, ..., v_n$.

On s'empresse de re-écrire (1.14) sous la forme

$$P^{-1} \bullet Y(t) = \mathrm{diag}\,(\lambda_1, ..., \lambda_n) \bullet [P^{-1} \bullet Y(t)] + P^{-1} \bullet B(t)\,,\ t \in I\,,$$

soit, en introduisant le vecteur de fonctions inconnues auxiliaires

$$t \in I \longmapsto Z(t) := P^{-1} \bullet Y(t)$$

et le vecteur colonne de fonctions

$$t \in I \longmapsto C(t) := P^{-1} \bullet B(t)\,,$$

$$Z'(t) = \mathrm{diag}\,(\lambda_1, ..., \lambda_n) \bullet Z(t) + C(t)\,.$$

Résoudre (1.14) sous les conditions initiales $Y(t_0) = Y_0$ se ramène donc à résoudre le système plus simple

$$Z'(t) = \mathrm{diag}\,(\lambda_1, ..., \lambda_n) \bullet Z(t) + \begin{pmatrix} c_1(t) \\ \vdots \\ c_n(t) \end{pmatrix}$$

sous les conditions initiales

$$Z(t_0) = P^{-1} \bullet Y_0 = \begin{pmatrix} z_{0,1} \\ \vdots \\ z_{0,n} \end{pmatrix}\,.$$

Utilisant l'outil de base (rappelé en tête de cette sous-section) qu'est la résolution sous conditions initiales d'une équation linéaire du premier ordre à coefficients constants de la forme (1.15), on trouve

$$Z(t) = \begin{pmatrix} \left(z_{0,1} e^{-\lambda_1 t_0} + \displaystyle\int_{t_0}^{t} c_1(\tau) e^{-\lambda_1 \tau}\, d\tau\right) e^{\lambda_1 t} \\ \vdots \\ \left(z_{0,n} e^{-\lambda_n t_0} + \displaystyle\int_{t_0}^{t} c_n(\tau) e^{-\lambda_n \tau}\, d\tau\right) e^{\lambda_n t} \end{pmatrix}\,,\ t \in I\,,$$

3. Notons que si A est réelle, les valeurs propres non réelles de A se groupent deux par deux (un nombre complexe et son conjugué) car $\overline{\lambda}$ est valeur propre dès que λ l'est.

et l'on s'empresse de revenir à la solution Y en posant

$$Y(t) = P \bullet \begin{pmatrix} \left(z_{0,1} e^{-\lambda_1 t_0} + \displaystyle\int_{t_0}^{t} c_1(\tau) e^{-\lambda_1 \tau}\, d\tau \right) e^{\lambda_1 t} \\ \vdots \\ \left(z_{0,n} e^{-\lambda_n t_0} + \displaystyle\int_{t_0}^{t} c_n(\tau) e^{-\lambda_n \tau}\, d\tau \right) e^{\lambda_n t} \end{pmatrix}, \ t \in I\,, \tag{1.16}$$

ce qui fournit l'unique solution de notre problème (résolution du système différentiel d'ordre 1 à coefficients constants (1.14) sous les conditions initiales $Y(t_0) = Y_0$).

Si A est une matrice réelle, B un vecteur de fonctions à valeurs réelles et si de plus le vecteur de données initiales Y_0 est un vecteur à coordonnées réelles, la solution $t \longmapsto Y(t)$ du problème est automatiquement une fonction à composantes réelles. Le passage dans ce cas *via* les complexes, s'il est en général un passage obligé pour pouvoir exploiter le fait que A soit diagonalisable (ceci étant, on l'a vu, beaucoup moins fréquent si l'on pense que l'ensemble des scalaires est $\mathbb{R}$), ne constitue qu'un passage intermédiaire, le retour au monde réel s'opérant ensuite de lui-même automatiquement.

1.6.4 Quid si A n'est pas diagonalisable sur $\mathbb{C}$?

Même si A n'est pas diagonalisable sur $\mathbb{C}$, le théorème de décomposition spectrale (théorème 1.3) permet d'exprimer A (dans une base de $\mathbb{C}^n$ dont les vecteurs colonnes correspondent aux colonnes de P) sous la forme

$$A = P \bullet T \bullet P^{-1}\,,$$

où T est une matrice "triangulaire supérieure", c'est-à-dire dont toutes les entrées strictement en dessous de la diagonale sont nulles (la diagonale étant occupée par les valeurs propres λ_j, chacune d'elles étant répétée avec sa multiplicité μ_j). La résolution du système

$$Z'(t) = T \bullet Z(t) + P^{-1} \bullet B(t)\,,\ t \in I\,,$$

sous les conditions initiales $Z(t_0) = P^{-1} \bullet Y_0$ se fait "en cascade" à partir de la dernière ligne (en remontant ligne par ligne de proche en proche) et le vecteur de fonctions

$$t \in I \longmapsto P \bullet Z(t)\,,$$

construit à partir du vecteur $t \longmapsto Z(t)$ obtenu au final, est le vecteur de fonctions solution du système

$$Y'(t) = A \bullet Y(t) + B(t)\,,\ t \in I\,,$$

sous les conditions initiales $Y(t_0) = Y_0$; ce vecteur est un vecteur de fonctions réelles si toutes les entrées du problème (entrées de A, valeurs prises par les composantes de B, coordonnées de Y_0) sont réelles.

1.6.5 À propos des équations linéaires d'ordre n à coefficients constants

Comme on l'a vu, la résolution d'une équation d'ordre n à coefficients du type (1.12)

$$y^{(n)}(t) = \alpha_1\, y^{(n-1)}(t) + \cdots + \alpha_{n-1}\, y'(t) + \alpha_n\, y(t) + \beta(t)\,,\ t \in I\,,$$

sous les conditions initiales $y^{(k)}(t_0) = y_{0,k}$, $k = 0, ..., n-1$, se ramène à la résolution du système (1.13) sous les conditions initiales

$$Y(t_0) = \begin{pmatrix} y_{0,1} \\ \vdots \\ y_{0,n} \end{pmatrix} .$$

Le polynôme caractéristique de la matrice (n, n)

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ \alpha_n & \alpha_{n-1} & \alpha_{n-2} & \cdots & \alpha_2 & \alpha_1 \end{pmatrix}$$

figurant dans (1.13) est, on le vérifie sans mal,

$$P(X) = X^n - \alpha_1 X^{n-1} - \cdots - \alpha_{n-1} X - \alpha_n \, ;$$

l'équation

$$X^n - \alpha_1 X^{n-1} - \cdots - \alpha_{n-1} X - \alpha_n = 0 \tag{1.17}$$

est dite *équation caractéristique* de l'équation différentielle homogène associée à l'équation (1.12).

Si l'on travaille avec $\mathbb{K} = \mathbb{C}$, la solution générale y_{gen} de l'équation homogène

$$y^{(n)}(t) = \alpha_1 \, y^{(n-1)}(t) + \cdots + \alpha_{n-1} \, y'(t) + \alpha_n \, y(t) \, , \ t \in I \, , \tag{1.18}$$

(dépendant, on l'a vu avec l'énoncé du théorème 1.5, de n degrés de liberté) est

$$y(t) = \sum_{j=1}^{p} \Big(\sum_{k=0}^{\mu_j - 1} \lambda_{j,k} \, t^k \Big) \, e^{\lambda_j t} \, ,$$

dès que

$$P(X) = \prod_{j=1}^{p} (X - \lambda_j)^{\mu_j}$$

($\lambda_1, ..., \lambda_p$ étant des nombres complexes distincts), les n degrés de liberté dont dépend y étant matérialisés par les scalaires complexes $\lambda_{j,k}$, $j = 1, ..., p$, $k = 0, ..., \mu_j - 1$.

Si les entrées α_j, $j = 1, ..., n$ sont <u>réelles</u> et que l'on travaille avec $\mathbb{K} = \mathbb{R}$, les $2q \leq n$ racines complexes distinctes non réelles $\lambda_1, ..., \lambda_{2q}$ de l'équation caractéristique (1.17) peuvent être couplées deux par deux $\xi_j \pm i\omega_j$, $j = 1, ..., q$, auquel cas P s'écrit

$$P(X) = \prod_{j=1}^{q} [(X - \xi_j - i\omega_j)(X - \xi_j + i\omega_j)]^{\mu_j} \times \prod_{j=2q+1}^{n} (X - \lambda_j)^{\mu_j} \, ,$$

(où les ξ_j, ω_j, $j = 1, ..., q$ et $\lambda_{2q+1}, ..., \lambda_n$ sont des scalaires réels cette fois) et la solution générale de l'équation homogène (1.18) (dépendant toujours de n degrés de liberté, réels cette fois) s'exprime

$$y(t) = \sum_{j=1}^{q} \Big(\sum_{k=0}^{\mu_j - 1} \rho_{j,k} t^k \cos(\omega_j t + \varphi_{j,k}) \Big) e^{\xi_j t} + \sum_{j=2q+1}^{n} \Big(\sum_{k=0}^{\mu_j - 1} \lambda_{j,k} \, t^k \Big) \, e^{\lambda_j t} \, ;$$

les n degrés de liberté sont dans ce cadre matérialisés par les amplitudes $\rho_{j,k} \geq 0$, $j = 1, ..., q$, $k = 1, ..., \mu_j$, les phases $\varphi_{j,k}$, $j = 1, ..., q$, $k = 0, ..., \mu_j - 1$ (modulo 2π), et les coefficients réels $\lambda_{j,k}$, $j = 2q+1, ..., n$, $k = 0, ..., \mu_j - 1$.

1.6.6 L'exemple des systèmes 2×2 ; discussion sur des exemples

On s'intéresse dans cette section à la description (dans le plan) de l'évolution d'un phénomène physique décrit par un couple de fonctions $(x(t), y(t))$, $t \geq 0$ assujetti à des conditions initiales $x(0) = x_0$ et $y(0) = y_0$ (avec $(x_0, y_0) \neq (0, 0)$) et dont l'évolution est régie par le système différentiel

$$\begin{pmatrix} dx/dt \\ dy/dt \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} \bullet \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{1.19}$$

(le système est pour simplifier supposé ici homogène et l'on suppose les entrées de la matrice $A = [a_{i,j}]$ toutes réelles).

Dans un premier temps, la matrice A est supposée diagonalisable sur $\mathbb{C}$ et ayant deux valeurs propres distinctes, λ_1 et λ_2, toutes les deux réelles. La solution du système est dans ce cas

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} \bullet \begin{pmatrix} X_0\, e^{\lambda_1 t} \\ Y_0\, e^{\lambda_2 t} \end{pmatrix}$$

avec

$$\begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix}^{-1} \bullet \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

$U_j = (u_j, v_j)$ désignant, pour $j = 1, 2$, un vecteur propre de la matrice A pour la valeur propre λ_j. Six cas sont à distinguer du point de vue de l'évolution du phénomène :

- on a $\lambda_2 < \lambda_1 < 0$;
- on a $\lambda_1 > \lambda_2 > 0$;
- on a $\lambda_1 > 0 > \lambda_2$;
- on a $\lambda_1 > \lambda_2 = 0$;
- on a $\lambda_1 = 0 > \lambda_2$;
- on a $\lambda_1 = \lambda_2 = 0$.

Le dernier cas retiendra peu l'attention car il correspond au cas où la trajectoire reste figée à son point initial (x_0, y_0). Les autres cas sont plus intéressants et nous les illustrerons.

Dans le premier cas, la trajectoire est "attirée" vers l'origine, quelque soient les conditions initiales, c'est-à-dire quelque soit le point source (x_0, y_0) dans le plan privé de l'origine.

Dans le second cas, les trajectoires s'échappent vers l'infini lorsque t tend vers $+\infty$ (quelque soit le point source (x_0, y_0) dans le plan privé de l'origine), suivant des branches paraboliques comme sur la figure 1.8 ci-dessous : l'origine joue le rôle de *répulsif*.

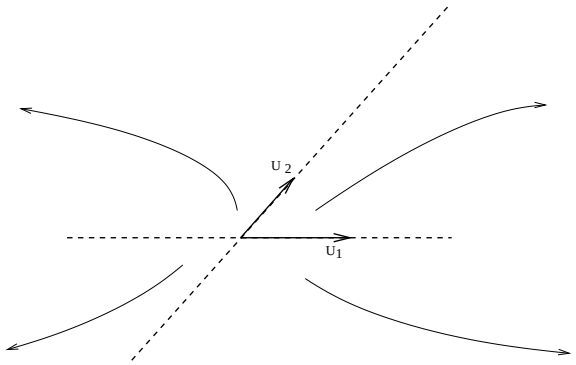

FIGURE 1.8 – Configuration de point "répulsif"

Dans le troisième cas (plus intéressant), on a la configuration de *point selle* ou *point col*; il y a un phénomène d'attraction dans la direction de U_2, de répulsion dans la direction de U_1, comme sur la figure 1.9 ci-dessous :

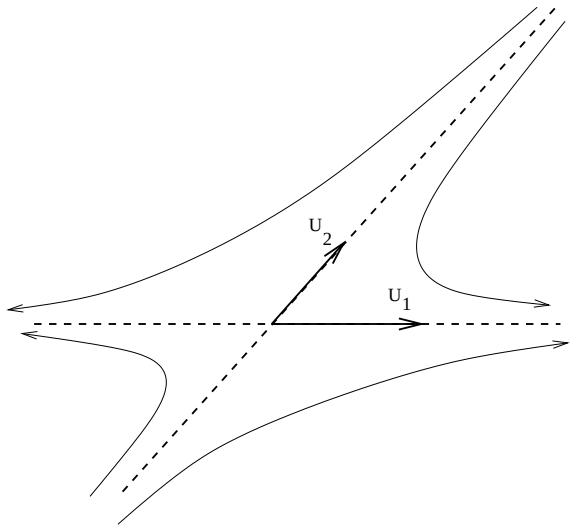

FIGURE 1.9 – Configuration de point "selle" ou "col"

Dans le quatrième cas, la trajectoire reste figée si le point initial est un point de la droite vectorielle dirigée par U_2. Sinon, on observe un phénomène de répulsion, les trajectoires étant parallèles à la droite vectorielle dirigée par le vecteur propre U_1 (figure 1.10 ci-dessous).

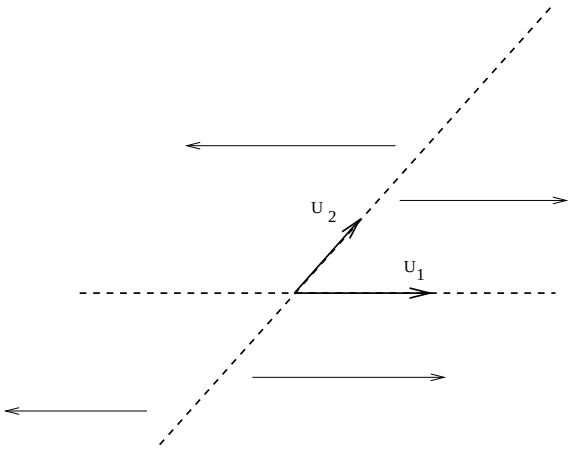

FIGURE 1.10 – Quatrième cas

Dans le cinquième cas, la droite vectorielle dirigée par U_1 "attire" toutes les trajectoires (figure 1.11).

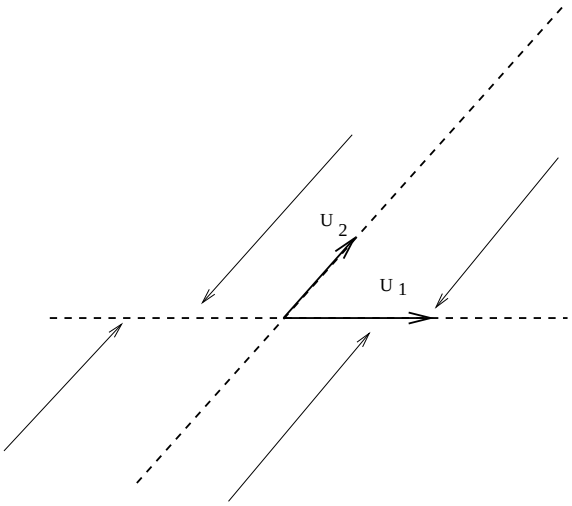

FIGURE 1.11 – Cinquième cas

Dans un second temps, A est toujours supposée diagonalisable sur $\mathbb{C}$ mais cette fois avec deux valeurs propres complexes conjuguées distinctes $\xi \pm i\omega$. La solution du système est dans ce cas toujours

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} \bullet \begin{pmatrix} X_0\, e^{\xi t + i\omega t} \\ Y_0\, e^{\xi t - i\omega t} \end{pmatrix}$$

avec

$$\begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix}^{-1} \bullet \begin{pmatrix} x_0 \\ y_0 \end{pmatrix},$$

mais $U_1 = (u_1, v_1)$ et $U_2 = (u_2, v_2)$ sont cette fois des vecteurs de $\mathbb{C}^2$ qu'il n'est plus possible de représenter dans le plan $\mathbb{R}^2$ comme sur les figures 1.8 à 1.11. On distingue trois cas :

– $\xi < 0$, auquel cas on observe un phénomène s'attraction vers l'origine, les trajectoires étant des spirales s'enroulant autour de l'origine (figure 1.12 (A)) ;
– $\xi = 0$, auquel cas les trajectoires se trouvent être des ellipses autour de l'origine ; c'est une configuration de stabilité, il n'y a ni attraction, ni répulsion, on "tourne" en situation d'équilibre autour de l'origine (figure 1.12 (B)) ;
– $\xi > 0$, auquel cas on observe un phénomène de répulsion depuis l'origine, les trajectoires étant des spirales s'échappant cette fois vers l'infini (figure 1.12 (C)).

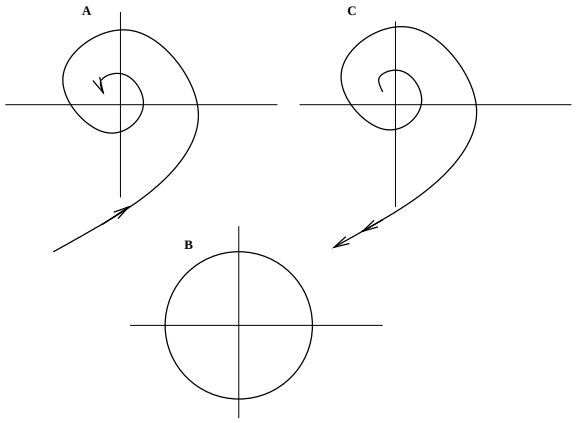

FIGURE 1.12 – Le cas des valeurs propres complexes conjuguées

On trouvera des illustrations concrètes de ces phénomènes dans les exemples détaillés dans le cours d'Alain-Yves LeRoux (section 6.2) auquel on se reportera (modèles empruntés à la météorologie, où l'on retrouve les phénomènes tourbillonnaires du type cyclones). Le cas "pathologique" où le polynôme caractéristique de A admet une racine réelle double λ_0 (et donc où A n'est diagonalisable que si $A = \lambda_0 I_2$) est laissé en exercice.

Notons que si $A = \lambda_0 I_2$ avec $\lambda_0 \neq 0$ on retrouve la première ou la seconde configuration (figures 1.8 pour le cas répulsif $\lambda_0 > 0$), les trajectoires étant des droites issues de l'origine. Si $A = 0$, les trajectoires sont des points et restent toutes stationnaires. Reste seulement le cas où le polynôme caractéristique de A admet une racine double mais où A n'est pas diagonalisable. L'étude des trajectoires se fait dans ce cas en montrant qu'il existe une base (u_1, v_1) et (u_2, v_2) de $\mathbb{R}^2$ telle que

$$A = \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} \bullet \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix} \bullet \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix}^{-1}.$$

On résoudra le système homogène (1.19) dans ce cas et on fera le dessin des trajectoires dans le repère $(0\,;\,(u_1, v_1), (u_2, v_2))$.

1.7 Formes quadratiques et hermitiennes ; orthogonalité et corrélation

1.7.1 Formes bilinéaires et quadratiques sur un $\mathbb{R}$-espace vectoriel

L'énergie, indicateur physique majeur, ne dépend pas de manière linéaire des entrées (l'énergie d'une combinaison linéaire d'entrées n'est pas la combinaison linéaire de leurs énergies !). Cette notion nous contraint à introduire le concept de forme bilinéaire, dont le "prototype" sur $\mathbb{R}^n$ est le *produit scalaire euclidien* :

$$\Big\langle (x_1, ..., x_n)\,,\, (y_1, ..., y_n) \Big\rangle := \sum_{j=1}^{n} x_j y_j\,.$$

On verra aussi que disposer de l'analogue d'un tel produit scalaire permet de conduire dans un espace vectoriel des raisonnements inspirés de la géométrie usuelle et fondés en particulier sur les idées sous-jacentes au théorème de Pythagore.

Définition 1.2 *Soit E un $\mathbb{R}$-espace vectoriel ; on appelle $\mathbb{R}$-forme bilinéaire sur E toute application de $E \times E$ dans $\mathbb{R}$ telle que :*

$$\begin{aligned} \forall\, v, \tilde{v}, w \in E\,, \forall\, \lambda, \tilde{\lambda} \in \mathbb{R}\,,\ \Theta(\lambda \cdot v + \tilde{\lambda} \cdot \tilde{v}\,,\, \vec{w}) &= \lambda\Theta(v\,,\, w) + \tilde{\lambda}\Theta(\tilde{v}\,,\, w) \\ \forall\, v, w, \tilde{w} \in E\,, \forall\, \lambda, \tilde{\lambda} \in \mathbb{R}\,,\ \Theta(v\,,\, \lambda \cdot w + \tilde{\lambda} \cdot \tilde{w}) &= \lambda\Theta(v\,,\, w) + \tilde{\lambda}\Theta(v\,,\, \tilde{w})\,. \end{aligned}$$

La forme Θ est de plus dite symétrique si et seulement si

$$\forall\, v, w \in E\,,\ \Theta(v, w) = \Theta(w, v)\,.$$

Exemples. Dans cette définition, nous ne sommes pas assujettis à supposer E de dimension finie ; aussi nous donnerons des exemples empruntés au cas où E est un espace de fonctions ou de suites :

- **1.** si E est le $\mathbb{R}$-espace des fonctions continues à valeurs réelles sur un intervalle $[a, b]$ de $\mathbb{R}$, alors
$$\Theta\ :\ (f, g) \to \int_a^b f(t)g(t)dt$$
est une forme bilinéaire (symétrique) ;
- **2.** si E est le $\mathbb{R}$-espace vectoriel des fonctions C^1 sur un intervalle $[a, b]$ de $\mathbb{R}$, à valeurs réelles, et telles que $f(a) = f(b) = 0$, alors
$$\Theta\ :\ (f, g) \to \int_a^b f(t)g'(t)dt$$
est une forme bilinéaire, mais non symétrique ($\Theta(f, g) = -\Theta(g, f)$ du fait de la formule d'intégration par parties) ;
- **4.** si E est un $\mathbb{R}$-espace vectoriel de dimension n rapporté à une base $\mathcal{B} = (e_1, ..., e_n)$ et si $\lambda_1, ..., \lambda_n$ sont n scalaires de $\mathbb{R}$, alors :
$$\Theta_{\mathcal{B},\lambda}\ :\ \Big(\sum_{i=1}^{n} x_i \cdot e_i\,,\, \sum_{i=1}^{n} y_i \cdot e_i\Big) \to \sum_{i=1}^{n} \lambda_i x_i y_i$$
est une forme bilinéaire symétrique ;
- **5.** Si $E = \mathbb{R}^3$, rapporté au repère orthonormé $(\vec{i}, \vec{j}, \vec{k})$, le produit scalaire ordinaire de deux vecteurs :
$$(x \cdot \vec{i} + y \cdot \vec{j} + z \cdot \vec{k}, x' \cdot \vec{i} + y' \cdot \vec{j} + z' \cdot \vec{k}) \to xx' + yy' + zz'$$
est bien sûr une forme bilinéaire symétrique ;

– **6.** Si E est l'espace-temps $\mathbb{R}^4$ (trois paramètres d'espace x, y, z, un paramètre temporel t) si important en mécanique non plus Newtonienne, mais relativiste, la <u>*forme de Lorentz*</u> :

$$\left(\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} , \begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix} \right) \to xx' + yy' + zz' - ctt'$$

(c est la vitesse de la lumière) est encore une forme bilinéaire symétrique.

Si E est un $\mathbb{R}$-espace vectoriel de dimension finie (que l'on peut donc rapporter à une base $\mathcal{B} = (e_1, ..., e_n)$, l'écriture matricielle nous permet de représenter (de manière biunivoque) toute forme bilinéaire. Nous avons en effet le résultat capital suivant, se déduisant immédiatement des propriétés de bilinéarité de Θ ; on remarque en effet immédiatement que

$$\begin{aligned} \Theta\Big(\sum_{i=1}^{n} x_i \cdot e_i \, , \, \sum_{j=1}^{n} y_j \cdot e_j \Big) &= \sum_{i=1}^{n} x_i y_j \Theta(e_i, e_j) \\ &= \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} \Theta(e_i, e_j)\, y_j \Big) . \end{aligned}$$

Proposition 1.3 *Soit E un $\mathbb{R}$-espace vectoriel de dimension finie, rapporté à une base $\mathcal{B} = (e_1, ..., e_n)$ et Θ une forme bilinéaire sur E. Il existe une unique matrice carrée $[a_{i,j}]_{1 \le i \le j \le n}$ telle que, pour tout $(x_1, ..., x_n) \in \mathbb{R}^n$, pour tout $(y_1, ..., y_n) \in \mathbb{R}^n$, on ait :*

$$\Theta\Big(\sum_{i=1}^{n} x_i \cdot e_i \, , \, \sum_{j=1}^{n} y_j \cdot e_j \Big) = \Big(x_1, x_2, ..., x_n \Big) \bullet \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \dots & a_{3,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & a_{n,3} & \dots & a_{n,n} \end{pmatrix} \bullet \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}$$

les coefficients $a_{i,j}$ de la matrice en question sont donnés simplement par les formules

$$a_{i,j} = \Theta(e_i, e_j) \, , \quad 1 \le i, j \le n \, .$$

La forme Θ est symétrique si et seulement si la matrice $[a_{i,j}]_{1 \le i,j \le n}$ est une matrice symétrique par rapport à sa diagonale principale. Cette matrice $[a_{i,j}]_{1 \le i,j \le n}$ (permettant d'identifier la forme bilinéaire une fois une base $\mathcal{B}$ de l'espace précisée) est dite <u>*matrice de la forme bilinéaire Θ dans la base $\mathcal{B}$*</u>

Si l'on remplace la base $\mathcal{B} = (e_1, ..., e_n)$ par la base $\widetilde{\mathcal{B}} = (\tilde{e}_1, ..., \tilde{e}_n)$ et si l'on note P la matrice de passage de $\widetilde{\mathcal{B}}$ à $\mathcal{B}$, la relation entre le système de coordonnées $(X_1, ..., X_n)$ d'un vecteur v dans la base $\widetilde{\mathcal{B}}$ en fonction du système $(x_1, ..., x_n)$ de ses coordonnées dans la base $\mathcal{B}$ est

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = P \bullet \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \, ;$$

on voit donc ainsi comment relier les matrices d'une forme bilinéaire Θ dans les bases $\mathcal{B}$ et $\widetilde{\mathcal{B}}$: la matrice de Θ lorsque E est rapporté à la base $\widetilde{\mathcal{B}}$ s'écrit donc

$$\widetilde{M_\Theta} = {}^t P \bullet M_\Theta \bullet P \, ,$$

où P désigne la matrice de passage de la base $\widetilde{\mathcal{B}}$ dans la base $\mathcal{B}$ et M_Θ la matrice de Θ lorsque E est rapporté à la base $(e_1, ..., e_n)$.

Remarque. On notera la différence avec la formule

$$\widetilde{M}_L = P^{-1} \bullet M_L \bullet P$$

lorsque L désigne une application $\mathbb{R}$-linéaire de E dans lui-même.

Si Θ est une forme bilinéaire symétrique sur un $\mathbb{R}$-espace vectoriel E, On appelle *forme quadratique* de forme polaire Θ l'application $\mathcal{Q} : E \longrightarrow \mathbb{R}$ définie (à partir de Θ) par

$$\forall\, v \in E\,,\ \mathcal{Q}(v) := \Theta(v, v)\,.$$

Remarquons que l'on peut retrouver la forme bilinéaire Θ (dite *polarisée* de $\mathcal{Q}$) *via* la relation

$$\Theta(v, w) = \frac{1}{2}(\mathcal{Q}(v + w) - \mathcal{Q}(v) - \mathcal{Q}(w))\,.$$

Si $\mathcal{Q}$ est une forme quadratique sur un $\mathbb{R}$-espace vectoriel E, on a, pour tous vecteurs v, w dans E, pour tous scalaires λ, μ dans $\mathbb{R}$,

$$\mathcal{Q}(\lambda \cdot v + \mu \cdot w) = \lambda^2 \mathcal{Q}(v) + \mu^2 \mathcal{Q}(w) + 2\lambda\mu\Theta(v, w)\,;$$

si $\lambda = \mu = 1$, le terme

$$2\Theta(v, w)$$

est qualifié de *terme d'interférence* et matérialise le défaut de linéarité de $\mathcal{Q}$ puisque

$$\mathcal{Q}(v + w) = \mathcal{Q}(v) + \mathcal{Q}(w) + 2\Theta(v, w)\,.$$

Exemples.

- si E est l'espace des fonctions d'un intervalle $[a, b]$ de $\mathbb{R}$ et à valeurs dans $\mathbb{R}$,

$$f \longmapsto \int_a^b f^2(t)\, dt$$

  est une forme quadratique sur E, dite *énergie* ;
- si $E = \mathbb{R}^N$ et si $p(1), ..., p(N)$ sont N nombres de $[0, 1]$ de somme 1, la forme

$$\begin{aligned} X = (x_1, ..., x_N) \longmapsto V(X) &= \sum_{j=1}^N p(j)\, x_j^2 - \Big(\sum_{j=1}^N p(j)\, x_j\Big)^2 \\ &= \sum_{j=1}^N p(j)\, \Big(x_j - \sum_{k=1}^N p(k) x_k\Big)^2 \end{aligned}$$

  est une forme quadratique amenée à jouer un rôle important : c'est la *variance de X, pondérée par le choix de poids* (*ou probabilités*) $p(1), ..., p(N)$; la polarisée de cette variance

$$(X, Y) \longmapsto \mathrm{cov}\,(X, Y) := \sum_{j=1}^N p(j)\, x_j y_j - \Big(\sum_{j=1}^N p(j)\, y_j\Big)\, \Big(\sum_{j=1}^N p(j)\, x_j\Big)$$

  est dite *covariance de X et Y, pondérée par le choix de poids* (*ou probabilités*) $p(1), ..., p(N)$. On retrouvera ces notions dans la seconde partie du cours.

Si E est un $\mathbb{R}$-espace vectoriel de dimension finie, la matrice d'une forme bilinéaire dans une base $(e_1, ..., e_n)$ est par définition la matrice (symétrique) de la forme bilinéaire associée lorsque E est rapportée à cette base. Si P est la matrice de passage de la base $\widetilde{\mathcal{B}}$ à la base $\mathcal{B}$, la matrice $\widetilde{M}_\mathcal{Q}$ d'une forme quadratique $\mathcal{Q}$ exprimée dans la base $\widetilde{\mathcal{B}}$ se déduit de la matrice $M_\mathcal{Q}$ de cette même forme quadratique exprimée dans la base $\mathcal{B}$ *via* la relation

$$\widetilde{M}_\mathcal{Q} = {}^t P \bullet M_\mathcal{Q} \bullet P\,.$$

1.7.2 Orthogonalité relative à une forme bilinéaire symétrique

Si E est un $\mathbb{R}$-espace vectoriel et Θ une forme bilinéaire symétrique sur E, on dit que deux vecteurs v et w de E sont *orthogonaux* relativement à Θ (on dit aussi "relativement à la forme quadratique $\mathcal{Q}$ de polarisée Θ") si et seulement si

$$v \perp_\Theta w \Longleftrightarrow \Theta(v,w) = \frac{1}{2}\Big(\mathcal{Q}(v+w) - \mathcal{Q}(v) - \mathcal{Q}(w)\Big) = 0\,.$$

Remarque. Il peut fort bien arriver que des vecteurs non nuls soient orthogonaux à eux mêmes : par exemple, dans l'espace-temps $\mathbb{R}^4$, les vecteurs du cône :

$$\{(x,y,z,t)\,;\, \sqrt{x^2+y^2+z^2} = \sqrt{c}\,t\}$$

(dit "frontière du cône du futur") sont orthogonaux à eux-mêmes relativement à la forme de Lorentz ; de tels vecteurs sont dits *isotropes*.

Une forme bilinéaire symétrique Θ sur un $\mathbb{R}$-espace vectoriel E est appelée *produit scalaire* si :

- $\Theta(v,v) = \mathcal{Q}(v) \geq 0$ pour tout $v \in E$ (on dit que la forme est *positive*) ;
- $\Theta(v,v) = \mathcal{Q}(v) = 0$ si et seulement si $v = 0$ (on dit que la forme est *définie*).

Exemples et non-exemples de produits scalaires.

- le fait qu'une forme soit définie positive exclut l'existence de vecteurs isotropes non nuls ; la forme de Lorentz ne correspond donc pas à un produit scalaire sur $\mathbb{R}^4$;
- la covariance relative à un choix $\{p(1),...,p(N)\}$ de probabilités sur $\{1,...,N\}$ correspond à une forme positive sur $\mathbb{R}^N$ mais cette forme n'est pas définie par les vecteurs $(a,...,a)$ (correspondant aux fonctions constantes sur $\{1,...,N\}$) sont de variance nulle ;
- en revanche, l'énergie correspond bien à un produit scalaire sur l'espace des fonctions continues sur $[a,b]$:

$$(f,g) \longmapsto \int_a^b f(t)g(t)dt$$

  (dit *corrélation*) ;
- sur l'espace $\mathbb{R}^N$, la forme

$$\Big(X = (x_1,...,x_N)\,,\, Y = (y_1,...,y_N)\Big) \longmapsto \Big\langle X\,,\,Y\Big\rangle := \sum_{j=1}^N x_j y_j$$

  définit un produit scalaire dit *produit scalaire canonique sur* $\mathbb{R}^N$.

Un $\mathbb{R}$-espace vectoriel équipé d'un produit scalaire $\langle\,,\,\rangle$ est dit *espace euclidien*.

Tout $\mathbb{R}$-espace de dimension finie hérite toujours d'une structure euclidienne : étant donnée une base $(e_1,...,e_n)$ de E, il suffit, pour construire un produit scalaire sur E, de se donner n scalaires strictement positifs $\lambda_1,...,\lambda_n$ et de considérer l'application

$$\Big(\sum_{i=1}^n x_i\cdot e_i\,,\,\sum_{j=1}^n y_j\cdot e_j\Big) \longmapsto \sum_{i=1}^n \lambda_i\, x_i y_i\,;$$

tous les produits scalaires sur E peuvent d'ailleurs être réalisés ainsi.

Si E est un $\mathbb{R}$-espace vectoriel euclidien (donc muni d'un produit scalaire), une *base orthonormée* de E (relativement à cette structure euclidienne) est une base $(v_1,...,v_n)$ telle que

$$\langle v_i\,,\,v_j\rangle = \begin{cases} 0 \text{ si } i \neq j \\ 1 \text{ si } i = j \end{cases}$$

On verra plus loin comment, étant donné un espace euclidien E, construire une base orthonormée $(v_1, ..., v_n)$ à partir d'une base $(e_1, ..., e_n)$ donnée. Tout espace euclidien de dimension finie admet donc toujours au moins une base orthonormée (relativement au produit scalaire induisant la structure euclidienne).

L'intérêt de disposer d'une base orthonormée est de disposer de vecteurs constituant des informations non-redondantes. Le point noir cependant en est le caractère "fragile" ; on verra que la construction d'une base orthonormée $(v_1, ..., v_n)$ à partir d'une base $(e_1, ..., e_n)$ quelconque (ce qui est toujours possible si E est de dimension finie) peut conduire à la construction de vecteurs v_j tels que $\langle v_j\,,\, v_j\rangle = 1$ certes, mais au prix parfois d'écritures

$$v_j = \sum_{k=1}^{n} \lambda_{j,k} \cdot e_k\,,$$

où certaines des coordonnées $\lambda_{j,k}$ peuvent être très grandes ! les termes se compensent, mais la structure $(v_1, ..., v_n)$ s'en trouve fragilisée : une petite détérioration des e_j peut très vite "casser" le caractère orthonormé de la base $(v_1, ..., v_n)$!

Plutôt que de travailler dans une base orthonormée, il peut s'avérer plus judicieux parfois de disposer de systèmes redondants (plus "robustes" que des systèmes orthonormés). On peut décomposer un vecteur suivant un tel système redondant en utilisant l'idée de *matching* familière aux informaticiens (voir l'exemple 1.8 plus loin).

Si toutefois le produit scalaire est intimement lié à une expérience physique d'où le vecteur v est issu, alors disposer d'une base orthonormée pour ce produit scalaire est un solide atout pour décomposer v de manière "économique".

1.7.3 Projections orthogonales et applications

Soit E un $\mathbb{R}$-espace vectoriel (que nous supposerons ici de dimension finie) équipé d'un produit scalaire. Ce produit scalaire induit la définition d'une *norme* dans E, définie par

$$\|v\| := \sqrt{\langle v\,,\, v\rangle}\,,$$

puis, associée à cette norme, d'une *distance* entre éléments de E :

$$d(v_1, v_2) := \|v_1 - v_2\|\,.$$

Outre l'inégalité triangulaire :

$$\forall\, u, v, w \in E\,,\ \|w - u\| \leq \|w - v\| + \|v - u\|\,,$$

on dispose du théorème de Pythagore :

$$\forall\, v, w \in E\,,\ |v - w|^2 = |v|^2 + |w|^2 - 2\langle v\,,\, w\rangle$$

et de l'inégalité de Cauchy-Schwarz :

$$\forall\, v, w \in E\,,\ |\langle v\,,\, w\rangle| \leq \|v\| \times \|w\|\,.$$

Ce sont les divers outils dont on dispose pour travailler dans le cadre euclidien $(E, \langle\,,\,\rangle)$.

Soit A un sous-ensemble de E, de la forme

$$A = \{v \in E \, ; v = v_0 + u \, , \ u \in F\} \, ,$$

où v_0 est un élément de E et F un sous-espace vectoriel de E ; un tel sous-ensemble A est dit *sous-espace affine de E, passant par v_0 et dirigé par le sous-espace vectoriel F.*

Il faut imaginer A comme un espace de "modèles" (le nombre de degrés de liberté étant la dimension du sous-espace F). Étant donné un vecteur v quelconque de E, il est naturel de se demander s'il existe un vecteur modèle v_A dans A qui "colle" au mieux à A, c'est-à-dire qui approche au mieux v au sens de la distance

$$d(u, v) := \|u - v\| \, .$$

Un fait majeur est le suivant :

Théorème 1.6 *Si v est un vecteur quelconque de E, il existe un <u>unique</u> vecteur $\mathrm{pr}_A(v)$ de l'ensemble A réalisant le minimum de la distance de v aux éléments de A. Si A est pensé comme un ensemble de vecteurs modèles (dépendant de $\dim F$ degrés de liberté), ce vecteur $\mathrm{pr}_A(v)$ réalise, parmi les vecteurs "modèles" de A, le modèle le plus proche de v, celui que l'on retiendra si l'on cherche à approcher v par un élément de A de la forme $v_0 + u$, $u \in F$. Le vecteur $\mathrm{pr}_A(v)$, caractérisé par la propriété*

$$\forall w \in A \, , \ \langle v - \mathrm{pr}_A(v) \, , \, w \rangle = 0 \, , \tag{1.20}$$

est appelé <u>projection orthogonale de v sur le sous-espace affine $A = v_0 + F$</u>.

<u>Recherche pratique d'une projection orthogonale :</u> Supposons que F soit le sous-espace de dimension p engendré par les p vecteurs $f_1, ..., f_p$ (définisant un système libre) et que l'on cherche $\mathrm{pr}_A(v)$ sous la forme

$$\mathrm{pr}_A(v) - v_0 = \sum_{j=1}^{p} \lambda_j \cdot f_j \, ,$$

où les scalaires $\lambda_1, ..., \lambda_p$ sont tels (si l'on veut que les contraintes (1.20) soient satisfaites) que, pour tout $i = 1, ..., p$,

$$\Big\langle f_i \, , \, \sum_{j=1}^{p} \lambda_j \cdot f_j \Big\rangle = \langle v - v_0, f_i \rangle \, .$$

Ceci nous conduit à la résolution du système de p équations à p inconnues :

$$\sum_{j=1}^{p} \lambda_j \, \langle f_i \, , \, f_j \rangle = \langle v - v_0 \, , \, f_i \rangle \, , \ i = 1, ..., p \, .$$

Ce système s'écrit aussi

$$\mathrm{Gram}\,(f_1, ..., f_p) \bullet \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix} = \begin{pmatrix} \langle v - v_0 \, , \, f_1 \rangle \\ \vdots \\ \langle v - v_0 \, , \, f_p \rangle \end{pmatrix} \, ,$$

où Gram $(f_1, ..., f_p)$ (dite *matrice de Gram* de $(f_1, ..., f_p)$) est la matrice symétrique inversible dont l'entrée en position (i, j) vaut $\langle f_i\,,\, f_j\rangle$.

Ce système se résout suivant la méthode du pivot de Gauss (ce qui s'avère d'autant plus efficace que la matrice de Gram $G(f_1, ..., f_p)$ est "creuse", c'est-à-dire contient le plus d'entrées nulles possibles). Il se trouve d'ailleurs (on verra cela plus loin) qu'une telle matrice de Gram $G(f_1, ..., f_p)$ est diagonalisable comme matrice symétrique réelle, de plus dans une base orthonormée de E (les sous-espaces propres étant orthogonaux entre eux).

Exemple 1.5. Considérons le $\mathbb{R}$-espace vectoriel E constitué des fonctions de $[0, 2^N]$ dans $\mathbb{R}$, avec noeuds aux points $0, 1, ..., 2^N$ et V_2 le sous-espace vectoriel de E constitué des fonctions affines par morceaux de $[0, 2^N]$ dans $\mathbb{R}$, avec noeuds aux points $0, 2, 4, ..., 2^N$. On munit le $\mathbb{R}$-espace E du produit scalaire "de l'énergie"

$$(f, g) \longmapsto \int_0^{2^N} f(t)g(t)dt\,.$$

Le $\mathbb{R}$-espace vectoriel E est de dimension $2^N + 1$ et admet comme base le système constitué des fonctions "triangle"

$$t \in [0, 2^N] \longmapsto \Delta_j(t) := \max(0, 1 - |t - j|)\,,\ j = 0, ..., 2^N\,;$$

cette base n'est pas orthonormée car

$$\langle \Delta_0\,,\, \Delta_1\rangle = \int_0^1 t(1-t)\,dt = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}\,.$$

Le sous-espace vectoriel F admet, lui, comme base le système constitué des $2^{N-1} + 1$ fonctions

$$t \in [0, 2^N] \longmapsto \widetilde{\Delta}_j(t) := \max(0, 1 - 2|t - j|)\,,\ j = 0, ..., 2^{N-1}\,.$$

La matrice de Gram de ce système est la matrice

$$\begin{pmatrix} 1/3 & 1/6 & 0 & \dots & 0 & 0 \\ 1/6 & 2/3 & 1/6 & 0 & \dots & 0 \\ 0 & 1/6 & \cdot & \cdot & \cdots & 0 \\ \vdots & \cdot & \cdot & \dots & \dots & \vdots \\ 0 & \dots & 0 & 1/6 & 2/3 & 1/6 \\ 0 & \dots & 0 & 0 & 1/6 & 1/3 \end{pmatrix}.$$

L'inversion (ou même la diagonalisation) de la matrice de Gram $G(\widetilde{\Delta}_0, \widetilde{\Delta}_1, ..., \widetilde{\Delta}_{2^{N-1}})$ ne sont pas nécéssaires ici pour résoudre le système

$$G(\widetilde{\Delta}_0, \widetilde{\Delta}_1, ..., \widetilde{\Delta}_{2^{N-1}}) \bullet \Lambda = \begin{pmatrix} \langle v\,,\, \widetilde{\Delta}_0\rangle \\ \langle v\,,\, \widetilde{\Delta}_1\rangle \\ \langle v\,,\, \widetilde{\Delta}_2\rangle \\ \vdots \\ \langle v\,,\, \widetilde{\Delta}_{2^{N-1}}\rangle \end{pmatrix}$$

donnant (en colonne) le vecteur $\Lambda = (\lambda_0, \lambda_2, \lambda_2, ..., \lambda_{2^{N-1}})$ des coordonnées de $\mathrm{pr}_{V_2}(v)$ dans la base

$$(\widetilde{\Delta}_0, \widetilde{\Delta}_1, \widetilde{\Delta}_2, ..., \widetilde{\Delta}_{2^{N-1}})$$

de V_2 ; ce système se résoud en effet aisément de proche en proche. La décomposition d'un vecteur v sout la forme

$$v = \mathrm{pr}_{V_2}(v) + (v - \mathrm{pr}_{V_2}(v))$$

fournit une décomposition orthogonale de la fonction v en une fonction $\mathrm{pr}_{V_2}(v)$ que l'on pourrait qualifier de *résumé de l'information v à l'échelle* 2 et une fonction $v - \mathrm{pr}_{V_2}(v)$ que l'on pourrait qualifier de *détails de l'information v à l'échelle* 1. Cette opération pourrait d'ailleurs se poursuivre et l'on pourrait projeter $\mathrm{pr}_{V_2}(v)$ sur le sous-espace V_4 (de dimension 2^{N-2}, inclus dans V_2) des

fonctions affines par morceaux avec noeuds aux points $0, 4, 8, 12, ...$ On verrait ainsi surgir le résumé $\mathrm{pr}_{V_4}(v) = \mathrm{pr}_{V_4}(\mathrm{pr}_{V_2}(v))$ de v à l'échelle $2^2 = 4$ et les détails $\mathrm{pr}_{V_2}(v) - \mathrm{pr}_{V_4}(v)$ de v à l'échelle 2, *etc.* On obtient ainsi une décomposition orthogonale temps-échelles de l'information $v : [0, 2^N] \longmapsto \mathbb{R}$. Sur la figure ci-dessous, on a représenté (en bas) l'enregistrement d'une secousse sismique (matérialisé par un signal digital de longueur $2^{10} + 1 = 1025$, et l'on a figuré les graphes de $S - \mathrm{pr}_{V_2}(S)$, $\mathrm{pr}_{V_2}(S) - \mathrm{pr}_{V_4}(S)$, *etc.*, successivement de bas en haut, soit les détails successifs aux échelles 2^0, 2^2, ..., ce qui permet de visualiser le contenu de l'information à la fois en temps et en échelles (comme si on se reculait progressivement pour voir le signal de plus en plus loin). Ces divers signaux s'ajoutent pour constituer une décomposition orthogonale de l'information S (le produit scalaire étant celui correspondant au produit scalaire canonique sur $\mathbb{R}^{1025}$).

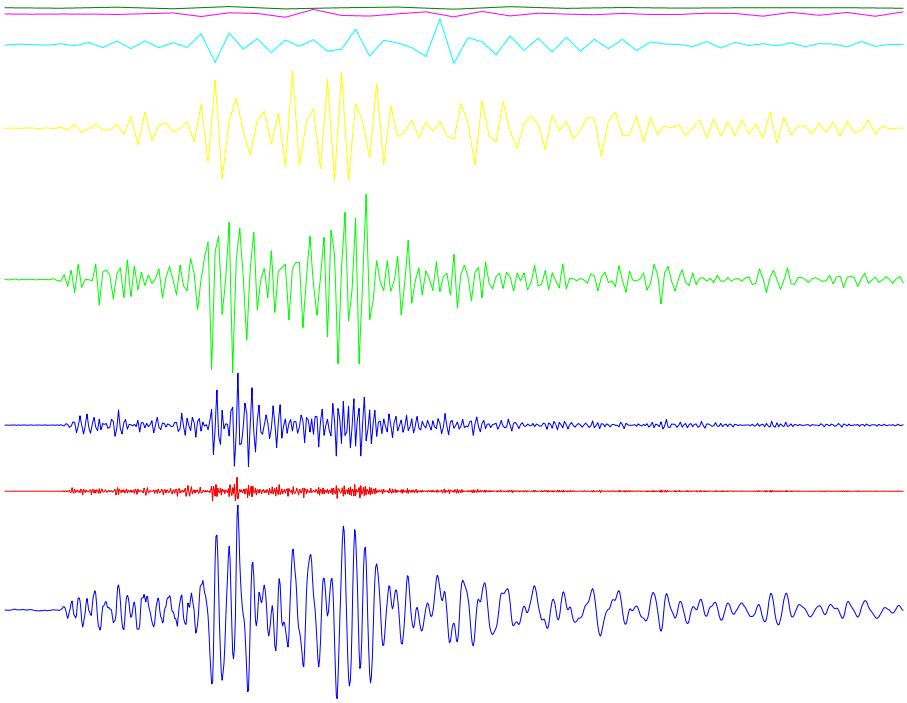

FIGURE 1.13 – Analyse temps-échelles d'une secousse sismique

Exemple 1.6. Soit E l'ensemble des fonctions de $\{1, ..., m + N\}$ dans $\mathbb{R}$ et

$$S = (s(1), ..., s(m), s(m + 1), ..., s(m + N))$$

un élément de E. On munit naturellement E du produit scalaire

$$\langle s_1 \, , \, s_1 \rangle := \frac{1}{N + m} \sum_{j=1}^{N+m} s_1(j) \, s_2(j) \, .$$

Il est souvent important (en particulier en théorie du signal ou du contrôle) de voir comment l'information $(s(1), ..., s(m), s(m + 1), ..., s(m + N))$, prolongée à droite par des zéros, est "auto-corrélée" avec elle-même dans le temps. Pour cela, on considère le sous-espace vectoriel $A = F$ de E engendré par les m vecteurs

$$S_j = (s(m + 1 - j), ..., s(m + N), 0, ..., 0) \, , \; j = 0, ..., m - 1$$

qui correspondent respectivement à l'information S, décalée dans le passé de m crans, de $m - 1$ crans,..., d'un cran. En cherchant la projection orthogonale

$$\mathrm{pr}_F(S) = \sum_{j=0}^{m-1} \alpha_j \cdot S_j \, ,$$

de S sur F, on cherche à déterminer les scalaires $\alpha_0, ..., \alpha_{m-1}$ tels que la quantité

$$\sum_{k=m+1}^{m+N} |S(k) - \alpha_{m-1}S(k-1) - \cdots - \alpha_0 S(k-m)|^2$$

soit minimale, c'est-à-dire tels que l'on puisse affirmer que l'information S soit au mieux assujettie à la règle d'autocorrélation

$$S(k) \simeq \alpha_{m-1}S(k-1) + \alpha_{m-2}S(k-2) + \cdots + \alpha_0 S(k-m)\,,\ k = m+1, ..., m+N\,.$$

La connaissance de tels paramètres α_j s'avère intéressante pour la réalisation d'appareils susceptibles de "décorréler" l'entrée S, c'est-à-dire de la transformer au mieux en ce que l'on appelle en théorie de l'information ou en télécommunications un *bruit blanc*. On retrouve de telle idées en traitement digital de l'information, par exemple en téléphonie mobile.

Exemple 1.7 : nuages de points dans le plan, droite de régression, principe des moindres carrés. Voici une autre application importante des idées "pythagoriciennes" dans le plan euclidien ; considérons, comme sur la figure ci-dessous, un "nuage de points",

$$\{(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)\}$$

(il y en a six sur notre figure) correspondant par exemple aux mesures simultanées de deux phénomènes physiques dont on veut savoir quel est le meilleur compromis possible pour les supposer linéairement dépendants (ce qu'ils ne sont bien sûr à première vue rigoureusement pas, comme la figure nous le confirme, sinon tous les points seraient alignés).

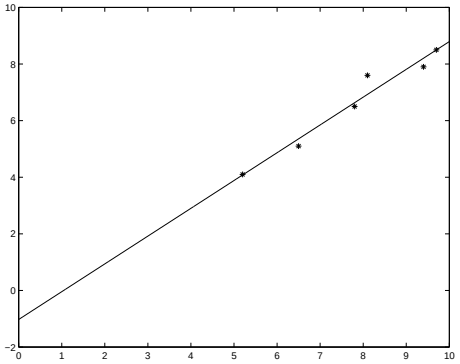

FIGURE 1.14 – Tracé d'une droite de régression linéaire

Une droite est intéressante à rechercher : c'est la droite d'équation affine $y - ax - b = 0$ telle que la quantité

$$F(a, b) := (y_1 - ax_1 - b)^2 + \cdots + (y_N - ax_N - b)^2$$

soit minimale (si elle existe). Introduisons les moyennes m_x et m_y des valeurs respectives des x_j et y_j, soit

$$m_x := \frac{x_1 + \cdots + x_N}{N}$$

$$m_y := \frac{y_1 + \cdots + y_N}{N}$$

et posons $x'_j = x_j - m_x$ et $y'_j = y_j - m_y$ pour $j = 1, ..., N$. On a $\sum_{j=1}^{N} x'_j = \sum_{j=1}^{N} y'_j = 0$. Si nous trouvons a' et b' minimisant la fonction

$$J(a', b') := (y'_1 - a'x'_1 - b')^2 + \cdots + (y'_N - a'x'_N - b')^2\,,$$

on en déduira que les valeurs de a et b minimisant $F(a,b)$ sont

$$a = a' \,, \quad b = b' + m_y - a' m_x \,.$$

Un calcul simple montre que

$$\begin{aligned} J(a',b') &= a'^2 \sum_{j=1}^N x_j'^2 - 2a \sum_{j=1}^N x_j' y_j' + N b'^2 + \sum_{j=1}^N y_j'^2 \\ &= \Big(a' \sqrt{\sum_{j=1}^N x_j'^2} - \frac{\sum_{j=1}^N x_j' y_j'}{\sqrt{\sum_{j=1}^N x_j'^2}} \Big)^2 + N b'^2 + \sum_{j=1}^N y_j'^2 - \frac{\Big(\sum_{j=1}^N x_j' y_j' \Big)^2}{\sum_{j=1}^N x_j'^2} \,; \end{aligned}$$

le minimum de cette fonction est donc atteint pour

$$a' = \frac{\sum_{j=1}^N x_j' y_j'}{\sqrt{\sum_{j=1}^N x_j'^2}} \,, \quad b' = 0$$

et vaut

$$\min J(a',b') = \min F(a,b) = \sum_{j=1}^N y_j'^2 - \frac{\Big(\sum_{j=1}^N x_j' y_j' \Big)^2}{\sum_{j=1}^N x_j'^2}$$

(qui, remarquons-le, est forcément une quantité positive ou nulle, d'après l'*inégalité de Cauchy-Schwarz*).

La droite affine réalisant le meilleur compromis concernant la dépendance linéaire de l'information y à partir de l'information x au sein du nuage de points est donc la droite affine d'équation

$$y - m_y = a'(x - m_x) \,;$$

cette droite importante est dite *droite de régression linéaire* et le nombre

$$\rho := \frac{\sum_{j=1}^n x_j' y_j'}{\sqrt{\sum_{j=1}^N x_j'^2} \sqrt{\sum_{j=1}^N y_j'^2}}$$

est dit *coefficient de corrélation entre les x_j et les y_j* au sein du nuage. Ainsi la droite de régression linéaire a-t-elle pour équation cartésienne

$$\frac{y - m_y}{\sqrt{\sum_{j=1}^N y_j'^2}} = \rho \frac{x - m_x}{\sqrt{\sum_{j=1}^N x_j'^2}} \,.$$

Ces notions jouent un rôle très important en théorie des probabilités et plus particulièrement dans l'étude des modèles statistiques (les enquêtes d'opinion par exemple).

Ce principe peut être généralisé si l'on prédit l'existence d'une liste de M (avec M en général très petit devant N) fonctions réelles "modèle" $\varphi_1, ..., \varphi_M$ et de paramètres scalaires $a_1, ..., a_M$ tels que :

$$y_j \simeq a_1 \varphi_1(x_j) + \cdots + a_M \varphi_M(x_j) \,, \ j = 1, ..., N \,.$$

Les vecteurs $v_k = (\varphi_k(x_1), ..., \varphi_k(x_N)) \in \mathbb{R}^N$, $k = 1, ..., M$, engendrent un sous-espace (ici vectoriel et non plus affine) $A = F$ de $\mathbb{R}^N$ sur lequel il s'avère judicieux de projeter orthogonalement le vecteur $Y = (y_1, ..., y_N)$; on a alors

$$\mathrm{pr}_F(Y) = \sum_{k=1}^N a_j \cdot v_k$$

et les scalaires $(a_1, ..., a_N)$ ainsi obtenus sont exactement ceux qui minimisent la fonctionnelle

$$(\lambda_1, ..., \lambda_N) \longmapsto J(\lambda_1, ..., \lambda_M) := \sum_{j=1}^{N} (y_j - a_1\varphi_1(x_j) - \cdots - a_N\varphi_N(x_j))^2 \,,$$

ce du fait de la propriété caractéristique (1.20) de la projection orthogonale sur le sous-espace (ici vectoriel) $A = F$. Si les vecteurs $v_1, ..., v_M$ forment une famille libre, la matrice de Gram $G(v_1, ..., v_M)$ est inversible et le calcul de $(a_1, ..., a_M)$ se fait *via* la formule matricielle

$$\begin{pmatrix} a_1 \\ \vdots \\ a_M \end{pmatrix} = [G(v_1, ..., v_M)]^{-1} \bullet \begin{pmatrix} \sum_{j=1}^{N} y_j\varphi_1(x_j) \\ \vdots \\ \sum_{j=1}^{N} y_j\varphi_M(x_j) \end{pmatrix} ,$$

jolie formule certes, mais à laquelle il est souvent préférable de substituer la méthode algorithmique du pivot de Gauss pour résoudre le système linéaire de M équations à M inconnues $(a_1, ..., a_M)$:

$$\begin{aligned} \sum_{k=1}^{M} a_k \Big(\sum_{j=1}^{N} \varphi_k(x_j)\varphi_1(x_j) \Big) &= \sum_{j=1}^{N} y_j\varphi_1(x_j) \\ \vdots &= \vdots \\ \sum_{k=1}^{M} a_k \Big(\sum_{j=1}^{N} \varphi_k(x_j)\varphi_M(x_j) \Big) &= \sum_{j=1}^{N} y_j\varphi_M(x_j) \,. \end{aligned}$$

Exemple 1.8 : modes propres, algorithmes de *matching***.** Supposons que l'on dispose d'une collection finie $(v_j)_{j=1,...,N}$ de vecteurs d'un espace euclidien E (de dimension finie n, supposée strictement plus grande que N). La matrice de Gram $G(v_1, ..., v_N)$ est une matrice symétrique réelle. Admettons (on verra plus loin que c'est toujours le cas) que cette matrice est diagonalisable (comme matrice à entrées réelles). Supposons aussi (ceci n'est pas toujours le cas, mais l'est avec une probabilité 1 si l'on s'octroie une marge d'erreur pour la détermination numérique des v_j dans une base donnée) que les valeurs propres (toutes réelles) soient de modules distincts, rangés en ordre décroissant :

$$|\lambda_1| > |\lambda_2| > \cdots |\lambda_N| \,.$$

Associé à chaque valeur propre λ_j, $j = 1, ..., N$, on dispose d'un vecteur propre u_j de norme 1, unique à un facteur signe près, correspondant à la valeur propre λ_j. À chacun de ces vecteurs

$$u_j = (u_{j,1}, ..., u_{j,N}) \in \mathbb{R}^N \,,$$

on peut associer un vecteur particulier du sous-espace vectoriel de E engendré par $v_1, ..., v_N$, à savoir le vecteur

$$w_j := \sum_{k=1}^{N} u_{j,k} \cdot v_k \,.$$

Compte-tenu du fait que λ_1 correspond à la plus grande valeur propre de la matrice de Gram $G(v_1, ..., v_N)$ (mesurant les corrélations entre les vecteurs v_j, $j = 1, ..., N$), le "motif" w_1 correspond au vecteur du sous-espace engendré par $v_1, ..., v_N$ qui (toujours au sens des moindres carrés, comme dans l'exemple 1.7) "colle" le mieux statistiquement à tous les v_j. Viennent ensuite, dans cet ordre, les vecteurs w_2, w_3,... Plus les "trous" entre les $|\lambda_j|$ sont profonds, moins redondante sera l'information contenue dans la liste $w_1, w_2, w_3, ...$ des premiers vecteurs w_j. Si non nuls, les vecteurs w_j, $j = 1, ..., N$, sont appelés *modes propres* de la famille $(v_j)_{j=1,...,N}$. Ces modes propres fournissent un moyen de concentrer en un "'dictionnaire" plus court tout un ensemble de données $v_1, ..., v_N$ (éventuellement parfois redondantes).

Pour donner un exemple, supposons que l'espace E corresponde à un espace vectoriel d'images médicales et que $v_1, ..., v_N$ corresponde à un dictionnaire d'images correspondant chacune à une pathologie (observée chacune sur un malade pris dans une liste de N individus). Pour tester si une image prise sur un nouveau patient (hors de la liste) est ou non une image pathologique et comment,

il est plus efficace de ne tester cette image que contre les premiers modes propres $w_1, w_2, ...$, de la famille d'images $v_1, ..., v_N$ (ces premiers modes propres ont déjà intégré les redondances de cette famille !).

Le test d'un vecteur f contre un dictionnaire de vecteurs $(e_1, e_2, ..., e_M)$ (tous de norme 1) se fait suivant un algorithme de *matching* cher aux informaticiens :

- on calcule tous les produits scalaires $\langle f\ ,\ e_j\rangle$, $j = 1, ..., M$ et on repère celui $(\langle f\ ,\ e_{j_1}\rangle)$ le plus grand en valeur absolue ;
- on forme $f_1 = f - \langle f\ ,\ e_{j_1}\rangle \cdot e_{j_1}$ (projection de f sur l'hyperplan vectoriel orthogonal à e_{j_1}), puis on recommence la première opération avec f_1 à la place de f...
- On voit ainsi apparaître une décomposition "évolutive" de f :

$$f = \langle f\ ,\ e_{j_1}\rangle \cdot e_{j_1} + \langle f\ ,\ e_{j_2}\rangle \cdot e_{j_2} + \cdots$$

qui fournit un "pistage" du vecteur f contre le dictionnaire $(e_1, ..., e_M)$ proposé.

Un tel algorithme, dit *glouton*, s'avère dans de nombreuses questions pratiques du monde expérimental, particulièrement efficace, surtout lorsque couplé avec la recherche préalable d'un dictionnaire moins redondant de modes propres associé à une famille (elle redondante) de vecteurs donnés dans un espace euclidien E.

Les algorithmes de projection orthogonale peuvent être itérés ; on donne ici deux exemples de résultats auquels ce type de démarche conduit.

Exemple 1.9. Soient F_1 et F_2 deux sous-espaces vectoriels d'un espace euclidien E de dimension finie. On suppose, comme sur la figure ci-dessous, que F_1 et $F_2^\perp$ (sous-espace constitué des vecteurs orthogonaux à F_2) font un angle non nul , c'est-à-dire, concrètement, que

$$\forall v_1 \in F_1\ ,\ \forall v_2 \in F_2^\perp\ ,\ |\langle v_1\ ,\ v_2\rangle| \leq \rho\, \|v_1\|\, \|v_2\|$$

avec $\rho < 1$. Alors, en "copiant" un argument inspiré de la géométrie euclidienne classique (dans le plan ou l'espace), on sait retrouver un vecteur inconnu s de F_1 en partant de sa projection orthogonale (supposée elle connue) s_0 sur F_2. C'est le mécanisme itératif $s_0 \longmapsto s_1 \longmapsto s_2 \longmapsto ...$ (utilisant alternativement les projections orthogonales sur F_1 et sur $s_0 + F_2^\perp = s + F_2^\perp$) décrit sur la figure ci-dessous qui rend compte de ce scénario d'usage fréquent dans nombre de problèmes pratiques.

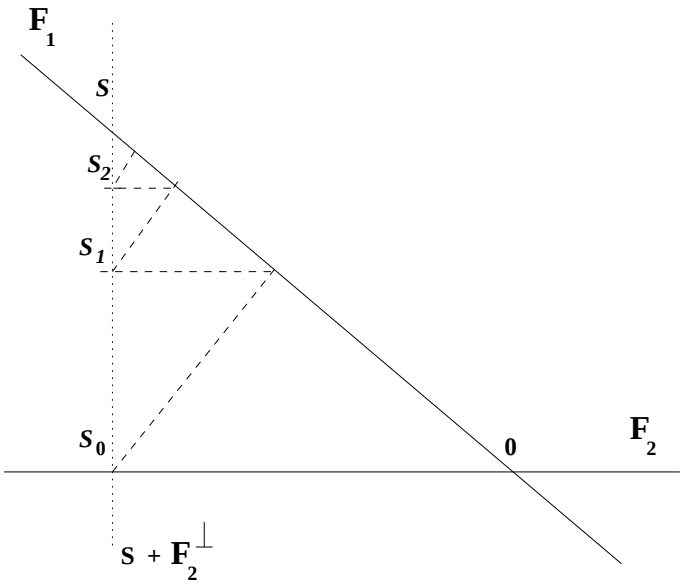

FIGURE 1.15 – Un scénario de "projections orthogonales itérées"

Exemples 1.10. Si $A_1, ..., A_N$ sont N sous-espaces vectoriels affines d'un $\mathbb{R}$-espace vectoriel eu-

clidien E de dimension finie, de la forme

$$A_j = \{v_0 + v\,;\, v \in F_j\}$$

et si l'on considère la suite issue d'un vecteur quelconque u de E

$$u \longmapsto Q(u) \longmapsto Q[Q(u)] \longmapsto \dots\,,$$

où

$$Q = \mathrm{pr}_{A_N} \circ \cdots \circ \mathrm{pr}_{A_1}\,,$$

approche la projection orthogonale de V sur le sous-espace affine

$$A = \{v_0 + v\,;\, v \in F_1 \cap F_2 \cdots \cap F_N\}\,;$$

pour peu que l'intersection de $F_1 \cap \cdots \cap F_N$ soit réduite au vecteur nul, on dispose ici d'un moyen d'atteindre v_0 à partir d'un scénario itératif utilisant les projections orthogonales sur les sous-espaces affines $A_1, ..., A_N$. La figure ci-dessous illustre ce scénario ("en spirale") lui aussi très utilisé dans les démarches expérimentales.

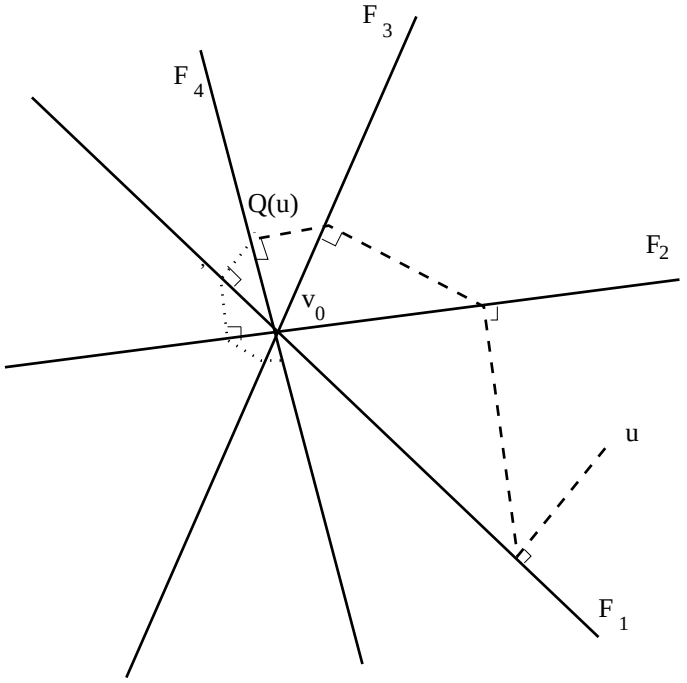

FIGURE 1.16 – Un second scénario de "projections orthogonales itérées" ($N = 4$)

Comment construire une base orthonormée à partir d'une base quelconque ?

Si E est un $\mathbb{R}$-espace vectoriel de dimension finie équipé d'un produit scalaire $\langle\ ,\ \rangle$ et si $(e_1, ..., e_n)$ est une base de E, voici comment on construit une base orthonormée $(v_1, ..., v_n)$ de E (relativement à ce produit scalaire) :

- on démarre avec $v_1 = e_1/\|e_1\|$ (on rappelle que $\|u\| := \sqrt{\langle u, u\rangle}$) ;
- on projette e_2 sur la droite vectorielle $\mathbb{R} \cdot v_1 = \mathbb{R} \cdot e_1$, soit

$$\mathrm{pr}_{\mathbb{R}\cdot v_1}(e_2) = \langle e_2\,,\, v_1\rangle \cdot v_1\,,$$

puis l'on pose

$$v_2 := \frac{e_2 - \mathrm{pr}_{\mathbb{R}\cdot v_1}(e_2)}{\|e_2 - \mathrm{pr}_{\mathbb{R}\cdot v_1}(e_2)\|} = \frac{e_2 - \langle e_2\,,\, v_1\rangle \cdot v_1}{\left\|e_2 - \langle e_2\,,\, v_1\rangle \cdot v_1\right\|}\ ;$$

– on projette e_3 sur le plan vectoriel engendré par v_1 et v_2 (qui d'ailleurs est aussi le plan vectoriel engendré par e_1 et e_2), ce qui donne :

$$\mathrm{pr}_{\mathbb{R}\cdot v_1+\mathbb{R}\cdot v_2}(e_3) = \langle e_3\,,\,v_1\rangle\cdot v_1 + \langle e_3\,,\,v_2\rangle\cdot v_2\,,$$

puis l'on pose

$$\begin{aligned} v_3 &:= \frac{e_3 - \mathrm{pr}_{\mathbb{R}\cdot v_1+\mathbb{R}\cdot v_2}(e_3)}{\|e_3 - \mathrm{pr}_{\mathbb{R}\cdot v_1+\mathbb{R}\cdot v_2}(e_3)\|} \\ &:= \frac{e_3 - \langle e_3\,,\,v_1\rangle\cdot v_1 - \langle e_3\,,\,v_2\rangle\cdot v_2}{\Big\|e_3 - \langle e_3\,,\,v_1\rangle\cdot v_1 - \langle e_3\,,\,v_2\rangle\cdot v_2\Big\|}\,; \end{aligned}$$

– on continue de la sorte jusqu'à la construction de $v_1, ..., v_n$.

C'est dans l'opération de normalisation décrite dès la seconde étape (obligeant à diviser un vecteur de norme pouvant être éventuellement très petite par sa norme) que peut résider la fragilité de la base orthogonale construite (les v_j peuvent s'exprimer dans la base $(e_1, ..., e_n)$ avec des coordonnées pouvant être en valeur absolue très grandes, même si $\|v_j\| = 1$ par "compensation" !) Ce procédé est néanmoins très utile : c'est le procédé *d'orthonormalisation de Gram-Schmidt*.

1.7.4 Le cas complexe : formes sesquilinéaires et hermitiennes, espaces hermitiens

Formes sequilinéaires sur un $\mathbb{C}$-espace vectoriel complexe

La forme bilinéaire canonique sur $\mathbb{C}^N$:

$$\Big((z_1, ..., z_N)\,,\,(w_1, ..., w_N)\Big) \longmapsto \sum_{j=1}^{N} z_j w_j$$

admet toujours des vecteurs isotropes (c'est-à-dire des vecteurs $(z_1, ..., z_n)$ tels que $z_1^2 + \cdots + z_n^2 = 0$). C'est le cas d'ailleurs de toute forme bilinéaire symétrique sur un $\mathbb{C}$-espace vectoriel de dimension finie. C'est là la raison majeure pour laquelle la notion de forme bilinéaire cesse d'être une notion intéressante lorsque l'on passe du cadre des $\mathbb{R}$-espaces vectoriels à celui des $\mathbb{C}$-espaces vectoriels. On lui préfère la notion de *forme sesquilinéaire*.

Une *forme sesquilinéaire* sur un $\mathbb{C}$-espace vectoriel E est une application

$$\Theta\ : E\times E \longrightarrow \mathbb{C}$$

qui est $\mathbb{C}$-linéaire par rapport à la première entrée, c'est-à-dire :

$$\forall\, v, \tilde{v}, w \in E\,, \forall\, \lambda, \tilde{\lambda} \in \mathbb{C}\,,\ \Theta(\lambda\cdot v + \tilde{\lambda}\cdot \tilde{v}) = \lambda\Theta(v,w) + \tilde{\lambda}\Theta(\tilde{v},w)$$

et $\mathbb{C}$-*anti-linéaire* par rapport à la seconde entrée, c'est-à-dire :

$$\forall\, v, w, \tilde{w} \in E\,,\ \forall\, \mu, \tilde{\mu} \in \mathbb{C}\,,\ \Theta(v, \mu\cdot w + \tilde{\mu}\cdot \tilde{w}) = \overline{\mu}\Theta(v,w) + \overline{\tilde{\mu}}\Theta(v,\tilde{w})\,,$$

où

$$\overline{\alpha + i\beta} := \alpha - i\beta$$

désigne la conjugaison complexe.

Une forme sesquilinéaire Θ sur un $\mathbb{C}$-espace vectoriel E vérifie la *symétrie hermitienne* si et seulement si

$$\forall\, v, w \in E\,, \Theta(w, v) = \overline{\Theta(v, w)}\,.$$

Toute forme sesquilinéaire Θ sur un $\mathbb{C}$-espace vectoriel E induit, dès qu'elle vérifie la symétrie hermitienne, ce que l'on appelle une *forme hermitienne* H sur E, suivant la relation

$$\forall\, v \in E\,,\ H(v) = \Theta(v, v)\,;$$

cette forme hermitienne est une application de E dans $\mathbb{R}$. On peut d'ailleurs reconstruire alors la forme sesquilinéaire Θ (dite *forme polaire* de la forme hermitienne H) *via* la relation :

$$\Theta(v, w) = \frac{H(v + w) - H(v - w) + iH(v + iw) - iH(v - iw)}{4}\,.$$

Les formes sesquilinéaires vérifiant la symétrie hermitienne sont le pendant (dans le cadre des $\mathbb{C}$-espaces vectoriels) des formes bilinéaires symétriques dans le contexte réel, tandis que les formes hermitiennes constituent, elles, le pendant (toujours dans le cadre des $\mathbb{C}$-espaces vectoriels) des formes quadratiques dans ce même contexte réel.

On appelle enfin *produit scalaire* sur un $\mathbb{C}$-espace vectoriel E toute forme sesquilinéaire $(z, w) \longmapsto \langle z\,,\, w\rangle$ sur E, vérifiant la symétrie hermitienne, et de plus telle que :

- $\langle v\,,\, v\rangle := \|v\|^2 \geq 0$ pour tout vecteur v de E (la forme est dite *positive*) ;
- $\|v\| = \sqrt{\langle v\,,\, v\rangle} = 0$ si et seulement si $v = 0$ (la forme est *définie*).

Un produit scalaire sur un $\mathbb{C}$-espace vectoriel E est donc une forme sesquilinéaire vérifiant la symétrie hermitienne et de plus définie positive.

Un $\mathbb{C}$-espace vectoriel E équipé d'un produit scalaire est dit $\mathbb{C}$-*espace vectoriel hermitien (ou hilbertien)*. Outre le fait que l'on y dispose de la formule de Pythagore :

$$\|v - w\|^2 = \|v\|^2 + \|w\|^2 - 2\mathrm{Re}\,\langle v\,,\, w\rangle\,,$$

l'inégalité de Cauchy-Schwarz :

$$\forall\, v, w \in E\,,\ |\langle v\,,\, w\rangle| \leq \|v\| \times \|w\|$$

(avec égalité uniquement lorsque les vecteurs v et w sont liés par une relation $\lambda \cdot v + \mu \cdot w = 0$) y joue un rôle essentiel. On peut, on le verra, conduire dans un tel cadre hilbertien, des raisonnements géométriques directement inspirés de la géométrie "à la Pythagore" dans le plan ou l'espace.

Exemples.

- La forme sesquilinéaire sur $\mathbb{C}^N$:

$$\Big((z_1, ..., z_N)\,,\, (w_1, ..., w_N)\Big) \longmapsto \sum_{j=1}^{N} z_j \overline{w}_j$$

vérifie la symétrie hermitienne ; c'est même un produit scalaire, dit *produit scalaire canonique* sur $\mathbb{C}^N$; si $\lambda_1, ..., \lambda_N$ sont N scalaires complexes, la forme sesquilinéaire sur $\mathbb{C}^N$

$$\Big((z_1, ..., z_N)\,,\,(w_1, ..., w_N)\Big) \longmapsto \sum_{j=1}^{N} \lambda_j z_j \overline{w}_j$$

ne définit un produit scalaire que si $\lambda_j > 0$ pour tout $j = 1, ..., N$.

– Si E désigne le $\mathbb{C}$-espace vectoriel des fonctions continues de $[a, b]$ dans $\mathbb{C}$ et $w : [a, b] \longrightarrow]0, +\infty[$ une fonction continue strictement positive sur $[a, b]$, l'application

$$(f, g) \longmapsto \int_a^b w(t) f(t) \overline{g(t)}\, dt$$

définit un produit scalaire sur E ; la forme hermitienne correspondante

$$f \longmapsto \int_a^b w(t)\, |f(t)|^2\, dt$$

correspond à l'énergie (pondérée par le "poids" w) ;

– Si E_N désigne le $\mathbb{C}$-espace vectoriel des polynômes trigonométriques

$$P(t) = \sum_{k=-N}^{N} c_k\, e^{ikt}$$

de degré inférieur ou égal à N et à coefficients complexes (espace engendré par les harmoniques fondamentales relatives à la période 2π en dessous du seuil de fréquence N), l'application

$$(P, Q) \longmapsto \sum_{k=-N}^{N} c_k(P)\, \overline{c_k(Q)}$$

est un produit scalaire sur le $\mathbb{C}$-espace vectoriel E_N ; la forme hermitienne correpondante associe à $P \in E_N$ la quantité $\sum\limits_{k=-N}^{N} |c_k(P)|^2$.

Projection orthogonale sur un sous-espace affine

Soit E un $\mathbb{C}$-espace vectoriel hermitien de dimnsion finie, dans lequel on note $\langle\ ,\ \rangle$ le produit scalaire et $\|\ \| = \sqrt{\langle\ ,\ \rangle}$ la norme associée.

Soit F un sous-espace vectoriel de E de dimension $p < \dim E$ et v_0 un élément de E. On a exactement le pendant du théorème 1.6 dans le contexte complexe cette fois, à savoir :

Théorème 1.7 *Soit v un vecteur de E n'appartenant pas au sous-espace affine*

$$v_0 + F = \{v_0 + u\,;\, u \in F\}\,.$$

Il existe un unique vecteur $\mathrm{pr}_{v_0+F}(v)$ du sous-espace affine $v_0 + F$ tel que

$$\|v - \mathrm{pr}_{v_0+F}(v)\| = \min_{u \in F} \|v - (v_0 + u)\|\,;$$

ce vecteur, dit projection orthogonale de v sur $v_0 + F$ est caractérisé par les conditions

$$\forall u \in F\,,\ \langle u\,,\, v - v_0\rangle = \langle u\,,\, \mathrm{pr}_{v_0+F}(v) - v_0\rangle\,;$$

si $(f_1, ..., f_p)$ est une base de F, les (uniques) scalaires $(\lambda_1, ..., \lambda_p)$ tels que

$$\mathrm{pr}_{v_0+F}(v) = v_0 + \sum_{j=1}^{p} \lambda_j \cdot f_j$$

sont solutions du système linéaire de p équations à p inconnues

$$G(f_1,...,f_p)\bullet\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_p\end{pmatrix}=\begin{pmatrix}\langle v-v_0\,,\,f_1\rangle\\ \vdots\\ \langle v-v_0\,,\,f_p\rangle\end{pmatrix},$$

où $G(f_1,...,f_p)$ est la matrice de Gram de $(f_1,...,f_p)$ définie par

$$G_{i,j}=\langle f_i\,,\,f_j\rangle\,,\ 1\leq i,j\leq p\,.$$

Si $v\in v_0+F$, alors on pose $\mathrm{pr}_{v_0+F}(v)=v$.

Ce théorème a exactement le même type de conséquences que le théorème 1.6 dans le cas des $\mathbb{R}$-espaces euclidiens (notons d'ailleurs que l'énoncé est identique). Suivant exactement le même scénario que dans un espace euclidien, on peut, étant donné un système libre $(e_1,...,e_p)$ de E, construire p vecteurs $v_1,...,v_p$ dans le sous-espace vectoriel engendré par $e_1,...,e_p$ tels que

$$\forall i,j\in\{1,...,p\}\,,\ \langle v_i\,,\,v_j\rangle=\begin{cases}1\text{ si }i=j\\ 0\text{ sinon}\end{cases};$$

on calcule pour cela les v_k de proche en proche en posant $v_1=e_1/\|e_1\|$ et

$$v_k=\frac{e_k-\mathrm{pr}_{\mathbb{C}\cdot e_1+\cdots+\mathbb{C}\cdot e_{k-1}}(e_k)}{\|e_k-\mathrm{pr}_{\mathbb{C}\cdot e_1+\cdots+\mathbb{C}\cdot e_{k-1}}(e_k)\|}=\frac{e_k-\sum_{j=1}^{k-1}\langle e_k\,,\,v_j\rangle\cdot v_j}{\left\|e_k-\sum_{j=1}^{k-1}\langle e_k\,,\,v_j\rangle\cdot v_j\right\|}\,.$$

Un tel système $(v_1,...,v_p)$ est dit *orthonormé* ; si de plus $p=\dim E$, on dit que c'est une *base orthonormée*. Le procédé algorithmique décrit ci-dessus pour construire un système orthonormé à partir d'un système libre est le *procédé de Gram-Schmidt*.

Matrice d'une forme sesquilinéaire dans une base ; matrices hermitiennes

Si E est un $\mathbb{C}$-espace vectoriel de dimension finie n, Θ une forme sesquilinéaire sur E et $\mathcal{B}=(e_1,...,e_n)$ une base de E, on vérifie aisément en utilisant la sesquilinéarité que

$$\Theta\Big(\sum_{i=1}^n z_i\cdot e_i\,,\,\sum_{j=1}^n w_j\cdot e_j\Big)=(z_1,...,z_n)\bullet M_{\Theta,\mathcal{B}}\bullet\begin{pmatrix}\overline{w_1}\\ \cdots\\ \overline{w_n}\end{pmatrix},$$

où $M_{\Theta,\mathcal{B}}$ est la matrice à entrées complexes dont l'entrée au carrefour de la i-ème ligne et de la j-ème colonne vaut $\Theta(e_i,e_j)$. Cette matrice est dite *matrice de Θ dans la base $\mathcal{B}$*. Si l'on change de base, on a

$$M_{\Theta,\widetilde{\mathcal{B}}}={}^tP\bullet M_{\Theta,\mathcal{B}}\bullet\overline{P}\,,$$

où P désigne la matrice de passage de la base $\widetilde{\mathcal{B}}$ dans la base $\mathcal{B}$ (les colonnes de P correspondent aux coordonnées des vecteurs de $\widetilde{\mathcal{B}}$, exprimés dans la base $\mathcal{B}$).

Dire que la forme Θ vérifie la symétrie hermitienne équivaut à dire que sa matrice $M_{\Theta,\mathcal{B}}$ dans une base quelconque est une *matrice hermitienne*, c'est-à-dire une matrice A de taille (n,n) à entrées complexes telle que

$$a_{j,i}=\overline{a_{i,j}}\,.$$

On appelle alors *matrice de la forme hermitienne H associée à Θ dans la base $\mathcal{B}$* précisément cette matrice hermitienne $M_{\Theta,\mathcal{B}}$.

Dualité dans un $\mathbb{C}$-espace hilbertien de dimension finie

Soit E un $\mathbb{C}$-espace hilbertien de dimension n où l'on note $\langle\ ,\ \rangle$ le produit scalaire. On se fixe (ceci est possible) une base orthonormée de E (relativement à ce produit scalaire), $(v_1, ..., v_n)$.

Soit $L\ : E \longrightarrow E$ une application $\mathbb{C}$-linéaire de E dans lui-même, de matrice $A = [a_{i,j}]_{1\leq i,j\leq n}$ dans la base $(v_1, ..., v_n)$, ce qui signifie

$$L(v_j) = \sum_{i=1}^{n} a_{i,j} \cdot v_i\,.$$

Comme le système est $(v_1, ..., v_n)$ est orthonormé, on a

$$\langle L(v_j)\,,\ v_i\rangle = a_{i,j}\ \forall\, i, j \in \{1, ..., n\}\,.$$

On peut définir un nouvel opérateur $L^*\ : E \longrightarrow E$ lié à L par les relations

$$\langle L(v_j)\,,\ v_i\rangle = \langle v_j\,,\ L^*(v_i)\rangle\,,\ 1 \leq i, j \leq n\,,$$

ou encore, si l'on ne souhaite plus faire apparaître la base $(v_1, ..., v_n)$, par la relation dite *relation de dualité* :

$$\forall\, u, v \in E\,,\ \langle L(u)\,,\ v\rangle = \langle u\,,\ L^*(v)\rangle\,. \tag{1.21}$$

Cette application $\mathbb{C}$-linéaire L^*, parfaitement déterminée par la relation de dualité (1.21) est dite *application linéaire adjointe de L*.

Règle pratique importante : *Si $(v_1, ..., v_n)$ est une base orthonormée de E (relativement au produit scalaire définissant la structure hilbertienne sur E), la matrice de L^* dans la base $(v_1, ..., v_n)$ se déduit de celle de L dans cette même base via la relation*

$$M_{L^*,(v_1,...,v_n)} = \overline{{}^t M_{L,(v_1,...,v_n)}} := M^*_{L,(v_1,...,v_n)}\,. \tag{1.22}$$

Prendre l'adjointe A^* d'une matrice (n, n) à entrées complexes, c'est à la fois la transposer et en conjuguer les coefficients [4]. Dans une base orthonormée, la matrice de l'application linéaire adjointe L^* de L est donc l'adjointe de la matrice de L (dans cette même base). Ceci cesse d'être vrai dans une base non orthonormée !

On a les règles suivantes :

$$\begin{aligned}
(L_1 + L_2)^* &= L_1^* + L_2^* \\
(L_2 \circ L_1)^* &= L_1^* \circ L_2^* \\
(\lambda \cdot L)^* &= \overline{\lambda} \cdot L^* \\
(L^*)^* &= L \\
\mathrm{Id_E}^* &= \mathrm{Id}_E\,.
\end{aligned}$$

4. Dans les logiciels de calcul scientifique comme MATLAB ou SCILAB, cette opération de "transconjuguaison" ou "prise d'adjoint" (pas nécessairement sur une matrice carrée) est notée $A \longrightarrow A'$.

Une application $\mathbb{C}$-linéaire $L : E \longrightarrow E$ est dite *autoadjointe* si et seulement si $L = L^*$; ceci équivaut à dire que la matrice de L dans une base orthonormée $(v_1, ..., v_n)$ de E est une matrice hermitienne.

Une application $\mathbb{C}$-linéaire $L : E \longrightarrow E$ est dite *normale* si et seulement si L commute avec son adjointe, c'est-à-dire $L \circ L^* = L^* \circ L$; ceci équivaut à dire que si A désigne la matrice de L dans une base orthonormée, alors

$$A \bullet A^* = A^* \bullet A \, .$$

Une application $\mathbb{C}$-linéaire $L : E \longrightarrow E$ est dite *unitaire* si et seulement si L est inversible, d'inverse son adjointe, c'est-à-dire $L \circ L^* = L^* \circ L = \mathrm{Id}_E$; ceci équivaut à dire que si A désigne la matrice de L dans une base orthonormée, alors

$$A \bullet A^* = A^* \bullet A = I_n$$

(une telle matrice est dite *unitaire*). Dire que L est unitaire équivaut encore à dire que

$$\forall \, v, w \in E \, , \ \langle L(v) \, , \, L(w) \rangle = \langle v \, , \, (L^* \circ L)(w) \rangle = \langle v \, , \, w \rangle \, ;$$

ceci équivaut aussi au fait que l'image par L d'une base orthonormée quelconque est encore une base orthonormée. Notons que l'on peut aussi caractériser le fait que L est unitaire par :

$$\forall \, v, w \in E \, , \ \langle L^*(v) \, , \, L^*(w) \rangle = \langle v \, , \, ((L^*)^* \circ L^*)(w) \rangle = \langle v \, , \, (L \circ L^*)(w) \rangle = \langle v \, , \, w \rangle \, .$$

Les matrices unitaires (matrices des opérateurs unitaires exprimés dans une base orthonormée) se reconnaîssent, elles, grâce à la règle suivante :

Une matrice (n, n) à entrées complexes est unitaire (i.e $A \bullet A^* = A^* \bullet A = I_n$) si et seulement si les vecteurs colonnes de A (resp. les vecteurs lignes de A) forment une base orthonormée pour le produit scalaire canonique

$$\Big((z_1, ..., z_n) \, , \, (w_1, ..., w_n) \Big) \longmapsto \sum_{j=1}^{n} z_j \overline{w_j}$$

sur $\mathbb{C}^n$.

On a les implications suivantes entre ces trois propriétés (autoadjoint, normal, unitaire) :

$$\begin{aligned} \text{AUTOADJOINT} &\implies \text{NORMAL} \\ \text{UNITAIRE} &\implies \text{NORMAL} \end{aligned}$$

Décomposition spectrale des opérateurs normaux

Soit E un $\mathbb{C}$-espace vectoriel équipé d'un produit scalaire $\langle \, , \, \rangle$ et $L : E \longrightarrow E$ une application linéaire normale (par exemple autoadjointe ou unitaire).

On sait que L admet au moins un vecteur propre v_1 (relatif à une valeur propre λ_1 racine complexe du polynôme caractéristique de L. Le fait que L soit normal

implique que $\overline{\lambda_1}$ est valeur propre de L^* : on a effet

$$\begin{aligned}\langle L^*(v_1) - \overline{\lambda_1}\cdot v_1\,,\, L^*(v_1) - \overline{\lambda_1}\cdot v_1\rangle &= \langle v_1\,,\, L(L^*(v_1))\rangle - \overline{\lambda_1}\langle v_1\,,\, L^*(v_1)\rangle \\ &\quad +|\lambda_1|^2\|v_1\|^2 - \lambda_1\langle L^*(v_1)\,,\, v_1\rangle \\ &= \langle v_1\,,\, L^*(L(v_1))\rangle - \lambda_1\langle v_1\,,\, L(v_1)\rangle \\ &= \overline{\lambda_1}\langle v_1\,,\, L^*(v_1)\rangle - \lambda_1\langle v_1\,,\, L(v_1)\rangle \\ &= \overline{\lambda_1}\langle L(v_1)\,,\, v_1\rangle - \lambda_1\langle v_1\,,\, L(v_1)\rangle \\ &= |\lambda_1|^2\,\|v_1\|^2 - |\lambda_1|^2\,\|v_1\|^2 = 0\,,\end{aligned}$$

ce qui prouve que v_1 est aussi vecteur propre de L^* (pour la valeur propre $\overline{\lambda_1}$). Si u est un vecteur de l'hyperplan vectoriel F orthogonal à v_1, on a

$$\langle L(u)\,,\, v_1\rangle = \langle u\,,\, L^*(v_1)\rangle = \langle u\,,\, \overline{\lambda_1}\cdot v_1\rangle = \lambda_1\langle u\,,\, v_1\rangle = 0\,;$$

la restriction de L à F est une application linéaire normale de F ($\mathbb{C}$-espace vectoriel de dimension $\dim E - 1$) dans lui-même. On peut lui appliquer le raisonnement que l'on vient d'appliquer à L. On prouve ainsi le très important résultat suivant :

Théorème 1.8 (décomposition spectrale des opérateurs normaux). *Si L est une application linéaire normale d'un $\mathbb{C}$-espace vectoriel hilbertien E de dimension finie dans lui-même (relativement au produit scalaire définissant la structure hermitienne), L est diagonalisable dans une base orthonormée (pour le produit scalaire définissant la structure hermitienne) de vecteurs propres. De plus, si v est vecteur propre de L pour la valeur propre λ, v est aussi vecteur propre de L^* pour la valeur propre $\overline{\lambda}$.*

Ce théorème peut être précisé dans le cas des applications linéaires autoadjointes ; dans ce cas en effet, outre que L est diagonalisable dans une base orthonormée, on déduit du théorème précédent que toute valeur propre λ de L est égale à sa conjuguée $\overline{\lambda}$, donc est réelle. On a ainsi le :

Théorème 1.9 (décomposition spectrale des opérateurs autoadjoints). *Si L est une application linéaire autoadjointe d'un $\mathbb{C}$-espace vectoriel hilbertien E de dimension finie dans lui-même (relativement au produit scalaire définissant la structure hermitienne), L est diagonalisable dans une base orthonormée (pour le produit scalaire définissant la structure hermitienne) de vecteurs propres. De plus, les valeurs propres de L sont toutes réelles.*

Diagonalisation des matrices réelles symétriques

Si A est une matrice réelle symétrique, on peut considérer A comme la matrice d'une application linéaire autoadjointe de $\mathbb{C}^n$ (muni du produit scalaire canonique) dans lui-même. D'après le théorème 1.9 de décomposition spectrale des opérateurs autoadjoints, il existe une base orthonormée de $\mathbb{C}^n$, $(v_1, ..., v_n)$, constituée de vecteurs propres pour $L\ : Z \in \mathbb{C}^n \longmapsto A \bullet Z$ dans laquelle la matrice de L est diagonale. Mais comme A est une matrice réelle, les vecteurs $v_1, ..., v_n$ peuvent être choisis dans $\mathbb{R}^n$. On a donc le résultat suivant :

Théorème 1.10 *Toute matrice symétrique réelle est diagonalisable dans une base orthonormée de $\mathbb{R}^n$ (pour le produit scalaire canonique sur $\mathbb{R}^n$), ce qui signifie que A s'écrit*

$$A = P \bullet D \bullet P^{-1}$$

où P est une matrice unitaire réelle (on dit aussi une matrice réelle orthogonale).

Réduction des formes quadratiques dans un espace euclidien

Soit E un $\mathbb{R}$-espace vectoriel équipé d'un produit scalaire et Q une forme quadratique $Q : E \longrightarrow \mathbb{R}$. La matrice de Q dans une base orthonormée $(e_1, ..., e_n)$ de E est une matrice réelle symétrique A. Il existe donc une matrice réelle orthogonale P et une matrice réelle diagonale D telles que

$$A = {}^tP \bullet D \bullet P\,.$$

L'image de la base $(e_1, ..., e_n)$ par P (qui est réelle orthogonale) est encore une base orthonormée $(v_1, ..., v_n)$ de E. Dans cette base, la matrice de la forme quadratrique Q est la matrice D.

Conséquence : *étant donnée une forme quadratique Q sur un $\mathbb{R}$-espace vectoriel E de dimension n, il existe toujours une base $(v_1, ..., v_n)$ dans laquelle la matrice du produit scalaire soit I_n (ce qui signifie que la base est orthonormée pour ce produit scalaire) et la matrice de Q soit diagonale réelle, i.e*

$$Q\Big(\sum_{j=1}^{n} X_j \cdot v_j\Big) = \sum_{j=1}^{n} \lambda_j X_j^2.$$

On appelle ce résultat (très utile dans la pratique dans le cadre réel) le théorème de *réduction simultanée des formes quadratiques* dans la mesure où il permet de réduire simultanément une forme quadratique provenant d'un produit scalaire et une forme quadratique quelconque.

Réduction des formes hermitiennes dans un espace hilbertien

Soit E un $\mathbb{C}$-espace vectoriel équipé d'un produit scalaire et H une forme hermitienne $H : E \longrightarrow \mathbb{R}$. La matrice de H dans une base orthonormée $(e_1, ..., e_n)$ de E est une matrice hermitienne A ; l'application $L : Z \in \mathbb{C}^n \longmapsto A \bullet Z$ est une application linéaire autoadjointe de $\mathbb{C}^n$ dans lui-même que l'on peut diagonaliser dans une base orthonormée de $\mathbb{C}^n$ (pour le produit scalaire canonique). Il existe donc une matrice à entrées complexes unitaire P et une matrice à entrées complexes diagonale D telles que

$$A = {}^tP \bullet D \bullet \overline{P}\,.$$

L'image de la base $(e_1, ..., e_n)$ par P (qui est unitaire) est encore une base orthonormée $(v_1, ..., v_n)$ de E. Dans cette base, la matrice de la forme hermitienne H est la matrice D.

Conséquence : *étant donnée une forme hermitienne H sur un $\mathbb{C}$-espace vectoriel E de dimension n, il existe toujours une base $(v_1, ..., v_n)$ dans laquelle la matrice du produit scalaire soit I_n (ce qui signifie que la base est orthonormée pour ce produit scalaire) et la matrice de H soit diagonale réelle, i.e*

$$H\Big(\sum_{j=1}^{n} z_j \cdot v_j\Big) = \sum_{j=1}^{n} \lambda_j |z_j|^2.$$

On appelle ce résultat (aussi utile que le précédent, dans le cadre complexe cette fois) le théorème de *réduction simultanée des formes hermitiennes* dans la mesure où il permet de réduire simultanément une forme hermitienne provenant d'un produit scalaire et une forme hermitienne quelconque.

Décomposition en valeurs singulières

Soit L une application linéaire surjctive de $\mathbb{R}^p$ dans $\mathbb{R}^n$, représentée par une matrice A lorsque $\mathbb{R}^p$ et $\mathbb{R}^n$ sont rapportés à leurs bases canoniques respectives. D'après la formule du rang, le noyau de L est un sous-espace vectoriel F de dimension $p-n$ de $\mathbb{R}^p$, que l'on peut rapporter à une base orthonormée $(e_{n+1}, ..., e_p)$ de $\mathbb{R}^p$ (pour le produit scalaire canonique). Si l'on complète (suivant le procédé de Gram-Schmidt) cette base en une base orthonormée $(e_1, ..., e_p)$ de $\mathbb{R}^p$, on constate que $(e_1, ..., e_n)$ constitue une base du sous-espace du sous-espace $F^\perp$ constitué des vecteurs de $\mathbb{R}^p$ orthogonaux au noyau de L. La restriction de L à $F^\perp$ est une application linéaire bijective entre $F^\perp$ et $\mathbb{R}^n$. Si $\mathbb{R}^p$ est rapporté à la base $(e_1, ..., e_p)$ et $\mathbb{R}^n$ à sa base canonique, la matrice de L dans ces bases s'écrit sous la forme :

$$\begin{pmatrix} & 0 & \cdots & \cdots & 0 \\ B & \vdots & \vdots & \vdots & \vdots \\ & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

où B est une matrice réelle $n \times n$ de déterminant non nul ; il existe donc une matrice orthogonale réelle V_1 de taille (p, p) telle que

$$A = \begin{pmatrix} & 0 & \cdots & \cdots & 0 \\ B & \vdots & \vdots & \vdots & \vdots \\ & 0 & \cdots & \cdots & 0 \end{pmatrix} \bullet V_1^* \,.$$

La matrice $B \bullet B^*$ (ici $B^* = {}^tB$ car B est réelle) est une matrice symétrique réelle que l'on peut donc écrire

$$B \bullet B^* = U \bullet \mathrm{diag}\,(\lambda_1, ..., \lambda_n) \bullet U^* \,,$$

où U est une matrice orthogonale réelle de taille (n, n) ; d'autre part, pour tout $v \in \mathbb{R}^n$,

$$\langle B \bullet B^*(v)\,,\, v\rangle = \langle B^*(v)\,,\, B^*(v)\rangle \geq 0\,,$$

ce qui implique que pour $j = 1, ..., n$, $\lambda_j = \mu_j^2 \geq 0$ (on appelle μ_j la racine positive du réel positif λ_j) ; on peut d'ailleurs faire en sorte que $\mu_1^2 \geq \mu_2^2 \geq \cdots \geq \mu_n^2$. Comme $\det B \neq 0$, ces nombres réels positifs μ_j^2, $j = 1, ..., n$ sont tous non nuls. Les nombres réels positifs $\mu_1, ..., \mu_n$ correspondants sont dits *valeurs singulières* de l'application linéaire L (il est clair que ces nombres ne dépendent que de L et non du choix des bases dans lesquelles L est exprimé).

Comme

$$B \bullet B^* = \Big(U \bullet \mathrm{diag}\,(\mu_1, ..., \mu_n)\Big) \bullet \Big(U \bullet \mathrm{diag}\,(\mu_1, ..., \mu_n)\Big)^* \,,$$

la matrice

$$B^* \bullet \left(\Big(U \bullet \mathrm{diag}\,(\mu_1, ..., \mu_n)\Big)^*\right)^{-1}$$

est une matrice orthogonale réelle W de taille (n, n) et l'on peut donc écrire

$$B^* = W \bullet \mathrm{diag}\,(\mu_1, ..., \mu_n) \bullet U^* \,,$$

soit

$$B = U \bullet \mathrm{diag}\,(\mu_1, ..., \mu_n) \bullet W^* .$$

On peut donc ainsi écrire

$$A = U \bullet \begin{pmatrix} \mu_1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \mu_n & 0 & \cdots & \cdots & 0 \end{pmatrix} \bullet V^*$$

où U et V sont respectivement des matrices orthogonales réelles de tailles (n, n) et (p, p).

Une telle décomposition est dite *décomposition en valeurs singulières* de la matrice A (ou de l'opérateur L) et on la retrouve sous un environnement de calcul scientifique (comme MATLAB ou SCILAB) sous la commande :

```
>>[U,D,V] = svd(A);
```

Cette transformation permet (quitte à effectuer des transformations orthogonales à la source et au but) de ramener l'étude d'une application linéaire surjective de $\mathbb{R}^p$ dans $\mathbb{R}^n$ à l'étude d'une application linéaire dont la matrice est une matrice diagonale de taille (n, n), complétée à droite par une matrice de zéros de taille $(n, p-n)$. C'est un outil très intéressant du point de vue pratique. On se doute en particulier du rôle important joué par les vecteurs colonnes de V, vecteurs "modèle" de l'espace source $\mathbb{R}^p$ lorsque'il est soumis à la transformation linéaire L de matrice A.

Si L n'est pas surjective, cette décomposition reste valide (mais $n - \text{rang}\,(L)$ valeurs singulières μ_j sont alors nulles).

Tout ceci peut se transposer au cadre complexe (L étant une application $\mathbb{C}$-linéaire de $\mathbb{C}^p$ dans $\mathbb{C}^n$) ; il suffit juste de remplacer "orthogonale" par "unitaire". Les valeurs singulières restent des nombres strictement positifs.

FIN DU CHAPITRE 1

Chapitre 2

Hasard, probabilités, statistique

2.1 Epreuve et ensemble d'évènements

Le lancer d'un dé à six faces constitue un exemple de ce que l'on appelle une *épreuve*; si le dé n'est pas pipé, tout non-mathématicien prévoira qu'il a une chance sur six de réaliser *six* lors d'une telle épreuve. Il étayera son résultat en vous faisant constater que, s'il itère N fois cette épreuve (avec N très grand), le quotient $N(6)/N$ (où $N(6)$ désigne le nombre de six obtenus lors de la série de N coups) tend asymptotiquement vers $1/6$ lorsque N tend vers l'infini. Son argumentation relève, on le verra, du point de vue qualifié de *statistique*.

Formulons mathématiquement le problème en considérant l'ensemble Ω de tous les résultats possibles de l'épreuve, ici

$$\Omega = \{1, \ldots, 6\}\,.$$

Le résultat de l'épreuve est *aléatoire*; l'épreuve consiste à choisir un élément dans l'ensemble Ω, espace que l'on qualifie ici d'*espace d'évènements* (on devrait en fait plutôt dire "d'évènements élémentaires"). Se donner la *loi* probabiliste à laquelle obéit cette épreuve aléatoire, à savoir ici le lancer du dé, consiste ici à se donner une collection de 6 nombres réels positifs $p_j, j = 1, \ldots, 6$ de somme 1. La quantité p_j représente ce que notre non-mathématicien interprétait comme *le nombre de chances de réaliser le résultat* j en lancant le dé.

Dans le cas de dés non pipés, il est naturel de travailler avec le modèle mathématique :

$$p_j = \frac{1}{6}\,,\ j = 1, \ldots 6.$$

Ce modèle est un cas particulier du cas où Ω est un ensemble fini et où chaque singleton (ou encore évènement élémentaire) $\{a\}$ de Ω est "chargé" avec la masse

$$P(\{a\}) = \frac{1}{\operatorname{card}\Omega}\,,$$

auquel cas, une partie A de Ω (que l'on appelle un évènement) est "chargée" avec la masse

$$P(A) = \frac{\operatorname{card} A}{\operatorname{card}\Omega}\,;$$

cette distribution de probabilité est dite *distribution uniforme* sur Ω.

Dans le cas de dés pipés, on doit envisager d'autres modèles, comme par exemple

$$p_1 = p_2 = p_3 = p_4 = \frac{1}{12}\,,\ p_5 = p_6 = \frac{1}{3}.$$

L'ensemble des évènements n'est pas nécéssairement fini, comme le montre l'exemple *géométrique* suivant : l'épreuve consiste à prendre au hasard deux points distincts sur un cercle Γ donné de rayon 1 et à tracer la corde qui les joint (comme sur la figure 2.1).

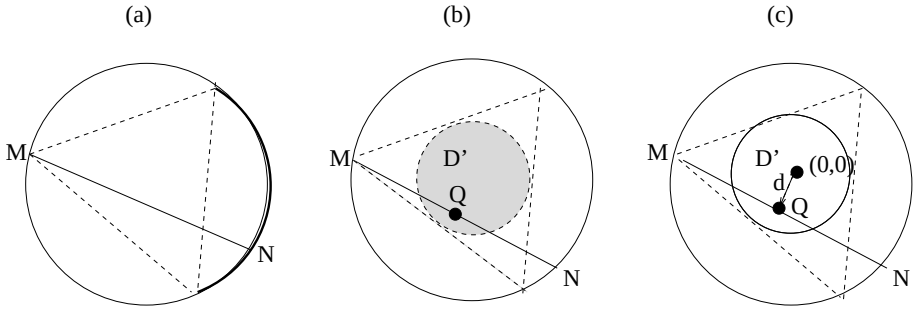

FIGURE 2.1 – Le paradoxe de Bertrand

On peut poser la question naïve suivante : les choix de points étant équiprobables sur $\Gamma \times \Gamma$ (on va préciser ceci ultérieurement), quel "nombre de chances" ou encore "probabilité" a-t'on que la longueur de la corde joignant deux points M et N de Γ soit supérieure à celle du coté d'un triangle équilatéral inscrit dans Γ ? (voir la figure 2.1 (a)). Notre non-spécialiste pourra répondre en disant qu'il y a une chance sur trois que ce qu'il souhaite se réalise lorsqu'il choisit ses deux points au hasard (et indépendamment l'un de l'autre) : son raisonnement s'appuie sur le fait que, dès que l'un des points est fixé, le second doit être dans un arc de cercle de longueur $\pi/3$. Ce calcul intuitif de probabilité est fondé sur le fait que l'ensemble d'évènements Ω choisi soit ici le produit $\Omega = \Gamma \times \Gamma$ du cercle par lui-même, la probabilité d'un produit d'intervalles cuvilignes $I \times J$ étant

$$\frac{\text{longueur}\,(I)}{\pi} \times \frac{\text{longueur}\,(J)}{\pi}\,.$$

Ceci est cependant ambigü, comme le faisait remarquer le mathématicien Joseph Bertrand en 1899. On peut en effet aussi repérer la corde par son milieu Q et considérer comme ensemble d'évènements Ω l'ensemble des positions possibles de ce milieu après tracé de la corde, soit le disque D dont Γ est le bord ; cette fois, l'ensemble des évènements n'est plus l'ensemble des paires de points de Γ, mais l'ensemble des points du disque fermé $D(0,1)$ (ces points repèrent le milieu de la corde, voir la figure 2.1 (b)) ; avec ce choix, le fait que la corde ait une longueur supérieure à celle du triangle équilatéral inscrit se rephrase en *le milieu de la corde est dans le disque D' de rayon $1/2$ et de centre l'origine* ; notre non-spécialiste pourrait tout autant affirmer qu'il y a une chance sur 4 (puisque $1/4$ représente le quotient de la surface de D' par la surface de $D(0,1)$) que son souhait se réalise. On pourrait aussi en repérant la corde par sa distance $d \in [0,1]$ à l'origine (voir la figure 2.1 (c)) et (en prenant comme ensemble des évènements $[0,1]$) dire que l'évènement que l'on cherche est celui qui correspond au fait que cette distance soit inférieure ou égale

à 1/2, soit l'intervalle $[0, 1/2]$, qui est de longueur $1/2$; comme la longueur de $[0, 1]$ est normalisée égale à 1, la probabilité de notre évènement suivant cette troisième interprétation vaudrait $1/2$!

Cet exemple montre bien que la mise sous forme mathématique d'un problème de nature probabiliste exige au préalable la définition de trois choses :

- un ensemble Ω d'évènements élémentaires (les points de Ω peuvent par exemple permettre de repérer les résultats d'une épreuve) ; c'est la non-définition claire de cet ensemble d'évènements qui explique le paradoxe de Bertrand ;
- une famille de sous-ensembles de Ω candidats à pouvoir être *mesurés* (la contrainte étant que le mathématicien n'a que l'accès au dénombrable pour faire ses calculs de mesure) ;
- une règle enfin pour mesurer, avec la convention que la mesure de Ω est 1 ; si Ω représente l'ensemble des résultats d'une certaine épreuve, alors la "mesure" de A sera interprétée par notre non-spécécialiste comme le pourcentage de chance qu'a le résultat de l'épreuve de tomber dans A.

Nous allons dans la section suivante définir proprement les notions mathématiques consolidant ces idées intuitives.

2.2 Tribus et probabilités

Une *tribu* $\mathcal{T}$ sur un ensemble abstrait Ω (appelé à être un ensemble d'éveènements, par exemple les résultats d'une épreuve) est une sous-famille $\mathcal{T}$ de l'ensemble des parties de Ω telle que :

- $\Omega \in \mathcal{T}$;
- si $A \in \mathcal{T}$, $\Omega \setminus A$ aussi (donc l'ensemble vide est en particulier dans $\mathcal{T}$) ;
- si $A_1, A_2 \in \mathcal{T}$, alors $A_1 \cap A_2 \in \mathcal{T}$;
- si $(A_n)_{n\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathcal{T}$, l'union de tous les A_n est encore dans $\mathcal{T}$.

Si Ω est un ensemble équipé d'une tribu, on dit que le couple $(\Omega, \mathcal{T})$ est un *espace probabilisable*.

Exemple 2.1

- un exemple basique est celui où Ω est un ensemble fini ou dénombrable et $\mathcal{T}$ la famille de toutes des parties de Ω (ensemble vide inclus) ;
- si $\Omega = \mathbb{R}^n$, la collection de tous les pavés du type $]a_1, b_1] \times \cdots \times]a_n, b_n]$ n'est pas une tribu, mais on appelle *tribu borélienne* (du nom du mathématicien français Emile Borel, 1871-1956) la plus petite tribu contenant tous ces pavés (c'est aussi la plus petite tribu contenant tous les ouvert de $\mathbb{R}^n$) et *tribu de Lebesgue* (du nom cette fois d'Henri Lebesgue, 1875-1941) la plus petite tribu contenant la tribu borélienne et tous les sous-ensembles de $\mathbb{R}^n$ de mesure extérieure nulle ; la tribu de Lebesgue est en fait constituée de tous les sous-ensembles quarrables de $\mathbb{R}^n$. (on l'admettra) ; l'axiome du choix en logique permet la construction de sous-ensembles de $\mathbb{R}^n$ non quarrables (mais ce n'est pas évident !), ce qui fait que ni la tribu borélienne, ni même la tribu de Lebesgue, ne forment la tribu de toutes les parties.

Signalons que la nécessité de travailler avec des ensembles d'évènements non dénombrables est très souvent chose courante en probabilité, même si le problème semble de nature discrète : par exemple, l'ensemble d'évènements attaché à une partie de pile ou face est $\{0, 1\}^{\mathbb{N}}$, ensemble des suites infinies de d'éléments de $\{0, 1\}$,

ensemble dont le cardinal a la puissance du continu (il y a correspondance entre les points de $[0,1]$ et leur écriture dans le système en base deux).

Une *mesure de probabilité* sur un espace probabilisable $(\Omega, \mathcal{T})$ est une application P de $\mathcal{T}$ dans $[0,\infty]$ telle que l'on ait $P(\Omega)=1$ et que, si les ensembles A_n, $n\in\mathbb{N}$, sont des éléments de $\mathcal{T}$ disjoints, on ait

$$P\Big(\bigcup_{n=0} A_n\Big) = \sum_{n=0}^{\infty} P(A_n)\,.$$

Ainsi, si A et B sont deux éléments de $\mathcal{T}$, on a la formule

$$P(A\cup B) = P(A)+P(B)-P(A\cap B)\,,$$

formule que l'on peut généraliser au cas plus général du calcul de la probabilité d'une des union d'éléments de $\mathcal{T}$ non nécessairement disjoints.

Les calculs de probabilités sur un ensemble fini (la distribution de probabilité étant la distribution uniforme) font intervenir des calculs de dénombrement. Si k et n sont deux entiers strictement positifs tels que $k\le N$, le nombre d'applications injectives d'un ensemble à k éléments dans un ensemble à n éléments est aussi égal au nombre de k-uplets $(n_1,...,n_k)$, avec $1\le n_j\le n$ et les n_j distincts (attention ! l'ordre des n_j est ici important) ; ce nombre, qui vaut

$$A_n^k = \frac{n!}{(n-k)!}$$

est aussi appelé *nombre d'arrangements* de l'ensemble à k éléments dans l'ensemble à n éléments. Le nombre de manières de choisir k éléments parmi N se déduit du précédent par la formule

$$C_n^k = \binom{n}{k} = \frac{A_n^k}{k!} = \frac{n!}{k!\,(n-k)!}$$

puisque le nombre de permutations d'un ensemble à k éléments est égal au cardinal du groupe symétrique $\mathcal{S}_k$, soit à k. Le nombre C_n^k (noté aussi $\binom{n}{k}$) est dit *nombre de combinaisons* de k éléments parmi n ; si $k=0$, on convient que $A_n^0=C_n^0=1$. La suite des nombres C_n^k est régie par les identités du *triangle de Pascal* :

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}\ ,\ k,n\in\mathbb{N}\,,$$

avec les conventions $\binom{m}{p}=0$ si $p>m$ ou si $p<0$. On a aussi la classique *formule du binôme* :

$$2^n = \sum_{k=0}^{N}\binom{n}{k}\ ,\ n\in\mathbb{N}\,.$$

qui relie entre eux les coefficients binômiaux (2^n est le nombre de parties d'un ensemble à n éléments et la formule du binôme s'obtint en comptant les parties suivant leur cardinal).

Exemple 2.2

- Sur $\Omega = 1, \ldots, n$ (ou plus généralement si Ω est un ensemble à n éléments), la probabilité affectant à chaque singleton le nombre $1/n$ est appelée *distribution de probabilité uniforme sur* Ω ; la probabilité de A vaut dans ce cas $\operatorname{card} A / \operatorname{card} \Omega$.
- Autre exemple plus sophistiqué : sur $\{0, \ldots, n\}$, on définit une probabilité P_τ associée à un paramètre $\tau \in [0, 1]$ en posant

$$P_\tau(k) := \binom{n}{k} \tau^k (1-\tau)^{n-k} .$$

c'est la *distribution de probabilité binomiale de paramètres* (n, τ), celle que l'on retrouve dans les problèmes de tirage avec remise : considérons une urne avec N boules, dont R rouges et B blanches ; on décide de tirer n boules de manière indépendante (ceci doit être précisé par la suite) et avec remise. La probabilité que l'on tire exactement k boules rouges (si on prend comme ensemble d'évènements $\Omega = \{1, \ldots, N\}^n$ et comme distribution de probabilité la loi uniforme, ce qui, on le verra ultérieurement rend compte de l'indépendance) est alors

$$P(R_k) = \frac{\binom{n}{k} R^k (N-R)^{n-k}}{N^n}$$

ce qui correspond à une loi de probablité binomiale avec $\tau = R/N$. Notons que c'est aussi le résultat que l'on trouve (avec $\tau = 1/2$) lorsque, dans une suite de N épreuves indépendantes de pile ou face avec une pièce honnête, on évalue la probabilité d'obtenir k fois pile lors de ces N épreuves. Ici encore, pour corroborer les hypothèses de l'expérience (l'*indépendance* des jets succesifs), on travaille avec $\Omega = \{pile, face\}^N$ et la distribution uniforme.
- Le dernier exemple que nous proposons sur un ensemble fini correspond, lui, aux résultats que l'on obtient lorsque l'on envisage des tirages hors d'une urne, mais cette fois sans remise. Considérons 3 nombres entiers R, N, n avec $R < N$ et $n \leq N$; on définit une probabilité sur $\{0, \ldots, n\}$ en posant, pour tout k dans $\{0, \ldots, n\}$ tel que

$$\max(0, n - (N-R)) \leq k \leq \min(n, R) ,$$

$$p_k = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}$$

et 0 sinon. Cette loi de probabilité est la loi *hypergéométrique* de paramètres $n, \tau_1 = R/N, \tau_2 = (N-R)/N$; cette définition peut être élargie au cas où le nombre des paramètres τ_j (avec $\sum \tau_j = 1$) dépasse 2 ; on a alors la loi *polyhypergéométrique*. On retrouve (on le verra en exercice dans la section 2.3) les p_k lorsque l'on effectue n tirages successifs (et indépendants) hors d'une urne contenant N boules, dont R rouges et $N - R$ blanches ; en considérant comme espace probabilisé l'espace de toutes les suites à n termes de 0, 1 possibles (1 si la boule tirée est rouge, 0 si elle est blanche), équipé de la distribution de probabilité uniforme, le calcul de la probabilité de l'évènement *il y a exactement k boules rouges tirées* donne le résultat p_k.
- On peut également donner des exemples dans le cas où $\Omega = \mathbf{N}$; se donner une distribution de probabilité (la tribu étant celle de toutes les parties) équivaut à se donner une suite (p_n) de réels positifs telle que la série $\sum_n p_n$ converge et ait pour somme 1 : par exemple la suite

$$p_n = (1-p)p^n, \quad 0 \leq p < 1$$

(loi *géométrique ou de Pascal* de paramètre p) ou encore

$$p_n = \frac{\lambda^n}{n!} exp(-\lambda), \quad \lambda \in \mathbf{R}^+$$

(loi *de Poisson* de paramètre λ).

Pour construire des probabilités sur $\mathbb{R}$ ou $\mathbb{R}^n$ équipés de la tribu borélienne, on utilise un résultat d'extension permettant de prolonger à la tribu borélienne une "mesure" sur la collection des pavés du type $]a_1, b_1] \times \cdots \times]a_n, b_n]$. Pourvu que la mesure obtenue soit telle que $P(\mathbb{R}^n) = 1$, on construit bien ainsi des distributions de probabilité. Le "prototype" de telles distributions de probabilité sur $\mathbb{R}^n$ repose sur la donnée d'une fonction intégrable $f : \mathbb{R}^n \to [0, +\infty]$ telle que

$$\int \cdots \int_{\mathbb{R}^n} f(x_1, ..., x_n)\, dx_1 dx_2 \cdots dx_n = 1\,.$$

Une telle fonction est dite *densité de probabilité* sur $\mathbb{R}^n$.

Remarque. Il faut noter que la valeur "ponctuelle" $f(x_1, ..., x_n)$) n'a pas de sens du point de vue physique car l'accès à un point spécifique de l'espace $\mathbb{R}^n$ (ou une valeur spécifique du temps si $n = 1$) est numériquement impossible. La fonction f doit être pensée comme une "densité de masse" répartie dans $\mathbb{R}^n$ ou dans un sous-ensemble de $\mathbb{R}^n$ d'une certaine manière ; on dit aussi en termes mathématiques que la fonction f est *définie presque partout*.

Si $[a, b]$ est par exemple un intervalle de $\mathbb{R}$ et si

$$f(t) = \frac{1}{b-a}\chi_{[a,b]}(t)\,,$$

où $\chi_{[a,b]}$ désigne la fonction valan 1 sur $[a, b]$, 0 ailleurs, alors f est une densité de probabilité et la probabilité qui correspond à cette densité est celle donnée par

$$P(]\alpha, \beta]) = \frac{\beta - \alpha}{b-a}\,;$$

on l'appelle *probabilité uniforme* sur $[a, b]$. On peut de même définir une distribution de probabilité uniforme sur un sous-ensemble quarrable borné A de $\mathbb{R}^n$ (de volume n-dimensionnel strictement positif) : c'est celle dont la densité est donnée par

$$f(x) = \frac{\chi_A(x)}{\int \ldots \int_A f(x_1, ..., x_n)\, dx_1 \ldots dx_n}$$

où χ_A désigne la fonction valant 1 sur A et 0 sur le complémentaire de A.

Plus généralement, si f est une telle fonction densité sur $\mathbb{R}^n$, la probabilité induite par f sur l'espace probabilisable $\mathbb{R}^n$ équipé de la tribu borélienne telle que, pour tout borélien de $\mathbb{R}^n$, on ait

$$P(A) = \int \ldots \int_A f(x_1, ..., x_n)\, dx_1 \ldots dx_n\,;$$

on dit que cette distribution de probabilité sur $\mathbb{R}^n$ est une *distribution de densité de probabilité* f.

Exemple 2.3 Un exemple majeur de telle distribution sur $\mathbb{R}$ est la *distribution de Gauss*, correspondant à la fonction

$$t \to f(t) = \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)$$

(dite *gaussienne réduite centrée*). Cette distribution joue un rôle crucial en physique car la gaussienne est la fonction réalisant le meilleur compromis dans l'incontournable principe d'incertitude (impossible de localiser simultanément une particule et son spectre !) ; on la retrouvera aussi dans

les théorèmes limite de la théorie des probabilités : la célèbre *expérience de Galton* consistant à laisser tomber un stock de billes au dessus d'une grille avec des plots (la bille pouvant tomber indifféremment à gauche ou à droite lorsqu'elle rencontre un plot), que l'on expliquera plus loin dans ce cours, montre que le tas de billes se répartit dans le réceptacle suivant une distribution de Gauss (voir la figure 2.2).

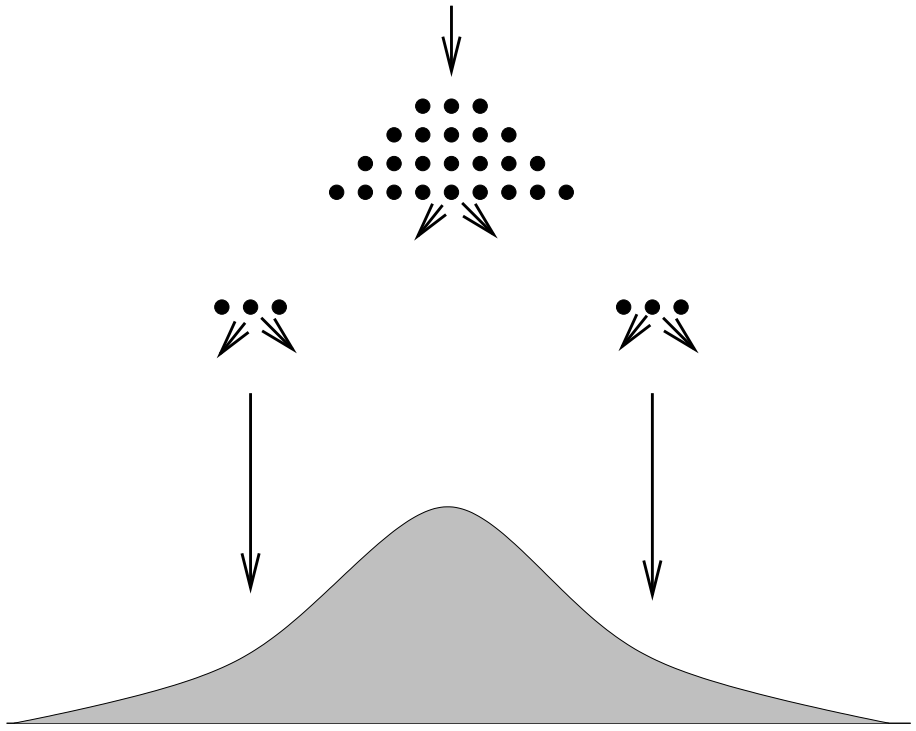

FIGURE 2.2 – Le triangle de Galton

Exemple 2.4 Autre exemple important (car lié, on le verra plus loin, au processus de désintégratio atomique), celui de la densité de probabilité sur $\mathbb{R}$ définie par

$$f(t) = 0 \text{ si } t < 0\,,\ f(t) = \lambda e^{-\lambda t} \text{ si } t \geq 0\,,$$

où λ désigne un paramètre srictement positif ; cette densité génère la distribution de probabilité suivant *une loi exponentielle* (de paramètre λ) sur $\mathbb{R}$.

Si P est une probabilité sur $\mathbb{R}$ (équipé de la tribu borélienne), la fonction

$$x \in \mathbb{R} \to P(]-\infty, x[)$$

est dite *fonction de répartition* de la distribution de probabilité f sur $\mathbb{R}$; cette fonction de répartition est une fonction croissante sur $\mathbb{R}$, continue à gauche, tendant vers 0 en $-\infty$ et vers 1 en $+\infty$. On peut d'ailleurs montrer que toute fonction F sur $\mathbb{R}$ ayant ces trois propriétés est la fonction de répartition d'une distribution de probabilité sur $\mathbb{R}$; s'il existe de plus une fonction positive intégrable f (d'intégrale 1) telle que

$$F(x) = \int_{]-\infty,x[} f(t)dt\,,$$

(par exemple, si F, en plus des trois propriétés mentionnées, est dérivable sur $\mathbb{R}$ et de dérivée continue), la distribution de probabilité correspondant à F est une distribution à densité, de densité $f = F'$.

Tout triplet $(\Omega, \mathcal{T}, P)$, où $\mathcal{T}$ est une tribu sur l'ensemble d'évènements Ω et P une probabilité sur $\mathcal{T}$, est dit *espace probabilisé*.

2.3 Notions de probabilité induite et conditionnelle ; indépendance

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et A un élément de $\mathcal{T}$ tel que $P(A) \neq 0$. La famille

$$\mathcal{T}_A = \{E \cap A,\ E \in \mathcal{T}\}$$

est encore une tribu, cette fois considérée comme famille de parties de A ; on définit une probabilité P_A sur $(A, \mathcal{T}_A)$ en posant

$$P_A(B) := \frac{P(B)}{P(A)},\ B \in \mathcal{T}_A.$$

On dit que $(A, \mathcal{T}_A)$ est ainsi équipé de la *probabilité induite par P sur A*.

On peut aussi définir une autre probabilité que P sur $(\Omega, \mathcal{T})$, liée cette fois à l'évènement A ; on la notera $P(\cdot \mid A)$ et on la définit par

$$P(B \mid A) := \frac{P(A \cap B)}{P(A)} ;$$

c'est la *probabilité conditionnelle attachée à P sous le conditionnement A*.

Si $A_1, ..., A_n$ sont n évènements d'un espace probabilisé $(\Omega, \mathcal{T}, P)$ tels que $P(A_1 \cap \cdots \cap A_n) > 0$, on prouve par récurrence la formule suivante :

$$P(A_1 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1}) .$$

Exemple 2.5 On dispose d'un jeu de 32 cartes ; on en tire 8 (et ce de manière indépendante (disons que ceci signifie, on le verra un peu plus loin, que l'espace des évènements est l'espace de tous les choix de 8 cartes possibles, la distribution de probabilité étant la distribution uniforme). L'évènement $A :=$ *on tire 2 valets* a pour probabilité

$$P(A) = \frac{\binom{4}{2}\binom{28}{6}}{\binom{32}{8}} ;$$

(on calcule son cardinal et l'on divise par le cardinal de l'ensemble des évènements, à savoir le nombre de combinaisons de 8 cartes parmi 32). L'évènement $B :=$*une carte tirée au moins est un valet* a pour probabilité

$$P(B) = 1 - \frac{\binom{28}{8}}{\binom{32}{8}}$$

et l'on a donc

$$P(A \mid B) = \frac{P(A)}{P(B)} = \frac{\binom{4}{2}\binom{28}{6}}{\binom{32}{8} - \binom{28}{8}}.$$

Une formule importante impliquant le conditionnement est la *formule de Bayes* ou encore *formule des probabilités totales*. Elle repose sur le fait très simple suivant ; si $(\Omega, \mathcal{T}, P)$ est un espace probabilisé et $(A_k)_{1\leq k\leq m}$ une collection de m éléments de $\mathcal{T}$ deux à deux dispoints et tels que

$$P(\Omega \setminus \bigcup_{k=1}^{m} A_k) = 0\,,$$

alors, pour tout évènement B de $\mathcal{T}$, on a

$$P(B) = \sum_{\{j,\, P(A_j)\neq 0\}} P(B \mid A_j)P(A_j)\,.$$

Couplée avec la formule de Bayes, on a aussi la *formule des preuves* que l'on peut énoncer ainsi :

Soit $(\Omega, \mathcal{T})$ un espace probabilisé et $B_1, \ldots, B_m, \ldots$ une suite d'éléments deux à deux disjoints de $\mathcal{T}$ dont l'union est de probabilité 1 (on dit encore un système complet d'évènements) ; on suppose tous les B_j de probabilité strictement positive ; alors, pour tout élément A de $\mathcal{T}$ de probabilité strictement positive, pour tout entier positif k, on a

$$P(B_k \mid A) = \frac{P(A \mid B_k)P(B_k)}{\sum_j P(A \mid B_j)B(B_j)}\,.$$

Voici l'interprétation de la formule des preuves, que l'on prouve en utilisant la formule de Bayes ainsi que le fait que

$$P(B_k \mid A) = P(A \cap B_k)/P(A) = P(A \mid B_k)P(B_k)/P(A)\,,$$

et à quoi elle sert concrètement : si A est un évènement de probabilité strictement positive et si les B_k, $k = 1, ..., m$, constituent un jeu d'hypothèses, le problème que l'on se pose est celui de savoir si la réalisation de A confirme ou infirme la liste d'hypothèses B_k, $k = 1, ..., m$. On connait les probabilités *a priori* des B_k, $k = 1, ..., m$, ainsi que les probabilités conditionnelles $P(A \mid B_k)$, $k = 1, ..., m$, et l'on souhaite calculer les probabilités *a posteriori* $P(B_k \mid A)$, $k = 1, ..., m$; le fait que $P(B_k \mid A)$ soit voisin de 1 est la preuve que la réalisation de A confirme l'hypothèse B_k ; le fait que $P(B_k \mid A)$ soit voisin de 0 est, en revanche, une indication que la réalisation de A infirme l'hypothèse B_k.

Exemple 2.6 (le processus de désintégration atomique) Supposons que les atomes d'un corps radioactif se désintègrent de manière aléatoire. On considère un espace probabilisé $(\Omega, \mathcal{T}, P)$ adapté aux règles imposées par le physicien, à savoir la règle suivante : la probabilité que, si un atome n'est pas désintégré à l'instant t_0, il se désintègre pas dans l'intervalle de temps $[t_0, t_0 + t]$, est une fonction *qui ne dépend que de t* (fonction que l'on appellera $F(t)$) ; regardons l'évolution d'un atome de la population, intact à l'instant $t = 0$; appelons A_s l'évènement *l'atome est encore intact à l'instant $t \geq 0$*. On a $P(A_s) = 1 - F(s)$ et

$$P(A_{s+t}) = P(A_{s+t} \mid A_s)P(A_s),\ s, t \geq 0$$

soit, en tenant compte de ce qui régit le processus,

$$1 - F(s + t) = (1 - F(s))(1 - F(t))\ s, t \geq 0.$$

En résolvant cette équation fonctionnelle classique (et en supposant F continue en $t = 0$), on voit qu'il existe un réel (forcément positif) λ tel que $F(t) = e^{-\lambda t}$, d'où

$$1 - F(t) = \exp(-\lambda t)\,.$$

On a $F(t) = 1 - e^{-\lambda t}$; ce type de processus est dit *processus exponentiel*.

Si A et B sont deux évènements d'un espace probabilité $(\Omega, \mathcal{T}, P)$ avec $P(A) > 0$, alors $P(A \mid B) = P(A)$ signifie intuitivement que la réalisation ou la non réalisation de B n'influe pas sur celle de A ; les évènements A et B sont alors dits indépendants. Plus généralement, on a la définition suivante :

Etant donné un espace probabilisé $(\Omega, \mathcal{T}, P)$, deux évènements A et B, éléments de $\mathcal{T}$ sont dits indépendants si et seulement si $P(A \cap B) = P(A)P(B)$.

Attention ! Il ne faut pas confondre cette notion d'indépendance (qui est une notion probabiliste faisant intervenir la probabilité P) avec la notion d'*incompatibilité*, qui elle est une notion purement ensembliste : deux évènements A et B sont dits *incompatibles* si et seulement si $A \cap B$ est vide.

Cette notion s'étend au cadre des familles d'évènements :

Étant donné un espace probabilisé $(\Omega, \mathcal{T}, P)$ et une collection finie ou dénombrable d'évènements $(A_j)_{j\in J}$, les A_j sont dits mutuellement indépendants si et seulement si, pour tout sous ensemble fini K dans J, on a

$$P\left(\bigcap_{j\in K} A_j\right) = \prod_{j\in K} P(A_j)\,.$$

Étant donnée une collection finie ou dénombrable $(\mathcal{T}_j)_{j\in J}$ de sous-tribus de $\mathcal{T}$, les $\mathcal{T}_j$ sont dites mutuellement indépendantes si et seulement si toute collection (A_j) avec $A_j \in \mathcal{T}_j$ est une collection d'évènements mutuellement indépendants.

Si $(\Omega, \mathcal{T}, P)$ est un espace probabilisé, deux évènements A et B de $\mathcal{T}$ sont indépendants si et seulement si leurs complémentaires A^c et B^c dans Ω sont indépendants : on écrit en effet

$$P(A^c)P(B^c) = (1 - P(A))(1 - P(B)) = 1 - P(A) - P(B) + P(A)P(B)\ ;$$

cette expression vaut $P((A \cup B)^c) = P(A^c \cap B^c)$ si et seulement si $P(A)P(B) = P(A \cap B)$, ce qui prouve l'assertion.

Etant donné un espace probabilisé $(\Omega, \mathcal{T}, P)$, deux évènements A et B tels que $P(A)P(A^c) > 0$ sont indépendants si et seulement si

$$P(B) = P(B \mid A) = P(B \mid A^c)\,.$$

Exemple 2.7 Considérons l'épreuve consistant à tirer un nombre au hasard entre 1 et n, tous les choix étant équiprobables. L'ensemble des évènements élémentaires est $\Omega = \{1, \ldots, n\}$, la probabilité est la loi uniforme. Si p est un diviseur premier de n, notons E_p l'évènement *le nombre tiré est divisible par p*. Il est facile de constater, en utilisant le lemme de Gauss (si un nombre premier divise un produit de facteurs, il divise automatiquement l'un au moins des facteurs) que la collection d'évènements (E_{p_j}), où $(p_j)_j$ est la famille des diviseurs premiers de n, est une collection d'évènements mutuellement indépendants. Alors il en est de même pour la collection des $(E^c_{p_j})$; ceci implique que la probabilité de l'évènement $A :=$ *le nombre tiré est premier avec n* vaut

$$P(A) = \prod_j (1 - \frac{1}{p})\ ;$$

en évaluant $P(A)$ comme le quotient du nombre de cas favorables sur le nombre de cas possibles, on retrouve ainsi la formule d'Euler : si $\varphi(n)$ désigne le nombre d'entiers entre 1 et n premiers avec n, alors

$$\varphi(n) = n \prod_{p|n, p\ premier} (1 - \frac{1}{p_j})\,.$$

2.4 Variables aléatoires.

2.4.1 Variables aléatoires discrètes

Soit un espace probabilisé $(\Omega, \mathcal{T}, P)$; on appelle *variable aléatoire discrète* sur $(\Omega, \mathcal{T}, P)$ toute application de Ω dans un ensemble E fini ou dénombrable telle que l'image réciproque de tout sous-ensemble de E soit un élément de $\mathcal{T}$.

On pourrait penser, à lire cette définition, que la probabilité P ne joue aucun rôle ; de fait, ceci n'est pas le cas, car la donnée d'une variable aléatoire présuppose toujours la donnée d'une distribution de probabilité P sur l'ensemble Ω initial ; en effet, la connaissance de cette distribution de probabilité P sur l'espace des évènements, distribution de probabilité conditionnant la règle du jeu, joue un rôle essentiel dans la définition de ce que l'on appelle la *loi de la variable aléatoire* X (définition que l'on donnera un peu plus loin). Or ce qui est important lorsque l'on considère une variable aléatoire n'est pas tant la fonction elle même que la loi de probabilité à laquelle cette fonction obéit (loi qui dépend d'une part de la fonction, d'autre part de la probabilité P sur l'espace Ω des évènements.

L'ensemble des évènements élémentaires est un ensemble correspondant à tous les résultats possibles de l'épreuve. Donnons deux exemples de variables aléatoires, attachés à deux épreuves souvent rencontrées : les tirages dans une urne avec remise et le jeu de pile ou face.

Exemple 2.8

- **1.** On considère une urne contenant B boules blanches et R boules rouges, hors de laquelle on puise n fois de suite une boule, avec remise à chaque fois, les divers tirages successifs étant indépendants (ce qui ne serait bien sûr pas le cas si les tirages s'effectuaient sans remise !) Le modèle d'espace d'évènements Ω associé à cette épreuve est l'espace $\{1, ..., B+R\}^n$, de cardinal $(B+R)^n$, la tribu étant la famille de toutes les parties de $\{1, B+R\}^n$ (les boules sont numérotées de 1 à $B+R$) ; la probabilité que l'on met sur cet espace Ω est la probabilité uniforme [1]. Un exemple de variable aléatoire discrète sur $(\Omega, \mathcal{P}(\Omega), P)$ est par exemple le nombre de boules rouges tirées au terme des n tirages successifs ; c'est une variable aléatoire à valeurs dans $\{0, ..., n\}$, ensemble fini à $n+1$ éléments.
- **2.** Soit $\Omega = \{0,1\}^{\mathbb{N}^*}$ l'ensemble de toutes les suites indexées par $\mathbb{N}^*$ constituées de 0 ou de 1, équipé de la plus petite tribu contenant les ensembles du type $E_1 \times E_2 \times E_3 \times \cdots$, où $E_k \subset \{0,1\}$ et $E_k = \{0,1\}$ sauf pour un nombre fini d'indices k. Cet ensemble Ω modélise par exemple l'ensemble de tous les résultats possibles d'u jeu sans fin à pile ou face. La probabilité que l'on met est celle qui consiste à poser

$$P(E_1 \times E_2 \times \cdots) = P(E_1) \times P(E_2) \times \cdots \ ;$$

  la règle du jeu est que les résultats des lancers correspondant à des jets différents sont des évènements indépendants et l'on se reporte à la définition de la mutuelle indépendance entre évènements introduite dans la section précédente pour justifier la définition que l'on vient de prendre pour la probabilité sur Ω ; on convient ensuite de ce que, pour un jet donné

$$\begin{aligned} P(\textit{on obtient pile}) &= p \\ P(\textit{on obtient face}) &= 1-p\,, \end{aligned}$$

  où $p \in]0,1[$ (la pièce peut être truquée si $p \neq 1/2$, mais le jeu de pile ou face n'est pas, lui, biaisé, au vu de la probabilité que l'on a mis sur $\{0,1\}^N$). Un exemple de variable

1. On pourrait d'ailleurs aussi prendre comme modèle d'espace d'évènements $\{0,1\}^{\mathbb{N}}$, avec cette fois $P(E_1 \times E_2 \cdots \times E_n) = P(E_1) \cdots P(E_n)$ si $E_k \subset \{0,1\}$, avec $P(\{0\}) = B/(B+R)$ d'une part, et $P(\{1\}) = R/(B+R)$ d'autre part.

aléatoire discrète sur $(\Omega, \mathcal{T}, P)$ (à valeurs dans $\mathbb{N} \cup \{+\infty\}$) est la variable aléatoire discrète correspondant au numéro du jet où l'on obtient *pile* pour la première fois.

Par définition, *la loi d'une variable aléatoire* $X : (\Omega, \mathcal{P}, P) \to E$ (où E est un ensemble fini ou dénombrable) est la probabilité sur E (équipé de la tribu de toutes ses parties) que l'on définit en transportant *via* X la probabilité P (il était donc bien indispensable de disposer d'une probabilité, c'est-à-dire en quelque sorte d'une règle du jeu) sur l'espace probabilisable $(\Omega, \mathcal{T})$. La loi de X est donc la distribution de probabilité P_X sur $(E, \mathcal{P}(E))$ définie par

$$P_X(A) := P(\{\omega \in \Omega\,;\, X(\omega) \in A\}) = P(X^{-1}(A))\,.$$

Reprenons les deux exemples ci-dessus :

Exemple 2.9

- **1.** Dans le premier exemple, on a

$$P(X = k) = \binom{n}{k} \left(\frac{R}{B+R}\right)^k \times \left(\frac{B}{B+R}\right)^{n-k}$$

car le dénombrement des cas favorables donne

$$\binom{n}{k} \times R^k \times B^{n-k}\,;$$

si l'on pose $\tau = R/(B+R)$, la loi de X est donnée par

$$P_X(\{k\}) = \binom{n}{k} \tau^k (1-\tau)^{n-k}\,,\ k = 0, ..., n\,.$$

(on vérifie que la somme de ces nombres vaut bien 1 à cause de la formule du binôme). Cette loi de probabilité très importante (régissant ici la variable aléatoire X du premier exemple) est dite *loi binomiale* de paramètre $\tau \in [0, 1]$.
- **2.** Dans le second exemple, on a

$$P(X = k) = (1-p)^{k-1} p\,,\ k = 1, 2, ...\,;$$

on note que

$$\sum_{k=1}^{\infty} (1-p)^{k-1} p = p \times \sum_{k=0}^{\infty} (1-p)^k = \frac{p}{1-(1-p)} = 1\,,$$

ce qui montre que $P(X = \infty) = 0$ (ce d'ailleurs quelque soit la valeur de p, c'est-à-dire que la pièce soit truquée ou non, ce qui intuitivement n'est pas évident !). La loi de probabilité P_X sur $\mathbb{N}^*$ définie par

$$P_X(\{k\}) = (1-p)^{k-1} p\,,\ k = 1, 2, ...$$

est dite *loi de Pascal* ou encore *loi géométrique* de raison $1 - p$; c'est la version discrète de la *loi exponentielle* que l'on retrouvera plus loin.

Une loi de probabilité sur $\mathbb{N}$ joue un rôle très important en physique ; cette loi est dite *loi de Poisson de paramètre* $\lambda > 0$; elle correspond à la distribution de probabilité sur $\mathbb{N}$ définie par

$$P(\{k\}) = \exp(-\lambda) \frac{\lambda^k}{k!}\,, \quad k = 0, 1, ...$$

Voici la raison pour laquelle cette loi est si importante, en même temps que sa modélisation comme loi d'une certaine variable aléatoire sur un espace probabilisé *ad*

hoc. Nous supposons que nous disposons d'une liste infinie (dénombrable) d'instants τ_i, $i \in \mathcal{I}$. Ces instants seront des instants dits "marqués" sur un intervalle de l'axe réel de longueut $T > 0$, par exemple $[0, T]$, ce marquage s'effectuant de manière stochastique. La règle initiale de marquage est simple : on décide que la probabilité, lorsque l'on marque N points de cet intervalle, pour que k d'entre eux soient dans l'intervalle (t_1, t_2), est

$$p_{t_1,t_2,N}(k) = \begin{pmatrix} N \\ k \end{pmatrix} \Big(\frac{t_2 - t_1}{T}\Big)^k \Big(1 - \Big(\frac{t_2 - t_1}{T}\Big)\Big)^{N-k},$$

ce qui correspond à l'idée que le marquage des instants s'effectue de manière uniforme. Si l'on suppose N très grand et $t_2 - t_1$ petit devant T, on a l'approximation classique de la loi binomiale

$$p_{t_1,t_2,N}(k) \simeq \exp\Big(-\frac{N(t_2 - t_1)}{T}\Big) \times \frac{1}{k!}\Big(\frac{N(t_2 - t_1)}{T}\Big)^k,$$

approximation qui résulte de la formule de Stirling

$$n\,! \simeq \sqrt{2\pi}\, n^{n+1/2} e^{-n}\,.$$

Si l'on suppose maintenant que N et T tendent vers l'infini, mais que le nombre N/T reste constant et égal à un paramètre $\lambda > 0$, l'approximation devient

$$p_{t_1,t_2}(k) \simeq e^{-\lambda(t_2 - t_1)} \frac{1}{k!} (\lambda(t_2 - t_1))^k\,.$$

Le paramètre λ peut être interprété comme la densité de l'ensemble des instants marqués. Ce nombre $p_{t_1,t_2}(k)$ s'interprète comme la probabilité que l'on marque k points dans l'intervalle (t_1, t_2), λ représentant la densité des points (ou particules) marquées. On retrouve la loi de Poisson dans les processus d'émission de particules. Les signaux enregistrés dans les techniques de CAT-Scanner ou d'échographie par exemple sont des bruits poissonniens, c'est-à-dire des signaux correspondant à des variables aléatoires suivant une loi de Poisson (en déterminant le paramètre, on retrouve la densité de rayonnement de l'organe).

2.4.2 Vecteurs de variables aléatoires discrètes ; lois marginales

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé ; un vecteur aléatoire discret (n-dimensionnel) est une application de Ω dans $E_1 \times \cdots \times E_n$, où $E_1, ..., E_n$ sont des ensembles finis ou dénombrables. Comme $E = E_1 \times \cdots \times E_n$ est encore un ensemble fini ou dénombrable, on peut faire entrer la notion de vecteur aléatoire discret dans le cadre de la notion de variable aléatoir discrète. On retiendra cependant une notion, celle de *loi marginale*.

Définition 2.1 *Soit $(X_1, ..., X_n)$ un vecteur de variables aléatoires discrète sur l'espace probabilisé $(\Omega, \mathcal{T}, P)$, à valeurs dans $E_1 \times \cdots \times E_n$, où*

$$E_j := \{x_{j,l}\,;\, l = 0, 1, 2, ...\}\,,\ j = 1, ..., n\,.$$

Pour chaque $j = 1, ..., n$, X_j est une variable aléatoire discrète sur $(\Omega, \mathcal{T}, P)$ et la loi de X_j est la distribution de probabilité P_{X_j} sur E_j donnée par

$$P_{X_j}(\{x_{jk}\}) = \sum_{l_1} \cdots \sum_{l_{j-1}} \sum_{l_{j+1}} \cdots \sum_{l_n} P_{X_1,...,X_n}(\{(x_{1,l_1}, ..., x_{j-1,l_{j-1}}, x_{jk}, x_{j+1,l_{j+1}}, ..., x_{n,l_n})\}),$$

où $P_{X_1,...,X_n}$ désigne la distribution de probabilité du vecteur $(X_1, ..., X_n)$. La loi de X_j ainsi construite est appelée loi marginale (d'indice j) de la loi du vecteur $(X_1, ..., X_n)$.

2.4.3 Variables aléatoires réelles ou vecteurs de variables aléatoires réelles

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé ; une variable aléatoire sur $(\Omega, \mathcal{T}, P)$ et à valeurs réelles (resp. à valeurs dans $\mathbb{R}^n$) est une application X de l'espace des évènements Ω dans $\mathbb{R}$ (resp. dans $\mathbb{R}^n$) telle que, pour tout sous-ensemble quarrable A de $\mathbb{R}$ (resp. $\mathbb{R}^n$), l'ensemble

$$\{\omega \in \Omega\,;\, X(\omega) \in A\}$$

appartienne à la tribu $\mathcal{T}$.

Une telle variable aléatoire sur $(\Omega, \mathcal{T}, P)$ induit par transport une probabilité sur $\mathbb{R}$ (resp. $\mathbb{R}^n$) équipé de la tribu des ensembles quarrables (tribu de Lebesgue). Cette distribution de probabilité P_X sur $\mathbb{R}$ (resp. $\mathbb{R}^n$) est dite *loi de la variable aléatoire réelle X* (resp. *loi du vecteur de variables aléatoires X*).

Une loi de probabilité sur $\mathbb{R}$ (modélisable comme la loi d'une certaine variable aléatoire) joue un rôle important en physique, la *loi de Gauss normale*, définie par

$$P(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-t^2/2}\, dt$$

si A est un sous-ensemble quarrable de $\mathbb{R}$; on a aussi une loi de probabilité sur $\mathbb{R}^n$ en posant

$$P(A) = (2\pi)^{-n/2} \int\!\!\ldots\!\!\int_A \exp(-(x_1^2 + \cdots + x_n^2)/2)\, dx_1 \cdots dx_n$$

pour tout sous-ensemble quarrable A de $\mathbb{R}^n$. Il s'agit d'une distribution de probabilité car

$$\left(\int_{\mathbb{R}} \exp(-t^2/2) dt \right)^2 = \iint_{\mathbb{R}^2} \exp(-(t^2 + u^2)/2)\, dtdu = \int_0^{+\infty} \int_0^{2\pi} e^{-r^2/2} r dr d\theta = 2\pi$$

grâce au théorème de Fubini et à la formule de changement de variables (passage aux coordonnées polaires).

L'expérience du triangle de Galton (figure 2.2) modélise la distribution normale de Gauss comme la loi d'une variable aléatoire ; cette expérience illustre d'ailleurs un théorème limite de la théorie des probabilités, le *théorème limite centrale* que nous verrons plus loin.

Une raison importante qui explique le rôle majeur de la distribution de Gauss en physique est que la fonction

$$t \to \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

se transforme par la prise de spectre essentiellement en elle-même, plus précisément en la fonction

$$\omega \to e^{-\omega^2/2}\,.$$

Cette fonction réalise le meilleur compromis dans le *principe d'incertitude* d'Heisenberg (impossibilité de localiser simultanément une particule et son spectre) et fournit donc un modèle raisonnable pour les particules en mécanique quantique (que ce soit sous l'angle déterministe ou sous l'angle probabiliste).

2.5 Variables aléatoires réelles à densité

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et X une variable aléatoire réelle (resp. un vecteur de variables aléatoires réelles, c'est-à-dire une variable aléatoire à valeurs dans $\mathbb{R}^n$).

On dit que X est une variable aléatoire à densité (resp. un vecteur variables aléatoires à densité) s'il existe une fonction $f : t \to f(t)$, positive, intégrable sur $\mathbb{R}$ et d'intégrale 1 sur $\mathbb{R}$ (resp. une fonction $f : (x_1, ..., x_n) \to f(x_1, ..., x_n)$) telle que la loi de X soit donnée par

$$P_X(A) = \int_A f(t)dt$$

pour tout sous-ensemble quarrable A de $\mathbb{R}$ (resp.

$$P_X(A) = \int\!\!\ldots\!\!\int_A f(x_1, ..., x_n)\, dx_1 \ldots dx_n$$

pour tout sous-ensemble quarrable A de $\mathbb{R}^n$). La fonction f est alors appelée *densité* de la variable aléatoire X (resp. *densité* du vecteur de variables aléatoires X). On dit que la variable aléatoire X (resp. le vecteur de variables aléatoires X) est une *variable aléatoire réelle à densité* (resp. un *vecteur de variables aléatoires réelles à densité*).

Bien sûr, toutes les variables aléatoires réelles ne sont pas à densité ; il est d'ailleurs très important de remarquer que si X est une variable aléatoire réelle à densité (resp. un vecteur de variables aléatoires réelles à densité) sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, alors, pour tout $t \in \mathbb{R}$, on a $P(\{\omega\,;\, X(\omega) = t\}) = 0$ (resp. pour tout $x \in \mathbb{R}^n$, $P(\{\omega\,;\, X(\omega) = x\}) = 0$). Dans le cas d'une variable aléatoire à densité f, il faut comprendre $f(t)$ à travers la formule "physique" $P(X \in [t, t + dt[) = f(t)dt$ et surtout pas $P(X = t) = f(t)$ car $P(X = t)$ vaut toujours 0 pour une telle variable puisque $P_X(\{t\}) = P(X = t) = \int_{\{t\}} f(u)du = 0$ (le singleton $\{t\}$ est de mesure nulle).

Une variable aléatoire (resp. un vecteur de variables aléatoires) prenant ses valeurs dans un sous-ensemble dénombrable de $\mathbb{R}$ (resp. de $\mathbb{R}^n$) ne saurait être à densité ; si par exemple X suit une loi de Poisson de paramètre $\lambda > 0$ sur $(\Omega, \mathcal{T}, P)$, on a, pour tout $n \in \mathbb{N}$, $P(X = n) = e^{-\lambda}\lambda^n/n! > 0$).

Exemple 2.10

- une variable aléatoire suivant une loi normale ($P_X(A) = (2\pi)^{-1/2} \int_A e^{-t^2/2}\, dt$) est une variable aléatoire à densité ; plus généralement, si $m_1, ..., m_n$ sont des nombres réels, $\sigma_1, ..., \sigma_n$ des nombres réels strictement positifs

$$f(x_1, ..., x_n) = (2\pi)^{-n/2}(\sigma_1 \cdots \sigma_n)^{-1/2} \exp\Big(- \sum_{j=1}^{n} \frac{(x_j - m_j)^2}{2\sigma_j} \Big),$$

  un vecteur de variables aléaoires ayant pour densité f est dit *vecteur gaussien* ; de même que les variables normales jouent un rôle clef en physique (particulièrement, on l'a vu, en physique quantique), les vecteurs gaussiens (qui en sont la généralisation naturelle après translation, dilatation, et passage d'une variable à plusieurs) ont aussi un rôle essentiel ;
- on retrouve la variable aléatoire réelle de densité

$$f(t) \ := \begin{cases} \lambda e^{-\lambda t} \text{ si } t \geq 0 \\ 0 \text{ si } t < 0 \end{cases}$$

  dans le processus de désintégration atomique (voir l'exemple 1.6).

2.6 Fonction de répartition d'une variable aléatoire réelle

Si X est une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, on préfère souvent retenir sa loi en retenant sa *fonction de répartition*, c'est-à-dire la fonction

$$t \to P_X(]-\infty, t[) = P(X \in]-\infty, t[)\,.$$

Cette fonction est croissante, de limite 0 en $-\infty$, de limite 1 en $+\infty$, et est continue à gauche ; d'ailleurs une fonction de $\mathbb{R}$ dans $[0, 1]$ ayant ces quatre propriétés est la fonction de répartition d'une variable aléatoire.

Si la fonction de répartition est de classe C^1 et de dérivée f , la variable aléatoire réelle correspondante est une variable à densité, de densité précisément f.

Exemple 2.11

- La fonction de répartition d'une variable aléatoire gaussienne suivant une loi normale est

$$t \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-u^2/2}\, du \ ;$$

  seule une table fournit les valeurs de cette fonction (inexprimable en termes de fonctions simples).
- La fonction de répartition d'une variable aléatoire à densité $t \to \lambda e^{-\lambda t}$ pour $t \geq 0$ est

$$t \to \lambda \int_0^t e^{-\lambda u} du = \Big[- e^{-\lambda u} \Big]_0^t = 1 - e^{-\lambda t}$$

  si $t \geq 0$ et vaut 0 si $t < 0$.

2.7 Espérance et variance d'une variable aléatoire

2.7.1 Le cas des variables aléatoires à valeurs dans un sous-ensemble fini ou dénombrable de $\mathbb{R}^n$

Si X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ et à valeurs dans un sous-ensemble fini de $\mathbb{R}^n$, on appelle *espérance* de X la "valeur

moyenne" que prend X ; comme X prend une valeur donnée x_k avec la probabilité $p_k = P(X = x_k) = P_X(\{x_k\})$, cette valeur moyenne vaut naturellement

$$E[X] := \sum_k x_k P(X = x_k)$$

(la somme des valeurs prises, ces valeurs étant "pondérées" de coefficients correspondant aux probabilités respectives avec lesquelles elles sont prises). On retrouve bien ici le procédé classique pour calculer une moyenne en tenant compte de coefficients.

Si maintenant X prend ses valeurs dans un sous-ensemble dénombrable de $\mathbb{R}^n$, $X(\Omega) = \{x_k\,;\, k \in \mathbb{N}\}$, on dit que X *a une espérance* si et seulement si la quantité

$$\sum_k \|x_k\|\, P(X = x_k)$$

est finie ; si c'est le cas, on peut poser sans ambigüité

$$E(X) = \lim_{N\to\infty} \sum_{k=0}^{N} x_k P(X = x_k)$$

et l'espérance (ou plutôt le *vecteur des espérances*) si X est à valeurs vectorielles est bien définie sur le modèle d'une valeur moyenne comme dans le cas où X prend ses valeurs dans un ensemble fini.

Dans le cas particulier où X est une variable aléatoire réelle positive sur un espace probabilisé $(\Omega, \mathcal{T}, P)$ et telle que

$$\sum_k |x_k|\, P(X = x_k) = \sum_k x_k\, P(X = x_k) = +\infty\,,$$

on peut encore définir l'espérance de X et l'on dit que X *est d'espérance infinie*. Si X est à valeurs vectorielles ou à valeurs réelles non positives, la clause de sécurité

$$\sum_k \|x_k\|\, P(X = x_k) < \infty$$

est essentielle pour povoir parler de l'espérance de X.

Exemple 2.12

- Si X suit une loi de Poisson de paramètre $\lambda > 0$, on a

  $$E[X] := \sum_{k=0}^{\infty} k \times e^{-\lambda}(\lambda^k/k!) = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \lambda^{k-1}/(k-1)! = \lambda e^{-\lambda} \times e^{\lambda} = \lambda\,;$$

  l'espérance est donc égale au pararamètre pour une telle loi.
- Supposons que X suive une loi binomiale de paramètres (n, τ) ; pour calculer l'espérance, on utilise une méthode classique empruntée à Polya, dite *méthode des séries génératrices* ; soit $\tau' = 1 - \tau$; on calcule

  $$(\tau x + \tau')^n = \sum_{k=0}^{n} \binom{n}{k} \tau^k (\tau')^{n-k} x^k$$

  puis on dérive, ce qui donne

  $$n\tau(\tau x + \tau')^{n-1} = \sum_{k=1}^{n} k \binom{n}{k} \tau^k (\tau')^{n-k} x^{k-1}\,;$$

  on évalue enfin en 1, ce qui donne $E[X] = \sum_{k=0}^{n} kP(\{X = k\}) = np$.

– Supposons que X suive une loi géométrique de paramètre p ; on écrit, pour $x \in]-1,1[$,

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \,,$$

puis on dérive, ce qui donne

$$\left(\frac{1}{1-x}\right)^2 = \sum_{k=1}^{\infty} k x^{k-1}$$

d'où

$$E[X] = (1-p)\sum_{k=1}^{\infty} k p^k = \frac{p}{1-p}\,.$$

– Le mécanisme est un peu plus complexe pour la loi hypergéométrique de paramètres rationnels $(n, \tau_1 = R/N, \tau_2 = (N-R)/N)$; on part de

$$(1+xy)^R (1+y)^{N-R} = \sum_{n=0}^{\infty} \Big(\sum_{k_1+k_2=n} \binom{R}{k_1} \binom{N-R}{k_2} x^{k_1} \Big)\, y^n$$

avec les conventions

$$\binom{R}{k_1} = 0,\ k_1 > R$$

et

$$\binom{N-R}{k_2} = 0,\ k_2 > N-R\,;$$

en dérivant par rapport à x, on obtient

$$yR(1+xy)^{R-1}(1+y)^{N-R} = \sum_{n=0}^{\infty} \left(\sum_{\substack{k_1+k_2=n \\ k_1>0}} k_1 \binom{R}{k_1} \binom{N-R}{k_2} x^{k_1-1} \right) y^n \ ;$$

faisant $x = 1$ et identifiant les coefficients de y^n, on trouve

$$E[X] = nR/N \ ;$$

la loi hypergéométrique de paramètres $(n, \tau, 1-\tau)$ a donc, comme la loi binomiale de paramètres n, τ, pour espérance $n\tau$ (que τ soit rationnel ou non).

Connaître l'espérance d'une variable aléatoire réelle (ou d'un vecteur de variables aléatoires réelles) prenant au plus une infinité dénombrable de valeurs ne suffit pas à l'analyse fine du résultat de l'épreuve dont cette variable aléatoire affiche le résultat. Il est naturel de concevoir, lorsque vous disposez par exemple de la totalité des notes d'une épreuve d'examen, que la moyenne de ces notes n'est pas une information suffisante sur le résultat : comment par exemple discerner, si la moyenne est 10/20 le cas où toutes les notes sont agglutinées à la moyenne et le cas où les copies se rangent en deux paquets, la moitié des notes se situant entre 0 et 5, l'autre moitié entre 15 et 20 ? Pour compléter l'information relative à une variable aléatoire réelle donnée (ou à un vecteur de variables alétoires réelles, les variables prenant toujours ici au plus une infinité dénombrable de valeurs), il faut définir, lorsque cela est possible, une autre notion, celle de *variance*.

Soit X une variable aléatoire réelle comme ci-dessus ou un vecteur de variables aléatoires réelles défini sur un espace probabilisé $(\Omega, \mathcal{T}, P)$. On dit que X *admet une variance* (ou encore est *à variance finie*) si et seulement si la variable positive $\|X\|^2$ admet une espérance ; ceci implique que le vecteur X a une espérance et l'on définit la variance de X par

$$V(X) := E\Big[\|X - E[X]\|^2\Big]\,.$$

L'*écart type* est alors la racine carrée de la variance.

$$\sigma(X) := \sqrt{E\Big[\parallel X - E[X] \parallel^2\Big]}\,.$$

L'écart type mesure donc ce en quoi diffère la variable X d'une variable prenant la valeur $E[X]$, soit la valeur moyenne de X, avec la probabilité 1.

Exemple 2.13

- Un calcul inspiré des techniques utilisant les fonctions génératrices (comme pour le calcul de l'espérance) donne la variance d'une loi binômiale de paramètres (n, τ) ; on a alors pour une telle loi $V(X) = n\tau(1-\tau)$.
- Une variable suivant une loi de Poisson de paramètre λ a pour variance $E[X] = \lambda$.

2.7.2 Le cas général des variables aléatoires réelles ou des vecteurs de variables aléatoires à densité

Supposons maitenant que X soit une variable alétoire réelle positive à densité sur un espace probabilisé $(\Omega, \mathcal{T}, P)$; pour chaque $t \in [0, \infty[$, on sait que $P(X = t) = 0$; par contre, la quantité infinitésimale $P(t \leq X \leq t + dt)$ vaut, elle, $f(t)\,dt$ (c'est précisément la définition de la densité). L'analogue continu de l'expression qui nous a permis la définition de l'espérance est donc

$$E[X] = \sum_{t \geq O} tP(X \in [t, t+dt[) \simeq \int_0^\infty t f(t) dt \in [0, \infty]\,.$$

Si maintenant X est une variable aléatoire réelle à densité, que l'on peut écrire $X = X^+ - X^-$, où $X^+ := \sup(X, 0)$ est à densité $f \times \chi_{[0,\infty]}$ et $X_- := \sup(-X, 0)$ est à densité $f(-t) \times \chi_{]0,\infty[}$, on dit que X admet une espérance si et seulement si les quantités $E[X^+]$ et $E[X^-]$ sont finies, ce qui équivaut à dire que

$$\int_{\mathbb{R}} |t|\, f(t)\, dt < \infty\,;$$

l'espérance de X est alors

$$E[X] = E[X^+] - E[X^-] = \int_{\mathbb{R}} t\, f(t)\, dt\,.$$

Si enfin $X = (X_1, ..., X_n)$ est un vecteur de variables aléatoire à densité sur $(\Omega, \mathcal{T}, P)$, de densité la fonction intégrable $f(x_1, ..., x_n)$, on dit que X admet une espérance si

$$\int\!\!\ldots\!\!\int_{\mathbb{R}^n} \|x\|\, f(x_1, ..., x_n)\, dx_1 \ldots dx_n < \infty\,;$$

on appelle alors espérance de $(X_1, ..., X_n)$ (ou plutôt *vecteur des espérances de* X) le vecteur de $\mathbb{R}^n$ suivant :

$$\left(\int\!\!\ldots\!\!\int_{\mathbb{R}^n} x_1\, f(x_1, ..., x_n)\, dx_1 \ldots dx_n\,, \ldots\,,\ \int\!\!\ldots\!\!\int_{\mathbb{R}^n} x_n\, f(x_1, ..., x_n)\, dx_1 \ldots dx_n \right).$$

Exemple 2.14

- Le vecteur des espérances d'un vecteur gaussien $(X_1, ..., X_n)$ de densité

$$f(x_1, ..., x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{\sigma_1 \cdots \sigma_n}} \exp\Big(- \sum_{j=1}^{n} \frac{(x_j - m_j)^2}{2\sigma_j} \Big)$$

  vaut $(m_1, ..., m_n)$; un tel vecteur modélise une particule vivant statistiquement près du point $(m_1, ..., m_n)$.
- Une variable aléatoire réelle de densité

$$f(t) = \frac{1}{\pi(1+t^2)}$$

  (loi de Poisson continue) n'a pas d'espérance car

$$\int_{\mathbb{R}} \frac{|t|}{1+t^2} dt = +\infty \,.$$

- Une variable aléatoire réelle à densité $\lambda e^{-\lambda t}$ pour $t > 0$ (comme dans le processus de désintégration atomique) a pour espérance

$$E[X] = \lambda \int_0^\infty t e^{-\lambda t}\, dt = \Big[- t e^{-\lambda t} \Big]_0^\infty + \int_0^\infty e^{-\lambda t}\, dt = 1/\lambda \,.$$

On se contentera dans ce cours de définir l'espérance des variables aléatoires réelles à densité, mais on notera néanmoins qu'en couplant ce sous-paragraphe avec le précédent, on peut définir l'espérance d'une variable aléatoire réelle se présentant comme la somme d'une variable aléatoire X_{cont} réelle à densité et d'une variable aléatoire réelle X_{disc} prenant au plus une infinité dénombrable de valeurs. L'espérance de X existe si et seulement si les quantités $E[|X_{\text{cont}}|]$ et $E[|X_{\text{disc}}|]$ sont finies et vaut alors

$$E[X] := E[X_{\text{cont}}] + E[X_{\text{disc}}] \,,$$

ces deux quantités ayant été définies respectivement dans ce qui précède et dans la sous-section précédente.

Ceci se trouve justifié par une propriété importante de l'espérance : le fait que la prise d'espérance soit une opération linéaire, c'est-à-dire que, si X et Y sont deux vecteurs de variables aléatoires réelles (de $(\Omega, \mathcal{T}, P)$ dans $\mathbb{R}^n$) et si $E[\|X\|]$ et $E[\|Y\|]$ sont des quantités finies, alors

$$E[\lambda X + \mu Y] = \lambda E[X] + \mu E[Y] \,, \quad \forall \lambda, \mu \in \mathbb{R} \,.$$

Prouvons juste cela au moins heuristiquement si $n = 1$. On a

$$\begin{aligned}
E[\lambda X + \mu Y] &\simeq \sum_{t,s} (\lambda t + \mu s) P(X \in [t, t+dt[, Y \in [s, s+ds[) \\
&\simeq \lambda \sum_t t \Big(\sum_s P(X \in [t, t+dt[, Y \in [s, s+ds[) \Big) \\
&\qquad + \mu \sum_s s \Big(\sum_t P(X \in [t, t+dt[, Y \in [s, s+ds[) \Big) \\
&\simeq \lambda \sum_t t P(X \in (t, t+dt[) + \mu \sum_s s P(X \in [s, s+ds[) \\
&\simeq \lambda E[X] + \mu E[Y] \,.
\end{aligned}$$

La linéarité de l'espérance est un fait très important. Si l'on y réfléchit bien, on voit que cette linéarité s'explique exactement comme la linéarité de l'intégrale, mais cette fois l'ensemble de départ des fonctions n'est plus $\mathbb{R}^n$ comme au chapitre 1, mais un ensemble abstrait Ω sur lequel, comme pour $\mathbb{R}^n$, on a le moyen de "mesurer" certains ensembles dits quarrables (ici en l'occurrence les ensembles éléments de la tribu $\mathcal{T}$).

Si X est un vecteur de variables aléatoires réelles à densité (ou plus généralement la somme d'un tel vecteur $X_{\text{cont}} : (\Omega, \mathcal{T}, P) \to \mathbb{R}^n$ et d'un vecteur X_{disc} de variables aléatoires de $(\Omega, \mathcal{T}, P)$ dans un sous-ensemble au plus dénombrable de $\mathbb{R}^n$), on dit que X *a une variance* si et seulement si $E[\|X_{\text{cont}}\|^2]$ et $E[\|X_{\text{disc}}\|^2]$ sont finies. La variance de X est alors

$$V(X) := E[\|X - E[X]\|^2] \, .$$

On remarque que si X est une variable aléatoire réelle ayant une variance, alors la linéarité de l'espérance implique :

$$\begin{aligned} V(X) = E[(X - E[X])^2] &= E[X^2 - 2E[X]X + (E[X])^2)] \\ &= E[X^2] - 2(E[X])^2 + (E[X])^2 = E[X^2] - (E[X])^2 \end{aligned}$$

puisque l'espérance d'une variable aléatoire constante $X = C$ vaut C (on applique ceci avec $C = E[X]$).

Contrairement à la prise d'espérance, la prise de variance n'est pas une opération linéaire ; comme la prise d'énergie en physique, c'est une opération quadratique ; on a d'ailleurs choisi la variance (dans le cadre des probabilités discrètes sur un ensemble fini à N éléments) parmi la liste d'exemples de formes quadratiques proposée dans la section 1.7.1 ; le calcul de la variance d'une somme de deux variables aléatoires réelles ayant toutes les deux une variance fait apparaître des termes croisés (ou encore d'interférence) que dans le vocabulaire des probabilités, on appelle *termes de corrélation* ; plus précisément, on a, si X et Y sont deux variables aléatoires réelles ayant toutes les deux une variance et définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$:

$$\begin{aligned} V(X+Y) &= E[(X+Y)^2] - (E[X] + E[Y])^2 \\ &= V(X) + V(Y) + 2E[(X - E[X])(Y - E[Y])] \, ; \end{aligned}$$

la quantité "mixte"

$$E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]\, E[Y]$$

(qui existe dès que X et Y sont deux variables aléatoires réelles telles que X^2 et Y^2 soient d'espérance finie), est dite *covariance* des variables X et Y.

La notion de covariance se généralise au cas d'un vecteur $X = (X_1, ..., X_n)$ de variables aléatoires réelles tels que $E[\|X\|^2]$ soit fini : la *matrice de covariance* de X symétrique [2]

$$\text{covar}\,(X_1, ..., X_n) := \Big(E[(X_i - E[X_i])(X_j - E[X_j])] \Big)_{1 \le i,j \le n} \, .$$

2. C'est en fait une matrice de Gram (voir la section 1.7.3) relativement cette fois non à un produit scalaire, mais à une forme bilinéaire symétrique.

La variance de $\lambda_1 X_1 + \cdots + \lambda_n X_n$ vaut

$$\mathrm{Var}\,(\lambda_1 X_1 + \cdots + \lambda_n X_n) = (\lambda_1, ..., \lambda_n) \bullet \mathrm{covar}\,(X_1, ..., X_n) \bullet \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} .$$

Si $X_1, ..., X_n$ sont à valeurs complexes, la matrice de covariance [3] est

$$\mathrm{covar}\,(X_1, ..., X_n) := \Big(E[(X_i - E[X_i])(\overline{X_j - E[X_j]})] \Big)_{1 \le i,j \le n}$$

et l'on a dans ce cas

$$\mathrm{Var}\,(\lambda_1 X_1 + \cdots + \lambda_n X_n) = (\overline{\lambda_1}, ..., \overline{\lambda_n}) \bullet \mathrm{covar}\,(X_1, ..., X_n) \bullet \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} , \qquad \forall \lambda \in \mathbb{C}^n .$$

2.7.3 Les inégalités de Markov et de Bienaymé-Tchebychev

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et X une variable aléatoire réelle positive sur cet espace ; alors, si λ est un nombre réel strictement positif donné, on a, formellement :

$$\lambda \sum_{t>\lambda} P(t \le X \le t+dt) \le \sum_{t>\lambda} t\, P(t \le X \le t+dt) \le \sum_{t} t\, P(t \le X \le t+dt) \simeq E[X] ;$$

mais le membre de gauche de cette inégalité est une manière approchée d'écrire $\lambda P(X > \lambda)$, d'où il résulte l'importante *inégalité de Markov* :

$$P(X > \lambda) \le \frac{E[X]}{\lambda} , \qquad \forall \lambda > 0 ,$$

valable pour toute variable aléatoire réelle positive sur un espace probabilisé.

Si X est une variable aléatoire réelle admettant une variance, on peut utiliser cette inégalité (en l'appliquant à la variable aléatoire réelle positive $|X - E[X]|^2$) pour contrôler quantitativement comment X diffère statistiquement de sa moyenne ; on a en effet

$$P(|X - E[X]|^2 > \lambda^2) = P(|X - E[X]| > \lambda) \le \frac{E[|X - E[X]|^2]}{\lambda^2} = \frac{V(X)}{\lambda^2} .$$

C'est la célèbre *inégalité de Bienaymé-Tchebychev*, ingrédient essentiel de l'argumentation probabiliste. Cette inégalité est intéressante pour les valeurs de λ inférieures à l'écart type de X.

2.8 Indépendance de variables aléatoires réelles

2.8.1 Indépendance de deux variables aléatoires réelles

Deux variables aléatoires réelles X et Y sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$ sont dites *indépendantes* si et seulement si, pour toute paire (A, B) de boréliens $\mathbb{R}$, on a

$$P(\{\omega \in \Omega\,;\, x \in A\,,\, y \in B\}) = P_X(A) \times P_Y(B) .$$

3. C'est toujours une matrice de Gram, mais associée cette fois à une forme sesquilinéaire vérifiant la symétrie hermitienne, voir la section 1.7.4.

Si les variables aléatoires X et Y sont à valeurs discrètes (respectivement dans E_1 et E_2), dire que X et Y sont indépendantes équivaut à dire que la loi du couple (X, Y) est le produit des lois marginales, c'est à dire

$$P_{X,Y}(\{(x_{1,l_1}, x_{2,l_2})\}) = P_{X_1}(\{x_{1,l_1}\}) \times P_{X_2}(\{x_{2,l_2}\})\,.$$

Si X et Y sont deux variables aléatoires réelles indépendantes à densités (de densités respectives f_X et f_Y, le vecteur (X, Y) est un vecteur de variables aléatoires réelle à densité, de densité

$$f_{(X,Y)}(x, y) = f_X(x) \times f_Y(y)\,.$$

Ceci résulte immédiatement de la définition de l'indépendance. Réciproquement, si (X, Y) est un couple de variables aléatoires réelles à densité $(x, y) \to f(x, y) = g(x)h(y)$, alors X et Y sont des variables aléatoires réelles indépendantes, de densités respectives $x \to g(x)$ et $y \to h(y)$.

En fait, pour des variables aléatoires réelles, on dispose du critère suivant (que l'on admettra) : deux variables aléatoires réelles X et Y sont indépendantes si et seulement si, pour tout choix de fonctions boréliennes bornées g et h, on a

$$E[g(X)h(Y)] = E[g(X)] \times E[h(Y)]\,.$$

Contentons nous ici de retrouver ce résultat sous l'hypothèse que le couple (X, Y) est un couple à densité, de densité $u(x, y)$. L'indépendance de X et Y équivaut, on l'a vu, au fait que u se scinde en

$$u(x, y) = f_1(x)f_2(y)$$

(presque partout), où f_1 doit alors être interprétée comme la densité de la loi de X et f_2 celle de la densité de la loi de Y. On a

$$\begin{aligned} E[g(X)] &\simeq \sum_x g(x)\, P(x \le X < x + dx) = \int_{-\infty}^{\infty} g(x)\, f_1(x)\, dx \\ E[h(Y)] &\simeq \sum_y h(y)\, P(y \le Y < y + dy) = \int_{-\infty}^{\infty} h(y)\, f_2(y)\, dy\,. \end{aligned}$$

D'autre part

$$\begin{aligned} E[g(X)h(Y)] &= \sum_x \sum_y g(x)h(y)\, P(x \le X < x + dx\,,\, y \le Y < y + dy) \\ &= \iint_{\mathbb{R}^2} g(x)\, h(y)\, u(x, y)\, dxdy \\ &= \iint_{\mathbb{R}^2} g(x)\, h(y)\, f_1(x)\, f_2(y)\, dxdy \\ &= \Big(\int_{-\infty}^{\infty} g(x)\, f_1(x)\, dx\Big) \times \Big(\int_{-\infty}^{\infty} h(y)\, f_2(y)\, dy\Big) \\ &= E[g(X)] \times E[h(Y)] \end{aligned}$$

à cause du théorème de Fubini permettant le calcul des intégrales multiples en intégrant successivement par rapport aux diverses variables.

Si X et Y sont indépendantes et ont toutes deux une variance, on a formellement

$$\begin{aligned} E[XY] &= \sum_t \sum_s tsP(X \in [t, t+dt[, Y \in [s, s+ds[) \\ &= \sum_t \sum_s tsP(X \in [t, t+dt[)\, P(Y \in [s, s+ds[) \\ &= \Big(\sum_t tP(X \in [t, t+dt[)\Big) \times \Big(\sum_s sP(Y \in [s, s+ds[)\Big) \\ &= E[X] \times E[Y]\,. \end{aligned}$$

Si deux variables aléatoires réelles ayant toutes les deux une variance sont indépendantes, on a

$$\mathrm{covar}\,(X, Y) = E[XY] - E[X]\,E[Y] = 0\,.$$

La réciproque de cette proposition est fausse : deux variables aléatoires réelles peuvent avoir une covariance nulle sans être indépendantes.

Si $(X_1, ..., X_n)$ sont n variables aléatoires réelles deux à deux indépendantes, la matrice de covariance $\mathrm{covar}\,(X_1, ..., X_n)$ est une matrice diagonale ; comme dans le cas de deux variables, la réciproque est fausse.

Si $X_1, ..., X_n$ sont n variables indépendantes deux à deux, on a, du fait des formules intervenant dans le calcul de variance,

$$V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)\,.$$

Exemple 2.15 Si X et Y sont deux variables indépendantes, à valeurs dans $\mathbb{N}$ et suivant toutes deux une loi de Poisson (de paramètres respectifs λ_X, λ_Y), on a, pour tout $n \in \mathbb{N}$,

$$P(\{X+Y=n\}) = \sum_{k=0}^{n} P(\{X=k\})P(\{Y=n-k\}) = e^{-\lambda_X-\lambda_Y}\,\frac{(\lambda_X+\lambda_Y)^n}{n!}\,,$$

ce qui montre que $X+Y$ suit aussi une loi de Poisson de paramètre $\lambda_X + \lambda_Y$. A propos de cet exemple, examinons un phénomène physique où intervient la loi de Poisson : il s'agit de l'émission de particules. Etant donnés deux instants $t < s$, notons $A_k(t, s)$ l'évènement "*k particules exactement sont émises pendant le laps de temps* $]t, s]$". L'expérience est censée obéir aux règles suivantes :

- (a) si les intervalles $]t_j, s_j]$ sont disjoints, les évènements correspondants $A_{k_j}(t_j, s_j)$ sont indépendants ;
- (b) pour tout k, il existe une fonction ϕ_k telle que $P(A_k(t, s)) = \phi_k(s - t)$ (c'est ce qu'on appelle une condition de stationnarité sur le processus) ;
- (c) si l'on note $B(t) := \bigcup_{k=2}^{\infty} P(A_k(0, t))$, on a

$$\lim_{\substack{t\to 0 \\ t>0}} \frac{P(B(t))}{t} = 0\,.$$

Si l'on suppose $\alpha = \phi_0(1) > 0$ (ce que nous ferons), il est facile de voir, compte tenu des hypothèses (a) et (b), que $\phi_0(n) = \alpha^n$ pour tout entier positif n ; pour tout rationnel positif ξ, on en déduit $\phi_0(\xi) = \alpha^\xi$. La continuité à gauche de ϕ_0 (propriété des mesures de probabilité) permet d'étendre ceci à tout ξ réel positif. Mais

$$\phi_1(t) = 1 - \alpha^t + o(t) = -\log(\alpha)\, t + o(t)$$

du fait de l'hypothèse (c). Mais, grâce à l'indépendance, on a, pour tout $k \in \mathbf{N}$, pour tout $t, s > 0$,

$$\phi_k(t+s) = \sum_{j=0}^{k} \phi_j(t)\phi_{k-j}(s)\,,$$

ce que l'on peut aussi écrire, toujours grâce à (c),

$$\phi_k(t+s) = \phi_k(t)\alpha^s + \phi_{k-1}(t)((-\log(\alpha)s + o(s)) + o_1(s) \ ;$$

on voit ainsi que ϕ_k est dérivable à droite sur $[0, \infty[$; la fonction étant croissante, elle est dérivable sur $]0, \infty[$ et solution de l'équation différentielle

$$\phi_k'(t) = -\log(\alpha)(\phi_{k-1}(t) - \phi_k(t)).$$

Par une induction immédiate (sur k), on voit que, si $\lambda = -\log(\alpha) > 0$, on a

$$\phi_k(t) = \frac{(\lambda t)^k}{k!}\exp(-\lambda t)$$

et l'on voit apparaitre la loi de Poisson de paramètre λ.

Si X et Y sont deux variables aléatoires réelles indépendantes à densités (de densités respectives f_X et f_Y, le vecteur (X, Y) est un vecteur de variables aléatoires réelle à densité, de densité

$$f_{(X,Y)}(x, y) = f_X(x) \times f_Y(y)\,.$$

Ceci résulte immédiatement de la définition de l'indépendance.

Si $X_1, ..., X_n$ sont n variables indépendantes deux à deux, on a, du fait des formules intervenant dans le calcul de variance,

$$V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)\,.$$

2.8.2 Une application importante : la loi faible des grands nombres

Soient $X_1, ..., X_n, ...$ une suite de variables indépendantes deux à deux, toutes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$ et ayant même loi, de moyenne m et de variance finie. Alors la suite de variables

$$Z_n := \frac{\sum_{k=1}^{n} X_k}{n} - m$$

satisfait la propriété suivante :

$$\forall \lambda > 0\,, \quad \lim_{n\to\infty} P(|Z_n| \geq \lambda) = 0\,.$$

On dit que la suite $(Z_n)_n$ *converge en probabilité* vers 0. Plus généralement, une suite de variables aléatoires réelles $(Z_n)_n$, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$ *converge en probabilité* (ou encore *converge stochastiquement*) vers une variable aléatoire réelle Z si et seulement si

$$\forall \lambda > 0\,, \quad \lim_{n\to\infty} P(|Z_n - Z| \geq \lambda) = 0\,.$$

Ce résultat, annonçant la *loi forte des grands nombres* que nous énoncerons ultérieurement, est connu comme la *loi faible des grands nombres*. Cette loi faible résulte de l'inégalité de Bienaymé-Tchebychev. En effet, si

$$Y_n = \frac{\sum_{j=1}^{n} X_j}{n}$$

on a $V(Y_n) = (1/n^2)(V(X_1) + \cdots + V(X_n)) = V(X_1)/n$ (puisque les X_j sont indépendantes deux à deux et que la prise de variance est une opération quadratique) et $E[Y_n] = m$ (d'après la linéarité de la prise d'espérance) ; on a donc

$$P(|Y_n - m| > \lambda) = P(|Z_n| > \lambda) \leq \frac{V(Y_n)}{\lambda^2} = \frac{V(X_1)}{n\lambda^2},$$

ce qui montre que cette quantité tend bien vers 0 lorsque n tend vers l'infini.

On reviendra sur cette notion de convergence (comparée avec d'autres) dans une section ultérieure.

2.8.3 Régression linéaire

Cette notion a déjà été introduite à l'occasion de l'utilisation du théorème de projection orthogonale (exemple 1.7, section 1.7.3) mais nous la rappelons ici dans ce nouveau contexte car elle joue un rôle important en probabilités ou en statistique. Etant données deux variables aléatoires réelles X et Y, toutes deux définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$, ayant toutes les deux une variance, on appelle *droite de régression de Y relativement à X* la droite d'équation $y = \alpha x + \beta$, où la variable $Z = \alpha X + \beta$ réalise la meilleure approximation de Y au sens des moindres carrés pour la fonctionnelle correspondant à la corrélation, c'est-à-dire est la combinaison linéaire des variables 1 et X telle que

$$E[|Y - \alpha X - \beta|^2]$$

soit minimale. Chercher la droite de régression de Y relativement à X, c'est tenter de déterminer les meilleurs paramètres α, β qui donnent la fonction affine en X dont la distribution statistique des valeurs approche *au mieux, et en tout cas en un sens stochastique*, la distribution statistique des valeurs de Y. Les conditions vérifiées par α et β doivent être :

$$E[Y - \alpha X - \beta] = E[(Y - \alpha X - \beta)X] = 0$$

Si $V(X)\,V(Y) \neq 0$ et si $\sigma(X)$, $\sigma(Y)$ désignent les écarts-types de X et Y, l'équation de la droite de régression s'écrit

$$\frac{y - E[Y]}{\sigma(Y)} = \left(\frac{\text{covar}\,(X,Y)}{\sigma(X)\sigma(Y)}\right)\frac{x - E[X]}{\sigma(X)}.$$

Le coefficient

$$\rho(x,y) := \frac{\text{covar}\,(X,Y)}{\sigma(X)\sigma(Y)}$$

est dit *coefficient de corrélation de X et Y* ; ce coefficient peut être nul quand bien même X et Y sont indépendantes !

2.8.4 Indépendance mutuelle d'une famille de variables indépendantes

Si X est une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{T}, P)$, la tribu $\mathcal{T}(X)$ est par définition la plus petite tribu contenant tous les ensembles $X^{-1}(A)$, où A est un borélien quelconque de $\mathbb{R}$.

Une collection finie ou dénombrable $(X_i)_{i\in I}$ de variables aléatoires réelles toutes définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$ est dite *collection de variables mutuellement indépendantes* si les tribus $\mathcal{T}(X_i)$, $i \in I$, forment une collection de tribus mutuellement indépendantes. La même définition vaut si les X_i sont des vecteurs de variables aléatoires réelles ; on dit que la collection des vecteurs de variables aléatoires réelles $(X_i)_i$ est une collection de vecteurs de variables aléatoires réelles mutuellement indépendants si les tribus $(\mathcal{T}(X_i))_i$ forment une collection de sous-tribus de $\mathcal{T}$ mutuellement indépendantes.

Si $X_1, ..., X_n$ sont des variables aléatoires discrètes, l'indépendance mutuelle de $(X_1, ..., X_n)$ équivaut au fait que la loi du vecteur $(X_1, ..., X_n)$ soit le produit des lois marginales $P_{X_1}, ..., P_{X_n}$, c'est-à-dire :

$$P_{X_1,...,X_n}(\{(x_{1,l_1}, ..., x_{n,l_n})\} = \prod_{j=1}^{n} P_{X_j}(\{x_{j,l_j}\})\,.$$

Si $X_1, ..., X_n$ sont des variables aléatoires réelles, l'indépendance mutuelle de $X_1, ..., X_n$ équivaut à ce que la densité $(x_1, ..., x_n) \to f(x_1, ..., x_n)$ du vecteur $(X_1, ..., X_n)$ s'exprime prsque partout sous la forme

$$f(x_1, ..., x_n) = f_1(x_1) \times \cdots \times f_n(x_n)\,,$$

chaque fonction $x \to f_j(x)$ étant alors une densité sur $\mathbb{R}$ ($x \to f_j(x)$ correspond à la loi de la variable aléatoire X_j).

On admettra aussi que le fait que $X_1, ..., X_n$ soient des variables aléatoires réelles mutuellement indépendantes équivaut au fait suivant : pour tout choix de fonctions boréliennes bornées $f_1, ..., f_n$ sur $\mathbb{R}$,

$$E[f_1(X_1) \cdots f_n(X_n)] = \prod_{j=1}^{n} E[f_j(X_j)]\,.$$

2.9 Les théorèmes limite de la théorie des probabilités

2.9.1 La notion de convergence en probabilité et la loi faible des grands nombres

Une suite $(X_n)_{n\geq 0}$ de variables aléatoires réelles toutes définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$ est dite <u>converger en probabilité</u> (on dit ausi *converge* <u>*stochastiquement*</u> vers une variable aléatoire X si et seulement si

$$\lim_{n\to+\infty} P(|X_n - X| > \epsilon) = 0 \qquad \forall \epsilon > 0\,.$$

On a déjà rencontré un premier théorème limite du calcul des probabilités, la *loi faible des grands nombres*, s'énonçant ainsi :

Théorème 2.1 (loi faible des grands nombres) *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires réelles ou complexes, toutes définies sur le même espace probabilisé.*

On suppose ces variables deux à deux indépendantes, ayant toutes même variance (finie) et même espérance $E[X]$ (c'est le cas si les variables ont même loi et ont une variance) ; alors, la suite de variables aléatoires

$$\frac{X_1 + \cdots + X_n}{n}$$

converge en probabilité vers la variable constante $E[X]$.

L'épithète "faible" tient au fait que l'on introduira plus loin une notion de convergence plus forte que la convergence stochastique (à savoir la convergence presque sûre) et que, sous certaines hypothèses, la suite de variables aléatoires

$$\frac{X_1 + \cdots + X_n}{n}$$

convergera toujours vers $E[X]$ au sens de cette nouvelle convergnce (on parlera alors de *loi forte* des grands nombres).

Remarque. Si $(X_n)_n$ est une suite de variables aléatoires ayant toutes une variance telles que

$$\begin{aligned} \lim_{n\to+\infty} V(X_n) &= 0 \\ \lim_{n\to+\infty} E[X_n] &= m \in \mathbb{C}\,, \end{aligned}$$

alors la suite $(X_n)_n$ converge stochastiquement vers la constante m. Il suffit en effet d'appliquer Bienaymé-Tchebychev :

$$\begin{aligned} P(|X_n - m| > \epsilon) &\leq P(|X_n - E[X_n]| > \epsilon/2) + P(|E[X_n] - m| > \epsilon/2) \\ &\leq \frac{V(X_n)}{\epsilon^2} + 0 \end{aligned}$$

si n est assez grand.

2.9.2 La notion de convergence en loi et le théorème de la limite centrale

Une suite $(X_n)_n$ de variables aléatoires réelles est dite <u>converger en loi</u> vers une variable aléatoire X si et seulement si

$$F_{X_n}(t) \to F_X(t)\,,$$

si F_{X_n} (resp. F_X) désigne la fonction de répartition de X_n (resp. de X), ce en tout point t où la fonction F_X est continue.

La convergence en loi est une notion plus faible que la convergence stochastique : si $(X_n)_n$ est une suite de variables aléatoires réelles convergeant en probabilité vers une variable X (*i.e* $\lim_{n\to\infty} P(|X_n - X| > \lambda) = 0$ pour tout $\lambda > 0$), alors la suite $(X_n)_n$ converge en loi vers X. On admettra ce résultat.

Une condition suffisante pour qu'une suite $(X_n)_n$ de variables aléatoires réelles converge en loi est qu'il existe une fonction Φ continue en $\omega = 0$, telle que, pour tout nombre réel ω,

$$\lim_{n\to\infty} E[e^{i\omega X_n}] = \Phi(\omega)\,;$$

la fonction Φ s'écrit alors

$$\Phi(\omega) = E[e^{i\omega X}]\,,$$

où X est une variable aléatoire réelle X ; cette variable X est alors la limite en loi de la suite $(X_n)_n$. Ce résultat important est le théorème de Paul Lévy (Paul Lévy fut l'un des initiateurs de la théorie des probabilités dans la première moitié du XX-ème siècle). La notion sous-jacente de *fonction caractéristique* (on appelle *fonction caractéristique* d'une variable aléatoire réelle X la fonction $\omega \in \mathbb{R} \to E[e^{i\omega X}]$) fait entrer la transformtion de Fourier (ou, pour les physiciens, la *prise de spectre*) au service de la théorie des probabilités. On notera aussi, ce qui est commode, que la fonction caractéristiques d'une somme de variables aléatoires indépendantes est le produit des fonctions caractéristiques ; ceci est un avatar de l'intérêt calculatoire de la transformation de Fourier qui transforme l'opération de convolution (passage à travers un filtre linéaire dont les paramètres restent immuables dans le temps, par exemple une cellule mécanique ou une cellule électrique) en l'opération autrement plus maniable de multiplication.

Voici une illustration de la convergence en loi, avec le *théorème de la limite centrale* qui précise la convergence des lois binomiales vers la loi de Gauss et éclaire l'expérience du triangle de Galton évoquée plus haut dans ce cours. Le résultat s'énonce ainsi :

Théorème 2.2 (théorème de la limite centrale) *Soit $(X_n)_n$ une suite de variables aléatoires réelles, que l'on suppose mutuellement indépendantes, toutes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$; on suppose que ces variables ont toutes même loi, avec pour espérance m et pour variance σ^2 ; Il existe alors une variable aléatoire X, elle aussi définie sur $(\Omega, \mathcal{T}, P)$, suivant une loi de Gauss de paramètres (0,1), telle que la suite de variables*

$$Y_n := \frac{\sum_{k=1}^{n} X_k - nm}{\sqrt{n}\,\sigma}$$

converge en loi vers X.

C'est ce théorème qui nous permet, quand $n\tau(1-\tau) > 10$, d'approcher raisonnablement bien la loi binomiale de paramètres (n, τ) par une loi gaussienne de paramètres $\mu = n\tau$, $\sigma = \sqrt{n}\tau(1-\tau)$, ce qui est une approximation très couramment utilisée.

Voici (proposée ici à titre d'exercice) une preuve rapide du théorème limite centrale : on commence par remarquer que si Y est une variable aléatoire centrée et de variance 1, on a, pour t fixé, en utilisant par exemple la formule de Taylor avec reste intégral à l'ordre 2 pour développer l'exponentielle,

$$E[\exp(itY/\sqrt{n})] = 1 - \frac{t^2}{2n} + o_t(1/n)\,;$$

en utilisant l'indépendance mutuelle des variables X_j, on a, pour tout n, pour tout t réel,

$$E\Big[\exp\Big(it\frac{\sum_{k=1}^{n} X_k - nm}{\sqrt{n}\,\sigma}\Big)\Big] = \Big(E\Big[\exp\Big(it\frac{X_1 - m}{\sqrt{n}\,\sigma}\Big)\Big]\Big)^n\,;$$

on a donc, pour tout t réel,

$$E[\exp(itY_n)] = \Big(1 - \frac{t^2}{2n} + o_t(1/n)\Big)^n \longrightarrow \exp(-t^2/2)\,, \qquad n \longrightarrow \infty\,;$$

comme la fonction $t \mapsto \exp(-t^2/2)$ est la fonction caractéristique d'une variable aléatoire suivant une loi gaussienne réduite centrée, la conclusion du théorème limite centrale résulte du théorème de Paul Lévy.

Dans le cadre des variables aléatoires discrètes, le concept de *fonction génératrice* remplace souvent avantageusement celui de fonction caractéristique.

Soit Y une variable aléatoire à valeurs dans $\mathbb{N}$ définie sur un espace probabilisé $(\Omega, \mathcal{T}, P)$; la fonction

$$z \in \overline{U} = \{z \in \mathbb{C};\ |\ z\ | \leq 1\} \mapsto G_Y(z) = E[z^Y] = \sum_{n=0}^{\infty} P(\{Y = n\}) z^n$$

est une fonction continue sur $\overline{U}$, de classe $\mathcal{C}^k$ sur $\overline{U}$ dès que $E[|\ Y\ |^k] < \infty$, que l'on appelle *fonction génératrice* de Y ; de plus, si $E[|\ Y\ |^k] < \infty$, pour tout $1 \leq l \leq k$,

$$(\frac{d}{dz})^l [G(z)] = \sum_{k=0}^{\infty} (k+1) \cdots (k+l-1)\, P(X = k)\,; \qquad (\dagger)$$

Pour des variables discrètes, on a la version suivante du théorème de Paul Lévy : une suite de variables aléatoires $(X_n)_n$ à valeurs dans $\mathbb{N}$, toutes définies sur $(\Omega, \mathcal{T}, P)$, converge en loi vers une variable aléatoire discrète X (aussi à valeurs dans $\mathbb{N}$) si et seulement si il existe une fonction G, définie sur $\overline{U}$ et continue en 1, telle que

$$\lim_{n \to \infty} G_{X_n}(z) = \Phi(z), \quad z \in \overline{U}\,.$$

Le concept de fonction génératrice s'avère un outil commode pour calculer espérance, variance, *etc.*, d'une variable aléatoire à valeurs dans $\mathbb{N}$: par exemple, en évaluant le second membre de $(\dagger)$ en $z = 1$, on trouve espérance, variance, *etc.* On notera aussi que (tout au moins formellement), comme c'est le cas pour les fonctions caractéristiques, la fonction génératrice d'une somme de variables aléatoires à valeurs dans $\mathbb{N}$ indépendantes est le produit des fonctions génératrices.

2.9.3 La convergence presque sûre et la loi forte des grands nombres

Une suite de variables aléatoires réelles $(X_n)_n$, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$ converge P-presque sûrement vers une variable aléatoire réelle X si l'ensemble A des $\omega \in \Omega$ où $X_n(\omega)$ ne converge pas vers $X(\omega)$ est tel que $P(A) = 0$.

La convergence presque sûre d'une suite de variables aléatoires $(X_n)_n$ vers une variable aléatoire X implique la convergence en probabilité de la suite $(X_n)_n$ vers cette même variable, donc *a fortiori* la convergence en loi de $(X_n)_n$ vers X. La convergence presque sûre est donc la plus forte des trois notions de convergence que nous avons introduit dans ce cours, la notion de convergence en loi étant, elle, la plus faible.

Avant de donner un exemple de convergence presque sûre (la loi forte des grands nombres), nous avons l'amélioration suivante de l'inégalité de Bienaymé-Tchebychev, due à Kolmogorov, dite *inégalité de Kolmogorov* : soient $X_1, \ldots, X_n$ n variables mutuellement indépendantes, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{T}, P)$, admettant chacune une variance et toutes de moyenne nulle ; alors, pour tout $\epsilon > 0$,

$$P\left[\left\{ \max_{1 \leq j \leq n} \left| \sum_{l=1}^{j} X_l \right| \geq \epsilon \right\}\right] \leq \frac{\sum_{k=1}^{n} V(X_k)}{\epsilon^2}\,.$$

La preuve de l'inégalité de Kolmogorov peut être considérée comme un intéressant exercice sur le conditionnement ; nous la mentionnons brièvement ici. On introduit les variables

$$Y_j := \sum_{l=1}^{j} X_l, \quad j = 1, \ldots, n\,;$$

on introduit aussi le système complet d'événements suivants $(A_q(\epsilon)), q = 0, \ldots, n$, où

$$A_0(\epsilon) = A_0 = \{\mid Y_1 \mid < \epsilon, \ldots, \mid Y_n \mid < \epsilon\}$$

et, pour q entre 1 et n,

$$A_q(\epsilon) = A_q = \{\mid Y_l \mid < \epsilon,\ l < q;\ \mid Y_q \mid \geq \epsilon\}\,.$$

Utilisant ce système complet et l'indépendance des X_j, on a

$$E[Y_n^2] = \sum_{k=1}^{n} V(X_k) = \sum_{q=0}^{n} E[Y_n^2 \chi_{A_q}] \geq \sum_{q=1}^{n} E[Y_n^2 \chi_{A_q}]\,. \qquad (*)$$

Soit q entre 1 et n ; si nous écrivons Y_n sous la forme

$$Y_n = Y_q + X_{q+1} + \cdots + X_n\,,$$

nous voyons en utilisant l'indépendance mutuelle des X_j (l'indépendance deux à deux ne suffit pas ici) que

$$E[Y_n^2 \chi_{A_q}] = E[Y_q^2\ \chi_{A_q}] + \sum_{l=q+1}^{n} E[X_l^2\ \chi_{A_q}] \qquad (**)$$

où χ_{A_q} désigne la fonction indicatrice de l'événement A_q, valant 1 si A_q est réalisé, 0 sinon. Mais, comme $E[Y_q^2 \chi_{A_q}] \geq \epsilon^2 P(A_q)$ (voir la définition de $A_q(\epsilon)$), on déduit de (**) que

$$E[Y_n^2 \chi_{A_q}] \geq \epsilon^2 P(A_q)\ ;$$

en ajoutant les inégalités ainsi obtenues, on déduit de $(*)$

$$\sum_{q=1}^{n} P(A_q(\epsilon)) \leq \frac{\sum_{k=1}^{n} V(X_k)}{\epsilon^2}$$

ce qui est la conclusion voulue si l'on se reporte à la définition des $A_q(\epsilon)$.

La *loi forte des grands nombres* est une conséquence de l'inégalité de Kolmogorov (comme la loi faible résultait de l'inégalité de Bienaymé-Tchebychev). La loi forte des grands nombres réalise l'articulation fondamentale entre la théorie des probabilités et le raisonnement fondant les statistiques. Voici, pour conclure cette brève initiation aux probabilités, l'énoncé de cette loi capitale :

Théorème 2.3 (loi forte des grands nombres) *Soit $(X_n)_n$ une suite de variables aléatoires réelles, mutuellement indépendantes, toutes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$ (conditionnant l'épreuve). On suppose que ces variables ont toutes même loi, avec pour espérance m et pour variance σ^2 (il faut penser X_k comme le résultat d'une épreuve, épreuve que l'on répète indéfiniment de manière indépendante). Alors la suite de variables*

$$Z_n := \frac{\sum_{k=1}^{n} X_k}{n}$$

converge P-presque sûrement vers la variable P-presque partout égale à m, qui est la valeur moyenne du résultat de l'épreuve.

Ainsi, si l'on jette un dé indéfiniment et si l'on divise le nombre de points obtenus par le nombre de coups, on approche presque sûrement lorsque le nombre de coups tend vers l'infini, la constante

$$\frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2} = 3.5$$

(qui correspond à la valeur moyenne de l'épreuve).

Voici, à titre d'exercice, le chemin logique menant de l'inégalité de Kolmogorov à la formulation de la loi forte des grands nombres. Il n'y a pas de restriction à supposer les variables X_j toutes centrées (sinon, on leur retranche leur espérance). On reprend les variables Y_l introduites lors de la preuve de l'inégalité de Kolmogorov. Posons, pour tout $\epsilon > 0$, pour tout $n \in \mathbb{N}$,

$$B_n(\epsilon) = \Big\{ \sup_{k \geq n} \Big| \frac{Y_k}{k} \Big| \geq \epsilon \Big\} .$$

On a

$$\begin{aligned} P(B_{2^j}(\epsilon)) &\leq \sum_{l=j}^{\infty} P\Big[\sup_{2^l \leq k < 2^{l+1}} \Big(\Big| \frac{Y_k}{k} \Big| > \epsilon \Big) \Big] \\ &\leq \sum_{l=j}^{\infty} P\Big[\sup_{1 \leq k < 2^{l+1}} \Big(\mid Y_k \mid \Big) \geq \epsilon 2^l \Big] \\ &\leq \sum_{l=j}^{\infty} \Big[\frac{1}{\epsilon^2 2^{2l}} (2^{l+1}\sigma^2) \Big] . \end{aligned}$$

On utilise ici (pour la dernière inégalité) l'inégalité de Kolmogorov. On voit ainsi que pour tout $\epsilon > 0$, on a

$$\lim_{j \to +\infty} P(B_{2^j}(\epsilon)) = 0 .$$

Ceci étant acquis pour tout $\epsilon > 0$, on en déduit que la suite Z_n converge vers 0 P-presque partout, ce qui achève la preuve de la loi forte des grands nombres.

2.9.4 La convergence en moyenne

Une autre notion "forte" de convergence est celle de convergence en moyenne.

On dit qu'une suite $(X_n)_{n \geq 0}$ de variables aléatoires réelles ou complexes (toutes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$ converge en moyenne (ou encore en norme L^1) vers une variable aléatoire X si toutes les variables X_n ainsi que X ont une espérance et si

$$\lim_{n \to +\infty} E[|X_n - X|] = 0 .$$

À cause de l'inégalité de Markov, la convergence en moyenne implique la convergence stochastique, donc *a fortiori* la convergence en loi. Par contre, on ne peut en général situer ce type fort de convergence par rapport au type (tout aussi fort) qu'est la convergence presque sûre.

Autre type de convergence, plus fort encore que la convergence en moyenne : on dit qu'une suite $(X_n)_{n \geq 0}$ de variables aléatoires réelles ou complexes (toutes définies sur un même espace probabilisé $(\Omega, \mathcal{T}, P)$ converge en moyenne quadratique (ou encore en norme L^2) vers une variable aléatoire X si toutes les variables X_n ainsi que X ont une variance et si

$$\lim_{n \to +\infty} E[|X_n - X|^2] = 0 .$$

Comme l'énergie est une quantité de nature quadratique en physique, cette dernière notion de convergence s'avère aussi importante. Le fait que la convergence en moyenne quadratique implique la convergence en moyenne est une conséquence de l'inégalité de convexité de Hölder suivant laquelle :

$$E[|X_n - X|] \leq E[|X_n - X|^p]^{1/p} E[1^q]^{1/q} = E[|X_n - X|^p]^{1/p}$$

pour $p \geq 1$ et $q \geq 1$ tels que $1/p + 1/q = 1$ (prendre ici $p = 2$).

2.10 Le raisonnement statistique

2.10.1 La notion d'estimateur

Soit X une variable aléatoire dont la loi (inconnue, mais de type connu) dépend d'un paramètre θ ; par exemple (comme c'est le cas dans les populations sondées à la veille d'un scrutin à deux choix), X peut être une variable de Bernouilli de paramètre θ ; ce peut aussi (analyse du rayonnement d'un tissu organique) être une loi de Poisson de paramètre (la densité du tissu) λ.

Un *échantillon* de taille n de X est par définition la donnée de n variables aléatoires $X_1, ..., X_n$, mutuellement indépendantes et de même loi, à savoir la loi de la variable inconnue X.

Un *estimateur* de θ est par définition la donnée, pour chaque n, d'une variable aléatoire réelle T_n s'écrivant comme une fonction (déterministe) des variables $X_1, ..., X_n$ ($T_n = f_n(X_1, ..., X_n)$). Une réalisation $(x_1, ..., x_n)$ de $(X_1, ..., X_n)$ (dite aussi *échantillon expérimental* ou *échantillon observé*) fournit, avec $t_n := f_n(x_1, ..., x_n)$, une *estimation* à l'ordre n de θ.

L'estimateur est dit converger vers θ si la la suite $(T_n)_{n\geq 1}$ converge en probabilité vers la constante θ lorsque n tend vers $+\infty$.

Exemple 2.16. Si

$$T_n = \frac{X_1 + \cdots + X_n}{n}$$

et que X est supposée avoir une variance, la suite $(T_n)_n$ définit un estimateur convergent vers θ (c'est la loi faible des grands nombres). Si X a une variance, il suffit, pour que $(T_n)_n$ définisse un estimateur convergent vers $E[X]$, que les T_n aient toutes une variance et que

$$\begin{aligned} \lim_{n\to+\infty} E[T_n] &= E[X] \\ \lim_{n\to+\infty} V[T_n] &= 0 \end{aligned}$$

(il suffit d'appliquer l'inégalité de Bienaymé-Tchebychev, voir la remarque à la fin de la section 2.9.1).

L'estimateur $(T_n)_n$ de θ est dit *sans biais* si $E[T_n] = \theta$ pour tout $n \geq 1$.

On dit qu'un estimateur $(T_n)_n$ de θ est *absolument correct* s'il est sans biais et converge vers θ.

Un estimateur $(T_n)_n$ de θ est dit *efficace* s'il est absolument correct et de plus tel que la quantité

$$E(|T_n - \theta|^2)$$

(dite aussi *risque quadratique*) soit minimale parmi toutes les quantités

$$E(|S_n - \theta|^2)\,,$$

où $(S_n)_n$ est un estimateur quelconque absolument correct de θ.

Parmi les estimateurs absolument corrects du paramètre θ, les estimateurs efficaces sont donc ceux qui minimisent le risque quadratique.

2.10.2 Exemples classiques d'estimateurs

Estimateurs de l'espérance

Supposons que le paramètre à estimer de la variable aléatoire inconnue X soit son espérance $E[X]$ (X étant supposée avoir une variance). Dans ce cas, on dispose d'un estimateur absolument correct pour $E[X]$, à savoir l'estimateur

$$T_n = \overline{X_n} := \frac{X_1 + \cdots + X_n}{n}\,.$$

Le fait que cet estimateur converge vers $E[X]$ résulte, on l'a vu, de la loi faible des grands nombres.

Estimateurs de la variance

a) *Lorsque $E[X] = m$ est connue*

Dans ce cas, on sait que la variance de X est l'espérance de $(X-m)^2$; comme les variables $(X_k - m)^2$, $k = 1, 2, ...$, sont mutuellement indépendantes lorsque les X_k le sont, la loi faible des grands nombres nous assure (pourvu que $(X-m)^2$ ait une variance, c'est-à-dire que $E[|X|^4] < +\infty$) que

$$S^2_{X,n} := \frac{1}{n}\sum_{k=1}^{n}(X_k - m)^2\,,\ n = 1, 2, ...$$

est un estimateur absolument correct de $V(X)$ (on applique le résultat précédent concernant l'estimateur $\overline{X_n}$ de l'espérance).

b) *Lorsque $E[X]$ est inconnue*

Si l'on pose

$$\overline{X_n} := \frac{1}{n}(X_1 + X_2 + \cdots + X_n)\,,$$

les variables

$$X_k - \overline{X_n}\,,\ k = 1, ..., n\,,$$

ne sont plus cette fois mutuellement indépendantes ; cependant, on peut extraire de ces n variables $n-1$ variables mutuellement indépendantes et

$$\widetilde{S}^2_{X,n} := \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X_n})^2$$

définit un estimateur convergent pour $V(X)$.

Estimateur du coefficient de corrélation entre deux variables X et Y

Si X et Y sont deux variables aléatoires réelles définies sur le même espace probabilisé et ayant toutes les deux une variance, le coefficient de régression (on dit aussi *de corrélation*) de X et de Y est défini par

$$\rho(X,Y) = \mathrm{corr}\,(X,Y) := \frac{E[(X - E[X])(Y - E[Y])]}{\sigma(X)\sigma(Y)}\,.$$

Un estimateur de ce coefficient de corrélation est donné par

$$C_{X,Y,n} := \frac{1}{n-1}\frac{\sum_{k=1}^{n}(X_k - \overline{X_n})\,(Y_k - \overline{Y_n})}{\widetilde{S}_{X,n}\,\widetilde{S}_{Y,n}}\,,$$

où

$$\begin{aligned}
\overline{X_n} &:= \frac{1}{n}\sum_{k=1}^{n} X_k \\
\overline{Y_n} &:= \frac{1}{n}\sum_{k=1}^{n} Y_k \\
\widetilde{S_{X,n}} &:= \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X_n})^2} \\
\widetilde{S_{Y,n}} &:= \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X_n})^2}\,.
\end{aligned}$$

2.10.3 Estimation par intervalle ; l'exemple des gaussiennes et le test de Student

Intervalle d'acceptation au risque α ; intervalle de confiance

Supposons que X voit une variable aléatoire réelle dépendant d'un paramètre θ et que T_n définisse un estimateur convergent vers θ.

La loi de T_n dépend bien sûr de θ (que l'on ne connaît pas) ; on introduit malgré tout la notion d'*intervalle d'acceptation de T_n au risque α* (α étant un seuil entre 0 et 1) ; un tel intervalle $I_{n,\alpha}$ est défini par

$$P(\{T_n \in I_{n,\alpha}\}) = 1 - \alpha\,.$$

Un tel intervalle d'acceptation de T_n au risque α n'est pas unique et de plus dépend de θ. Par contre, si $\beta \in [0,1]$, l'intervalle $[t_{1,\alpha,\beta}(\theta), t_{2,\alpha,\beta}(\theta)]$ défini par les deux conditions

$$\begin{aligned}
P(\{T_n < t_{1,\alpha,\beta}\}) &= \alpha\beta \\
P(\{T_n > t_{2,\alpha,\beta}\}) &= (1-\alpha)\beta
\end{aligned}$$

est un intervalle d'acceptation de T_n au risque α parfaitement déterminé (mais dépendant toujours bien sûr de n et de θ). On se limitera souvent au cas $\beta = 1/2$ en disant que ceci revient à choisir l'intervalle d'acceptation au risque α dans une configuration où les risques sont également partagés.

Définition 2.2 *Soit X une variable aléatoire dépendant d'un paramètre θ, $(T_n)_n$ un estimateur convergent de θ ; on appelle intervalle de confiance de θ au risque α (et à l'ordre n), les risques étant partagés, l'intervalle constitué des valeurs estimées $t_n = f_n(x_1, ..., x_n)$ de θ (à l'ordre n) telles que*

$$t_n \in [t_{\alpha,1}(\theta), t_{\alpha,2}(\theta)]\,,$$

où $[t_{\alpha,1}(\theta), t_{\alpha,2}(\theta)]$ est l'intervalle d'acceptation de θ au risque α (les risques étant partagés).

Remarque. Un intervalle de confiance au risque α (risques partagés) est aussi appelé *fourchette de vraisemblance*. On dit souvent (mais ce n'est qu'une formulation heuristique) pour définir une telle fourchette de vraisemblance : "*le paramètre estimé θ appartient à la fourchette avec une probabilité (équilibrée) $1-\alpha$*".

Exemple 2.17 (intervalle de confiance pour la moyenne d'une variable gaussienne d'écart type connu)

Supposons que X suive une loi normale de moyenne θ et d'écart type σ, ce qui signifie que X est une variable à densité, de densité

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}}\,.$$

Considérons l'estimateur de θ à l'ordre n

$$\overline{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k\,.$$

La variable $\overline{X}_n$ est une somme de n variables gaussiennes indépendantes de moyenne θ et de variance σ^2/n. La somme de deux variables gaussiennes indépendantes de moyennes respectives m_1 et m_2 et d'écarts type respectifs σ_1 et σ_2 est une gaussienne de moyenne m_1+m_2 et d'écart type $\sqrt{\sigma_1^2+\sigma_2^2}$: en effet, pour tout $t \in \mathbb{R}$,

$$\int_{\mathbb{R}} \Big(\frac{1}{\sqrt{2\pi}\,\sigma_1} e^{-(u-m_1)^2/(2\sigma_1^2)}\Big)\Big(\frac{1}{\sqrt{2\pi}\,\sigma_2}\, e^{-(t-u-m_2)^2/(2\sigma_2^2)}\Big)\,du = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2+\sigma_2^2}} e^{\frac{(t-m_1-m2)^2}{2(\sigma_1^2+\sigma_2^2)}}$$

(on peut le voir en prenant par exemple les transformées de Fourier de ces deux fonctions de t et en utilisant le fait que la transformation de Fourier, c'est-à-dire la prise de spectre, échange les opérations de convolution et de multiplication). La variable $\overline{X}_n$, qui est une somme de n gaussiennes indépendantes de moyenne θ/n et d'écart type σ/n, est donc une variable gaussienne de moyenne $n \times \theta/n = \theta$ et d'écart type $\sigma/\sqrt{n}$. Si l'on choisit $\alpha = 5\%$, l'intervalle d'accepation de T_n au risque α (les risques étant partagés) est l'intervalle $[t_1(\theta), t_2(\theta)]$ est donné par les conditions :

$$\begin{aligned} P\Big(N < \frac{t_1(\theta)-\theta}{\sigma/\sqrt{n}}\Big) &= 0.025 = 2.5/100 \\ P\Big(N > \frac{t_2(\theta)-\theta}{\sigma/\sqrt{n}}\Big) &= 0.025 = 2.5/100\,, \end{aligned}$$

où N désigne une variable aléatoire suivant une loi de Gauss normale (de moyenne nulle et d'écart type 1). En se reportant à une table de la loi de Gauss, on voit que

$$\begin{aligned} t_1(\theta) &= \theta - 1.96\,\frac{\sigma}{\sqrt{n}} \\ t_2(\theta) &= \theta + 1.96\,\frac{\sigma}{\sqrt{n}}\,. \end{aligned}$$

L'intervalle de confiance est

$$\Big[\frac{x_1+\cdots+x_n}{n} - 1.96\frac{\sigma}{\sqrt{n}}, \frac{x_1+\cdots+x_n}{n} + 1.96\frac{\sigma}{\sqrt{n}}\Big]\,.$$

Si n tend vers $+\infty$, $\sigma/\sqrt{n}$ tend vers 0 et le diamètre de l'intervalle de confiance tend vers 0, ce qui montre que l'estimation est d'autant plus précise que n est grand.

Les lois du χ^2 et de Student

Une variable aléatoire réelle suit une *loi du χ^2* (*chi-deux*) à p degrés de liberté si elle se réalise comme une somme

$$\chi^2_p = \sum_{k=1}^{p} Z_k^2 \,,$$

où $Z_1, ..., Z_p$ sont p variables gaussiennes réduites centrées et mutuellement indépendantes. Une telle variable est une variable à densité, de densité

$$f_p(t) = \frac{1}{2^{p/2}\Gamma(p/2)} \exp(-t/2)\, t^{p/2-1} \,, \qquad p \in \mathbb{N} \,.$$

L'espérance d'une telle variable aléatoire est p, sa variance vaut $2p$.

Une variable aléatoire réelle suit une *loi de Student* à p degrés de liberté si elle se réalise sous la forme

$$\mathcal{S}_p = \frac{Z}{\sqrt{\frac{\chi^2_p}{p}}} \,,$$

où Z suit une loi normale (gaussienne réduite centrée) et χ^2_p une loi du chi-deux à p degrés de liberté, les variables Z et χ^2_p étant supposées indépendantes.

Exemple 2.18. Si X et une variable gaussienne d'espérance θ et d'écart-type σ, $X_1, ..., X_n$ n variables mutuellement indépendantes de loi la loi de X, alors l'estimateur canonique de σ (l'espérance θ n'étant pas supposée connue) est

$$S_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X_n})^2} \,,$$

où

$$\overline{X_n} := \frac{1}{n} \sum_{k=1}^{n} X_k$$

est l'estimateur de l'espérance. La variable

$$\frac{\overline{X}_n - \theta}{\frac{S_n}{\sqrt{n}}}$$

suit alors une loi de Student à $n-1$ degrés de liberté.

Le test de Student (intervalle de confiance pour la moyenne d'une variable gaussienne de moyenne et écart type inconnus)

Supposons que X suive une loi gaussienne de moyenne θ (inconnue) et d'écart type σ (inconnu) ; la variable

$$\mathcal{S}_{n-1} = \frac{\overline{X_n} - \theta}{\frac{S_n}{\sqrt{n}}}$$

suit (voir l'exemple 2.18) une loi de Student à $n-1$ degrés de liberté (ce quelque soit la valeur de σ). On exploite alors ce fait crucial pour déterminer un intervalle de confiance pour le calcul de moyenne.

Supposons que α soit un seuil de risque et que τ_α soit tel que

$$P(|\mathcal{S}_{n-1}| > \tau_\alpha) = \alpha$$

(on se reporte à une table de loi de Student à $n-1$ paramètres pour la détermination de τ_α). Si t_n est l'estimation

$$t_n = \frac{x_1 + \cdots + x_n}{n}$$

de la moyenne de X à partir d'un échantillon observé et

$$s_n = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(x_k - t_n)^2}$$

celui de l'écart-type de X, l'intervalle de confiance au risque α (les risques étant partagés) pour la moyenne est l'intervalle constitué des nombres θ tels que

$$\left|\frac{t_n - \theta}{\frac{s_n}{\sqrt{n}}}\right| \leq \tau_\alpha \,,$$

c'est-à-dire l'intervalle

$$\left[t_n - \tau_\alpha \frac{s_n}{\sqrt{n}}, t_n + \tau_\alpha \frac{s_n}{\sqrt{n}}\right].$$

2.10.4 Intervalles de confiance et théorème de la limite centrale

Soit X une variable de Bernouilli (liée par exemple à un certain évènement A par $X(\omega) = 1$ si $\omega \in A$, $X(\omega) = 0$ sinon). La variable X dépend d'un paramètre $\theta = P(A)$ qui est aussi $E[X]$.

Un estimateur sans biais convergent vers θ est

$$\overline{X_n} = \frac{X_1 + \cdots + X_n}{n} \,.$$

D'après le théorème de la limite centrale, la suite de variables

$$\frac{\overline{X_n} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}}$$

converge en loi, lorsque n tend vers l'infini, vers une loi de Gauss centrée réduite $\mathcal{N}(0,1)$. Pour n assez grand ($n \geq 30$) et sous les conventions $n\theta \geq 5$ et $n(1-\theta) \geq 5$, on peut d'ailleurs assimiler la loi de la variable

$$\frac{\overline{X_n} - \theta}{\sqrt{\frac{\theta(1-\theta)}{n}}}$$

à la loi normale $\mathcal{N}(0,1)$.

Soit $\alpha \in [0,1]$ un seuil et τ_α défini par

$$P(|\mathcal{N}(0,1)| > \tau_\alpha) = \alpha$$

(on se reporte aux tables). Si f_n est la fréquence de réalisations de A estimée à partir de n observations, c'est-à-dire

$$f_n := \frac{x_1 + \cdots + x_n}{n}\,,$$

l'intervalle d'acceptation de

$$T_n = \overline{X_n}$$

au risque α (les risques étant comme toujours partagés) est

$$A_\alpha = \{t\,;\, \frac{|t-\theta|}{\sqrt{\frac{f_n(1-f_n)}{n}}} < \tau_\alpha\}$$

et l'intervalle de confiance pour θ est donc

$$\Big[f_n - \tau_\alpha\sqrt{\frac{f_n(1-f_n)}{n}}, f_n + \tau_\alpha\sqrt{\frac{f_n(1-f_n)}{n}}\Big]\,,$$

f_n représentant la fréquences des réalisations observées sur n épreuves (les contraintes $n \geq 30$ et $n\min(\theta, 1-\theta) \geq 5$ étant supposées *a priori* remplies).

FIN DU CHAPITRE 2 ET DU COURS