Algèbre Linéaire
(Version provisoire)

Alain-Yves LeRoux

1 Espaces vectoriels

1.1 La définition

Définition 1.1 *Un espace vectoriel V sur $\mathbb{R}$ ou $\mathbb{C}$ est la donnée d'un ensemble V dont les éléments sont appelés* **vecteurs**, *muni des deux opérations :*

$$\textit{de } V \times V \textit{ dans } V \quad , \quad (u,v) \longrightarrow u+v \qquad (l'addition) \quad ,$$

$$\textit{de } \mathbb{R} \times V \textit{ dans } V \ , \quad (\lambda, u) \longrightarrow \lambda u \qquad (\textit{la multiplication par un scalaire})$$

telles que
- l'addition est associative, commutative, possède un élément neutre noté 0 et tout élément u de V a un symétrique noté $-u$,
- pour tout $\lambda, \mu \in \mathbb{R}$ et pour tout $u \in V$, on a

$$\lambda(\mu u) = \lambda\mu u \quad , \quad (\lambda+\mu)u = \lambda u + \mu u \quad , \quad \lambda(u+v) = \lambda u + \lambda v \quad , \quad 1u = u \ .$$

Les éléments de $\mathbb{R}$ sont appelés des **scalaires**.

Exemples
1 ° - $V = \mathbb{R}^2$, les éléments sont notés $x = (x_1, x_2)$ ou $y = (y_1, y_2)$. L'élément neutre est $0 = (0,0)$, le symétrique de x est $-x = (-x_1, -x_2)$, l'addition est définie par $x + y = (x_1 + y_1, x_2 + y_2)$ et la multiplication par un scalaire par $\lambda x = (\lambda x_1, \lambda y_1)$. On définit de la même façon $\mathbb{R}^3$ et $\mathbb{R}^n$ pour tout entier $n \geq 1$.
2 ° - L'espace des fonctions continues sur un intervalle $[0, L]$ avec $L > 0$. Il est noté $C^0([0, L])$, avec les opérations classiques d'addition et de multiplication par un scalaire.
3 ° - L'espace des fonctions continues sur $[0, L]$, nulles en $x = 0$ et en $x = L$. On le note

$$V_0 \ = \ \{ \ v \in C^0([0,L]) \quad | \quad v(0) = 0 \quad , \quad v(L) = 0 \ \ \} \ .$$

1.2 Les sous-espaces vectoriels

Soit W un sous ensemble de V un espace vectoriel donné.

Définition 1.2 *W est un sous espace vectoriel de V lorsque $W \neq \emptyset$ et si l'addition de vecteurs et la multiplication par un scalaire sont des lois internes pour W, c'est à dire*

$$\forall u, v \in W \ , \ \forall \lambda \in \mathbb{R}, \quad u + v \in W \quad , \quad \lambda u \in W \ .$$

Exemples 1
1 °- Toute droite passant par l'origine est un sous espace vectoriel de $\mathbb{R}^2$. C'est encore vrai pour $\mathbb{R}^n$.
2 °- L'espace V_0 défini ci-dessus est un sous espace de $C^0([0, L])$.

Remarque 1.3 *: Tout sous espace vectoriel est un espace vectoriel.*

Proposition 1.4 *Un ensemble $W \neq \emptyset$, inclus dans un espace vectoriel V est un sous espace vectoriel de V si et seulement si*

$$\forall \lambda, \mu \in \mathbb{R}, \quad \forall u, v \in W \quad \lambda u + \mu v \in W \ .$$

1.3 Les bases

Définition 1.5 *Soit V un espace vectoriel, $u_1, u_2, .., u_m$ des vecteurs de V et $\lambda_1, \lambda_2, ...\lambda_m$ des réels. Le vecteur $v = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_m u_m$ est une* **combinaison linéaire** *des vecteurs $u_1, u_2, ..., u_n$ de V.*

Proposition 1.6 *Soit $u_1, u_2, ..., u_m$ une famille de vecteurs de V ; on pose*

$$W \ = \ \{ \ u \in V \ \mid \ u = \sum_{j=1}^{m} \lambda_j u_j \ \ , \ \lambda_j \in \mathbb{R} \ \} \ .$$

Alors W est un sous espace vectoriel de V.

Remarque 1.7 *Il s'agit du sous espace engendré par le* **système générateur** $u_1, u_2, ..., u_m$.

Définition 1.8 *Soit $u_1, u_2, ..., u_m$ une famille de vecteurs de V. On dit que les vecteurs $u_1, u_2, ..., u_m$ sont* **indépendants** *lorsque*

$$\lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_m u_m = 0 \quad \Longrightarrow \quad \lambda_1 = \lambda_2 = ... = \lambda_m = 0 \ .$$

La famille $u_1, u_2, ..., u_m$ est appelée une **famille libre** *de vecteurs. Dans le cas contraire on parle de* **vecteurs liés**.

Définition 1.9 *Une famille $u_1, u_2, ..., u_n$ constitue une* **base** *de V si les vecteurs $u_1, u_2, ..., u_m$ sont indépendants et constituent un système générateur de V.*

Exemple 1.10
1 ° - Pour $V = \mathbb{R}^2$, *les vecteurs* $u_1 = (1, 0)$ *et* $u_2 = (0, 1)$ *constituent une base, dite* **base canonique**, *de* $\mathbb{R}^2$.
2 ° - Les fonctions

$$u_j(x) = sin\left(j\frac{\pi x}{L}\right) \quad , \quad 1 \leq j \leq m$$

constituent un système libre de V_0. *En effet, si pour tout*

$$x \in [0, L] \;\;,\;\; \sum_{j=1}^{m} b_j \; sin\left(j\frac{\pi x}{L}\right) \;=\; 0 \;,$$

on peut interprêter cette combinaison linéaire comme une série de Fourier (finie) dont la somme est nulle, les coefficients de Fourier sont donc nuls, et pour tout $j = 1, 2, ..., m$, *on a* $b_j = 0$.

Théorème 1.11 *Soit* $\{u_1, u_2, ..., u_n\}$ *une base de* V *un espace vectoriel. Alors*

$$\forall v \in V \quad \exists \lambda_1, \lambda_2, ..., \lambda_n \quad uniques \quad tels\ que \quad v \;=\; \sum_{j=1}^{n} \lambda_j u_j \;.$$

Démonstration *Soit* $v \in V$; *le système* $\{u_1, u_2, ..., u_n\}$ *étant générateur, il existe au moins un jeu de coefficients* $\{\lambda_1, \lambda_2, ..., \lambda_n\}$ *tels que* $v \;=\; \sum_{j=1}^{n} \lambda_j u_j$. *Il reste à montrer l'unicité. Supposons qu'il existe un autre un jeu de coefficients* $\{\mu_1, \mu_2, ..., \mu_n\}$ *tels que* $v \;=\; \sum_{j=1}^{n} \mu_j u_j$.*Alors*

$$\sum_{j=1}^{n} (\lambda_j - \mu_j)\, u_j \;=\; 0 \;,$$

et comme les u_j *sont indépendants, on obtient pour tout* j , $\lambda_j = \mu_j$, *d'où l'unicité.*

Théorème 1.12 (admis) *Soit* V *un espace vectoriel admettant une base à* n *vecteurs. Alors toute autre base de* V *aura aussi* n *vecteurs.*

Définition 1.13 *Ce nombre* n *est appelé* **dimension** *de* V, *et est noté* $dimV$.

Remarque 1.14 *La dimension peut être infinie. C'est le cas pour l'espace* V_0.

Exemple 1.15 *On considère l'espace vectoriel* $\mathbb{R}^2$ *et les trois vecteurs*

$$u_1 = (1, 1) \;\;,\;\; u_2 = (2, 1) \;\;,\;\; u_3 = (3, 1) \;\;.$$

Soit $x = (x_1, x_2)$ *un vecteur de* $\mathbb{R}^2$ *; on recherche une décomposition de* x *sur cette donnée de trois vecteurs* u_1, u_2, u_3, *de la forme* $x = \lambda_1 u_1 + \lambda_2 u_2 + \lambda_3 u_3$. *On trouve par exemple* $\lambda_1 = 2x_2 - x_1$, $\lambda_2 = x_1 - x_2$, $\lambda_3 = 0$, *mais aussi* $\lambda_1 = 0$, $\lambda_2 = 3x_2 - x_1$, $\lambda_3 = x_1 - 2x_2$. *Il n'y a pas unicité de la décomposition, et donc la donnée des trois vecteurs* u_1, u_2, u_3 *ne constitue pas une base de* $\mathbb{R}^2$. *Par contre, pris deux à deux ils constituent chaque fois une base de* $\mathbb{R}^2$ *qui est bien de dimension 2.*

Proposition 1.16 (admise) *Soit V un espace vectoriel de dimension n, et p vecteurs $u_1, u_2, ..., u_p$.*
-1- si $\{u_1, u_2, ..., u_p\}$ constitue un système générateur de V, on peut en extraire une base de V et nécessairement $p \geq n$. Si $p = n$ il s'agit d'une base de V.
-2- si les vecteurs $u_1, u_2, ..., u_p$ sont indépendants, on peut compléter cette donnée par $n - p$ vecteurs $u_{p+1}, .., u_n$ choisis judicieusement, de manière à obtenir une base de V . Nécessairement $n \geq p$, et si $n = p$, les vecteurs $u_1, u_2, ..., u_p$ constituent une base de V.

Exemple 1.17 *On considère l'espace V_0 constitué des fonctions continues sur $[0, L]$, nulles en $x = 0$ et en $x = L$. On peut envisager un sous espace vectoriel de V_0 constitué des séries de Fourier tronquées (cf.Exemple 1.10). On peut envisager d'autres types de bases. On se donne un entier N, assez grand en pratique, puis on pose*

$$h = \frac{L}{N+1}$$

puis $x_j = jh$, $j = 0, 1, 2.., N + 1$, afin de découper l'intervalle $[0, L]$ en $N + 1$ petits intervalles $[x_j, x_{j+1}]$ avec $j = 0, 1, .., N$. On introduit le sous espace V_h de V_0 défini par

$$V_h = \{ \ v \in V_0 \ \mid \ v \ affine \ sur \ chaque \ intervalle \ [x_j, x_{j+1}] \ \} \ .$$

On construit une base de V_h constituée des fonctions ϕ_k telles que

$$\phi_k(x_j) = \begin{cases} 1 & si \ j = k \\ 0 & si \ j \neq k \end{cases} \quad pour \ k = 1, .., N \ \ et \ j = 0, 1, .., N + 1 \ .$$

La partie non nulle du graphe de ϕ_k représente un triangle de hauteur 1 et de base $[x_{k-1}, x_{k+1}]$. Chaque fonction ϕ_k est continue et nulle en $x = 0$ et en $x = L$, pour $1 \leq k \leq N$, d'où $\phi_k \in V_0$. Soit $u \in V_0$; on considère la fonction $u_h \in V_h$ telle que

$$u_h(x_j) = u(x_j).$$

*On dit que u_h est l'***interpolée** *$\mathbb{P}_1$ de u aux points x_j. On démontre immédiatement*

$$u_h = \sum_{k=1}^{N} u(x_k)\phi_k$$

en testant chaque point x_k. Les vecteurs ou fonctions ϕ_k constituent une base de V_h. La dimension de V_h est donc N. L'espace vectoriel de dimension finie V_h approche l'espace vectoriel V_0 qui est de dimension infinie.

2 Les applications linéaires

2.1 Définition, propriétés élémentaires

Définition 2.1 *Soit V et W deux espaces vectoriels sur $\mathbb{R}$ et f une application de V dans W. On dit que f est* **linéaire** *lorsque*

$$\forall \lambda, \mu \in \mathbb{R}, \ \forall u, v \in V \ , \quad f(\lambda u + \mu v) = \lambda f(u) + \mu f(v) \ .$$

Remarque 2.2 *On a nécessairement* $f(0) = 0$ *, lorsque* f *est linéaire.*

Définition 2.3 *On note* $\mathcal{L}(V, W)$ *l'ensemble des applications linéaires de* V *dans* W*, puis*

$$\mathcal{I}mf \ = \ \{ \ f(u) \in W \ \mid \ u \in V \ \} \quad , \quad \mathcal{K}erf \ = \ \{ \ u \in V \ \mid \ f(u) = 0 \ \} \ .$$

Remarque 2.4 *L'ensemble* $\mathcal{I}mf$ *est un sous espace vectoriel de* W *, et* $\mathcal{K}erf$ *est un sous espace vectoriel de* V*.*

Exemple 2.5 *On prend* $V = \mathbb{R}^3$ *et* $W = \mathbb{R}^2$*, et* $f(x) = (x_1 + x_2, x_1 - x_3)$*. Alors* $\mathcal{I}mf = \mathbb{R}^2$ *et* $\mathcal{K}erf$ *est une droite de l'espace* $\mathbb{R}^3$ *définie par l'intersection des plans* $x_1 + x_2 = 0$*,* $x_1 - x_3 = 0$*.*

Soit $f, g \in \mathcal{L}(V, W)$, $\lambda \in \mathbb{R}$. On définit les applications $f + g$ et λf par

$$\forall v \in V \quad (f+g)\,(u) \ = \ f(u) + g(u) \quad , \quad (\lambda f)\,(u) \ = \ \lambda \ f(u) \ .$$

On vérifie immédiatement que ces applications appartiennent à $\mathcal{L}(V, W)$ et que, muni de ces opérations, $\mathcal{L}(V, W)$ est un espace vectoriel sur $\mathbb{R}$.

2.2 Les matrices

Soit f une application linéaire de V, un espace vectoriel de dimension n, dans W, un espace vectoriel de dimension m. On utilise une base $\{v_1, v_2, ..., v_n\}$ de V et une base $\{w_1, w_2, ..., w_m\}$ de W. Ainsi un vecteur $v \in V$ et sa transformée $f(v) \in W$ s'écrivent respectivement

$$v \ = \ \sum_{k=1}^{n} \alpha_k \ v_k \quad , \quad f(v) \ = \ \sum_{i=1}^{m} \beta_i \ w_i \quad .$$

Il existe une relation entre les β_i et les α_k. On pose

$$f(v_k) \ = \ \sum_{i=1}^{m} B_{ik} \ w_i \ .$$

On obtient

$$f(v) \ = \ f\left(\sum_{k=1}^{n} \alpha_k v_k\right) \ = \ \sum_{k=1}^{n} \alpha_k \ f(v_k) \ = \ \sum_{k=1}^{n} \alpha_k \ \sum_{i=1}^{m} B_{ik} \ w_i \ = \ \sum_{i=1}^{m} \ \left(\sum_{k=1}^{n} B_{ik} \alpha_k\right) \ w_i \ ,$$

D'où, pour tout $i = 1, 2, ..., m$,

$$\sum_{k=1}^{n} B_{ik} \alpha_k \ = \ \beta_i \quad .$$

Prenons maintenant un autre espace vectoriel U de dimension finie p, sur lequel on utilise la base $\{u_1, u_2, ..., u_p\}$, et une applications g de U dans V . On considère un vecteur $u \in U$ et sa décomposition

$$u \;=\; \sum_{j=1}^{p} \lambda_j \; u_j \;.$$

On pose comme précédemment

$$g(u_j) \;=\; \sum_{k=1}^{n} A_{kj} \; v_k \;,$$

pour obtenir

$$g(u) \;=\; g\left(\sum_{j=1}^{p} \lambda_j u_j\right) \;=\; \sum_{j=1}^{p} \lambda_j \; g(u_j) \;=\; \sum_{k=1}^{n} \left(\sum_{j=1}^{p} A_{kj} \lambda_j\right) \; v_k \;.$$

D'où, pour tout $k = 1, 2, ..., n$,

$$\sum_{j=1}^{p} A_{kj} \lambda_j \;=\; \alpha_k$$

si on prend $v = g(u)$. On peut maintenant vouloir représenter $f\,(g(u))$. On obtient

$$f\,(g(u)) \;=\; \sum_{i=1}^{m} \left(\sum_{k=1}^{n} \; B_{ik} \; \sum_{j=1}^{p} A_{kj} \lambda_j\right) \; w_i \;=\; \sum_{i=1}^{m} \sum_{j=1}^{p} \; \left(\sum_{k=1}^{n} B_{ik} A_{kj}\right) \; \lambda_j \; w_i \;.$$

On pose

$$C_{ij} \;=\; \sum_{k=1}^{n} B_{ik} A_{kj} \;.$$

On a ainsi introduit trois tableaux de coefficients : le premier contient les coefficients A_{kj} et peut se mettre sous la forme de n lignes et p colonnes ; il sera noté A, et représente l'application linéaire g. Le second contient les coefficients B_{ik} et peut se mettre sous la forme de m lignes et n colonnes ; il sera noté B et représente l'application linéaire f. Le troisième contient les coefficients C_{ij} et peut se mettre sous la forme de m lignes et p colonnes ; il sera noté C et représente l'application linéaire composée $f \circ g$.

Définition 2.6 *Une* **matrice** *est la donnée d'un tableau de coefficients rangées en p et n colonnes ; elle est dite* **carrée** *lorsque $p = n$.*

Exemple 2.7 *On considère l'espace vectoriel V constitué des polynômes de degré inférieur ou égal à deux. L'application*

$$p \;\longrightarrow f(p) \;=\; \int_{-1}^{1} (x - t)^2 \, p(t) \; dt$$

est bien une application de V dans V. On veut la représenter par une matrice. Pour cela, il nous faut d'abord utiliser une base de V; prenons les polynômes 1, x et x^2. On pose

$$p(t) \;=\; \lambda_0 + \lambda_1 t + \lambda_2 t^2.$$

Alors

$$f(p)(x) \;=\; \int_{-1}^{1} t^2 p(t) dt \;-\; \left(\int_{-1}^{1} 2t\ p(t)\ dt\right)\ x \;+\; \left(\int_{-1}^{1} p(t)\ dt\right)\ x^2\ .$$

On obtient

$$f(p)(x) \;=\; \frac{2}{3}\ \lambda_0 \;+\; \frac{2}{5}\lambda_2 \;+\; \left(\frac{4}{3}\ \lambda_1\right)\ x \;+\; \left(2\lambda_0 + \frac{2}{3}\ \lambda_2\right)\ x^2,$$

ce qui donne la matrice :

$$A \;=\; \begin{pmatrix} \frac{2}{3} & 0 & \frac{2}{5} \\ 0 & \frac{4}{3} & 0 \\ 2 & 0 & \frac{2}{3} \end{pmatrix}\ .$$

Proposition 2.8 *L'ensemble des matrices à p lignes et n colonnes est un espace vectoriel de dimension pn sur $\mathbb{R}$ si on le munit des deux opérations*

$$(A+B)_{ij} \;=\; A_{ij} \;+\; B_{ij} \quad , \quad (\lambda A)_{ij} \;=\; \lambda\ A_{ij}$$

pour toutes matices A, B à p lignes et n colonnes et $\lambda \in \mathbb{R}$.

On peut aussi effectuer des produits de matrices pour représenter la composition des applications linéaires, comme dans l'exemple ci-dessus, où une matrice C a été définie par

$$C_{ij} \;=\; \sum_{k=1}^{n} B_{ik} A_{kj}\ .$$

La matrice A avait n lignes et p colonnes, la matrice B avait m lignes et n colonnes, pour obtenir une matrice C de m lignes et p colonnes. Pour obtenir le coefficient C_{ij} on effectue les produits, composante par composante, de la ligne N$^\circ$ i de la première matrice citée (ici B) par la colonne N$^\circ$ j de la seconde matrice citée (ici A). Pour cela, il est indispensable que le nombre de colonnes de la première matrice citée (ici B) soit exactement le même que le nombre de lignes de la seconde matrice citée (ici A).

Ces distinctions entre lignes et colonnes font que, dans le cas des matrices carrées ($p = n$) le produit de matrice n'est pas commutatif en général, c'est à dire

$$BA \neq AB \quad (en\ général).$$

Le produit BA n'est pas défini lorsque $n \neq p$. Lorsque $n = p$, ce produit possède un éléményt neutre, qui est la matrice **identité** :

$$I \;=\; \begin{pmatrix} 1 & 0 & 0 & .. & 0 \\ 0 & 1 & 0 & .. & 0 \\ .. & .. & .. & .. & .. \\ 0 & 0 & 0 & .. & 1 \end{pmatrix}\ .$$

Pour toute matrice carrée A, on a $AI \;=\; IA \;=\; A$.
La **diagonale** d'une matrice carrée est la donnée des éléments A_{jj} à la ligne N ° j et à la colonne de même N ° j. Une matrice carrée est dite **matrice diagonale** lorsque tous ses éléménts hors diagonale sont nuls : $A_{ij} = 0 \;\; si \; i \neq j$. Ainsi, la matrice identité est une matrice diagonale.

Remarque 2.9 *On peut définir la puissance d'une matrice A^m. Dans ce cas, il y a commutativité :*

$$A \;(A^m) \;=\; (A^m) \;A \;=\; A^{m+1} \;.$$

On peut donc aussi définir des séries entières de matrices, et des fonctions analytiques (développables en séries entières) de matrices $f(A)$, et conserver la commutativité.

Proposition 2.10 *Soit f une application linéaire de V, muni de la base $\{v_1, v_2, .., v_n]$ dans W muni de la base $\{w_1, w_2, ..., w_m\}$. On écrit les décompositions de chaque $f(v_k)$ sur la base des w_i, c'est à dire*

$$f(v_k) \;=\; \sum_{i=1}^{m} B_{ik}\; w_i \;.$$

Alors les coefficients B_{ik} constituent la matrice de l'application linéaire f.

Remarque 2.11 *L'ordre des indices i, k doit absolument être respecté.*

Définition 2.12 *Une matrice carrée A est* **symétrique** *lorsque pour tout i, j on a $A_{ij} = A_{ji}$.*

Remarque 2.13 *Les matrices symétiques se retrouvent fréquemment dans les applications.*

2.3 Un exemple d'application

On considère l'intervalle $[O, L]$ comme représentant une étagère de longueur L. Cette étagère est posée sur appuis à ses extrémités. On pose dessus quelques bouquins, ce qui représente une charge qu'on notera f. Il s'agit d'une fonction bornée, constante par morceaux. Cette charge va provoquer un petit déplacement u de l'étagère par rapport à sa position "sans charge", qu'on représente par le déplacement nul : $u = 0$. Le déplacement u se traduit par une fonction continue (l'étagère ne se brise pas) nulle en $x = 0$ et en $x = L$, c'est à dire sur les appuis. Les déplacements possibles sont biens décrits par l'espace V_0 décrit plus haut.

Le calcul théorique du déplacement u en fonction de la charge f se fait en résolvant l'équation différentielle

$$-u'' \;+\; K\; u \;=\; f \quad , \quad u(0) = u(L) = 0 \;.$$

Le coefficient K représente la raideur du matériau. On veut approcher le déplacement u. Pour cela, on va utiliser l'espace V_h également introduit plus haut. On rencontre un premier problème : on ne peut pas dériver deux fois un élément de V_h. Par contre, on peut le dériver une fois, sauf au points $x_j,\; j = 0, 1, .., N+1$, et obtenir une fonction constante par morceaux qui est intégrable. On prend un vecteur ϕ_k de la base $\{\phi_1, \phi_2, ..., \phi_N\}$, on multiplie l'équation précédente par ϕ_k et on intègre par parties pour obtenir

$$\int_0^L \;(u'(x)\phi_k'(x) \;+\; K\; u(x)\; \phi_k(x))\; dx \;\;=\;\; \int_0^L \; f(x)\phi_k(x)\; dx \;.$$

Il reste à poser $u(x) = \sum_{j=1}^{N} \lambda_j \; \phi_j(x)$ pour obtenir une expression de la forme

$$\sum_{j=1}^{N} A_{kj} \; \lambda_j \; = \; b_k$$

avec

$$b_k = \int_0^L f(x)\phi_k(x) \; dx \quad , \quad A_{kj} \; = \; \int_0^L \; \left(\phi_j'(x)\phi_k'(x) + K\phi_j(x)\phi_k(x)\right) \; dx \; .$$

On peut noter que la matrice A est symétrique et bien entendu carrée, et que les coefficients A_{kj} se calculent facilement (au pire des polynômes de degré 2). On peut aussi constater que

$$A_{kj} \; = \; 0 \quad si \; |j-k| \geq 2 \; ,$$

c'est à dire que la matrice est tridiagonale.
La matrice A représente l'application linéaire qui fait correspondre le déplacement à la charge. En l'inversant, ce qui sera précisé plus loin, on obtient le déplacement en fonction de la charge. Ce calcul permet de prévoir si l'étagère va résister à la charge qu'on lui impose. Bien sûr, on ne se lance pas dans de tels calculs dans nos bricolages domestiques, mais ils sont indispensables pour la conception d'un pont routier ou du toît d'une grande surface et sa résistance au poids de neige.

3 L'inversion d'une matrice carrée

Dans l'exemple précédent, la relation entre la charge $b = \sum_{k=1}^{N} b_k\phi_k$ et la déformation $u = \sum_{j=1}^{N} \lambda_j \; \phi_j$ se traduisait par

$$\sum_{j=1}^{N} A_{kj} \; \lambda_j \; = \; b_k$$

pour chaque valeur de k. On note A la matrice des coefficients A_{kj} afin d'obtenir l'écriture matricielle

$$Au = b \; .$$

Ceci permet, connaissant la déformation u, de trouver la charge b, mais offre peu d'intérêt. On est en pratique plus intéressé par le calcul de la déformation u connaissant la charge b. L'équation $Au = b$ constitue alors un système de N équations linéaires à N inconnues, les coefficients λ_j. La question est de savoir si la matrice A admet une matrice **inverse** A^{-1}de coefficients B_{ik} tels que

$$\sum_{k=1}^{N} B_{ik}A_{kj} \; = \; \delta_{ij} \; = \; \begin{cases} 1 & si \; i = j, \\ 0 & si \; i \neq j. \end{cases}$$

La matrice des coefficients δ_{ij} est la matrice identité. La relation ci-dessus s'écrit aussi

$$A^{-1}A \; = \; I \; .$$

Le calcul de A^{-1} correspond à résoudre N^2 équations à N^2 inconnues, les B_{ik}. Bien entendu on ne va pas a priori calculer A^{-1} (c'est à dire résoudre N^2 équations) pour ensuite résoudre le système

$Au = b$ qui n'a que N équations. On a simplement besoin de savoir si A^{-1} existe ou non, c'est à dire si A est inversible ou non. En effet, en multipliant par A^{-1} on obtient

$$A^{-1}Au \;=\; A^{-1}b \quad , \quad d'où \quad u \;=\; A^{-1}b \;.$$

On cherche maintenant un critère qui nous indique si la matrice inverse A^{-1} existe ou non.

3.1 Les déterminants en dimensions 1 et 2

Dans le cas de l'espace vectoriel $\mathbb{R}$, de dimension 1 sur $\mathbb{R}$ bien entendu, l'équation

$$Au \;=\; b$$

est une équation scalaire et la solution évidente est

$$u \;=\; \frac{b}{A} \;,$$

à condition d'avoir $A \neq 0$. Si $A = 0$, on ne peut évidemment pas résoudre de façon unique cette équation : si $b \neq 0$ il n'y a pas de solution, et si $b = 0$ il y en a une infinité.

Dans le cas de la dimension 2, en considérant par exemple l'espace vectoriel $\mathbb{R}^2$ et sa base canonique, la matrice A est de la forme

$$A \;=\; \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

et le système $Au = b$ correspond aux deux équations

$$\alpha\lambda_1 + \beta\lambda_2 \;=\; b_1 \quad , \quad \gamma\lambda_1 + \delta\lambda_2 \;=\; b_2 \;.$$

La résolution donne

$$\lambda_1 \;=\; \frac{\delta b_1 - \beta b_2}{\alpha\delta - \beta\gamma} \quad , \quad \lambda_2 \;=\; \frac{\alpha b_2 - \gamma b_1}{\alpha\delta - \beta\gamma} \quad ,$$

à condition que les dénominateurs soient non nuls, c'est à dire

$$\alpha\delta - \beta\gamma \neq 0 \;.$$

Ce nombre est appelé le **déterminant** de la matrice A.

Définition 3.1 *Le* **déterminant** *de la matrice*

$$A \;=\; \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

est le nombre

$$det(A) \;=\; \alpha\delta - \beta\gamma \;,$$

aussi noté

$$det(A) \;=\; |A| \;=\; \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} \;.$$

Remarque 3.2 *La matrice inverse de*

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

est

$$A^{-1} = \frac{1}{det(A)} \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} .$$

Si $det(A)$ est non nul, la matrice A^{-1} existe. Si $det(A) = 0$, il n'y a pas de matrice inverse.

3.2 Les déterminants en dimension supérieure à 2

On considère en dimension 3 la matrice

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} .$$

Définition 3.3 *Le déterminant de la matrice A est le nombre réel*

$$det(A) = A_{11} \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} - A_{21} \begin{vmatrix} A_{12} & A_{13} \\ A_{32} & A_{33} \end{vmatrix} + A_{31} \begin{vmatrix} A_{12} & A_{13} \\ A_{22} & A_{23} \end{vmatrix} .$$

Ce déterminant est obtenu en développant par rapport à la première colonne. On prend chaque élément de cette colonne, affecté d'un signe, et multiplié par la matrice 2×2 obtenue en éliminant la colonne concernée et la ligne de l'élément concerné. Ainsi, un facteur A_{ij} sera multiplié par la matrice 2×2 obtenue en retirant la ligne i et la colonne j de la matrice 3×3 initiale, notée A.

On peut opérer à partir de n'importe quelle colonne ou n'importe quelle ligne. La règle des signes est celle-ci :

$$Le\ facteur\ A_{ij}\ est\ affecté\ du\ signe\ (-1)^{i+j} .$$

On peut aussi la représenter dans un tableau

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

Pour la dimension supérieure ou égale à 4, on conserve la même règle. Ainsi, pour $n = 4$, on développe par rapport à une ligne ou une colonne, en calculant une combinaison de déterminants d'ordre 3. Ensuite, pour $n = 5$, on procède de même en combinant des déterminants d'ordre 4, et pour une dimension $n+1$ on procède à une combinaison de déterminants d'ordre n, etc... La règle des signe reste la même, ainsi, pour $n=6$, par exemple, elle est représentée par le tableau

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

Théorème 3.4 (admis) *Soit A et B deux matrices carrées de même ordre. Alors*

$$det\,(AB) \;=\; det\,(A)\;\; det\,(B) \quad .$$

Proposition 3.5 *Si la matrice A est inversible, alors*

$$det\left(A^{-1}\right) \;=\; \frac{1}{det\,(A)} \;.$$

Proposition 3.6 *Le système homogène*

$$Au \;=\; 0$$

admet la solution triviale (et unique) $u=0$ si et seulement si $det\,(A) \neq 0$.

Remarque 3.7 *Si $Au=0$ admet une solution $u\neq 0$ alors nécessairement $det\,(A) \;=\; 0$.*

Proposition 3.8 *Soit $v_1, v_2, ..., v_n$ la donnée de n vecteurs d'un espace vectoriel V de dimension n , et $e_1, e_2, ..., e_n$ une base de V. On pose pour chaque valeur de $j = 1, 2, ..., n$,*

$$v_j \;=\; \sum_{i=1}^{n} A_{ij}\; e_i \;.$$

Alors les vecteurs v_j forment une base de V si et seulement si $det\,(A) \neq 0$.

Démonstration 1° - On suppose $det\,(A) \neq 0$; il nous faut montrer que $v_1, v_2, ..., v_n$ forment une base. Les vecteurs $v_1, v_2, ..., v_n$ sont indépendants : si on dispose de coefficients $\lambda_1, \lambda_2, ..., \lambda_n$ tels que

$$\sum_{j=1}^{n} \lambda_j v_j \;=\; 0 \;,$$

alors

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij} \lambda_j \right) \; e_i \;=\; 0,$$

et donc, les e_i constituant une base, le vecteur u de composantes λ_j est une solution de $Au = 0$ avec $det\,(A) \neq 0$, donc $u = 0$ et tous les λ_j sont nuls.
Le système $v_1, v_2, ..., v_n$ est générateur : Soit w un vecteur de V. Il peut s'écrire

$$w \;=\; \sum_{i=1}^{n} \mu_i e_i$$

et on cherche ensuite un vecteur u de composantes λ_j telles que pour tout $i = 1, 2, .., n$, on ait

$$\sum_{j=1}^{n} A_{ij}\lambda_j \;=\; \mu_i \;.$$

Puisque $det\,(A) \neq 0$, le vecteur u existe et est unique ; on obtient

$$w \;=\; \sum_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}\lambda_j\right) e_i \;=\; \sum_{j=1}^{n}\left(\lambda_j \sum_{i=1}^{n} A_{ij} e_i\right) \;=\; \sum_{j=1}^{n} \lambda_j v_j \;.$$

Le vecteur w a été reconstitué sur le système $v_1, v_2, ..., v_n$ qui constitue donc bien un système générateur. D'après la définition1.9, les vecteurs $v_1, v_2, ..., v_n$ constituent bien une base.
2 ° - On suppose que les vecteurs $v_1, v_2, ..., v_n$ constituent une base de V. On considère un vecteur $u \in V$. Il se décompose de façon unique sur la base des v_j et aussi sur la base des u_i, c'est à dire qu'il existe un seul jeu de composantes $\lambda_1, \lambda_2, ..., \lambda_n$ et un seul jeu de composantes $\mu_1, \mu_2, ..., \mu_n$ tels que

$$u \;=\; \sum_{j=1}^{n} \lambda_j \; v_j \quad , \quad u \;=\; \sum_{i=1}^{n} \mu_i e_i \;.$$

On en déduit que les composantes $\lambda_1, \lambda_2, ..., \lambda_n$ et $\mu_1, \mu_2, ..., \mu_n$ sont liées par les relations

$$\sum_{j=1}^{n} A_{ij}\lambda_j \;=\; \mu_i \;, \quad i = 1, 2, ..., n \;.$$

L'unicité des deux jeux de composantes implique que $det\,(A) = 0$.

Remarque 3.9 *La démonstration précédente est instructive à plus d'un titre. Dans la seconde partie, les λ_j sont les composantes de u dans la base $v_1, v_2, ..., v_n$ et les μ_i sont les composantes de u dans la base $e_1, e_2, ..., e_n$. Ils sont liés par les relations*

$$\sum_{j=1}^{n} A_{ij}\lambda_j \;=\; \mu_i \;, \quad i = 1, 2, ..., n \;.$$

La matrice A permet donc de passer des composantes λ_j dans la base $v_1, v_2, ..., v_n$ aux composantes μ_i dans la base $e_1, e_2, ..., e_n$. Il s'agit de la matrice de **changement de base.** *La matrice inverse A^{-1} permet de revenir à la base $v_1, v_2, ..., v_n$.*

Proposition 3.10 *Si deux vecteurs colonnes de A sont colinéaires (l'une est proportionnelle à l'autre), alors $det(A) = 0$.*

Démonstration : cela signifie que dans la liste des vecteurs $v_1, v_2, ..., v_n$ il y a un v_j et un v_k tels qu'il existe un scalaire γ pour lequel $v_j = \gamma v_k$. De ce fait les v_i ne constituent pas une base, et donc $det(A) = 0$.

Définition 3.11 *Soit A une matrice carrée de coefficients A_{ij}. La matrice A' de coefficients*

$$A'_{ij} \; = \; A_{ji}$$

est appelée **matrice transposée** *de A. Elle obtenue en échangeant les lignes et les colonnes.*

Théorème 3.12 *Soit A une matrice carrée. On a*

$$det(A) \; = \; det(A') \quad .$$

Ce résultat (admis) indique que tous les résultats obtenus sur les colonnes s'appliquent aussi aux lignes, et en particulier celui-ci.

Proposition 3.13 *Si deux vecteurs lignes de A sont colinéaires (l'une est proportionnelle à l'autre), alors $det(A) = 0$.*

3.3 Changement de base

On s'intéresse maintenant à l'influence d'un changement de base sur la matrice d'une application linéaire f de V dans V. On considère un vecteur u qui se décompose ainsi

$$u \; = \; \sum_{j=1}^{n} \lambda_j \, v_j \quad , \quad u \; = \; \sum_{i=1}^{n} \mu_i e_i \quad ,$$

sur les deux bases précédentes. On a les relations

$$\textit{pour tout } j \;\;,\;\; v_j \; = \; \sum_{i=1}^{n} A_{ij} e_i \quad , \quad \textit{et pour tout } i \;,\;\; \mu_i \; = \; \sum_{j=1}^{n} A_{ij} \lambda_j \;.$$

On note M la matrice représentant f dans la base $v_1, v_2, ..., v_n$ et H la matrice représentant f dans la base $e_1, e_2, ..., e_n$. On a d'une part

$$f(u) \; = \; \sum_{j,k} \; M_{kj} \lambda_j v_k \; = \; \sum_{i,j,k} \; M_{kj} \lambda_j A_{ik} e_i \quad ,$$

et d'autre part

$$f(u) \; = \; \sum_{i,k} H_{ik} \mu_k e_i \; = \; \sum_{i,j,k} H_{ik} A_{kj} \lambda_j e_i \;.$$

par identification des composantes sur une base, il vient pour tout i, j,

$$\left(\sum_{k=1}^{n} A_{ik} M_{kj} \right) \lambda_j \; = \; \left(\sum_{k=1}^{n} H_{ik} A_{kj} \right) \lambda_j$$

et en prenant pour vecteur u tel que $\lambda_j \neq 0$, il reste $\left(\sum_{k=1}^{n} A_{ik} M_{kj}\right) = \left(\sum_{k=1}^{n} H_{ik} A_{kj}\right)$. En écriture matricielle, cela devient

$$A\ M\ =\ H\ A\ ,$$

ce qui constitue la relation de passage de H à M et inversement. On a

$$H\ =\ A\ M\ A^{-1} \quad ou \quad M\ =\ A^{-1} H\ A\ .$$

La matrice de changement de base A est aussi appelée **matrice de passage**, et dans ce cas, notée P . Elle est relativemant facile à construire. On se fixe un indice k entre 1 et n, et on prend $u = v_k$, le $k-ième$ vecteur de la base $v_1, v_2, ..., v_n$, ce qui revient à prendre $\lambda_k = 1$, $\lambda_j = 0\ si\ j \neq k$. On a alors, pour chaque composante μ_i, $i = 1, 2, .., n$ du vecteur v_k,

$$\mu_i\ =\ \sum_{j=1}^{n} A_{ij} \lambda_j\ =\ A_{ik}\ .$$

On constate que la $k - ième$ colonne de la matrice de passage A est constituée des composantes μ_i de v_k dans la base $v_1, v_2, ..., v_n$.

Remarque 3.14 *Ceci était déjà implicitement écrit en posant*

$$v_k\ =\ \sum_{i=1}^{n}\ A_{ik} e_i\ \ .$$

4 Diagonalisation d'une matrice carrée

Soit A une matrice carrée opérant sur un espace vectoriel V de dimension n. On recherche les directions privilégiées pour lesquelles A se comporte comme une simple homothétie, c'est à dire qu'un cherche un vecteur $u \neq 0$ tel qu'il existe un scalaire λ pour lequel

$$Au\ =\ \lambda u\ .$$

Définition 4.1 *Si une telle situation existe, le paramètre λ est appelée* **valeur propre** *de A et le vecteur u est appelé* **vecteur propre** *associé à la valeur propre λ.*

Remarque 4.2 *Il est fondamental d'exiger $u \neq 0$, sinon tout réel est valeur propre. On remarque aussi que si u est vecteur propre, tout vecteur de la forme βu avec $\beta \neq 0$,est aussi vecteur propre associé à la même valeur propre λ.*

Supposons qu'on ait trouvé n vecteurs propres $u_1, u_2, ..., u_n$ d'une application linéaire f, constituant une base de V. La matrice de passage d'une base donnée $e_1, e_2, ..., e_n$, dans la base des vecteurs propres $u_1, u_2, ..., u_n$ est notée P . On note A la matrice représentant f dans la base $e_1, e_2, ..., e_n$, et D la matrice représentant f dans la base $v_1, v_2, ..., v_n$. On a

$$D \;=\; P^{-1} A\, P \quad , \quad ou \quad A \;=\; P\; D\; P^{-1} \quad ,$$

comme précédemment, avec

$$u_j \;=\; \sum_{i=1}^{n} P_{ij} e_i \; .$$

Soit u un vecteur de V. Il peut s'écrire

$$u \;= \sum_{j=1}^{n} \alpha_j \; u_j \quad , \quad u \;=\; \sum_{i=1}^{n} \beta_i \; e_i$$

avec, pour chaque $i = 1, 2, ..., n$, la relation

$$\sum_{j=1}^{n} P_{ij} \alpha_j \;=\; \beta_i \quad .$$

On veut écrire la matrice D. On a

$$f(u) \;=\; f\left(\sum_{j=1}^{n} \alpha_j u_j\right) \;=\; \sum_{j=1}^{n} \alpha_j \; f\left(u_j\right) \; ,$$

et en notant λ_j la valeur propre associée au vecteur propre u_j, donc telle que

$$f(u_j) \;=\; \lambda_j \; u_j \; ,$$

on obtient

$$f(u) \;=\; \sum_{j=1}^{n} \lambda_j \alpha_j \; u_j \; , \; \textit{à comparer avec } f(u) \;=\; \sum_{j=1}^{n} \left(\sum_{k=1}^{n} D_{jk} \alpha_k\right) \; u_j \; ,$$

qui est la formule attendue. On en déduit, pour chaque $j = 1, 2, ..., n$, la relation

$$\sum_{k=1}^{n} D_{jk} \alpha_k \;=\; \lambda_j \; \alpha_j \; .$$

On prend par exemple, pour j fixé, le vecteur $u = u_j$, c'est à dire $\alpha_j = 1, \;\; \alpha_k = 0 \;\; si \;\; k \neq j$ et on obtient

$$D_{jj} \;=\; \lambda_j \; .$$

Ensuite, on fixe encore j et aussi un autre indice $i \neq j$ et on prend $u = u_j + u_i$, pour avoir $\alpha_j = \alpha_i = 1$, les autres valeurs α_k étant nulles. Il reste

$$D_{jj} \;+\; D_{ji} \;=\; \lambda_j \quad ,$$

et on en déduit

$$D_{ji} = 0 \quad , \quad si \;\; i \neq j \; ,$$

c'est à dire que la matrice représentant f est une **matrice diagonale** dans la base des vecteurs propres, les éléments diagonaux étant exactement les valeurs propres $\lambda_1, \lambda_2, ..., \lambda_n$, dans l'ordre de classement des vecteurs propres .

Théorème 4.3 *Soit A une matrice carrée opérant sur un espace vectoriel V de dimension n, admettant n valeurs propres* **distinctes**. *Alors les n vecteurs propres associés contituent une base de V et exprimée dans cette base de vecteurs propres, cette matrice est une matrice diagonale, la diagonale étant constituée des valeurs propres $\lambda_1, \lambda_2, ..., \lambda_n$ rangée dans l'ordre des vecteurs propres de la base.*

Démonstration- Il suffit de montrer que les vecteurs propres $u_1, u_2, ..., u_n$ sont indépendants. Supposons qu'il existe des coefficients $\alpha_1, \alpha_2, ..., \alpha_n$ tels que

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

Montrons que ces coefficients α_k sont tous nuls. On commence par s'intéresser au premier, c'est à dire α_1. On introduit la matrice

$$B_1 \ = \ (A - \lambda_2 I) \ (A - \lambda_3 I) \ ... \ (A - \lambda_n I) \ ,$$

qui correspond à un produit de matrices. On a nécessairement

$$B_1 \left(\sum_{k=1}^{n} \alpha_k u_k \right) \ = \ 0 \ , \ \textit{ou encore} \quad \sum_{k=1}^{n} \alpha_k \ B_1 u_k \ = \ 0 \ .$$

On calcule chaque $B_1 u_k$. On a

$$B_1 u_k \ = \ \left(\prod_{j=2}^{n} (\lambda_k - \lambda_j) \right) \ u_k \ ,$$

et on observe que ce produit est toujours nul si $k \geq 2$. En effet, le facteur $(\lambda_k - \lambda_j)$ devient nul lorsque $j = k$. Il reste donc

$$\sum_{k=1}^{n} \alpha_k \ B_1 u_k \ = \ \alpha_1 \ \left(\prod_{j=2}^{n} (\lambda_1 - \lambda_j) \right) \ u_1 \ = \ 0 \ ,$$

où le produit est non nul, les valeurs propres étant distinctes, où $u_1 \neq 0$ en tant que vecteur propre, et il ne reste que la possibilité $\alpha_1 = 0$. Pour montrer que la coefficient α_k est nul, on introduit la matrice

$$B_k \ = \ \prod_{j \neq k} (A - \lambda_j I)$$

et on évalue $B_k u_i$ pour $i \neq k$. On obtient

$$B_k u_i \ = \ \left(\prod_{j \neq k} (\lambda_i - \lambda_j) \right) u_i \ ,$$

qui est nul, le facteur $(\lambda_i - \lambda_j)$ devenant nul pour $j = i$. Il reste

$$\sum_{i=1}^{n} \alpha_i \ B_k u_i \ = \ \alpha_k \ \left(\prod_{j \neq k} (\lambda_k - \lambda_j) \right) \ u_k \ = \ 0 \ ,$$

où le produit est non nul, les valeurs propres étant distinctes, où $u_k \neq 0$ en tant que vecteur propre, et il ne reste que la possibilité $\alpha_k = 0$. Ainsi tous les coefficients α_k sont nuls, les vecteurs propres $u_1, u_2, ..., u_n$ sont indépendants et forment donc une base car ils génèrent nécessairement V, étant du même nombre que la dimension.

4.1 Le calcul des valeurs propres

L'introduction d'une base de vecteurs propres d'une application linéaire est ainsi très intéressante car dans cette base la matrice représentant cette application linéaire est diagonale. Le théorème précédent indique qu'il suffit de disposer de n valeurs propres distinctes $\lambda_1, \lambda_2, ..., \lambda_n$. La question qui se pose est maintenant de savoir caractériser ces valeurs propres, pour les calculer. Soit λ un nombre ; il s'agit d'une valeur propre de A lorsqu'il existe un vecteur propre u ($\neq 0$) tel que

$$(A - \lambda I)\, u \; = \; 0 \; .$$

D'après le remarque 3.7, ceci impose

$$det\,(A - \lambda I) \; = \; 0 \; .$$

On introduit donc la fonction P de $\mathbb{R}$ dans $\mathbb{R}$, définie par

$$P(\lambda) \; = \; det\,(A - \lambda I) \quad ,$$

et en constatant que le calcul du déterminant d'une matrice carrée $n \times n$ ne fait intervenir que des combinaisons linéaires de produits de n termes chacun, on se rend compte que $P(\lambda)$ est nécéssairement un polynôme de degré inférieur ou égal à n . Il est en fait exactement de degré n, le calcul du coefficient de λ^n étant assez facile à obtenir ; il est égal à $(-1)^n$, et il est bien différent de zéro.

Définition 4.4 *Le polynôme* $P(\lambda) \; = \; det(A - \lambda I)$ *est appelé* **polynôme caractéristique** *de la matrice* A.

On vient de constater le résultat suivant.

Théorème 4.5 *Les valeurs propres de* A *sont les racines de son polynôme caractéristique.*

Ainsi, les valeurs propres sont distinctes lorque le polynôme caractéristique n'admet pas de racines multiples. Ce polynôme étant de degré n, et si ses racines sont toutes simples, on dispose bien de n racines distinctes, et donc de n vecteurs propres.

Conclusion 1 *Lorsque les valeurs propres sont distinctes, il existe une base de vecteurs propres dans laquelle la matrice* A *est diagonale. Le calcul des valeurs propres correspond à l'extraction des racines d'un polynôme. Le calcul des vecteurs propres revient à chercher une solution non nulle d'un système de* $n-1$ *équations à* n *inconnues. En effet, il s'agit de résoudre*

$$(A - \lambda I) \;\; u \; = \; 0 \; ,$$

où les équations sont liées. On doit en enlever une et fixer arbitrairement une des composantes du vecteur u, *ce qui ne pose pas de problème puisque, selon la remarque 4.2,* u *n'est déterminé qu'à une constante multiplicative près.*

Remarque 4.6 *Il n'est jamais exclu que certaines valeurs propres soient des nombres complexes ; dans ce cas, les composantes du vecteur propre associé peuvent l'être également.*

Remarque 4.7 *Lorsqu'une valeur propre est double, deux cas peuvent se présenter, suivant le nombre de vecteurs propres associés. Dans le cas où il y a deux vecteurs propres, notons-les* u_j *et* u_{j+1}*, il suffit de citer deux fois la valeur propre dans la liste, une fois avec la notation* λ_j *et l'autre fois avec la notation* λ_{j+1}*, et de procéder exactement comme dans le cas précédent. Dans le cas où on ne trouve qu'un seul vecteur propre, on est contraint de compléter la liste des* u_j *pour en faire une base, par un autre vecteur qui ne soit pas une combinaison linéaire des précédents. Dans ce cas, on n'obtiendra pas une matrice diagonale dans la nouvelle base, mais on peut "bien choisir" le vecteur additionnel pour la rendre* **triangulaire***, c'est à dier telle que*

$$A_{ij} \ = \ 0 \ si \ i < j \quad (triangulaire \ inf\acute{e}rieure)$$

ou telle que

$$A_{ij} \ = \ 0 \ si \ i > j \quad (triangulaire \ sup\acute{e}rieure) \ \ .$$

On reprend la même idée pour une valeur propre triple ou lorsqu'il y a plusieurs valeurs propres multiles. D'un point de vue académique, cette situation est assez courante, par qu'elle permet de proposer des exercices "moins triviaux" aux étudiants. D'un point de vue pratique, cette situation signifie surtout que le modèle mathématique est surdimensionné, et qu'il est peut-être utile d'enlever une équation qui vient parasiter les autres....

Théorème 4.8 *Soit* A *une matrice carrée de valeurs propres* $\lambda_1, \lambda_2, ..., \lambda_n$ *chacune étant citée autant de fois que sa multiplicité (par exemple une valeur propre double est citée deux fois). Alors*

$$det(A) \ = \ \prod_{j=1}^{n} \lambda_j \quad ,$$

c'est à dire que le déterminant est égal au produit des valeurs propres.

Démonstration Dans la nouvelle base des vecteurs propres éventuellement complétée, la matrice A devient diagonale ou triangulaire, c'est à dire qu'en utilisant la matrice de passage P, on a

$$D \ = \ P^{-1} A \ P \ \ .$$

Or le déterminant de D, qui est soit diagonale soit triangulaire, est égal au produit des termes diagonaux, d'après la règle de calcul consistant à toujours développer suivant la première colonne. Il est donc égal au produit des valeurs propres. Il vient

$$\prod_{j=1}^{n} \lambda_j \ = \ det\,(D) \ = \ det\,(P) \ \ det\,(A) \ \ det\left(P^{-1}\right) \ = \ \frac{det\,(P)}{det\,(P)} \, det\,(A) \ = det\,(A) \quad ,$$

d'où le résultat.

Corollaire 4.9 *Une matrice inversible ne peut pas avoir de valeur propre nulle.*

Exemple : On considère la matrice

$$A \ = \ \begin{pmatrix} 8 & 0 & 9 \\ -3 & -1 & -3 \\ -6 & 0 & -7 \end{pmatrix}$$

dans une base e_1, e_2, e_3 donnée. Les valeurs propres sont -1 (double) et 2. Il y a deux vecteurs propres indépendants associés à la valeur propre -1 : $u_1 = e_1 - e_3$, et $u_2 = e_2$ et pour la valeur propre 2, on a $u_3 = 3e_1 - e_2 - 2e_3$. Dans la base des vecteurs propres, la matrice devient diagonale :

$$D = P^{-1} A \, P ,$$

avec

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} , \quad P = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ -1 & 0 & -2 \end{pmatrix} \quad et \quad P^{-1} = \begin{pmatrix} -2 & 0 & -3 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} .$$

5 Quelques exemples d'applications

5.1 Les systèmes d'équations différentielles linéaires

On travaille dans l'espace vectoriel $V = \mathbb{R}^n$, muni de la base canonique $e_1, e_2, ..., e_n$. Soit A une matrice $n \times n$, admettant n vecteurs propres $u_1, u_2, ..., u_n$ constituant une base de V. On note P la matrice de passage de la base $e_1, e_2, ..., e_n$ dans la base $u_1, u_2, ..., u_n$. On sait que la $j-ième$ colonne de P est constituée des composantes du vecteur u_j dans la base $e_1, e_2, ..., e_n$. On sait aussi que dans la base $u_1, u_2, ..., u_n$, la forme de la matrice est diagonale :

$$D = P^{-1} A \, P .$$

On se donne ensuite une fonction F définie sur $[0, +\infty[$ et à valeurs dans V et un vecteur $X_0 \in V$.

On cherche une fonction X définie sur $[0, +\infty[$ et à valeurs dans V, solution du système d'équations différentielles

$$X'(t) = A \, X(t) + F(t) \quad , \quad X(0) = X_0 \ .$$

Ce système peut aussi s'écrires

$$X'(t) = A \, P \, P^{-1} X(t) + F(t) \quad , \quad X(0) = X_0 \ .$$

On le multiplie par la matrice P^{-1} , pour obtenir

$$\left(P^{-1} X'(t)\right) = \left(P^{-1} A \, P\right) \left(P^{-1} X(t)\right) + \left(P^{-1} F(t)\right) \quad ,$$

et en posant

$$Y(t) = P^{-1} X(t) \ , \ G(t) = P^{-1} F(t) \ ,$$

on obtient

$$Y'(t) = D \, Y(t) + G(t) \ ,$$

où la matrice D est diagonale. Il s'agit en fait maintenant d'un système de n équations **indépendantes** de la forme

$$y_j'(t) = \lambda_j \, y_j(t) + g_j(t) \ ,$$

en notant λ_j la valeur propre associée au vecteur propre u_j, et $g_j(t)$ la $j-ème$ composante de $G(t)$, dans la base des vecteurs propres. L'intégration de cette équation donne une solution de la forme

$$y_j(i) \ = \ B_j \, e^{\lambda_j t} \ + \ z_j(t) \ ,$$

où $z_j(t)$ est une solution particulière de l'équation avec second membre, et B_j une constante d'intégration.

Il reste à revenir à $X(t)$. On a, pour chaque $i = 1, 2, ..., n$, dans la base $e_1, e_2, ..., e_n$,

$$X_i(t) \ = \ \sum_{j=1}^{n} P_{ij} \ B_j \ e^{\lambda_j t} \ + \ \sum_{j=1}^{n} P_{ij} z_j(t) \ ,$$

en séparant la solution du système homogène et la solution particulière du système avec second membre. Pour fixer les constantes d'intégration, on doit satisfaire, pour chaque $i = 1, 2, ..., n$,

$$\sum_{j=1}^{n} P_{ij} \ \ (B_j + z_j(0)) \ = \ X_i(0) \ = \ X_{0,i} \ .$$

Il s'agit d'un système linéaire.

On a aussi, en travaillant dans la base des vecteurs propres $u_1, u_2, ..., u_n$,

$$X(t) \ = \ \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij} B_j \ e^{\lambda_j t} \right) \ e_i \ + \ \sum_{i=1}^{n} \left(\sum_{j=1}^{n} P_{ij} z_j(t) \right) \ e_i \ \ ,$$

qui devient

$$X(t) \ = \ \sum_{j=1}^{n} B_j \ e^{\lambda_j t} \ u_j \ + \ \sum_{j=1}^{n} \ z_j(t) \ u_j \ .$$

Pour obtenir les constantes B_j il suffit maintenant de faire $t = 0$, et d'identifier les composantes (cette fois ci dans la base $u_1, u_2, ..., u_n$). On obtient

$$B_j + z_j(0) \ = \ \overline{X}_j(0) \ = \ \overline{X}_{0,j}$$

en notant ainsi

$$X(t) \ = \ \sum_{j=1}^{n} \overline{X}_j(t) \ u_j$$

les composantes de $X(t)$ dans la bas $u_1, u_2, ..., u_n$.

Remarque 5.1 *Lorsque la matrice A dépend de t, le système s'écrit*

$$X'(t) \ = \ A(t) \ X(t) \ + \ F(t) \quad , \quad X(0) \ = \ X_0 \ .$$

On veut procéder de la même façon, les valeurs propres et les vecteurs propres pouvant dépendre de t. En posant $X(t) \ = \ P(t)Y(t)$,il vient

$$P(t)Y'(t) \ + \ P'(t)Y(t) \ = \ A(t)P(t)Y(t) \ + \ F(t) \ ,$$

et en multipliant par $P^{-1}(t)$ on obtient

$$Y'(t) \;=\; \big(D(t) - P^{-1}(t)P'(t)\big)\,Y(t) \;+\; G(t)\;,$$

en conservant les mêmes notations. On perd l'avantage de la matrice diagonale, sauf dans un cas, lorsque les vecteurs propres sont des constantes. Dans ce cas, $P'(t) = 0$, et la matrice diagonale est conservée.

Remarque 5.2 *Lorqu'il y a des valeurs propres multiples, et qu'on ne dipose pas d'une base de vecteurs propres, on peut procéder de la même façon, avec une matrice D ou $D(t)$ qui est triangulaire. Les vecteurs de base additionnels viendront perturber le terme (appelé* **Source***) $F(t)$. On ne développera pas de mèthode générale pour ce cas, qui se résoud le plus souvent au cas par cas.*

6 Exercices et Exemples

6.1 Un exercice académique

On considère la matrice

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

où ϵ est un paramètre réel.
1 ° - Pour quelles valeurs de ϵ la matrice A est-elle inversible ? (Rep.$\epsilon \neq 1$),
2 ° - Ecrire la matrice A^{-1}. Rep.

$$A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ \frac{1}{\epsilon-1} & 1 & \frac{1}{1-\epsilon} \\ \frac{1}{1-\epsilon} & 0 & \frac{1}{\epsilon-1} \end{pmatrix} \quad .$$

3 ° - Déterminer ϵ pour que $\lambda = -1$ soit valeur propre de A.
4 ° - Déterminer le vecteur propre associé.
5 ° - Montrer que les autres valeurs propres sont positives.
6 ° - Déterminer la seule solution bornée du système différentiel

$$x'(t) \;=\; x(t) + y(t) + z(t)\;,\;\; y'(t) \;=\; y(t) + z(t) + 1\;,\;\; z'(t) \;=\; x(t) + y(t) + \epsilon z(t)\;,$$

telle $y(0) = 1$. Rep. $x(t) = e^{-t} + 1$, $y(t) = 2e^{-t} - 1$, $z(t) = -4e^{-t}$, Il n'y a pas de composantes sur les autres vecteurs propres car elles ne seraient pas bornées ($e^{\lambda t} \longrightarrow \infty \; si \; \lambda > 0$).

6.2 Quelques exemples en météorologie

On considère le système différentiel

$$\begin{array}{rcl} x'(t) & = & u(t) \ , \\ y'(t) & = & v(t) \ , \\ u'(t) & = & -k \ u(t) \ + \ \omega \ v(t) \ - a \\ v'(t) & = & -k \ v(t) \ - \ \omega \ u(t) \ - \ b \end{array}$$

avec des données initiales

$$x(0) \ = \ x_0 \ , \ y(0) \ = \ y_0 \ , \ u(0) \ = \ u_0 \ , \ v(0) \ = \ v_0$$

données. Ce système représente les trajectoires $x = x(t), \ y = y(t)$ de particules ou de phénomènes portées par un vent de vitesse $u(t), v(t)$. Il peut s'agir d'un nuage toxique, de vents de sables, d'un nuage de sauterelles, d'un tourbillon ou même d'une tornade ou d'un cyclone. Le paramètre k ($< 10^{-4}$) est un coefficient de friction. On a toujours $k \geq 0$. Le paramètre ω représente la force géostrophique, ou force de Coriolis ; on a

$$\omega \ = \ \frac{\pi \ sin\phi}{21600}$$

en un point de latitude ϕ, avec $\phi > 0$ dans l'hémisphère nord, $\phi < 0$ dans l'hémisphère sud, et $\phi = 0$ à l'équateur. Les termes a et b représentent le gradient de pression atmosphérique ; ils sont en général positif dans la zone tropicale de l'Atlantique (ITCZ). Le système est écrit pour une carte en coordonnées de Mercator.

La matrice du système s'écrit

$$\begin{pmatrix} 0 & 0 & u & 0 \\ 0 & 0 & 0 & v \\ 0 & 0 & -k & \omega \\ 0 & 0 & -\omega & -k \end{pmatrix} \ ,$$

et on obtient ses valeurs propres en calculant les racines de

$$det \begin{pmatrix} 0-\lambda & 0 & u & 0 \\ 0 & 0-\lambda & 0 & v \\ 0 & 0 & -k-\lambda & \omega \\ 0 & 0 & -\omega & -k-\lambda \end{pmatrix} \ = \ 0 \ ,$$

ce qui donne l'équation

$$\lambda^2 \left((\lambda + k)^2 + \omega^2 \right) \ = \ 0 \ .$$

Les racines sont

$$\begin{array}{rcll} \lambda & = & 0 & \textit{racine double}, \\ \lambda & = & -k - i\omega & \textit{racine simple} \\ \lambda & = & -k + i\omega & \textit{racine simple} \end{array} \ .$$

On constate cependant que le système constitué des deux dernières équations est indépendant, et qu'une fois résolu en $u(t)$ et en $v(t)$, les deux premières équations ne correspondront plus qu'à une simple recherche de primitives. On ne rencontrera donc aucun problème avec la valeur propre double $\lambda = 0$.

Le système en u, v fait apparaître les valeurs propres complexes

$$\lambda \;=\; -\, k \;-\; i\omega$$

et $\lambda \;=\; -k \;+\; i\omega$, et un terme source **constant**. On obtient une solution particulière constante $u_*,\; v_*$ solution du système linéaire

$$-\, k\, u \;+\; \omega\, v \;=\; a \quad , \quad -\, k\, v \;-\; \omega\, u \;=\; b\;.$$

On obtient

$$u_* \;=\; -\,\frac{\omega b + ka}{\omega^2 + k^2} \quad , \quad v_* \;=\; \frac{\omega a - kb}{\omega^2 + k^2}\;.$$

Par ailleurs, on sait que

$$e^{i\omega t} \;=\; cos(\omega t) \;+\; i\; sin(\omega t)$$

et on cherche une expression pour $u(t)$ qui soit de la forme

$$u(t) \;=\; A\; e^{-kt} cos(\omega t) \;+\; B\; e^{-kt} sin(\omega t) \;+\; u_*\;,$$

où u_* est la solution particulière que l'on vient de déterminer. On obtient ensuite v en utilisant la même équation, mise sous la forme

$$v(t) \;=\; \frac{1}{\omega}\;(u'(t) + ku(t) + a)$$

et on obtient, en utilisant

$$v_* \;=\; \frac{1}{\omega}\;(ku_* + a)$$

l'expression de $v(t)$,

$$v(t) \;=\; B\; e^{-kt} cos(\omega t) \;-\; A\; e^{-kt} sin(\omega t) \;+\; v_*\;.$$

Il reste à utiliser les conditions initiales pour déterminer A et B. On obtient immédiatement

$$A \;=\; u_0 - u_* \quad , \quad B \;=\; v_0 - v_* \quad .$$

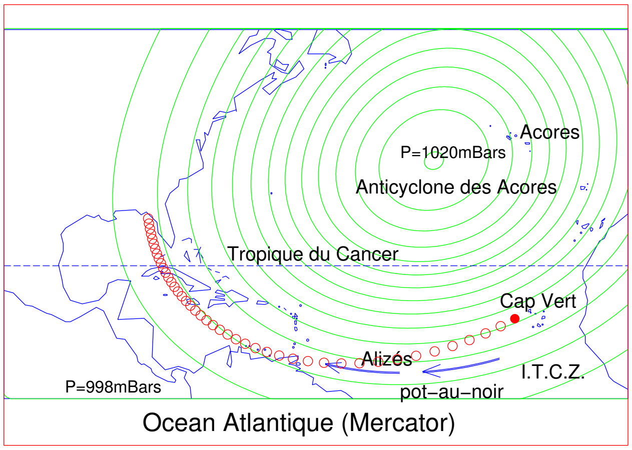
Les trajectoires $x(t)$ et $y(t)$ sont obtenues en intégrant. On obtient

$$x(t) = x_0 + u_* t + \frac{k\,(u_0 - u_*) + \omega\,(v_0 - v_*)}{\omega^2 + k^2}\,\big(1 - e^{-kt} cos(\omega t)\big) + \frac{\omega\,(u_0 - u_*) - k\,(v_0 - v_*)}{\omega^2 + k^2}\; e^{-kt} sin(\omega t)\;,$$

et

$$y(t) = y_0 + v_* t + \frac{k\,(v_0 - v_*) - \omega\,(u_0 - u_*)}{\omega^2 + k^2}\,\big(1 - e^{-kt} cos(\omega t)\big) + \frac{\omega\,(v_0 - v_*) + k\,(u_0 - u_*)}{\omega^2 + k^2}\; e^{-kt} sin(\omega t)\;.$$

Nous allons maintenant, à partir de ces formules, étudier quelques situations.
La zone tropicale : Au sud de l'anticyclone des Açores, en zone intertropicale, la pression varie très peu en fonction de la longitude (variable x) et on peut négliger la composante a par rapport à



b. De plus, si on n'est pas trop près de l'équateur, la friction k est petite devant ω. On obtient que v_* devient négligeable devant u_*, et es trajectoires sont très étirées d'est en ouest. Ceci correspond aux vents **Alizés**, que recherchent les navigateurs pour se rendre aux Caraibes. On observe que la direction (u_*, v_*) est très proche de la tangente à l'isobare, orthogonale à la direction (a, b) ce qui s'observe nettement sur les cartes météorologiques. Ces mêmes vents vont entraîner les cyclones qui se forment près des îles du Cap Vert en été, vers les Caraïbes et la Golfe du Mexique.

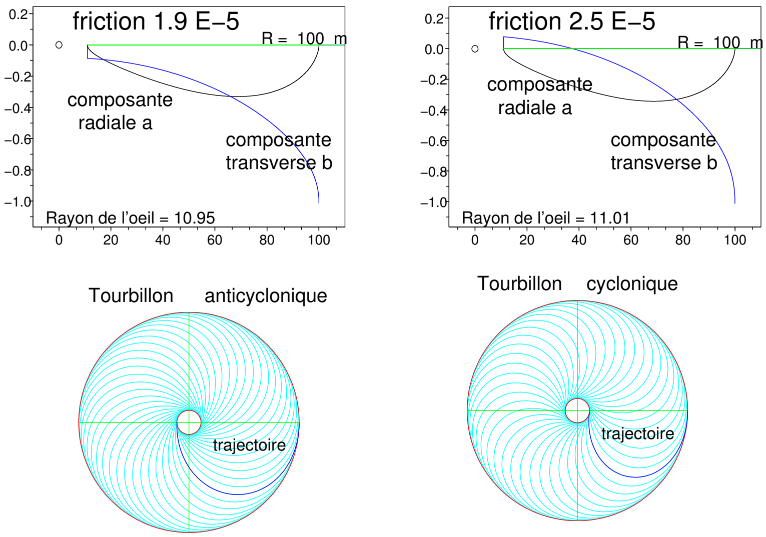
La zone équatoriale : Ici, la latitude ϕ est proche de zéro, et la force géostrophique (Coriolis) devient infime lorsque la pression devient telle que $b \simeq 0$, c'est à dire que la pression ne dépend pas de la longitude. Les composantes u_* et v_* de la vitesse sont proches de

$$u_* \simeq -\frac{ka}{\omega^2 + k^2} \ , \ v_* \simeq \frac{\omega a}{\omega^2 + k^2} \ .$$

Il est intéressant de remarquer qu'en négligeant la friction, on aurait eu $v_* \simeq \frac{a}{\omega}$, c'est à dire des vents très violents, ce qui n'est pas du tout le cas... En effet, les navigateurs se retrouvent quelquefois encalminés, donc immobilisés, dans une zone sans aucun vent, appelée la zone du **pot-au-noir**, en référence au récipient servant à contenir le cirage (noir) utilisé autrefois pour l'entretien des cuirs sur les navires, et qui traduit l'ambiance de cafard et d'angoisse que connaissaient à l'époque les marins affamés dans de telles circonstances. Cette zone est toujours redoutée par les skippers qui remontent l'Atlantique depuis le Cap Horn.

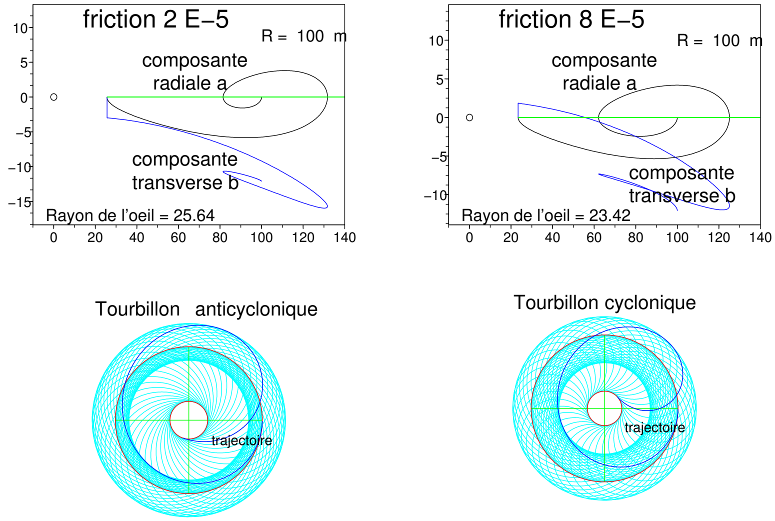
Les tourbillons : dans les deux exemples précédents, la variation en t restait très petite. On se place maintenant à une échelle plus réduite. On prend $x_0 = R, \ y_0 = 0, \ u_0 = 0, \ v_0 = -z_0$, avec

$z_0 > 0$. Le paramètre R correspond au rayon d'un tourbillon atmosphérique.



Le calcul précédent donne la trajectoire des particules depuis le rayon externe du troubillon, vers le centre, où elle vient s'enrouler autour d'un oeil central , dans un sens déterminé par la valeur du coefficient de friction. Comme chaque point du cercle de rayon R génère une trajectoire, on observe un vent tournant (tourbillon) qui ramène tout vers le centre et ensuite vers le haut. Le terme géostrophique est ici augmenté par rapport aux exemples à grande échelle, traduisant, en combinaison avec le terme source constant, l'aspiration vers le centre du phénomène. Ce phénomène explique les envols subits de foin que l'on observe en été, et est utilisé au Sahel, par exemple par les criquets pélerins pour s'élever en altitude, et migrer vers de nouvelles contrées. Le phénomène suit alors les vents imposés par les valeurs u_* et v_* des paramètres.

Les tornades : il s'agit d'une version plus violente du phénomène précédent, la force géostrophique est rendue plus forte par le phénomène d'aspiration. On observe un évasement du phénomène près du sol, puis un rétrécissement en altitude.
Les différentes figures montrent une famille de trajectoires reproduites d'après le calcul précédent de notre système d'équations différentielles. Ici encore la valeur du terme de friction oriente l'enroulement des trajectoires autour de l'oeil. Il existe une valeur critique de la friction pour laquelle cet oeil est réduit à un seul point.



Remarque 6.1 *Les* **cyclones** *ont un comportement analogue, en ce qui concerne l'enroulement autour de l'oeil. Le départ des trajectoires à partir du cercle externe est cependant très différent, il est complètement radial, et non transverse, ce qui implique un comportement non linéaire très différent de celui décrit ici. Les échelles sont également très différentes.*

Remarque 6.2 *Ces différents exemples suggèrent l'interprétation suivante : la solution particulière de l'équation avec second membre décrit la trajectoire du phénomène global, de basse fréquence. La solution générale de l'équation homogène décrit le détail du phénomène, de plus haute fréquence.*

7 Orthogonalité

7.1 Produit scalaire

Soit V un espace vectoriel sur $\mathbb{R}$.

Définition 7.1 *Une* **forme bilinéaire** *sur V est une fonction définie sur $V \times V$ à valeurs dans $\mathbb{R}$, linéaire en chacune de ses variables :*

$$(u, v) \longrightarrow a(u, v)$$

.

On a ainsi, pour tout $u, v \in V$, et tout $\lambda_1, \lambda_2, \mu_1, \mu_2 \in \mathbb{R}$,

$$a(\lambda_1 u_1 + \lambda_2 u_2, \mu_1 v_1 + \mu_2 v_2) = \lambda_1 \mu_1 a(u_1, v_1) + \lambda_2 \mu_1 a(u_2, v_1) + \lambda_1 \mu_2 a(u_1, v_2) + \lambda_2 \mu_2 a(u_2, v_2) \ .$$

Définition 7.2 *Une* **forme linéaire symétrique** *sur V est une forme linéaire sur V telle pour tout $u, v \in V$, $a(u,v) = a(v,u)$.*

On a alors

$$a(\lambda_1 u_1 + \lambda_2 u_2, \mu_1 v_1 + \mu_2 v_2) = \lambda_1 \mu_1 a(u_1, v_1) + (\lambda_2 \mu_1 + \lambda_1 \mu_2)\, a(u_1, v_2) + \lambda_2 \mu_2 a(u_2, v_2) \ .$$

Définition 7.3 *La* **forme quadratique** *associée à la forme linéaire $a(u,v)$ est l'application J de V dans $\mathbb{R}$ définie par*

$$\forall u \in V \quad J(u) \ = \ a(u,u) \ .$$

Proposition 7.4 *On a la relation suivante, pour tout $u, v \in V$,*

$$J(u+v) \ = \ J(u) \ + \ 2\ a(u,v) \ + \ J(v) \ .$$

Remarque 7.5 *On en déduit*

$$a(u,v) \ = \ \frac{1}{2} \ (J(u+v) - J(u) - J(v)) \quad ,$$

ce qui signifie qu'étant donnée une forme quadratique $J(u)$ on peut toujours lui associer une forme bilinéaire symétrique.

Définition 7.6 *Une forme bilinéaire symétrique est* **positive** *lorsque la forme quadratique associée est toujours positive, c'est à dire*

$$\forall u \in V \quad , \quad J(u) \ \geq \ 0 \ .$$

Elle est **définie positive** *lorsqu'elle est positive et que*

$$J(u) \ = \ 0 \ \Longrightarrow u = 0 \ .$$

Définition 7.7 *Un* **produit scalaire** *sur V est la donnée d'une forme bilinéaire symétrique définie positive sur V.*

Définition 7.8 *Un* **Espace Euclidien** *est un espace vectoriel de dimension finie, muni d'un produit scalaire. Le produit scalaire est souvent noté simplement (u,v) au lieu de $a(u,v)$.*

Exemples 1 *1 °-L'espace* $\mathbb{R}^n$ *muni du produit* $a(u,v) = \sum_{j=1}^{n} u_j v_j$ *est un espace euclidien. La forme quadratique associée induit la norme euclidienne*

$$\|u\| = \sqrt{J(u)} = \sqrt{\sum_{j=1}^{n} u_j^2} \ ,$$

qui elle même induit la distance euclidienne

$$\|u-v\| = \sqrt{J(u-v)} = \sqrt{\sum_{j=1}^{n} (u_j - v_j)^2} \ .$$

2 °- L'espace V_h *introduit précédemment, muni du produit scalaire*

$$a(u,v) = \int_0^L (u'(x)v'(x) + Ku(x)v(x)) \ dx \ .$$

3 °- L'espace vectoriel des polynômes de degré $\leq n$*, muni du produit scalaire*

$$a(p,q) = \int_{-1}^{1} w(x) \ p(x)q(x) \ dx \ ,$$

où $w(x)$ *est une fonction positive donnée (appelée fonction poids).*

Définition 7.9 *Deux vecteurs* u *et* v *sont* **orthogonaux** *lorsque leur produit scalaire est nul :*

$$a(u,v) = 0 \ .$$

Définition 7.10 *Une base* $u_1, u_2, ..., u_n$ *d'un espace Euclidien est une* **base orthogonale** *lorsque ses vecteurs sont deux à deux orthogonaux, c'est à dire*

$$(u_j, u_k) = 0 \quad si \ \ j \neq k \ .$$

L'utilisation d'une base orthogonale simplifie beaucoup les calculs. Ainsi, lorsque dans un espace V *la base* $u_1, u_2, ..., u_n$ *est orthogonale, le produit scalaire du vecteur* $u = \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n$ *par le vecteur* $v = \beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n$ *donne*

$$(u,v) = \alpha_1 \beta_1 \|u_1\|^2 + \alpha_2 \beta_2 \|u_2\|^2 + ... + \alpha_n \beta_n \|u_n\|^2 \ .$$

Il n'y a pas de termes en (u_j, u_k) *avec* $j \neq k$ *.*

Définition 7.11 *Une base orthogonale* $u_1, u_2, ..., u_n$ *est une* **base orthonormée** *lorsque pour tout* $j = 1, 2, ..., n$ *on a*

$$\|u_j\| = 1 \ .$$

Remarque 7.12 *Dans une base orthonormée* $u_1, u_2, ..., u_n$ *le produit scalaire du vecteur* $u = \alpha_1 u_1 + \alpha_2 u_2 + ... + \alpha_n u_n$ *par le vecteur* $v = \beta_1 u_1 + \beta_2 u_2 + ... + \beta_n u_n$ *donne simplement*

$$(u,v) = \alpha_1 \beta_1 + \alpha_2 \beta_2 + ... + \alpha_n \beta_n \ .$$

7.2 Application : la projection

Soit E un espace vectoriel muni d'un produit scalaire $a(u,v)$ et V un sous espace vectoriel de E, de dimension finie. On utilise une base $v_1, v_2, ..., v_n$ de V. On cherche un élément $u \in V$ rendant minimal la quantité $J(u-f)$.

Définition 7.13 *Cet élément u rendant $J(u-f)$ minimal est appelé* **projection** *de f sur V.*

Soit $i \in \{1, 2, ..., n\}$. On écrit que la quantité

$$J(u - f + \lambda v_i)$$

est minimale en $\lambda = 0$. S'agissant d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$, il suffit de dériver et d'écrire que la dérivée est nulle en $\lambda = 0$. On a

$$J(u - f + \lambda v_i) \; = \; J(u-f) \; + \; 2\lambda a(u-f, v_i) \; + \; \lambda^2 J(v_i) \; ,$$

de dérivée

$$\frac{d}{d\lambda} \, J(u - f + \lambda v_i) \; = \; 2 \, a(u-f, v_i) \; + \; 2\lambda \, J(v_i)$$

qui est nulle en $\lambda = 0$,c'est à dire que pour tout $i \in \{1, 2, ..., n\}$, on a

$$a(u, v_i) \; = \; a(f, v_i) \; .$$

On pose $b_i \; = \; a(f, v_i)$ et on décompose u sur la base : $u \; = \; \sum_{j=1}^{n} \mu_j v_j$, pour obtenir

$$\sum_{j=1}^{n} a(v_j, v_i) \; \mu_j \; = \; b_i \; .$$

Il s'agit d'un système linéaire dont la matrice de coefficients $A_{ij} \; = \; a(v_i, v_j)$ est symétrique. On montre facilement que $det(A) \neq 0$, d'où l'existence et l'unicité de la projection u de f sur V.

Remarque 7.14 *Dans l'exemple de l'étagère, le rôle de V étant joué par V_h, le déplacement u correspond à la projection de la charge f selon le produit scalaire*

$$a(u, v) \; = \; \int_0^L \left(u'(x)v'(x) + K u(x) v(x) \right) \, dx \; .$$

La matrice A obtenue est symétrique, définie positive, et tridiagonale, donc aisément inversible.

Exemple 7.15

On s'intéresse à l'espace H constitué des fonctions v définies sur l'intervalle $[0, 2]$, à valeurs dans $\mathbb{R}$, telles que

$$v \; \in \; C^1([0, 2]) \; \; ,$$

$$v \textit{ est un polynôme de degré } \leq 2 \textit{ sur } [0,1],$$
$$v \textit{ est un polynôme de degré } \leq 2 \textit{ sur } [1,2] \ .$$

L'espace H est un sous espace vectoriel de $C^1([0,2])$. En effet, si u et v sont des éléménts de H et λ, μ deux réels, on a bien $\lambda u + \mu v \in C^1([0,2])$, et $\lambda u + \mu v$ est bien un polynôme de degré ≤ 2 sur $[0,1]$ et aussi sur $[1,2]$. Les vecteurs sont de la forme

$$v(x) \;=\; \begin{cases} a \;+\; b(x-1) \;+\; c(x-1)^2 & si\ 0 < x < 1 \ , \\ a' \;+\; b'(x-1) \;+\; d(x-1)^2 & si\ 1 < x < 2 \ . \end{cases}$$

La continuité impose $a = a'$, et la continuité de la dérivée impose $b = b'$.
On considère les fonctions $v_1,\ v_2,\ v_3,\ v_4$ définies sur $[0,2]$ par :

$$\begin{aligned} v_1(x) &= 1 \ (fonction\ constante) \\ v_2(x) &= \sqrt{5}\ (x-1)^2 \\ v_3(x) &= \sqrt{3}\ (x-1) \\ v_4(x) &= \sqrt{5}\ (x-1)\ |x-1| \ . \end{aligned}$$

Ces quatre fonctions constituent une base de H.

En effet, il suffit de montrer l'indépendance linéaire. On considère un élémént de la forme $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4$, et on écrit que pour tout x,

$$\lambda_1 v_1(x) + \lambda_2 v_2(x) + \lambda_3 v_3(x) + \lambda_4 v_4(x) \;=\; 0.$$

En faisant $x = 0$, on obtient $\lambda_1 = 0$. Ensuite, en divisant par $\sqrt{5}(x-1)$, il reste

$$\lambda_2\,(x-1) + \;\lambda_3 \sqrt{\frac{3}{5}} \;+\; \lambda_4\,|x-1| \;=\; 0 \ .$$

On fait $x = 1$ pour établir $\lambda_3 = 0$, puis $x = 0$ et $x = 2$ pour obtenir le système

$$-\lambda_2 + \lambda_4 = 0 \ , \ \ \lambda_2 + \lambda_4 = 0 \ ,$$

d'où $\lambda_2 = \lambda_4 = 0$.
La matrice de projection sur cette base est notée A, dont les coefficients sont définis par

$$a_{ij} \;=\; \int_0^2 v_i(x)\ v_j(x)\ dx \qquad (i,j \in \{1,2,3,4\}) \ .$$

La matrice A est symétrique, définie positive. Le calcul des coefficients a_{ij}, pour $i,j \in \{1,2,3,4\}$, donne

$$A \;=\; \begin{pmatrix} 2 & \frac{2\sqrt{5}}{3} & 0 & 0 \\ \frac{2\sqrt{5}}{3} & 2 & 0 & 0 \\ 0 & 0 & 2 & \frac{\sqrt{15}}{2} \\ 0 & 0 & \frac{\sqrt{15}}{2} & 2 \end{pmatrix} \ .$$

Le choix particulier de la base permet d'utiliser des arguments de symétrie pour réaliser

$$a_{13} = a_{31} = a_{14} = a_{41} = a_{23} = a_{32} = a_{24} = a_{42} = 0 \ .$$

Le calcul des valeurs propres et des vecteurs propres d'une matrice de la forme

$$\begin{pmatrix} 2 & \alpha \\ \alpha & 2 \end{pmatrix}$$

avec $\alpha \in \mathbb{R}$, utilise

$$Det \begin{bmatrix} 2-\lambda & \alpha & \\ \alpha & 2 & -\lambda \end{bmatrix} = (2-\lambda)^2 - \alpha^2$$

, dont les racines sont $\lambda_1 = 2-\alpha$ et $\lambda_2 = 2+\alpha$, d'où les valeurs propres, et les vecteurs propres correspondants $R_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $R_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

La matrice A étant diagonale par blocs, d'après ce qui précède, avec $\alpha = \frac{2\sqrt{5}}{3}$ puis $\alpha = \frac{\sqrt{15}}{2}$, ses valeurs propres sont

$$\lambda_1 = 2 - \frac{2\sqrt{5}}{3} \ , \ \lambda_2 = 2 + \frac{2\sqrt{5}}{3} \ , \ \lambda_3 = 2 - \frac{\sqrt{15}}{2} \ , \ \lambda_4 = 2 + \frac{\sqrt{15}}{2} \ ,$$

et les vecteurs propres correspondants

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} \ , \ u_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \ , \ u_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \quad u_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} .$$

Le déterminant de A est égal au produit des valeurs propres, et vaut

$$\left(4 - \frac{20}{9}\right)\left(4 - \frac{15}{4}\right) = \frac{4}{9} .$$

La base u_1, u_2, u_3, u_4 est orthogonale, la matrice de passage P dans cette base et son inverse P^{-1} sont

$$P = \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \quad , \quad P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} .$$

Dans cette base, la matrice de projection est diagonale :

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix} .$$

Soit f une fonction intégrable sur $[0,2]$. On veut projeter f sur H, c'est à dire chercher le vecteur $u \in H$ le plus proche de f. On commence par évaluer

$$f_j = \int_0^2 f(x)\ v_j(x)\ dx \ , \quad j = 1,2,3,4 \ ,$$

qui sont les composantes d'un vecteur $b = f_1 v_1 + f_2 v_2 + f_3 v_3 + f_4 v_4$ que l'on réécrit sur la base des vecteurs propres : $b = b_1 u_1 + b_2 u_2 + b_3 u_3 + b_4 u_4$ .Dans cette base de vecteurs propres, le calcul des composantes de : $u = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 + \beta_4 u_4$ est immédiat :

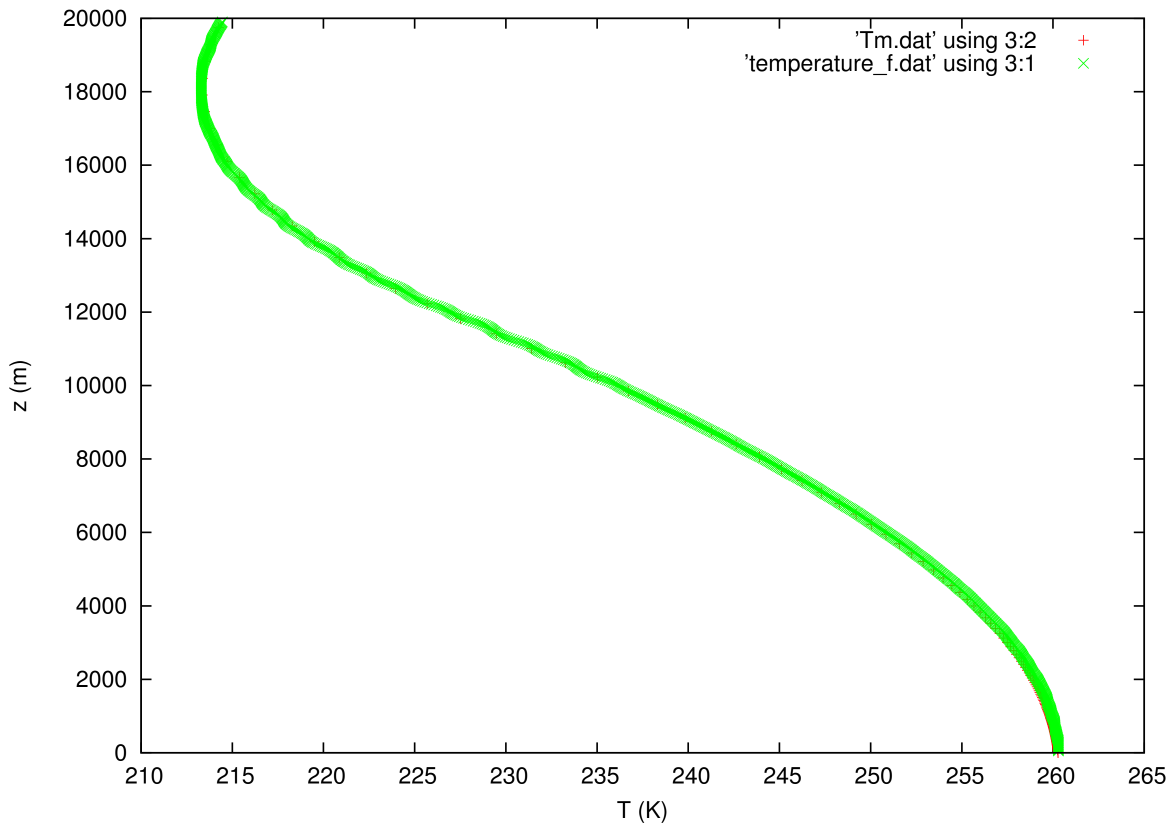
$$\beta_j \;=\; \frac{b_j}{\lambda_j} \quad , \quad j = 1, 2, 3, 4 \; .$$

Dans la base initiale, on obtient : $u = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \alpha_4 v_4$ avec

$$\alpha_1 \;=\; \frac{1}{2\lambda_1}\,(f_1 - f_2) \;+\; \frac{1}{2\lambda_2}\,(f_1 + f_2) \quad , \quad \alpha_2 \;=\; -\frac{1}{2\lambda_1}\,(f_1 - f_2) \;+\; \frac{1}{2\lambda_2}\,(f_1 + f_2) \quad ,$$

et c

$$\alpha_3 \;=\; \frac{1}{2\lambda_3}\,(f_3 - f_4) \;+\; \frac{1}{2\lambda_4}\,(f_3 + f_4) \quad , \quad \alpha_4 \;=\; -\frac{1}{2\lambda_3}\,(f_3 - f_4) \;+\; \frac{1}{2\lambda_4}\,(f_3 + f_4) \quad .$$



Cet exemple intervient lors de l'exploitation des mesures des radiosondes utilisées pour obtenir des mesures de températures à différentes altitudes. En pratique, l'intervalle $[0, 2]$ correspond à un intervalle allant du sol ($z = 0$) jusqu'à $z_0 = 20000$ *mètres*. La mi hauteur, $\frac{z_0}{2}$, et ici $x = 1$, correspond à la tropopause où on observe une inflexion de la courbe des températures, entre la troposphère (en dessous) et la stratosphère (au dessus). Le cas présenté ici correspond à une région de latitude assez proche du pôle, la température au sol étant de 260 ° K, soit -13 ° C. Dans une telle région, la tropopause se situe à 10 km d'altitude environ. Il n'y a pas de symétrie entre les deux côtés de la tropopause, ce qui exclut l'utilisation de polynômes de degré ≤ 3 par exemple. Pour des mesures donnant

$$f_1 = 471.49071 \ , \ f_2 = 352.31732 \ , \ f_3 = -32.7621 \ , \ f_4 = \ -30.221767 \ ,$$

on obtient $\alpha_1 = 235$, $\alpha_2 = 1$, $\alpha_3 = -28$, $\alpha_4 = 12$. On obtient l'approximation de la température au sol en faisant $x = 0$, ce qui donne $u(0) \ = \ \beta_1 + \sqrt{5}\ \beta_2 - \sqrt{3}\ \beta_3 - \sqrt{5}\ \beta_4 = 258.9\ K$
On observe que ce n'est pas exactement la valeur attendue, et comme la mesure d'une température au sol est très facile à réaliser directement, on peut remplacer une des quatre équations précédentes, par exemple la dernière par une équation liant les coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ et la température au sol mesurée (ici 260 ° K). Cette équation s'écrit $\alpha_1 + \sqrt{5}\ \alpha_2 - \sqrt{3}\ \alpha_3 - \sqrt{5}\ \alpha_4 = 260$. Lors de la résolution du nouveau système ainsi obtenu, les valeurs de α_1 et de α_2 sont inchangées. Il reste à résoudre le système de deux équations $-\sqrt{3}\ \alpha_3 \ - \ \sqrt{5}\ \alpha_4 \ = \ 260 \ - \ \alpha_1 \ - \ \sqrt{5}\alpha_2$, $2\ \alpha_3 \ + \ \frac{\sqrt{15}}{2}\ \alpha_4 \ = \ -32.7621$, ce qui donne $\alpha_3 \ = \ -26.096$, $\alpha_4 \ = \ 10.03347$. On observe une nette variation avec les résultats précédents, ce qui illustre un mauvais conditionnement de la matrice A, révélé par un rapport $\frac{min(\lambda_j)}{max(\lambda_k)}$ relativement petit. Notons que la matrice correspondant au second système linéaire n'est pas symétrique, et que la construction d'une base othogonale n'est pas immédiate.

7.3 La régression linéaire

On dispose de mesures expérimentales

$$(x_1, y_1) \ , \ (x_2, y_2) \ , \ \ldots \ , \ (x_n, y_n)$$

et on cherche à en tirer une loi d'état de la forme

$$y = ax + b \ ,$$

c'est à dire évaluer les paramètres a et b. On suppose qu'au moins deux x_j sont différents, c'est à dire que $n \geq 2$ et qu'il existe un indice j et un indice k tels que $x_j \neq x_k$. La **méthode des moindres carrés** consiste à rechercher a et b réalisant le minimum de

$$J(a,b) \ = \ \sum_{j=1}^{n} (ax_j + b - y_j)^2 \ .$$

Ce minimum est caractérisé par les racines communes des dérivées partielles $\frac{\partial J}{\partial a}(a,b)$ et $\frac{\partial J}{\partial b}(a,b)$, c'est à dire

$$\frac{\partial J}{\partial a}(a,b) \ = \ 2 \ \sum_{j=1}^{n} x_j \,(ax_j + b - y_j) \ = \ 0 \ \ , \ \ \frac{\partial J}{\partial b}(a,b) \ = \ 2 \ \sum_{j=1}^{n} (ax_j + b - y_j) \ = \ 0 \ .$$

Ces deux équations correspondent au système linéaire :

$$\left(\sum_{j=1}^{n} x_j^2 \right) \ a \ + \ \left(\sum_{j=1}^{n} x_j \right) \ b \ = \ \sum_{j=1}^{n} x_j y_j \ \ ,$$

$$\left(\sum_{j=1}^{n} x_j\right) a + n\, b = \sum_{j=1}^{n} y_j \ .$$

Le déterminant de la matrice est égal à

$$D = n \sum_{j=1}^{n} x_j^2 - \left(\sum_{j=1}^{n} x_j\right)^2 > 0 \ .$$

En effet, l'expression

$$\sum_{j=1}^{n} (x_j - x)^2$$

ne peut pas être négative ou nulle (si elle est nulle, on a $x = x_j$ pour tout j ce qui signifierait que tous les x_j sont égaux, ce qui n'est pas notre hypothèse). Comme il s'agit d'un polynôme du second degré en la variable x , qui s'écrit

$$n\, x^2 - 2 \left(\sum_{j=1}^{n} x_j\right) x + \sum_{j=1}^{n} x_j^2 > 0 \ ,$$

on sait que son discréminant est strictement négatif, d'où

$$\left(\sum_{j=1}^{n} x_j\right)^2 < n \sum_{j=1}^{n} x_j^2 \ .$$

On obtient, en posant

$$\overline{x} = \frac{1}{n} \sum_{j=1}^{n} x_j \ , \ \overline{y} = \frac{1}{n} \sum_{j=1}^{n} y_j \ , \ \sigma^2 = \frac{1}{n} \sum_{j=1}^{n} x_j^2 \ , \ \rho = \frac{1}{n} \sum_{j=1}^{n} x_j y_j \ ,$$

les valeurs de a et b,

$$a = \frac{\rho - \overline{x}\ \overline{y}}{\sigma^2 - \overline{x}^2} \quad , \quad b = \frac{\sigma^2 \overline{y} - \rho \overline{x}}{\sigma^2 - \overline{x}^2} \ .$$

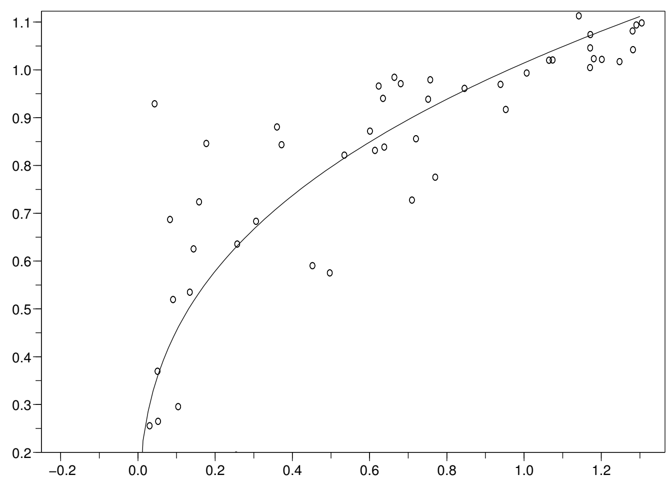
Il s'agit de la **droite des moindres carrés**.

On peut tout aussi bien rechercher une loi de la forme

$$y = a_1 \phi_1(x) + a_2 \phi_2(x) + ... + a_m \phi_m(x) \ ,$$

avec $m \leq n$. Ceci peut intervenir dans une recherche de loi de type exponentiel par exemple $z = A\, x^\gamma$, illustrée par une figure ci-dessous, avec une cinquantaine de mesures. On aura $ln(z) = ln(\ A) + \gamma\ ln(x)$, et on prendra $y = ln(z)$, $\phi_1(x) = 1$, $\phi_2(x) = ln(x)$ et bien entendu $a_1 = ln(A)$, $a_2 = \gamma$. On cherche à rendre minimal l'expression

$$J(a_1, a_2, ..., a_m) = \sum_{j=1}^{n} \left(a_1 \phi_1(x_j) + a_2 \phi_2(x_j) + ... + a_m \phi_m(x_j) - y_j\right)^2 \ .$$



On écrit que les dérivées partielles sont nulles au minimum, c'est à dire pour tout $i = 1, 2, ..., m$,

$$\frac{\partial J}{\partial a_i}(a_1, a_2, ..., a_m) \;=\; 2\sum_{j=1}^{n} \phi_i(x_j)\;\; (a_1\phi_1(x_j) + a_2\phi_2(x_j) + ... + a_m\phi_m(x_j)) \;=\; 0\ ,$$

ce qui conduit à un système linéaire

$$A \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ a_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_m \end{pmatrix}\ ,$$

avec

$$b_i \;=\; \frac{1}{n}\sum_{j=1}^{n} \phi_i(x_j) y_j$$

et les coefficients de la matrice

$$A_{ik} \;=\; \frac{1}{n}\sum_{j=1}^{n}\; \phi_i(x_j)\phi_k(x_j)\quad .$$

On remarque que cette matrice est symétrique, définie positive lorsque les m vecteurs

$$u_i = \begin{pmatrix} \phi_i(x_1) \\ \phi_i(x_2) \\ . \\ . \\ \phi_i(x_n) \end{pmatrix}$$

constituent un système libre de $\mathbb{R}^n$. Dans ce cas, la matrice A est inversible, et on peut calculer les coefficients $a_{1,} a_2, ..., a_m$.

8 Les formes hermitiennes et la dualité

On considère un espace vectoriel V sur $\mathbb{C}$ et non plus sur $\mathbb{R}$. Rappelons le nombre complexe conjugué du nombre complexe $z = x + iy$ est le nombre complexe $\overline{z} = x - iy$. On a

$$\overline{\overline{z}} = z, \ \overline{zw} \ = \ \overline{z}\ \overline{w} \ , \ z\overline{z} = x^2 + y^2 = |z| \ ,$$

et si $\overline{z} = z$, alors z est un réel.

8.1 Les formes hermitiennes

Définition 8.1 *Une fonction f de V dans $\mathbb{C}$ est* **antilinéaire** *lorsque,*

$$\textit{pour tout } u, v \in V \ , \ \textit{tout } \lambda \in \mathbb{C} \ , \quad f(u+v) \ = \ f(u) + f(v) \quad , \quad f(\lambda u) \ = \ \overline{\lambda} \ f(u) \ .$$

Définition 8.2 *Une fonction a définie sur $V \times V$ et à valeurs dans $\mathbb{C}$ est* **sesquilinéaire** *lorsqu'elle est linéaire par rapport au premier argument et antilinéaire par rapport au second.*

Ainsi la fonction a vérifie, pour tout $u, v \in V$, et tout $\lambda, \mu \in \mathbb{C}$,

$$a(\lambda u, \mu v) \ = \ \lambda \overline{\mu} a(u, v) \ .$$

Définition 8.3 *Une fonction a définie sur $V \times V$ et à valeurs dans $\mathbb{C}$ est une* **forme hermitienne** *lorsqu'elle est sesquilinéaire et vérifie, pour tout $u, v \in V$,*

$$a(u, v) \ = \ \overline{a(v, u)} \ .$$

Proposition 8.4 *Soit a une forme hermitienne sur V. Alors*

$$\forall u \in V \quad a(u, u) \ \in \ \mathbb{R} \ .$$

La démonstration est immédiate.

Définition 8.5 *Une forme hermitienne est* **positive** *lorsque*

$$\forall u \in V \quad a(u,u) \;\geq\; 0 \; .$$

Elle est **définie positive** *si elle est positive et vérifie*

$$a(u,u) \;=\; 0 \;\Longrightarrow u = 0 \quad .$$

Dans ce cas, la forme hermitienne est appelée **produit scalaire** *sur* V *, et la quantité*

$$\|u\| \;=\; \sqrt{a\,(u,u)}$$

est appelée **norme de** u *, associée au produit scalaire* $a\,(.,.)$ *.*

Exemples 1 *1 ° - Pour* $V = \mathbb{C}$, *on pose, pour* $z = x + iy$ *et* $w = u + iv$,

$$a(z,w) \;=\; z\overline{w} \;=\; (xu + yv) \;+\; i\;\;(yu - xv) \;\; .$$

On a bien $a(z,z) = z\overline{z} = x^2 + y^2$ *et la norme associée correspond au module* $|z| \;=\; \sqrt{z\overline{z}}$ *.*
2 ° - Si V *est un espace de polynômes à valeurs complexes de la variable réelle* $x \in I$ *un intervalle réel, on pose*

$$a(p,q) \;=\; \int_I \; w(x)\; p(x)\; \overline{q(x)}\; dx \; ,$$

où w *est une fonction (poids) donnée, strictement positive sur* I *. La norme est définie par*

$$\|p\| \;=\; \sqrt{\int_I w(x) \;\; |p(x)|^2 \, dx} \quad .$$

3 ° - On considère deux séries de Fourier complexes,

$$u(x) \;=\; \sum_{n=-\infty}^{+\infty} a_n \; e^{inx} \quad , \quad v(x) \;=\; \sum_{n=-\infty}^{+\infty} b_n \; e^{inx} \; ,$$

et on pose

$$a(u,v) \;=\; \sum_{n=-\infty}^{+\infty} a_n \overline{b_n} \; .$$

Il s'agit d'un produit scalaire définie sur l'ensemble des séries de Fourier convergentes. La norme est la racine carrée de la somme de la série absolument convergente

$$\|u\| \;=\; \sqrt{\sum_{n=-\infty}^{\infty} |a_n|^2} \quad .$$

Définition 8.6 *Un* **espace hermitien** *est un espace vectoriel sur $\mathbb{C}$, de dimension* **finie** *et muni d'un produit scalaire.*

Dans les trois exemples précédents, on est en présence d'un espace hermitien à condition de limiter à un entier n le degré des polynômes, ou de filtrer les hautes fréquences des séries de Fourier, en imposant $|n| \leq N$, pour un entier N fixé. Le produit scalaire est souvent noté

$$\langle u, v \rangle \; = \; a\,(u, v) \; .$$

Nous conserverons la notation $a\,(u, v)$.

Définition 8.7 *Soit a une forme hermitienne sur V. La fonction J de V dans $\mathbb{R}$ définie par*

$$J(u) \; = \; a(u, u)$$

est appelée **forme quadratique** *associée à a.*

Remarquons que si la forme a était seulement sesquilinéaire, la forme quadratique associée serait complexe. Dans le cas hermitien, on peut envisager des projections comme dans le cas Euclidien.

Soit H un espace Hermitien, de dimension n, et $e_1, e_2, ..., e_n$ une base de H. On note $a(.,.)$ le produit scalaire de H, et on considère deux vecteurs u et v de H, dont les développements sur la base sont notés

$$u \; = \; \sum_{j=1}^{n} \alpha_j e_j \quad , \quad v \; = \; \sum_{k=1}^{n} \beta_k e_k \; .$$

On calcule $a(u, v)$. On obtient

$$a(u, v) \; = \; \sum_{j=1}^{n} \sum_{k=1}^{n} \; \alpha_j \overline{\beta_k} \; a(e_j, e_k) \; .$$

On note A la matrice de coefficients $A_{jk} \; = \; a(e_j, e_k)$. On remarque qu'elle vérifie la propriété

$$A_{jk} \; = \; \overline{A_{kj}} \; ,$$

c'est à dire que sa transposée s'identifie avec sa complexe conjuguée.

Définition 8.8 *Un* **matrice hermitienne** *est une matrice dont la transposée est égale à sa complexe conjuguée.*

Ceci généralise la notion de matrice symétrique. On constate immédiatement qu'une matrice hermitienne à coefficients réels est symétrique.
On s'intéresse maintenant aux effets d'un changement de base. On introduit une nouvelle base $f_1, f_2, ..., f_n$ de H , et la matrice de passage ou de changement de base est notée P avec

$$f_k \; = \; \sum_{j=1}^{n} P_{jk} e_j \; .$$

Dans cette nouvelle bases les décompositions de nos deux vecteurs seront notées

$$u \;=\; \sum_{l=1}^{n} \gamma_l \; f_j \quad , \quad v \;=\; \sum_{m=1}^{n} \delta_m f_m \quad .$$

On a, pour tout j, k les relations

$$\alpha_j \;=\; \sum_{l=1}^{n} P_{jl}\gamma_l \quad , \quad \beta_k \;=\; \sum_{m=1}^{n} P_{km}\delta_m \quad .$$

D'où

$$a(u,v) = \sum_{j,k} \alpha_j \overline{\beta_k} A_{jk} \;=\; \sum_{j,k,l,m} (P_{jl}\gamma_l) \; \overline{(P_{km}\delta_m)} \; A_{jk} \;=\; \sum_{l,m} \gamma_l \overline{\delta_m} \left(\sum_{j,k} P_{jl} \overline{P_{km}} A_{jk} \right) \; ,$$

et en posant

$$B \;=\; P' \; A \; \overline{P} \; ,$$

on obtient

$$a(u,v) \;=\; \sum_{l,m} \gamma_l \overline{\delta_m} B_{lm} \quad .$$

Proposition 8.9 *La transformée d'une matrice Hermitienne A par une matrice de passage P est*

$$B \;=\; P' \; A \; \overline{P} \quad .$$

Remarque : Cette formule ne fait pas intervenir P^{-1}.

Théorème 8.10 *Les valeurs propres d'une matrice hermitienne sont* **réelles.** Si la forme hermitienne associée est positive, elles sont **non négatives**. *Si la forme hermitienne associée est un produit scalaire, elles sont* **strictement positives**.

Démonstration Soit u un vecteur propre d'une matrice hermitienne A, associé à une valeur propre λ. On a $u \neq 0$. On peut considérer A comme la matrice d'une forme hermitienne a sur $\mathbb{C}^n \times \mathbb{C}^n$ exprimée dans la base canonique $e_1, e_2, ..., e_n$, c'est à dire que l'on a

$$A_{jk} \;=\; a(e_j, e_k) \; .$$

De la même façon on peut développer u sur cette base :

$$u \;=\; \sum_{j=1}^{n} \alpha_j e_j \quad .$$

On calcule maintenant $a(\overline{u}, \overline{u})$ qui est un réel strictement positif, en utilisant

$$\sum_{k=1}^{n} a(e_j, e_k) \; \alpha_k \;=\; \lambda \; \alpha_j \;\; , \;\; j = 1, 2, ..., n$$

qui traduit $Au = \lambda u$, pour obtenir, en multipliant par $\overline{\alpha_j}$,

$$a(\overline{u},\overline{u}) \;=\; \sum_{j,k} \overline{\alpha_j}\ \alpha_k\ a(e_j,e_k) \;\;=\; \lambda \sum_{j=1}^{n}\ \overline{\alpha_j}\ \alpha_j \;=\; \lambda\ \sum_{j=1}^{n} |\alpha_j|^2\ ,$$

dont on déduit que λ est réelle et non négative si $a(\overline{u},\overline{u}) \geq 0$. Si la forme hermitienne associée est un produit scalaire, elle est strictement positive car le vecteur propre est non nul, par nature.

Théorème 8.11 *Soit a une forme hermitienne sur V un espace vectoriel de dimension finie. Alors il existe une base orthogonale $u_1, u_2, ..., u_n$ relativement à a, c'est à dire vérifiant*

$$a(u_j,u_k) \;=\; 0 \quad si\ j{\neq}k\ . \tag{8.1}$$

La démonstration se fait par construction de cette base. Dans cette base, la matrice de a est de la forme

$$A \;=\; \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ . & . & \dots & . \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}\ .$$

Si $u = \sum_{j=1}^{n} \alpha_j u_j$, la forme quadratique associée est

$$J(u) \;=\; \sum_{j=1}^{n}\ \lambda_j\ |\alpha_j|^2\ \ .$$

Ce résultat induit le résultat suivant, en posant $v_j = \frac{1}{\sqrt{\lambda_j}}\ u_j$ de façon à avoir une base **orthonormale**, c'est à dire telle que

$$a(v_j,v_k) \;=\; \begin{cases} 1 & si\ j=k\ , \\ 0 & si\ j{\neq}k\ . \end{cases}$$

Théorème 8.12 *Tout espace hermitien H admet une base orthonormale. Dans cette base, le produit scalaire hermitien est le produit scalaire canonique, c'est à dire*

$$a(u,v) \;=\; \sum_{j=1}^{n}\ \alpha_j\overline{\beta_j} \quad ,\quad pour\ u = \sum_{j=1}^{n}\alpha_j v_j\ ,\quad v \;=\; \sum_{j=1}^{n}\ \beta_j v_j\ ,$$

en notant $v_1, v_2, ..., v_n$ cette base.

8.2 La dualité

Soit H un espace hermitien, et f une application linéaire de H dans H. On choisit de représenter H par une base orthonormée $e_1, e_2, ..., e_n$.

Définition 8.13 *On note f^* l'unique application linéaire de H dans H, caractérisée par*

$$\forall u, v \in H \ , \quad a(f(u), v) \ = \ a(u, f^*(v)) \ \ .$$

L'application f^* est appelée **adjointe** *de f.*

Proposition 8.14 *Soit f, g deux applications linéaires sur H, et $\alpha \in \mathbb{C}$. Alors*

$$(f+g)^* \ = \ f^* + g^* \quad , \quad (f \circ g)^* \ = \ g^* \circ f^* \quad , \quad (\alpha f)^* \ = \ \overline{\alpha} f^* \quad ,$$

$$(f^*)^* \ = \ f \quad , \quad Id^* \ = \ Id \quad , \quad \left(f^{-1}\right)^* \ = \ (f^*)^{-1} \ \ .$$

On note A la matrice de f.

Proposition 8.15 *La matrice A^* de f^* est donnée par*

$$A^* \ = \ \overline{A'} \ ,$$

la conjuguée complexe de la transposée de A. Elle est appelée **matrice adjointe** *de A.*

Pour le démontrer, il suffit de développer

$$a(Au, v) \ = \ a(u, A^* v) \ \ .$$

Définition 8.16 *Soit H un espace hermitien, f une application linéaire de H dans H, représentée par la matrice M dans une base orthonormée $e_1, e_2, ..., e_n$.*
On dit que f est **normal** *lorsque $f \circ f^* = \ f^* \circ f$,*
On dit que f est **autoadjoint** *lorsque $f^* = \ f$,*
On dit que f est **unitaire** *lorsque $f \circ f^* = \ f^* \circ f \ = \ Id$,*
On dit que M est une matrice **normale** *lorsque $MM^* = M^*M$,*
On dit que M est une matrice **hermitienne** *lorsque $M^* = M$,*
On dit que M est une matrice **unitaire** *lorsque $MM^* = M^*M = Id$.*

Proposition 8.17 *L'application linéaire f vérifie :*
f est unitaire si et seulement si $f^* = f^{-1}$,
si f est autoadjoint, alors f est normal,
si f est unitaire, alors f est normal.

Proposition 8.18 *La matrice M vérifie :*
si M est hermitienne, alors M est normale,
si M est symétrique, alors M est normale,
si M est unitaire, alors M est normale,
si M est orthogonale, alors M est normale.

Proposition 8.19 *Soit f une application linéaire normale. Alors, pour tout $u \in H$, on a*

$$a\left(f(u), f(u)\right) \;=\; a\left(f^*(u), f^*(u)\right) \quad .$$

Démonstration : On calcule

$$a\left(f(u), f(u)\right) \;=\; a\left(u, f^*(f(u))\right) \;=\; a\left(u, f(f^*(u))\right) \;=\; a\left(f^*(u), f^*(u)\right) \;.$$

Proposition 8.20 *Si u est un vecteur propre de f , une application linéaire normale, associé à la valeur propre λ, alors u est aussi vecteur propre de f^* associé à la valeur propre $\overline{\lambda}$.*

Démonstration : On calcule

$$a\left(f^*(u) - \overline{\lambda}u, f^*(u) - \overline{\lambda}u\right) \;=\; a\left(f^*(u), f^*(u)\right) - \overline{\lambda}a\left(u, f^*(u)\right) - \lambda a\left(f^*(u), u\right) + \lambda\overline{\lambda}a\left(u, u\right) \;,$$

où

$$a\left(f^*(u), f^*(u)\right) \;=\; a\left(f(u), f(u)\right) \;,$$

d'après la proposition précédente, et

$$\overline{\lambda}a\left(u, f^*(u)\right) \;=\; a\left(f(u), \lambda u\right) \quad , \quad \lambda a\left(f^*(u), u\right) \;=\; a\left(\lambda u, f(u)\right) \quad .$$

On en déduit

$$a\left(f^*(u) - \overline{\lambda}u, f^*(u) - \overline{\lambda}u\right) \;=\; a\left(f(u) - \lambda u, f(u) - \lambda u\right) \;=\; 0 \;, \quad car\ f(u) = \lambda u \;,$$

et donc, a étant un produit scalaire,

$$f^*(u) \;=\; \overline{\lambda}u \;,$$

d'où le résultat.

Proposition 8.21 *Deux vecteurs propres associés à deux valeurs propres distinctes d'une application linéaire normale sont orthogonaux.*

Démonstration : On note f l'application linéaire, et u, v les vecteurs propres associées respectivement aux valeurs propres λ et μ. On a $f(u) = \lambda u \;\;,\;\; f(v) = \mu v$. On a

$$\lambda\ a(u,v) \;=\; a(\lambda u, v) \;=\; a(f(u), v) \;=\; a(u, f^*(v)) \;=\; a(u, \overline{\mu}v) \;=\; \mu\ a(u,v) \;,$$

d'où

$$(\lambda - \mu)\ \ a(u,v) \;=\; 0 \;,$$

et comme $\lambda{\neq}\mu$ on a nécessairement $a(u,v) = 0$.

Théorème 8.22 *Soit H un espace Hermitien et f une application linéaire normale sur H. Alors il existe une* **base orthonormale** *de H constituée des vecteurs propres de f.*

Démonstration (partielle) : On se place dans le cas où toutes les valeurs propres sont simples. Alors il existe bien une base de vecteurs propres $u_1, u_2, ..., u_n$, et d'après la proposition précédente, cette base est orthogonale. Il suffit de la rendre orthonormale, en remplaçant u_j par

$$v_j = \frac{1}{\sqrt{a(u_j, u_j)}} u_j .$$

Lorsqu'il y a des valeurs propres multiples, le résultat reste vrai, et on l'admettra.

Théorème Spectral 1 *Une matrice hermitienne est diagonalisable, ses valeurs propres sont réelles et il existe une matrice de passage* **unitaire** *qui la diagonalise, c'est à dire,*

$$M = P \, D \, P' ,$$

en notant M la matrice hermitienne, D la matrice diagonale et P la matrice de passage, qui est unitaire.

Ce résultat traduit en terme de matrices le théorème précédent sur les applications linéaires.

Proposition 8.23 *Soit f une application linéaire sur H un espace hermitien. Alors*

$$f \text{ est unitaire} \iff \forall u, v \in H \quad a(f(u), f(v)) = a(u, v) \quad \iff \forall u \in H \quad \|f(u)\| = \|u\| \quad .$$

Démonstration : Si f est unitaire, pour tout

$$u, v \in H, \; a(u, v) = a(f^*(f(u)), v) = a(f(u), f(v)) \; .$$

Inversement si $a(u, v) = a(f(u), f(v))$ pour tout $u, v \in H$, alors

$$\forall u, v \in H \; a(u, v) = a(f^*(f(u)), v) \; ,$$

et donc

$$a(u - f^*(f(u)), v) = 0 \; ,$$

pour tout v et en particulier pour $v = u - f^*(f(u))$,d'où $\|u - f^*(f(u))\| = 0$, et pour tout $u \in H$, $f^*(f(u)) = u$. On obtient $f(f^*(v)) = v$,pour tout $v \in H$ en inversant les rôles de u et de v. Pour la même hypothèse, en faisant $u = v$ on obtient $\|f(u)\| = \|u\|$. Réciproquement, on a

$$4a(f(u), f(v)) = \|f(u+v)\|^2 - \|f(u-v)\|^2 + i\,\|f(u+iv)\|^2 - i\,\|f(u-iv)\|^2$$

et donc, en utilisant plusieurs fois $\|f(w)\| = \|w\|$, avec $w = u + v$, puis $w = u - v$, $w = u + iv$, $w = u - iv$, on a

$$4\, a(f(u), f(v)) = \|u+v\|^2 - \|u-v\|^2 + i\,\|u+iv\|^2 - i\,\|u-iv\|^2 = 4\, a(u, v) \quad ,$$

d'où le résultat.

Proposition 8.24 *Soit f une application linéaire sur H un espace hermitien. Alors*

$$f \text{ unitaire} \Longleftrightarrow \text{sa matrice } A \text{ dans une base orhogonormale est unitaire} \quad .$$

Démonstration : On a

$$f \text{ unitaire} \Longleftrightarrow f^* = f^{-1} \Longleftrightarrow \overline{A'} \;=\; A^* \;=\; A^{-1} \Longleftrightarrow A \text{ unitaire} \; .$$

Proposition 8.25 *Soit f une application linéaire sur H un espace hermitien. Si f est* **unitaire***, alors*

$$|det(A)| \;=\; 1 \; .$$

Démonstration : On a $det(A^*) \;=\; det(\overline{A'}) \;=\; \overline{det(A')} \;=\; \overline{det(A)}$, puisque A' et A ont le même déterminant. Comme $f^* \circ f = Id$, on a $A^*A = Id$, et donc

$$1 \;=\; det(Id) \;=\; det(A^*A) \;=\; det(A^*)\, det(A) \;=\; \overline{det(A)}\, det(A) \;=\; |det(A)|^2 \;\; ,$$

d'où le résultat.

Proposition 8.26 *Soit M une matrice sur $\mathbb{C}^n$. Les propositions suivantes sont équivalentes :*

$$\begin{array}{ll}
(1) & MM^* \;=\; Id \\
(2) & M^*M \;=\; Id \\
(3) & M \text{ est inversible et } M^* = M^{-1} \\
(4) & \text{les vecteurs colonnes de } M \text{ forment une famille orthonormale de } \mathbb{C}^n \\
(5) & \text{les vecteurs lignes de } M \text{ forment une famille orthonormale de } \mathbb{C}^n
\end{array}$$

De plus, le déterminant d'une matrice unitaire est de module 1.

Proposition 8.27 *Soit f une application linéaire sur H un espace hermitien. Pour que f soit autoadjointe, il faut et il suffit que sa matrice dans une base orthonormale soit hermitienne.*

Démonstration : En notant A la matrice, on a

$$f \text{ autoadjointe} \Longleftrightarrow f = f^* \Longleftrightarrow A \;=\; \overline{A'} \Longleftrightarrow A \text{ hermitienne} \; .$$

Proposition 8.28 *Soit f une application linéaire sur H un espace hermitien. Si f est autoadjointe, alors ses valeurs propres sont toutes réelles*

Démonstration : On note λ la valeur propre et u le vecteur propre associé : $f(u) = \lambda u$. On a

$$a\left(f(u), u\right) \;=\; a\left(u, f^*(u)\right) \;=\; a\left(u, f(u)\right) \;,$$

c'est à dire

$$a\left(\lambda u, u\right) \;=\; a\left(u, \lambda u\right)$$

ou

$$\left(\lambda - \overline{\lambda}\right)\; a\left(u, u\right) \;=\; 0 \;\Longrightarrow \lambda = \overline{\lambda}\;,$$

la valeur propre λ est réelle.

Proposition 8.29 *Soit A une matrice normale sur $\mathbb{C}^n$. Alors il existe une matrice unitaire P telle que $P^{-1}AP$ soit diagonale. Elle s'écrit aussi P^*AP.*

Démonstration immédiate à partir du théorème 8.22 .