Université Paris I, Panthéon - Sorbonne

LICENCE M.A.S.S.

Cours d'Algèbre S4

JEAN-MARC BARDET (UNIVERSITÉ PARIS 1, SAMM)

U.F.R. 27 et Equipe SAMM (Statistique, Analyse et Modélisation Multidisiplinaire)
Université Panthéon-Sorbonne, 90 rue de Tolbiac, 75013 Paris.

Plan du cours

1. Rappels.
2. Produits scalaires et applications
3. Une application en analyse: les séries de Fourier.
4. Formes linéaires et espace dual.
5. Application linéaire adjointe et réduction des matrices symétriques et orthogonales.
6. Formes bilinéaires symétriques et formes quadratiques.

Bibliographie

Livres pour commencer...

- Azoulay, E. et Avignant, J. *Mathématiques. Rappels de cours et exercices résolus IUT-BTS*, Tome 2 - Algèbre et géométrie. Dunod.
- Bouzitat, C. et Pradel, J. *Mathématiques : suites, diagonalisation, formes quadratiques.* Cujas.
- Hayek, N et Leca, J.-P. *Mathématiques pour l'économie.* Dunod.
- Guerrien, B. *Algèbre linéaire pour économistes.* Economica.
- Lecoutre, J.-P. et Pilibossian, P. *Algèbre : Exercices corrigés avec rappels de cours.* Masson.

Livres de niveau plus élevé

- *Mathématiques L2.* Pearson Education.
- Monier, J.-M. *Algèbre et géométrie PC-PSI-PT - Cours, méthodes, exercices corrigés.* Dunod.
- Ramis, E., Deschamps, C. et Odoux, J. *Cours de mathématiques, tome 2 : Algèbre et applications à la géométrie.* Dunod.
- Ramis, J.P. et Warusfel, A. *Mathématiques Tout-en-un pour la Licence, Niveau L2 : Cours complets avec applications et 760 exercices corrigés.* Dunod.

1 Rappels

1.1 Espaces vectoriels

1.1.1 Espaces et sous-espaces vectoriels

Définition. $(E, +, \cdot)$ *est un* $\mathbb{R}$*-espace vectoriel (e.v.) si* $(E, +)$ *est un groupe commutatif, c'est-à-dire:*

- $\forall u, v \in E, \quad u + v = v + u.$
- $\forall u, v, w \in E, \quad u + (v + w) = (u + v) + w.$
- $\exists 0_E \in E$ *("vecteur nul de E"),* $\forall u \in E, \quad u + 0_E = u.$
- $\forall u \in E, \ \exists v \in E, \quad u + v = 0_E.$

et la multiplication par un scalaire, $\lambda \cdot u \in E$ *pour* $\lambda \in \mathbb{R}$ *et* $u \in E$*, vérifie:*

- $\forall u \in E, \quad 1 \cdot u = u.$
- $\forall \lambda \in \mathbb{R}, \ \forall u, v \in E, \quad \lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v.$
- $\forall \lambda, \mu \in \mathbb{R}, \ \forall u \in E, \quad (\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u \quad \text{et} \quad \lambda \cdot (\mu \cdot u) = (\lambda\mu) \cdot u.$

Exemple.

Vecteurs du plan, de l'espace, de $\mathbb{R}^n$, de fonctions, de polynômes...

Définition. $F \subset E$ *est un sous-espace vectoriel (s.e.v.) si muni des mêmes lois,* $(F, +, \cdot)$ *reste un* $\mathbb{R}$*-espace vectoriel, ce qui s'avère vrai si et seulement si:*

$$\begin{cases} 0_E \in F \\ \text{et } \forall \lambda \in \mathbb{R}, \ \forall u, v \in F, \quad \lambda \cdot u + v \in F. \end{cases}$$

Exemple.

Droites, plans vectoriels...

Propriété. *L'intersection de deux s.e.v. est un s.e.v.*

Définition.

- *Somme de sous-espaces vectoriels.*
  Soit $F_1, \ldots, F_n$ *n s.e.v. de E, leur somme* $F_1 + \ldots + F_n$ *est l'ensemble des vecteurs* $v \in E$ *pouvant s'écrire* $v = u_1 + \ldots + u_n$*, où* $u_i \in F_i$ *pour* $i = 1, \ldots, n$*. Ce sous-ensemble de E s'avère être aussi un s.e.v. de E.*
- *Combinaison linéaire de vecteurs.*
  Soit $H \subset E$ *un sous-ensemble de vecteurs. Le vecteur* $v \in E$ *est une combinaison linéaire de vecteurs de H si il peut s'écrire*
  $$v = \sum_{i=1}^{n} \alpha_i \cdot u_i,$$
  où $n \in \mathbb{N}$*,* $\alpha_i \in \mathbb{R}$ *pour* $i = 1, \ldots, n$ *et* $u_i \in H$ *pour* $i = 1, \ldots, n$*.*
  A l'usage le $\cdot$ *est souvent omis et on écrit* $v = \sum \alpha_i u_i$*.*
- *Sous-espace engendré.*
  H étant un sous-ensemble quelconque de vecteurs, l'ensemble des combinaisons linéaires de vecteurs de H s'avère être un s.e.v. de E, noté $< H >$ *et appelé sous-espace engendré par H.*

Propriété.

Si $E_1, E_2, \ldots, E_n$ *sont n s.e.v. de E, alors* $E_1 + \ldots + E_n = < E_1 \cup \ldots \cup E_n >$*.*

Définition.

- *Somme directe de s.e.v.*
  Soit $F_1, \ldots, F_n$ *n s.e.v. de E. Si pour tout* $v \in F_1 + \ldots + F_n$*, il existe un unique n-uplet* $(u_1, \ldots, u_n) \in F_1 \times \ldots \times F_n$ *(et non seulement au moins un, comme dans la définition ci-dessus de la somme simple) tel que* $v = u_1 + \ldots + u_n$*, on dit que* $F_1, \ldots, F_n$ *sont en somme directe, ce qui se précise en remplaçant* $F_1 + \ldots + F_n$ *par* $F_1 \oplus \ldots \oplus F_n$*.*
- *Sous-espaces supplémentaires.*
  Si $F_1 \oplus \ldots \oplus F_n = E$ *alors* $F_1, .., F_n$ *sont des s.e.v. supplémentaires de E.*

Propriété. *Cas* $n = 2$*.*
Deux s.e.v. F_1 *et* F_2 *sont en somme directe si et seulement si* $F_1 \cap F_2 = \{0_E\}$*.*

1.2 Bases d'un espace vectoriel en dimension finie

Définition. *Dimension finie.*
Un e.v. E est dit de dimension finie s'il peut être engendré par un nombre fini de vecteurs, soit $E =< H >$ pour H fini. Il est alors plus pratique de décrire ces ensembles H par des familles (finies) $(u_1, \dots, u_n)$.

Définition.
- *Le s.e.v. engendré par une famille $(u_1, \dots, u_n)$ est l'ensemble des vecteurs $v \in E$ pouvant s'écrire $v = \alpha_1 u_1 + \dots + \alpha_n u_n$, où $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Il s'écrit $< u_1, \dots, u_n >$ et correspond à $< \{u_1, \dots, u_n\} >$ avec la définition de $< H >$ ci-dessus.*
- *On dit aussi que la famille est génératrice d'un s.e.v. F quand F est son s.e.v. engendré, et simplement génératrice quand elle engendre tout l'e.v. E.*
- *Une famille $(u_1, \dots, u_n)$ est dite libre si, $(\alpha_1, \dots, \alpha_n)$ étant un n-uplet de réels,*

$$\alpha_1 u_1 + \dots + \alpha_n u_n = 0_E \quad \Longrightarrow \quad \alpha_1 = \dots = \alpha_n = 0.$$

- *Une famille de vecteurs libre et génératrice est une base.*

Propriété. *Coordonnées d'un vecteur dans une base.*
Soit $(u_1, \dots, u_n)$ une base de E. Pour tout $v \in E$, il existe un unique n-uplet de réels $(\alpha_1, \dots, \alpha_n)$ tel que $v = \alpha_1 u_1 + \dots + \alpha_n u_n$. Ces $(\alpha_1, \dots, \alpha_n)$ sont les coordonnées de v dans la base $(u_1, \dots, u_n)$.

Proposition. *Dimension d'un e.v. Soit E un e.v. de dimension finie.*
- *E admet des bases et elles ont toutes le même cardinal n, qui est appelé la dimension de E: $n = \dim E$.*
- *Une famille libre de E est alors constituée au plus de n vecteurs.*
- *Une famille génératrice de E est constituée au moins de n vecteurs.*
- *Une famille de n vecteurs de E est une base de E si et seulement si elle est libre ou elle est génératrice de E.*

Tout le langage défini ici pour un e.v. E s'applique aussi à un s.e.v. puisqu'un s.e.v. est lui-même un e.v.

Propriété. *Soit E un e.v. de dimension finie.*
- *Si F est un s.e.v. de E alors $\dim F \leq \dim E$.*
- *Si F est un s.e.v. de E et $\dim F = \dim E$ alors $E = F$.*
- *Si E_1 et E_2 sont des s.e.v. de E, alors $\dim(E_1 + E_2) = \dim E_1 + \dim E_2 - \dim(E_1 \cap E_2)$.*

Définition. *Rang d'une famille de vecteurs.*
Le rang d'une famille $(u_1, \dots, u_n)$ est la dimension du sous-espace $F =< u_1, \dots, u_n >$ qu'elle engendre. On peut calculer ce rang en éliminant des vecteurs de cette famille jusqu'à ce qu'elle soit libre. En effet, si $u_n \in < u_1, \dots, u_{n-1} >$, alors $F =< u_1, \dots, u_n >=< u_1, \dots, u_{n-1} >$, et si $(u_1, \dots, u_n)$ est libre, alors son rang est n.

1.3 Applications linéaires

1.3.1 Définitions - Propriétés

Définition.
- *Application linéaire de E dans F.*
  Soit E et F deux e.v. Une application f de E dans F est dite linéaire si elle vérifie

$$\begin{cases} \forall u, v \in E, & f(u+v) = f(u) + f(v) \\ \text{et } \forall \lambda \in \mathbb{R},\ \forall u \in E, & f(\lambda u) = \lambda f(u). \end{cases}$$

  En pratique, on peut vérifier ces deux propriétés simultanément en vérifiant que $\forall \lambda \in \mathbb{R}$, $\forall u, v \in E$, $f(\lambda u + v) = \lambda f(u) + f(v)$.
- *L'ensemble des applications linéaires de E dans F est noté $\mathcal{L}(E, F)$.*
- *Si $f \in \mathcal{L}(E, E)$, on dit que f est un endomorphisme de E.*

Exemple.

Homothéties de rapport $\mu \in \mathbb{R}$: $f(u) = \mu u$.
Projecteurs: Si $E = F \oplus G$, alors la projection p sur F parallèlement à G est l'endomorphisme de E qui à tout vecteur $v \in E$ associe $p(v) = u_1$, où u_1 est déterminé ainsi: Comme la somme est directe, v s'écrit de manière unique $v = u_1 + u_2$ où $u_1 \in F$ et $u_2 \in G$.

Propriété.
- *Si $f, g \in \mathcal{L}(E, F)$ alors $f + g \in \mathcal{L}(E, F)$.*
- *Si $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$ alors $g_O f \in \mathcal{L}(E, G)$.*
- *Si $f \in \mathcal{L}(E, F)$ et f bijective alors $f^{-1} \in \mathcal{L}(F, E)$.*
- *Si $f \in \mathcal{L}(E, F)$ et A s.e.v. de E alors $f(A) = \{v \in F\ ;\ \exists u \in A, f(u) = v\}$ est un s.e.v. de F.*
- *Si $f \in \mathcal{L}(E, F)$ et B s.e.v. de F alors $f^{-1}(B) = \{u \in E\ ;\ \exists v \in B, f(u) = v\}$ est un s.e.v. de E.*

1.3.2 Noyaux et Images

Définition. *Noyau et image d'une application linéaire $f \in \mathcal{L}(E, F)$.*

- *Le noyau* $\ker f$ *est le s.e.v. (voir propriété ci-dessus)* $f^{-1}(\{0_F\})$*, c'est-à-dire* $u \in \ker f \Leftrightarrow f(u) = 0_F$.
- *L'image* Imf *est le s.e.v.* $f(E)$.

Proposition. *Pour* $f \in \mathcal{L}(E, F)$, $\quad$ *(f injective)* $\Longleftrightarrow$ *(*$\ker f = \{0_E\}$*).*

Cas de la dimension finie

Propriété. *Soit E et F deux e.v. de dimension finie. Pour $f \in \mathcal{L}(E, F)$, on a*

- $\dim(\ker f) + \dim(Imf) = \dim E$ *(mais en général, on n'a pas* $\ker f \oplus Imf = E$*).*
  La dimension $\dim(Imf)$ *est aussi appelée rang de l'application et notée* $rgf = \dim(Imf)$.
- *Si* $\dim E = \dim F$ *alors* $\quad$ *(f bijective)* $\Longleftrightarrow$ *(*$\ker f = \{0_E\}$*).*
- *Si e est une base de E, f est entièrement définie par les images des vecteurs de e par f.*
  En effet soit $e = (u_1, \dots, u_n)$, $v \in E$ *et* $(\alpha_1, \dots, \alpha_n)$ *les coordonnées de v dans e,*

$$f(v) = f(\alpha_1 u_1 + \ldots + \alpha_n u_n) = \alpha_1 f(u_1) + \ldots + \alpha_n f(u_n).$$

  Donc $f(v)$ est bien déterminé par les $f(u_1), \dots, f(u_n)$.

1.4 Matrices et déterminants

1.4.1 Définitions et propriétés des matrices complexes

Définition. • *Une matrice complexe (réelle) d'ordre (n, p) est une famille de nombres complexes (réels) indicée sous la forme* $M = (m_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}}$. *Une présentation conventionnelle de M est la suivante :*

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1p} \\ m_{21} & m_{22} & \cdots & m_{2p} \\ \vdots & \vdots & \ldots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{np} \end{pmatrix}.$$

- *L'ensemble des matrices complexes (réelles) d'ordre (n, p) est un espace vectoriel noté* $\mathcal{M}_{n,p}(\mathbb{C})$ *(resp.* $\mathcal{M}_{n,p}(\mathbb{R})$*).*

Exemple.

Donner des exemples de matrices $(2, 3)$, $(1, 5)$ et $(6, 1)$, puis déterminer des exemples de matrices extraites pour chacune de ces matrices.

Définition. • *Une matrice complexe d'ordre (n, n) est appelée matrice carrée d'ordre n.*

- *L'ensemble des matrices complexes (réelles) carrée d'ordre n est appelé* $\mathcal{M}_n(\mathbb{C})$ *(resp.* $\mathcal{M}_n(\mathbb{R})$*).*
- *Soit* $M = (m_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}} \in \mathcal{M}_{n,p}(\mathbb{C})$. *On appelle matrice extraite de M une matrice N contenant certaines lignes et certaines colonnes de M, c'est-à-dire qu'il existe* $I \subset \{1, \cdots, n\}$ *et* $J \subset \{1, \cdots, p\}$ *tels que* $= (m_{ij})_{i \in I, j \in J}$.
- *Soit* $M = (m_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}} \in \mathcal{M}_{n,p}(\mathbb{C})$. *On appelle transposée de M la matrice notée ${}^tM$ d'ordre (p, n) telle que* ${}^tM = (m'_{ij})_{\substack{1 \le i \le p \\ 1 \le j \le n}}$ *avec* $m'_{ij} = m_{ji}$ *pour* $1 \le i \le p$ *et* $1 \le j \le n$,

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1p} \\ m_{21} & m_{22} & \cdots & m_{2p} \\ \vdots & \vdots & \ldots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{np} \end{pmatrix} \quad \Longrightarrow \quad {}^tM = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ m_{p1} & m_{p2} & \cdots & m_{pn} \end{pmatrix}.$$

Exemple.

Transposer les exemples de matrices précédentes.

Définition. *Soit une matrice carrée* $M = (m_{ij})_{1 \le i,j \le n} \in \mathcal{M}_n(\mathbb{C})$.

- *On dit que M est matrice symétrique lorsque* ${}^tM = M$.

- *On dit que M est une matrice diagonale lorsque $m_{ij} = 0$ dès que $i \neq j$.*
- *On dit que M est une matrice triangulaire supérieure (resp. inférieure) lorsque $m_{ij} = 0$ pour $i < j$ (resp. $j < i$).*
- *La trace de la matrice M est la somme des termes diagonaux, soit Trace$(M) = m_{11} + \cdots + m_{nn}$.*

Exemple.

Donner des exemples de matrices symétriques, diagonales et triangulaires supérieures, et calculer leurs traces.

Définition. *On définit les opérations algébriques suivantes sur les matrices :*

- *Multiplication par une constante : pour $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{C})$ et $\lambda \in \mathbb{C}$, alors $\lambda.A = (\lambda.a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{C})$.*
- *Addition : pour $A = (a_{ij})$ et $B = (b_{ij})$ deux matrices de $\mathcal{M}_{n,p}(\mathbb{C})$, on définit $A + B = (a_{ij} + b_{ij})$, matrice de $\mathcal{M}_{n,p}(\mathbb{C})$.*
- *Multiplication : pour $A = (a_{ij}) \in \mathcal{M}_{n,p}(\mathbb{C})$ et $B = (b_{ij}) \in \mathcal{M}_{n,p}(\mathbb{C})$, on définit $A.B = \left(\displaystyle\sum_{k=1}^{p} a_{ik}b_{kj}\right)$, matrice de $\mathcal{M}_{n,q}(\mathbb{C})$.*

Exemple.

Donner deux exemples de matrices d'ordre $(2, 3)$, puis d'ordre $(2, 2)$ et enfin d'ordre $(3, 1)$ et effectuer leurs additions et multiplications lorsque cela est possible.

Remarque.

Attention, même si les matrices A et B sont carrées, on n'a pas en général, $A.B = B.A$.

Propriété.

- *Pour $A, B, C \in \mathcal{M}_{n,p}(\mathbb{C})$, alors $A + B = B + A$ et $A + B + C = A + (B + C) = (A + B) + C$.*
- *Pour $A \in \mathcal{M}_{n,p}(\mathbb{C})$, $B \in \mathcal{M}_{p,q}(\mathbb{C})$ et $C \in \mathcal{M}_{q,r}(\mathbb{C})$, alors $A.B.C = (A.B).C = A.(B.C)$.*
- *On a ${}^t(A + B) = {}^tA + {}^tB$ et ${}^t(A.B) = {}^tB.{}^tA$.*

<u>**Cas des matrices carrées**</u>

Définition.

- *On appelle matrice identité d'ordre n la matrice diagonale I_n telle que :*

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

- *Soit $M \in \mathcal{M}_n(\mathbb{C})$. S'il existe une matrice N telle que $M.N = N.M = I_n$, on dit que M est inversible et $N = M^{-1}$ est sa matrice inverse.*

Exemple.

Déterminer, si elles existent, les matrices inverses de $A = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ et $C = \begin{pmatrix} -1 & 2 \\ 2 & -4 \end{pmatrix}$.

Propriété. *Si $A \in \mathcal{M}_n(\mathbb{C})$ et $B \in \mathcal{M}_n(\mathbb{C})$ sont inversibles alors $A.B$ et $B.A$ sont inversibles et $(A.B)^{-1} = B^{-1}.A^{-1}$.*

1.4.2 Matrices complexes et applications linéaires

Définition. *Soit $u \in \mathcal{L}(E, F)$, où E et F sont deux e.v. de dimensions finies et de bases respectives $e = (e_1, .., e_p)$ et $f = (f_1, .., f_n)$. On appelle matrice de u dans les bases (e, f) la matrice $Mat(u, e, f) = (m_{ij})$ telle que $u(e_j) = \displaystyle\sum_{i=1}^{n} m_{ij} f_i$ pour tout $j \in \{1, 2, .., p\}$, ce que l'on peut encore noter :*

$$Mat(u, e, f) = \begin{array}{c|c} \begin{matrix} u(e_1) & u(e_2) & \cdots & u(e_p) \end{matrix} & \\ \hline \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1p} \\ m_{21} & m_{22} & \cdots & m_{2p} \\ \vdots & \vdots & \ldots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{np} \end{pmatrix} & \begin{matrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{matrix} \end{array}$$

Attention! La matrice de u dépend des bases choisies.

Exercice.

Soit $E = \mathbb{R}_3[X]$ l'ensemble (espace vectoriel) des polynômes réels de degré inférieur ou égal à 3. Soit u l'application qui a un polynôme P de E associe P' (dérivé de P). Montrer que $u \in \mathcal{L}(\mathbb{R}_3[X], \mathbb{R}_2[X])$. Déterminer $Mat(u, e, f)$, où $e = (1, X, X^2, X^3)$ et $f = (1, X, X^2)$, bases canoniques.

Conséquence.

Soit x un vecteur de E. On associe un vecteur colonne X composé des coordonnées de x dans e et $y = u(x)$ a pour vecteur colonne Y (coordonnées dans f) tel que $Y = A.X$.

Exercice.

Soit $u \in \mathcal{L}(\mathbb{R}^2)$ tel que $u(e_1) = 2.e_1 - 6.e_2$ et $u(e_1) = -e_1 - 3.e_2$. Calculer l'image par u du vecteur $(1, 1)$. Retrouver le même résultat matriciellement.

Propriété. *Soit $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$, où E,F et G sont des e.v. de dimensions finies et de bases respectives e, f et g. Alors $Mat(v \circ u, e, g) = Mat(v, f, g).Mat(u, e, f)$.*

Définition. *Le rang de A, matrice d'ordre (n, p) est le rang de $u \in \mathcal{L}(E, F)$ (c'est-à-dire la dimension de l'image de u) tel que $A = Mat(u, e, f)$.*

Proposition. *Soit $e = (e_1, .., e_p)$ et $e' = (e'_1, .., e'_p)$ deux bases d'un e.v. E. Alors il existe une matrice d'ordre p inversible $P = (p_{ij})_{1 \leq i,j \leq p}$ telle que $e'_j = \sum_{i=1}^{p} p_{ij} e_i$ pour tout $i \in \{1, 2, .., p\}$. P est la matrice de passage de e dans e', que l'on peut encore noter*

$$P = \begin{array}{c|c} \begin{matrix} e'_1 & e'_2 & \cdots & e'_p \end{matrix} & \\ \hline \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1p} \\ p_{21} & p_{22} & \cdots & p_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{np} \end{pmatrix} & \begin{matrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{matrix} \end{array} .$$

Conséquence.

Soit x un vecteur de E, de vecteur colonne X dans e et de vecteur colonne X' dans e'. Alors $X = P.X'$ soit encore $X' = P^{-1}.X$.

Conséquence.

Soit $u \in \mathcal{L}(E)$ et $A = Mat(u, e, e)$, $B = Mat(u, e', e')$. Alors $B = P^{-1}.A.P$.

Exercice.

Soit $E = \mathbb{R}^2$. Soit $e = (e_1, e_2)$ base canonique de E et soit $f = (f_1, f_2)$, avec $f_1 = e_1 + e_2$ et $f_2 = e_1 - e_2$.

1. Montrer que f est une base de E et donner la matrice de passage P de e dans f.
2. Soit x un vecteur de coordonnées $(2, -3)$ dans e. Déterminer ses coordonnées dans f.
3. Soit $u \in \mathcal{L}(E)$ tel que $u(e_1) = -e_1 + 3.e_2$ et $u(e_2) = 3.e_1 - e_2$. Déterminer $Mat(u, e)$.
4. Déterminer $Mat(u, f)$.

1.4.3 Déterminant d'une matrice carrée

Définition. *Le déterminant d'une matrice carrée est une somme de différents produits de coefficients de la matrice.*

Exemple.

Calcul du déterminant d'une matrice $(2, 2)$: $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$.

Propriété. *Soit A, B deux matrices carrées d'ordre n.*

- $\det({}^tM) = \det(M)$ *et* $\det(\lambda M) = \lambda^n \det(M)$.

- $\det(A.B) = \det(A).\det(B) = \det(B.A)$.
- *On ne change rien au déterminant d'une matrice si on rajoute à une ligne (colonne) une combinaison linéaire des autres lignes (colonnes).*

Attention! En général, on n'a pas $\det(A+B) = \det(A) + \det(B)$.

Exemple.

Calcul de $\det(A)$, $\det(B)$, $\det(-2.A)$, $\det(A+B)$, $\det(A.B)$ et $\det(B.A)$ pour $A = \begin{pmatrix} 2 & -1 \\ 4 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 1 \\ -3 & -2 \end{pmatrix}$.

Définition. *On appelle cofacteur de la ligne i et de la colonne j, Δ_{ij}, d'une matrice carrée $M = (m_{ij})_{1\leq i,j\leq n}$ est $(-1)^{i+j}\det(M^{(ij)})$ où $M^{(ij)}$ est la matrice extraite de M à laquelle on a retiré la ligne i et la colonne j.*

Proposition. *Le déterminant d'une matrice $M = (m_{ij})_{1\leq i,j\leq n}$ est la somme des produits 2 à 2 des éléments d'une rangée par les cofacteurs correspondants (et ce pour n'importe quelle rangée). Ainsi, avec les notations précédentes :*

$$\det(M) = \sum_{j=1}^{n} m_{i_0 j}.\Delta_{i_0 j} = \sum_{i=1}^{n} m_{i j_0}.\Delta_{i j_0} \quad \textit{pour n'importe quel ligne } i_0 \textit{ ou colonne } j_0.$$

Conséquence.

Calcul du déterminant de n'importe quelle matrice par itération de l'utilisation du développement suivant une rangée, jusqu'à arriver à des déterminants d'ordre 2.

Exemple.

Calcul de $\det(A)$ où $A = \begin{pmatrix} 2 & -1 & 5 \\ 4 & 0 & 2 \\ -3 & -1 & 2 \end{pmatrix}$.

Conséquence.

- Soit $A = \begin{pmatrix} M_1 & M_2 \\ 0 & M_3 \end{pmatrix}$, où M_1, M_2 et M_3 sont des matrices carrées. Alors $\det(A) = \det(M_1)\det(M_3)$.
- Soit A matrice diagonale ayant sur sa diagonale les complexes $\lambda_1, \lambda_2, .., \lambda_n$. Alors $\det(A) = \lambda_1\lambda_2 ... \lambda_n$.

1.4.4 Applications

Déterminant de n vecteurs:

Définition. *Le déterminant de n vecteurs $(x_1, \cdots, x_n)$ dans une base $e = (e_1, \cdots, e_n)$ d'un espace vectoriel E tel que* $\dim E = n$ *est le déterminant de la matrice des coordonnées des vecteurs dans la base e, soit :*

$$\det(x_1, \cdots, x_n) = \begin{array}{c|c} \begin{array}{cccc} x_1 & x_2 & \cdots & x_n \end{array} & \\ \hline \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_n^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_n^{(2)} \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{(n)} & x_2^{(n)} & \cdots & x_n^{(n)} \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \\ \vdots \\ e_n \end{array} \end{array}.$$

Propriété. *On considère un espace vectoriel E tel que* $\dim E = n$.

- *Le déterminant de n vecteurs de E dépend de la base de E choisie, mais s'il est non nul dans une base de E, il est non nul dans n'importe quelle autre base de E.*
- *On a $(e_1, \cdots, e_n)$ base de E* $\Longleftrightarrow \det(e_1, \cdots, e_n) \neq 0$.

Exemple.

Soit $E = \mathbb{R}^3$ et $e = (e_1, e_2, e_3)$ base de E. Soit $f = (f_1, f_2, f_3)$, avec $f_1 = e_1 + e_2 + e_3$ et $f_2 = e_1 - e_2$ et $f_3 = e_2 - e_3$. La famille f est-elle une base de E ?

Déterminant d'un endomorphisme:

Définition. *Le déterminant d'un endomorphisme de E, espace vectoriel de dimension finie, est le déterminant de $Mat(u,e)$, où e est une base quelconque de E.*

Propriété.
- *Le déterminant d'un endomorphisme ne dépend pas de la base choisie.*
- *Si $u \in \mathcal{L}(E,F)$ $u \in \mathcal{L}(F,G)$ avec E, F et G des espaces vectoriels de même dimension finie, alors*

$$\det(v \circ u) = \det(u \circ v) = \det(u).\det(v).$$

Proposition.
- *Le rang d'un endomorphisme u, c'est-à-dire la dimension de son image, est également l'ordre de la plus grande matrice extraite de $Mat(u,e)$ dont le déterminant est non nul (e base quelconque). C'est aussi le rang de la matrice $Mat(u,e)$.*
- *Si $\det(Mat(u,e)) \neq 0$ alors $Ker(u) = \{0\}$, sinon $dim(Ker(u)) \geq 1$.*

Exemple.

Déterminer le rang de la matrice $A = \begin{pmatrix} 2 & -1 & 5 \\ 4 & 0 & 2 \\ 8 & -2 & 12 \end{pmatrix}$.

<u>Inverse d'une matrice:</u>

Définition. *On appelle comatrice de la matrice carrée M d'ordre n, la matrice carrée $Com(M)$ d'ordre n telle que $Com(M) = (\Delta_{ij})_{1 \leq i,j \leq n}$, les Δ_{ij} étant les cofacteurs de la matrice M.*

Proposition. *Soit A matrice carrée d'ordre n. Alors $A.{}^t Com(A) = \det(A).I_n$.*

Cas particulier.

Si $\det(A) \neq 0$, alors A est inversible et $A^{-1} = \dfrac{1}{\det(A)}\, {}^t\mathrm{Com}(A)$.

Exemple.

Calculer l'inverse de la matrice $A = \begin{pmatrix} 2 & -1 & 5 \\ 4 & 0 & 2 \\ 5 & -1 & 7 \end{pmatrix}$, après avoir vérifier que A est bien inversible.

1.5 Diagonalisation de matrices carrées

1.5.1 Vecteurs et valeurs propres réelles

Définition. *Soit E un espace vectoriel. On considère $u \in \mathcal{L}(E)$, endomorphisme sur E.*

- *On dit que $\lambda \in \mathbb{R}$ est une valeur propre réelle de u s'il existe un vecteur $v \in E \setminus \{0\}$ appelé vecteur propre associé à λ tel que $u(v) = \lambda.v$.*
- *On appelle spectre des valeurs propres de u l'ensemble de $\mathbb{R}$ constitué par les valeurs propres de u (qui peut aussi être $\emptyset$).*

Exemple.

Soit E un espace vectoriel et soit u tel que $\forall x \in E$, $u(x) = -3x$. Déterminer les valeurs et vecteurs propres.

Propriété. *Soit E un espace vectoriel et $u \in \mathcal{L}(E)$. Si λ est une valeur propre de u, alors il existe un sous-espace vectoriel E_λ appelé sous-espace propres associé à λ, tel que $\forall x \in E_\lambda,\ u(x) = \lambda.x$.*

Exercice.

Soit E un espace vectoriel et soit E_1 et E_2 deux sous-espaces vectoriels de E tels que $E = E_1 + E_2$ et $E_\cap E_2 = \{0\}$. Soit u l'application tel que $\forall x \in E_1$, $u(x) = x$ et $\forall x \in E_2$, $u(x) = 0$. Déterminer les valeurs et sous-espaces propres de u.

Définition. *Soit E un espace vectoriel et $u \in \mathcal{L}(E)$. On dit que u est un endomorphisme diagonalisable lorsqu'il existe une base de E constituée de vecteurs propres de u.*

Propriété. *Soit $u \in \mathcal{L}(E)$ où E est un espace vectoriel. Soit $Id \in \mathcal{L}(E)$, application identité telle que $\forall x \in E$, $Id(x) = x$. Alors le sous-espace vectoriel propre associé à une valeur propre λ de u est $E_\lambda = \ker(\lambda.Id - u)$.*

1.5.2 Cas de la dimension finie et polynôme caractéristique

Définition. *Une valeur propre $\lambda \in \mathbb{R}$ et un vecteur propre $X \in \mathbb{R}^n \setminus \{0\}$ d'une matrice $M \in \mathcal{M}_n(\mathbb{R})$ vérifient l'équation matricielle $M.X = \lambda.X$.*

Exemple.

Déterminer les valeurs et sous-espaces propres de $M = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix}$.

Propriété. *Soit $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension finie. Soit M la matrice de u dans une base quelconque de E. Alors les vecteurs et valeurs propres de M sont les mêmes que ceux de u.*

Définition. *On appelle polynôme caractéristique de*

1. *$u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension n, le polynôme $\chi_u(x) = \det(u - x.Id)$, où $x \in \mathbb{R}$.*
2. *d'une matrice $M \in \mathcal{M}_n(\mathbb{R})$, le polynôme $\chi_M(x) = \det(M - x.I_n)$, où $x \in \mathbb{R}$.*

Exemple.

Déterminer le polynôme caractéristique de $M = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix}$.

Propriété. *Si $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension n de base quelconque $e = (e_1, \cdots, e_n)$, et si $M = Mat(u, e)$, alors $\chi_u(x) = \chi_M(x)$ pour tout $x \in \mathbb{R}$.*

Propriété. *λ est une valeur propre de $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension n (respectivement de $M \in \mathcal{M}_n(\mathbb{R})$) si et seulement si le polynôme caractéristique de u (respectivement de M) s'annule en λ.*

Exemple.

Vérifier la propriété pour $M = \begin{pmatrix} 2 & -3 \\ -4 & 6 \end{pmatrix}$.

Propriété. *Si $M \in \mathcal{M}_n(\mathbb{R})$, alors le polynôme caractéristique $\chi_M(x)$ est un polynôme de degré n, de coefficients $(-1)^n$ pour le degré n, $(-1)^{n-1}.Tr(M)$ pour le degré $n-1$ et $\det(M)$ pour le degré 0.*

Exemple.

Vérifier la propriété pour $M = \begin{pmatrix} 1 & 0 & -3 \\ -2 & 1 & 6 \\ -3 & -2 & 1 \end{pmatrix}$.

Conséquence.

Le polynôme caractéristique $\chi_M(x)$ admet au moins une racine et au plus n dans $\mathbb{C} \Longrightarrow M$ admet au plus n valeurs propres dans $\mathbb{R}$.

Propriété.

- *Si $A \in \mathcal{M}_n(\mathbb{R})$ alors $\chi_A = \chi_{{}^tA}$.*
- *Soit $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension n, et soit e et f deux bases de E. On note $A = Mat(u, e)$ et $B = Mat(u, f)$. Alors $\chi_A = \chi_B$.*

1.5.3 Diagonalisation en dimension finie

Propriété. *Soit $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension finie. Alors (u diagonalisable) $\Longleftrightarrow$ (Il existe e base de E telle que $Mat(u, e)$ soit une matrice diagonale) $\Longleftrightarrow$ (Il existe $P \in \mathcal{M}_n(\mathbb{R})$ inversible telle que $P^{-1}.M.P = D$ avec D une matrice diagonale)*

Propriété. *On suppose que χ_u est scindé dans $\mathbb{R}$, soit $\chi_u(\lambda) = \prod_{i=1}^{p} (\lambda - \lambda_i)^{m(\lambda_i)}$, les λ_i étant les p valeurs propres réelles de u et les $m(\lambda_i)$ leur multiplicité. Alors $1 \leq \dim(E_{\lambda_i}) \leq m(\lambda_i)$.*

Théorème. *Soit $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension finie. Alors:*
(u diagonalisable) $\Longleftrightarrow$ ($E = E_{\lambda_1} \oplus E_{\lambda_2} \oplus .. \oplus E_{\lambda_p}$, les λ_i étant les p valeurs propres réelles de u) $\Longleftrightarrow$ (χ_u est scindé dans $\mathbb{R}$ et $\dim E_{\lambda_i} = m(\lambda_i)$).

Cas particulier.

Si $A \in \mathcal{M}_n(\mathbb{R})$ admet n valeurs propres distinctes alors A est diagonalisable.

Technique de diagonalisation d'une matrice carrée :

Soit A une matrice d'ordre n à diagonaliser.

1. On calcule χ_A et on détermine les racines réelles λ_i de χ_A: ce sont les valeurs propres de A.
2. S'il y a n racines λ_i distinctes, alors A est diagonalisable. On détermine une base de vecteurs propres $(v_1, \cdots, v_n)$ en résolvant les n équations $(A - \lambda_i.I_n)X_i = 0$.
3. Sinon, pour chaque racine multiple λ_i, on détermine la dimension du sous-espace vectoriel propre E_{λ_i} associé à cette racine : cette dimension est $n - rg(A - \lambda_i.I_n)$. Si pour toute racine multiple cette dimension est égale à la multiplicité, alors A est diagonalisable (sinon, elle ne l'est pas). Pour chaque racine multiple λ_i, une base de vecteurs propres de E_{λ_i} se trouve en résolvant $(A - \lambda_i.I_n)X = 0$.
4. Si A est diagonalisable, si $v = (v_1, \cdots, v_n)$ est la base de vecteurs propres associés aux valeurs propres $(\lambda_1, \cdots, \lambda_p)$, alors $D = P^{-1}.A.P$, ou encore $A = P.D.P^{-1}$ avec P et D les matrices suivantes :

$$P = \begin{array}{c|c} \begin{array}{cccc} v_1 & v_2 & \cdots & v_n \end{array} & \\ \hline \begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{22} & \cdots & v_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{pmatrix} & \begin{array}{c} e_1 \\ e_2 \\ \vdots \\ e_p \end{array} \end{array} \quad \text{et} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}$$

Exemple.

Diagonaliser la matrice $M = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$.

1.5.4 Applications de la diagonalisation

Puissances de matrices

Définition. *On appelle puissance n-ème d'une matrice carrée M, le produit $M^n = M \times M \times \cdots \times M$, n fois.*

Propriété. *Si $u \in \mathcal{L}(E)$ où E est un espace vectoriel de dimension finie de base e, et $M = Mat(u, e)$ alors $M^n = Mat(u \circ u \circ \cdots \circ u, e)$, où l'endomorphisme u est composé n fois.*

Conséquence.

Si M est une matrice diagonalisable, alors $M^n = P.D^n.P^{-1}$, où D est la matrice diagonale composée des valeurs propres de M, P la matrice de passage dans la base des vecteurs propres et ainsi,

$$D^n = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \lambda_p^n \end{pmatrix}$$

Exemple.

Calculer M^n pour $M = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 0 \\ -2 & -2 & -1 \end{pmatrix}$.

Résolution de systèmes linéaires

1. Systèmes homogènes

   Soit le système S_0 : $\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & .. & + & a_{1p}x_p & = 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & .. & + & a_{2p}x_p & = 0 \\ : & : & : & : & & : & : & : \\ a_{n1}x_1 & + & a_{n2}x_2 & + & .. & + & a_{np}x_p & = 0 \end{array} \right.$

   C'est un système de n équations et p inconnues. On a:

$S_0 \Longleftrightarrow A.X = 0$ avec $\left\{ \begin{array}{l} A = (a_{ij}) \;\; \text{matrice d'ordre } (n,p). \\ X \;\; \text{vecteur colonne des } x_i. \end{array} \right.$

L'ensemble solution de S_0 est donc un sous-espace vectoriel $\mathcal{S}_0 = \{X, A.X = 0\}$. Si u application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$ de matrice A dans une base quelconque, alors $\mathcal{S}_0 = \ker(u) \Longrightarrow \dim \mathcal{S}_0 = p - rg(A)$ et $rg(A) = r$ est déterminé par l'ordre d'une plus grande matrice carrée A_r extraite de A de déterminant non nul.
Les équations et variables utilisées dans A_r sont appelées principales (les autres sont appelées non-principales et $\mathcal{S}_0$ s'exprime en écrivant les r variables principales en fonction des $p - r$ variables non-principales.

2. Systèmes avec second membre

Soit le système S : $\left\{ \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1p}x_p & = b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2p}x_p & = b_2 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{np}x_p & = b_n \end{array} \right.$

C'est un système de n équations et p inconnues. On a:

$S \Longleftrightarrow A.X = B$ avec $\left\{ \begin{array}{l} A = (a_{ij}) \;\; \text{matrice d'ordre } (n,p). \\ X \;\; \text{vecteur colonne des } x_i. \\ B \;\; \text{vecteur colonne des } b_i \end{array} \right.$

(a) Si $n = p$ et $\det(A) \neq 0$: Système de Cramer.
Le système S admet une unique solution $\mathcal{S} = X_0$, où $X_0 = A^{-1}.B$ est un vecteur de $\mathbb{R}^p$.

(b) Si $rg(A) = r < p$ alors ou $\mathcal{S} = \emptyset$, ou $\mathcal{S}$ s'écrit à l'aide des $p - r$ variables non-principales.

Exemple.

Résoudre les systèmes $(S_1) : \left\{ \begin{array}{ccccc} x & + & y & = & 0 \\ -x & + & y & = & 3 \\ x & + & y & = & -2 \end{array} \right.$, $(S_2) : \left\{ \begin{array}{ccccccc} x & + & y & + & z & = & 0 \\ -x & + & y & + & z & = & 3 \end{array} \right.$ et

$(S_3) : \left\{ \begin{array}{ccccccc} x & + & y & + & z & = & 0 \\ -x & + & y & + & z & = & 3 \\ x & + & y & - & z & = & -2 \end{array} \right.$.

Résolution de systèmes homogènes de suites numériques récurrentes linéaires

1. Systèmes homogènes

Soit le système (S_H) : $\left\{ \begin{array}{ccccccccc} u_{n+1} = & a_{11}u_n & + & a_{12}v_n & + & \cdots & + & a_{1p}z_n \\ v_{n+1} = & a_{21}u_n & + & a_{22}v_n & + & \cdots & + & a_{2p}z_n \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ z_{n+1} = & a_{p1}u_n & + & a_{p2}v_n & + & \cdots & + & a_{pp}z_n \end{array} \right.$

C'est un système de p équations et p inconnues. On a:

$$(S_H) \iff \text{Equation matricielle } \; X_{n+1} = A.X_n \; \text{ avec } \left\{ \begin{array}{l} A = (a_{ij}) \;\; \text{matrice d'ordre } (p,p). \\ X_n = {}^t(u_n, v_n, \cdots, z_n). \end{array} \right.$$

Proposition. *L'ensemble solution de (S_H) est un sous-espace vectoriel $\mathcal{H} = \{A^n.X_0, X_0 \in \mathbb{R}^p\}$ de dimension p.*

Conséquence.

Si A est diagonalisable (ou si l'on sait calculer A^n), alors on connaît les expressions de u_n, v_n,..., z_n en fonction de n et de u_0, v_0 et z_0.

Exemple.

Déterminer l'expression générale des suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ définies par :

$$\left\{ \begin{array}{ccc} u_{n+1} & = & 3u_n + v_n \\ v_{n+1} & = & -2u_n - v_n \end{array} \right. \quad \text{avec } \; u_0 = 1, \; v_0 = 0.$$

2. Cas des suites récurrentes linéaires

On suppose que (u_n) est une suite réelle suivant l'équation de récurrence $u_{n+p} + a_{p-1}u_{n+p-1} + \cdots + a_0.u_n = 0$, où $(a_0,\cdots,a_{p-1})\in\mathbb{R}^p$, et avec $(u_0,\cdots,u_{p-1})\in\mathbb{R}^p$. On peut alors écrire que (u_n) vérifie le système de suites récurrentes suivant :

$$\left\{\begin{array}{rcccccccccc} u_{n+p} = & -a_{p-1}u_{n+p-1} & - & a_{p-2}u_{n+p-2} & - & \cdots & - & a_1u_{n+1} & - & a_0u_n \\ u_{n+p-1} = & u_{n+p-1} & + & 0 & + & \cdots & + & 0 & + & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ u_{n+1} = & 0 & + & 0 & + & \cdots & + & u_{n+1} & + & 0 \end{array}\right.$$

Ce système peut encore s'écrire, pour tout $n\in\mathbb{N}$:

$$X_{n+1}=A.X_n \quad \text{avec} \quad X_n=\begin{pmatrix} u_{n+p-1} \\ u_{n+p-2} \\ \vdots \\ u_n \end{pmatrix} \quad \text{et} \quad A=\begin{pmatrix} -a_{p-1} & -a_{p-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

On peut montrer que A est une matrice diagonalisable dans $\mathbb{C}$, dont le polynôme caractéristique est l'équation caractéristique associée à la suite (u_n) : on retrouve les résultats déjà obtenus.

Exemple.

En utilisant la résolution d'un système matriciel, déterminer l'expression générale de la suite $(u_n)_{n\in\mathbb{N}}$ telle que $4.u_{n+2}=5.u_{n+1}-u_n$ avec $u_0=0$ et $u_1=1$.

2 Produits scalaires et applications

On s'intéresse ici à certaines propriétés pouvant être vérifiées par des espaces vectoriels. En particulier on revoit de manière algébrique la notion d'orthogonalité. Dans toute la suite, on supposera que les espaces vectoriels sont des $\mathbb{R}$-e.v. c'est-à-dire des espaces vectoriels pour lesquels les scalaires sont des réels (et non des complexes par exemple).

2.1 Le cas de $\mathbb{R}^n$

Ici on suppose que $E = \mathbb{R}^n$. C'est un $\mathbb{R}$-e.v. de dimension n dont on connaît une base canonique: $e = (e_1, \cdots, e_n)$ avec $e_1 = (1, 0, \cdots, 1)$, $e_2 = (0, 1, 0, \cdots, 1)$, ..., $e_n = (0, 0, \cdots, 0, 1)$ (Attention! ce ne sont pas des coordonnées mais les valeurs des n-uplets qui ne dépendent d'aucune base; ces n-uplets correspondent cependant aux coordonnées dans le cas de la base canonique de $\mathbb{R}^n$). On peut alors définir entre deux vecteurs $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$ ce que l'on appelle le produit scalaire "canonique" sur $\mathbb{R}^n$:

$$< x, y >:= \sum_{j=1}^{n} x_j\, y_j = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

On peut vérifier sur des exemples dans $\mathbb{R}^2$ que:

- Si x et y sont deux vecteurs orthogonaux (au sens géométrique du terme) alors $< x, y >= 0$ et réciproquement;
- La norme habituelle d'un vecteur x est $\sqrt{< x, x >}$;
- Pour deux vecteurs, $< x, y >= \sqrt{< x, x >}\sqrt{< y, y >}\cos(\theta)$ où θ est l'angle (pas forcément orienté) entre les deux vecteurs;
- Pour deux vecteurs x et y, si $y^\perp$ est le projeté orthogonal de y sur x, alors $< x, y >$ est l'aire (algébrique) du parallélogramme défini par y, $y^\perp$ et $y + y^\perp$ et on a $| < x, y > | = \sqrt{< x, x >}\sqrt{< y^\perp, y^\perp >}$.

Ces liens entre produit scalaire canonique et géométrie se généralisent sur $\mathbb{R}^n$. Mais on peut aussi généraliser la notion de produit scalaire sur $\mathbb{R}^n$ et sur d'autres $\mathbb{R}$-e.v. ce qui permettra de définir un outil très puissant avec de nombreuses conséquences.

2.2 Produit scalaire

Définition. *Soit E un espace vectoriel. On dit que l'application telle que $\forall (x, y) \in E^2 \mapsto < x, y >\in \mathbb{R}$ est un produit scalaire si:*

1. $\forall (x, y_1, y_2) \in E^3,\ \forall (\lambda_1, \lambda_2) \in E^2,\ \ < x, \lambda_1.y_1 + \lambda_2.y_2 >= \lambda_1 < x, y_1 > + \lambda_2 < x, y_2 >$;
2. $\forall (x, y) \in E^2$, $< x, y >=< y, x >$;
3. $\forall x \in E$, $< x, x >\geq 0$;
4. $\forall x \in E$, $< x, x >= 0 \iff x = 0$.

Lorsque l'on peut définir un produit scalaire sur E, on dit que $(E, < ., . >)$ est un espace préhilbertien. Lorsque E est de dimension finie, alors on dira que $(E, < ., . >)$ est un espace euclidien.

Exemple.

- sur $\mathbb{R}^n$: pour $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$ deux vecteurs de $\mathbb{R}^n$, le produit scalaire canonique est $< x, y >= \sum_{j=1}^{n} x_j\, y_j$, ou un autre produit scalaire: $< x, y >= \sum_{j=1}^{n} p_j\, x_j\, y_j$ avec $p_j > 0$ pour $i = 1, \cdots, n$ (à vérifier!).
- sur $\ell^2 := \left\{(u_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}},\ \sum_{n=0}^{\infty} u_n^2 < \infty\right\}$, on peut définir le produit scalaire $< u, v >:= \sum_{n=0}^{\infty} u_n\, v_n$ pour $u = (u_n)_{n\in\mathbb{N}} \in \ell^2$ et $v = (v_n)_{n\in\mathbb{N}} \in \ell^2$ (justification un peu plus bas...).

- sur $\mathcal{C}^0([a,b])$, ensemble des fonctions réelles continues sur $[a,b]$, on peut définir le produit scalaire $< f, g >:= \int_a^b f(x)\, g(x) dx$ pour $f,\, g \in \mathcal{C}^0([a,b])$ (montrer que c'est bien un produit scalaire!).
- sur $\mathcal{C}^{0pm}([-\pi,\pi])$, ensemble des fonctions réelles continues sur $[-\pi,\pi]$, on peut définir le produit scalaire $< f, g >:= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, g(x) dx$ pour $f,\, g \in \mathcal{C}^{0pm}([-\pi,\pi])$.

Théorème (Théorème de Cauchy-Schwartz)**.** *Soit $(E, < .,. >)$ un espace préhilbertien. Alors :*

1. $\forall (x,y) \in E^2$, $| < x, y > | \leq \sqrt{< x, x >}\sqrt{< y, y >}$;
2. $\forall (x,y) \in E^2$, $\big(| < x, y > | = \sqrt{< x, x >}\sqrt{< y, y >}\big) \Longleftrightarrow$ *(x et y sont liés).*

Proof. Soit $(x,y) \in E^2$. Pour $\lambda \in \mathbb{R}$, par définition du produit scalaire, $< x + \lambda y, x + \lambda y > \geq 0$. Mais $< x + \lambda y, x + \lambda y > = < x, x > +2\lambda < x, y > + \lambda^2 < y, y >$., polynôme du second degré en λ. Pour qu'un tel polynôme soit positif quelque soient x et y, il faut que son discriminant soit négatif ou nul. Ce qui revient à écrire que $4 < x, y >^2 - 4 < x, x >< y, y > \leq 0$, d'où l'inégalité de Cauchy-Schwartz.
Pour montrer le point 2., on reprend le développement précédent qui conduit à écrire que $< x + \lambda y, x + \lambda y > = 0$. D'après la propriété 4. du produit scalaire, cela induit que $x + \lambda y = 0$, donc x et y sont liés. □

Conséquence.

Soit $(E, < .,. >)$ un espace préhilbertien. Alors $\forall (x,y) \in E^2$, $\sqrt{< x + y, x + y >} \leq \sqrt{< x, x >} + \sqrt{< y, y >}$ et s'il y a égalité alors x et y sont liés.

Proof. Soit $(x,y) \in E^2$. 0n a $< x + y, x + y > = < x, x > +2 < x, y > + < y, y >$, donc d'après l'Inégalité de Cauchy-Schwartz, $< x + y, x + y > \leq < x, x > +2\sqrt{< x, x >< y, y >} + < y, y > \leq (\sqrt{< x, x >} + \sqrt{< y, y >})^2$. En prenant la racine de cette inégalité (ceci est possible car tout est positif), on obtient bien l'inégalité triangulaire voulue. De plus, s'il y a égalité, cela signifie qu'il y a égalité dans l'Inégalité de Cauchy-Schwartz, donc x et y sont liés. □

Définition. *Si $(E, < .,. >)$ est un espace préhilbertien, alors l'application $x \in E \mapsto ||x|| = \sqrt{< x, x >} \in \mathbb{R}_+$ est une norme sur E.*

Propriété. *Soit $(E, < .,. >)$ un espace préhilbertien. Alors $\forall (x,y) \in E^2$,*

1. $\|x + y\| \leq \|x\| + \|y\|$ *(inégalité triangulaire) et* $| < x, y > | \leq \|x\|\, \|y\|$.
2. $\|x + y\|^2 = < x + y, x + y > = \|x\|^2 + \|y\|^2 + 2 < x, y >$, *donc* $< x, y > = \frac{1}{2}\big(\|x + y\|^2 - \|x\|^2 - \|y\|^2\big)$.
3. $| < x, y > | \leq \frac{1}{2}\big(\|x\|^2 + \|y\|^2\big)$.

Remarque.

Attention! le point 2. de la propriété précédente permet de définir un produit scalaire de manière unique à partir d'une norme. Mais toute norme n'est pas issue d'un produit scalaire (en fait c'est même plutôt l'inverse en général). Par exemple, pour $x = (x_1, \cdots, x_n) \in \mathbb{R}^n$, $\|x\|_\infty = \max_{1 \leq i \leq n}(|x_i|)$ est une norme mais elle ne définit pas un produit scalaire dès la dimension 2 (à montrer...).

Exercice.

Soit $(u_n)_{n \in \mathbb{N}} \in \ell^2$. On pose $U_n = \frac{1}{n}\sum_{p=0}^n u_p$. Montrer que $(U_n)_{n \in \mathbb{N}} \in \ell^2$.

2.3 Orthogonalité

Définition. *Soit $(E, < .,. >)$ un espace préhilbertien. Si pour $(x,y) \in E^2$, on a $< x, y > = 0$ alors on dit que x et y sont orthogonaux.*

Théorème (Théorème de Pythagore)**.** *Soit $(E, < .,. >)$ un espace préhilbertien. Alors :*

$$(||x + y||^2 = ||x||^2 + ||y||^2) \iff (x \text{ est orthogonal à } y).$$

Proof. Soit $(x,y) \in E^2$. 0n a $< x + y, x + y > = < x, x > +2 < x, y > + < y, y >$, donc $< x + y, x + y > = < x, x > + < y, y >$ si et seulement si $< x, y > = 0$. □

Définition. *Soit $(E, < .,. >)$ un espace préhilbertien et soit $A \subset E$. On appelle ensemble orthogonal de A dans E l'ensemble $A^\perp = \{x \in E,\ \forall a \in A,\ < x, a > = 0\}$ et $A^\perp$ est un sous-espace vectoriel de E.*

Proof. Il faut montrer que $A^\perp$ est bien un s.e.v. de E. Il est bien clair que $A^\perp$ est inclus dans E. Par ailleurs $0 \in A^\perp$, car $< 0, x >= 0$ pour tout $x \in E$. Enfin, si $x, y \in A^\perp$, et si $\lambda \in \mathbb{R}$, alors $x + \lambda y \in A^\perp$, car pour tout $a \in A$, $< x + \lambda y, a >=< x, a > +\lambda < y, a >= 0$ puisque $< x, a >= 0$ et $< y, a >= 0$ du fait que $x, y \in A^\perp$. $\square$

Exemple.

Soit $E = \mathbb{R}^n$ et $< .,. >$ le produit scalaire canonique sur E. Soit $a = (a_1, \cdots, a_n) \in E$. Déterminer $a^\perp = \{a\}^\perp$.

Propriété. *Soit $(E, < .,. >)$ un espace préhilbertien et soit $A \subset E$. Alors :*

1. $E^\perp = \{0\}$*;*
2. $A \cap A^\perp \subset \{0\}$*;*
3. *si $A \subset B \subset E$, alors $B^\perp \subset A^\perp$;*
4. $A \subset (A^\perp)^\perp$*.*

Proof. 1. Si $a \in E^\perp$, alors $\forall x \in E$, $< x, a >= 0$. En particulier, pour $x = a$, on obtient $< a, a >= 0$, ce qui n'est possible que si $a = 0$. Dans l'autre sens, on vérifie aisément que $< 0, x >= 0$ pour tout $x \in E$.
2. Si $a \in A \cap A^\perp$, alors $a \in A$. Mais on a aussi $a \in A^\perp$, ce qui signifie que $\forall x \in A$, $< x, a >= 0$, donc en particulier pour $x = a$, ce qui signifie que $a = 0$. Si $0 \notin A$, $A \cap A^\perp = \emptyset$.
3. Si $x \in B^\perp$, alors $\forall y \in B$, $< x, y >= 0$. Mais comme $A \subset B$, on a donc $\forall y \in A$, $< x, y >= 0$, ce qui signifie que $x \in A^\perp$.
4. Si $a \in A$, alors $\forall y \in A^\perp$, tel que $< a, y >= 0$. Mais pour tout $x \in (A^\perp)^\perp$, alors $\forall y \in A^\perp$, $< x, y >= 0$. Donc forcément $a \in (A^\perp)^\perp$. $\square$

Définition. *Soit $(E, < .,. >)$ un espace préhilbertien. On dit qu'une famille $(e_1, \cdots, e_n) \in E^n$ est une famille orthogonale de E si $e_i \neq 0$ pour tout $i = 1, \cdots, n$ et pour tout $i \neq j$, $< e_i, e_j >= 0$. De plus, si pour tout i, $||e_i|| = 1$, alors on dit que $(e_1, \cdots, e_n)$ est une famille orthonormale de E.*

Exemple.

1. Soit $E = \mathbb{R}^3$ et $< .,. >$ le produit scalaire canonique sur E. Déterminer une base orthonormale $e' = (e'_1, e'_2, e'_3)$ de E qui ne soit pas la base canonique $e = (e_1, e_2, e_3)$ (on donnera les coordonnées de e' dans e, ainsi que la matrice de passage P).
2. Exemple important: soit $e = (e_k)_{k \in \mathbb{N}}$ telle que $e_0 = 1$ et pour $k \in \mathbb{N}^*$, $e_{2k} = \sqrt{2}\cos(kx)$ et $e_{2k-1} = \sqrt{2}\sin(kx)$. Alors cette famille est orthonormale dans $\mathcal{C}^{0pm}_{2\pi}([-\pi, \pi])$ muni du produit scalaire $< f, g >= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)g(t)dt$.

Propriété. *Soit $(E, < .,. >)$ un espace préhilbertien.*

1. *Une famille orthogonale $(e_1, \cdots, e_n) \in E^n$ est une famille libre de E;*
2. *Si E est de dimension n, alors une famille orthogonale (respectivement orthonormale) $(e_1, \cdots, e_n)$ de E est une base dite orthogonale (resp. orthonormale) de E.*

Proof. 1. Soit $(e_1, \cdots, e_n) \in E^n$ une famille orthogonale. Alors $\forall (\lambda_1, \cdots, \lambda_n) \in \mathbb{R}^n$ tel que $\lambda_1 e_1 + \cdots + \lambda_n e_n = 0$, on a $< e_i, \lambda_1 e_1 + \cdots + \lambda_n e_n >= \lambda_i \|e_i\|^2 = 0$ donc $\lambda_i = 0$ pour tout $i = 1, \cdots, n$ car $\|e_i\| \neq 0$. La famille est bien libre.
2. C'est clair car dans un espace de dimension n toute famille libre de n vecteurs est une base. $\square$

Théorème (Procédé d'othogonalisation de Gram-Schmidt)**.** *Soit $(E, < .,. >)$ un espace euclidien (donc de dimension finie). Il existe une base orthonormale de E.*

Proof. On sait qu'il existe une base $(e_1, \cdots, e_n) \in E^n$ de E. On peut donc déjà choisir $u_1 = e_1/\|e_1\|$ comme premier vecteur. Ensuite $e_2 - < u_1, e_2 > u_1$ est orthogonal à u_1, mais n'est pas un vecteur nul (car e_2 et u_1 ne sont pas liés). Ainsi $u_2 = (e_2 - < u_1, e_2 > u_1)/\|e_2 - < u_1, e_2 > u_1\|$ est orthogonal à u_1 et de norme 1. On itère le procédé: par exemple pour u_3, on choisit $u_3 = (e_3 - < u_1, e_3 > u_1 - < u_2, e_3 > u_2)/\|e_3 - < u_1, e_3 > u_1 - < u_2, e_3 > u_2\|$... $\square$

Extension.

Le procédé de Gram-Schmidt s'étend aux espaces préhilbertiens de dimension dénombrable pour permettre de définir une famille orthonormale à partir d'une famille libre.

Propriété. *Soit $(E, <.,.>)$ un espace euclidien de dimension n et soit $(e_1, \cdots, e_n)$ une base orthonormale de E. Alors pour tout $x \in E$, $x = \sum_{j=1}^n < x, e_j > e_j$. De plus si $(x_1, \cdots, x_n)$ est une famille de vecteurs de E, alors la matrice $(< x_i, x_j >)_{1 \le i,j \le n}$ est appelée Matrice de Gram et*

$$det\big((< x_i, x_j >)_{1 \le i,j \le n}\big) = \big| det_e(x_1, \cdots, x_n)\big|^2.$$

Proof. On vérifie aisément que $x = \sum_{j=1}^n < x, e_j > e_j$ car x est décomposable de façon unique dans $(e_1, \cdots, e_n)$, donc $x = \sum_{j=1}^n \alpha_j e_j$ et donc $< x, e_i >= \alpha_i$ pour tout i car $(e_1, \cdots, e_n)$ est orthonormale. Comme pour tout $i = 1, \cdots, n$, $x_i = \sum_{j=1}^n < x_i, e_j > e_j$, on a $< x_i, x_j >= \sum_{k=1}^n \sum_{k'=1}^n < x_i, e_k >< x_j, e_{k'} >< e_k, e_{k'} >= \sum_{k=1}^n < x_i, e_k >< x_j, e_k >$. Mais $\big| \det_e(x_1, \cdots, x_n)\big|^2 = \det_e\big[{}^t(x_1, \cdots, x_n)(x_1, \cdots, x_n)\big] = \det_e\Big[\Big(\sum_{k=1}^n < x_i, e_k >< x_j, e_k >\Big)_{ij}\Big]$ car $< x_i, e_k >$ est la $k$ème coordonnée de x_i dans e. □

Exercice.

Soit $M = (m_{pq})_{1 \le p,q \le n}$ la matrice carrée telle que $m_{pq} = (e^{p+q} - 1)/(p+q)$. Montrer que $\det(M) > 0$.

Propriété. *Soit $(E, <.,.>)$ un espace euclidien (donc de dimension finie) et soit A un s.e.v. de E. Alors:*

- *$A \oplus A^\perp = E$;*
- *$dim A + dim A^\perp = dim E$;*
- *$A = (A^\perp)^\perp$.*

Proof. 1. En premier lieu, comme vu précédemment, $A \cap A^\perp = \{0\}$. De plus, si $(e_1, \cdots, e_p)$ est une base orthonormale de A (que l'on sait construire), alors il est aussi possible de trouver $(e_{p+1}, \cdots, e_n)$ telle que $(e_1, \cdots, e_n)$ soit une base de E (que l'on suppose de dimension n). Mais $e_i \in A^\perp$ pour $p+1 \le i \le n$ car $< e_i, e_j >= 0$ pour $1 \le j \le p$. Par suite, tout x peut s'écrire dans $(e_1, \cdots, e_n)$ donc comme la somme d'un vecteur de A et d'un vecteur de $A^\perp$. D'où $E = A + A^\perp$.
2. Cela se déduit immédiatement de la preuve précédente.
3. On sait déjà que $A \subset (A^\perp)^\perp$, mais comme $\dim A^\perp = n - \dim A$, on en déduit que $\dim (A^\perp)^\perp = n - \dim A^\perp = \dim A$. Or deux s.e.v. de E de même dimension et tels que l'un soit inclus dans l'autre sont égaux. □

Remarque.

Le premier point de la remarque précédente n'est pas nécessairement vrai en dimension non finie: l'orthogonal n'est pas forcément un supplémentaire. Ainsi si on considère $E = \mathbb{R}[X]$ l'ensemble des polynômes à coefficients réels et si on prend pour produit scalaire le produit scalaire tel que pour $P(X) = \sum_{k=0}^p a_k X^k$ et $Q(X) = \sum_{k=0}^q b_k X^k$, $< P, Q >:= \sum_{k=0}^{\max(p,q)} a_k b_k$ alors on peut montrer que l'hyperplan (pourquoi?) H engendré par $(1+X, 1+X^2, \cdots, 1+X^n, \cdots)$ est tel que $H^\perp = \{0\}$ (pourquoi?).

Propriété. *Soit $(E, <.,.>)$ un espace préhilbertien et soit F un sous-espace vectoriel de E. Pour $a \in E$, on définit, s'il existe, le vecteur $a' \in F$ tel que:*

$$\|a - a'\| := \inf_{x \in F}\{\|a - x\|\} = \inf_{x \in F}\{d(a, x)\} = d(a, F).$$

Si a' existe, alors $a - a' \in F^\perp$, a' est unique et on dira que a' est le projeté orthogonal de a sur F, noté $p_F(a)$.

Proof. Si a' existe, alors pour tout $x \in F$ et $\lambda \in \mathbb{R}$, $\|a - a' + \lambda x\|^2 = \|a - a'\|^2 + \|\lambda x\|^2 + 2 < a - a', \lambda x >$. Mais d'après la définition de a', on a $\|a - a' + \lambda x\|^2 \ge \|a - a'\|^2$. Par suite, $0 \le \|\lambda x\|^2 + 2 < a - a', \lambda x >$ soit pour tout $\lambda > 0$, $|\lambda| \|x\| \ge 2| < a' - a, x > |$. Par passage à la limite lorsque $\lambda \to 0^+$, on a donc $2| < a' - a, x > | = 0$. Ainsi $a - a' \in F^\perp$.
Montrons que a' est unique. Soit $a'' \in F \setminus \{a'\}$ vérifiant $\|a - a'\| = \|a - a''\|$. Mais $\|a - a''\|^2 = \|a - a' + a' - a''\|^2 = \|a - a'\|^2 + \|a' - a''\|^2$, le produit scalaire $< a - a', a' - a'' >= 0$ car $a - a' \in F^\perp$ et $a' - a'' \in F$. D'où la contradiction, car alors $\|a' - a''\| = 0$. □

Théorème. *Soit $(E, <.,.>)$ un espace préhilbertien et soit F un sous-espace vectoriel de E tel que $E = F \oplus F^\perp$ (toujours vérifié si E est euclidien de dimension finie ou si F est lui-même de dimension finie). Alors pour tout $a \in E$ il existe un unique projeté orthogonal de a sur F, qui vérifie $a = p_F(a) + p_{F^\perp}(a)$.*

Proof. Soit $(e_1, \cdots, e_p)$ une base orthonormale de F. On peut alors construire $(e_{p+1}, \cdot)$ une famille orthonormale de vecteurs telle que $(e_i)_i$ soit une base orthonormale de E. D'où, comme dans la preuve précédente, $(e_{p+1}, \cdot)$ est une base orthonormale de $F^\perp$ et $E = F \oplus F^\perp$. Donc pour tout $a \in E$, a s'écrit de manière unique $a = a_F + a_{F^\perp}$ où $a_F \in F$ et $a_{F^\perp} \in F^\perp$. Mais $a - a_F$ est la projection orthogonale sur $F^\perp$ car $a - (a - a_F) \in (F^\perp)^\perp = F$. On en déduit bien que $a_F = p_F(a)$ et $a_{F^\perp} = p_{F^\perp}(a)$. □

Propriété. *Soit $(E, <.,.>)$ un espace préhilbertien et soit F un sous-espace vectoriel de E de dimension finie tel que $(e_1, \cdots, e_p)$ base orthonormale de F. Alors pour tout $a \in E$, $p_F(a) = \sum_{k=1}^{p} < a, e_k > e_k$ et $p_{F^\perp}(a) = a - \sum_{k=1}^{p} < a, e_k > e_k$.*

Proof. On sait tout d'abord que comme $p_F(a) \in F$ alors $p_F(a) = \sum_{i=1}^{p} a_i e_i$. Pour que $p_F(a)$ soit bien le projeté orthogonal de a sur F, alors $a - p_F(a) \in F^\perp$, ce qui est équivalent au fait que $< a - p_F(a), e_i >= 0$ pour tout $i = 1, \cdots, p$. En conséquence comme la base $(e_1, \cdots, e_p)$ est orthonormale, $< a, i > -a_i < e_i, e_i >= 0$ pour tout $i = 1, \cdots, p$, donc $a_i =< a, e_i >$ et ainsi $p_F(a) = \sum_{i=1}^{p} < a, e_i > e_i$. □

Exercice.

Soit $E = \mathbb{R}^n$ muni du produit scalaire canonique, $X = (x_1, \cdots, x_n) \in E$ et $Y = (y_1, \cdots, y_n) \in E$. Déterminer $a_0 \in \mathbb{R}$ tel que $a_0 = \inf_{a \in \mathbb{R}} \{\sum_{i=1}^{n} (y_i - a)^2\}$, puis $(a_1, b_1) \in \mathbb{R}^2$ tels que $(a_1, b_1) = \inf_{(a,b) \in \mathbb{R}^2} \{\sum_{i=1}^{n} (y_i - (a \cdot x_i + b))^2\}$.

Propriété. *Soit $(E, <.,.>)$ un espace préhilbertien et soit F un sous-espace vectoriel de E de dimension finie. Alors l'application $x \in E \mapsto p_F(x) \in F$ est une application linéaire et vérifie $\|p_F(x)\| \le \|x\|$: elle est donc continue.*

Proof. Du fait de l'écriture de la projection orthogonale de p_F dans une base orthonormale de F, on en déduit que p_F est une application linéaire (du fait de linéarité du produit scalaire). DE plus, on a $\|x\|^2 = \|p_F(x)\|^2 + \|p_{F^\perp}(x)\|^2$ d'après Pythagore, d'où $\|p_F(x)\| \le \|x\|$. □

3 Une application en analyse: les séries de Fourier

3.1 Introduction

L'origine des séries de Fourier se trouve à la fin du XVIIIème siècle et au début du XIXème siècle avec la résolution de problèmes de physique: les oscillations d'une corde (étudié notamment par Bernoulli) et la conduction de la chaleur (étudié notamment par Fourier). Dans les deux cas, des polynômes trigonométriques permettent de trouver des solutions à ces équations.
En termes mathématiques, les séries de Fourier donnent un exemple intéressant d'approximation convergente d'une fonction suffisamment régulière sur un compact par une suite de polynômes trigonométriques. Nous nous intéresserons plutôt ici à une vision algébrique des séries de Fourier, permettant d'obtenir une base orthonormale dans un espace de dimension infinie. Ici nous nous placerons de le cadre des fonctions à valeurs réelles, mais il est possible de tout écrire à l'aide d'exponentielles complexes pour des fonctions à valeurs complexes.

3.2 Coefficients de Fourier

Définition. *Soit $f \in \mathcal{C}^{0pm}([-\pi, \pi])$, ensemble des fonctions réelles continues par morceaux sur $[-\pi, \pi]$. On peut définir:*

- *Les coefficients de Fourier réels de f: $a_n(f) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) dt$ et $b_n(f) := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) dt$ pour $n \in \mathbb{N}$.*
- *Le polynôme trigonométrique d'ordre n associé à f: $S_f^{(n)}(x) := \frac{a_0(f)}{2} + \sum_{k=1}^{n} (a_k(f) \cos(kx) + b_k(f) \sin(kx))$.*

Exercice.

Calculer les coefficients de Fourier de f telle que $f(x) = x$ sur $[-\pi, \pi]$.

Propriété. *Pour $f \in \mathcal{C}^{0pm}([-\pi, \pi])$,*

- *si f est paire, on a $b_n(f) = 0$ et $a_n(f) = \frac{2}{\pi} \int_0^{\pi} f(t) \cos(nt) dt$.*
- *si f est impaire, on a $a_n(f) = 0$ et $b_n(f) = \frac{2}{\pi} \int_0^{\pi} f(t) \sin(nt) dt$.*

Proof. La preuve est immédiate. Par exemple, si f est paire avec le changement de variable $t' = -t$,

$$\begin{aligned} a_n(f) &:= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)dt = \frac{1}{\pi}\int_{-\pi}^{0} f(t)\cos(nt)dt + \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)dt \\ &= -\frac{1}{\pi}\int_{\pi}^{0} f(-t')\cos(-nt')dt' + \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)dt = \frac{1}{\pi}\int_{0}^{\pi} f(t')\cos(nt')dt' + \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)dt \\ &= \frac{2}{\pi}\int_{0}^{\pi} f(t)\cos(nt)dt. \end{aligned}$$

□

Proposition. *Si $f \in \mathcal{C}^0([-\pi,\pi])$ et $< f_1, f_2 >= \frac{1}{\pi}\int_{-\pi}^{\pi} f_1(t)f_2(t)dt$, alors $S_f^{(n)}$ est la projection orthogonale de f sur $\mathcal{P}_n$, le sous-espace vectoriel de $\mathcal{C}^0([-\pi,\pi])$ engendré par $(e_k)_{0\le k\le 2n}$, où $e = (e_k)_{k\in\mathbb{N}}$ est telle que $e_0 = 1/\sqrt{2}$ et pour $k \in \mathbb{N}^*$, $e_{2k} = \cos(kx)$ et $e_{2k-1} = \sin(kx)$, et ainsi $\forall P \in \mathcal{P}_n$, $\|f - S_f^{(n)}\| \le \|f - P\|$.*

Proof. On montre d'abord que la famille $e = (e_k)_{k\in\mathbb{N}}$ est une famille orthonormale pour le produit scalaire (pour ce faire, on utilise des formules trigonométriques comme $2\cos(p)\cos(q) = \cos(p+q)\cos(p-q)$ ou bien $2\sin^2 x = 1 - \cos(2x)$). Ensuite, on utilise la formule de la projection sur une base orthonormale: $P_{\mathcal{P}_n}(f) = \sum_{j=0}^{2n} < f, e_j > e_j$. Enfin, par définition de la projection orthogonale, $\forall P \in \mathcal{P}_n$, $\|f - S_f^{(n)}\| \le \|f - P\|$. □

Exercice.

Calculer les coefficients de Fourier de $S_f^{(n)}$.

Proposition. *Si $f \in \mathcal{C}^{0pm}([-\pi,\pi])$, alors les séries $\sum |a_k(f)|^2$ et $\sum |b_k(f)|^2$ sont convergentes.*

Proof. On a $\|S_f^{(n)}\|^2 = \|\sum_{j=0}^{2n} < f, e_j > e_j\|^2 \sum_{j=0}^{2n} < f, e_j >^2 = a_0^2(f)/2 + \sum_{j=1}^{n} a_n^2(f) + b_n^2(f)$. Mais pour tout $n \in N$, grâce à Pythagore, $\|P_{\mathcal{P}_n}(f)\|^2 \le \|f\|^2 < \infty$ car f continue par morceaux sur $[-\pi,\pi]$. Donc les séries $\sum a_n^2(f)$ et $\sum b_n^2(f)$ convergent. □

Proposition (Théorème de Riemann-Lebesgue). *Si $f \in \mathcal{C}^{0pm}([-\pi,\pi])$, alors $a_n(f) \underset{n\to+\infty}{\longrightarrow} 0$ et $b_n(f) \underset{n\to+\infty}{\longrightarrow} 0$.*

Proof. Comme f est continue par morceaux, alors $\sum a_n^2(f) + b_n^2(f) < \infty$ d'où le résultat. □

Théorème (Théorème de Bessel). *Si $f \in \mathcal{C}^0([-\pi,\pi])$ alors:*

$$\frac{1}{2}|a_0(f)|^2 + \sum_{k=1}^{+\infty}(|a_k(f)|^2 + |b_k(f)|^2) = \frac{1}{\pi}\int_{-\pi}^{\pi}|f(t)|^2 dt.$$

Proof. Il est clair d'après la proposition précédente que $\|p_n(f)\|^2 \le \|f\|^2$ et que ceci est vrai également quand $n \to \infty$. Pour montrer l'égalité, il nous faut utiliser un résultat annexe, l'approximation uniforme de toute fonction continue par un polynôme trigonométrique (on connaît mieux un tel résultat pour les polynômes; c'est le Théorème de Stone-Weierstrass). Donc pour tout f continue sur $[-\pi,\pi]$, il existe une suite $(Q_n(f))_n$ telle que pour $\lim_{n\to\infty}\|f - Q_n(f)\|_\infty = \lim_{n\to\infty}\sup_{x\in[-\pi,\pi]}|f - Q_n(f)| = 0$. Donc pour $\varepsilon > 0$, il existe $N_0 \in \mathbb{N}$, tel que pour tout $n \ge N_0$, $\|f - Q_n(f)\|_\infty \le \varepsilon$. Mais $\|g\|_\infty \ge \|g\|$ (avec la norme associée au produit scalaire). Or $\|f - p_n(f)\| \le \|f - Q_n(f)\|$ car $p_n(f)$ est la projection orthogonale de f sur $\mathcal{P}_n$ et $Q_n(f) \in \mathcal{P}_n$. On en déduit donc que $\|f - p_n(f)\| \le \varepsilon$: il y a bien convergence de $p_n(f)$ vers f pour la norme associée au produit scalaire (convergence quadratique). □

Exercice.

Appliquer ce résultat à la fonction f telle que $f(x) = x$ sur $[-\pi,\pi]$. En déduire la valeur de $\sum_{k=1}^{\infty}\frac{1}{k^2}$.

Théorème (Théorème de Parceval). *Si f, $g \in \mathcal{C}^0([-\pi,\pi])$ alors:*

$$\frac{1}{2}a_0(f)a_0(g) + \sum_{k=1}^{+\infty}(a_k(f)a_k(g) + b_k(f)b_k(g)) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)g(t)dt =< f, g > .$$

Proof. On utilise le fait que $xy = \frac{1}{2}(x^2 + y^2 - (x-y)^2)$. On applique le Théorème de Bessel ci-dessus pour les fonctions f, g et $f - g$ et comme on a une convergence absolue on peut sommer les séries. □

3.3 Convergence des séries de Fourier

Définition. *Pour $f \in \mathcal{C}^{0pm}([-\pi,\pi])$,*

- *La série de Fourier associée à f est $S_f(x) := \dfrac{a_0(f)}{2} + \displaystyle\sum_{k=1}^{\infty}(a_k(f)\cos(kx) + b_k(f)\sin(kx))$ (Attention! cette série n'existe pas forcément (même si f est continue), mais on peut montrer qu'elle existe presque partout).*
- *Une fonction f sur $[-\pi,\pi]$ est dite développable en série de Fourier si $f(x) = S_f(x)$ pour tout $x \in [-\pi,\pi]$.*

Théorème (Théorème de Dirichlet). *Si $f \in \mathcal{C}^{1pm}([-\pi,\pi])$ (donc de classe $\mathcal{C}^1$ par morceaux sur $[-\pi,\pi]$) alors:*

1. *La série $S_f(x)$ de Fourier de f converge sur $[-\pi,\pi]$.*
2. *$\forall x \in \mathbb{R}$, on a $S_f(x) = \dfrac{f(x^+) + f(x^-)}{2}$, où $f(x^-)$ (resp. $f(x^+)$) désigne la limite à gauche (resp. à droite) de f en x.*

Proof. Trop difficile à montrer en quelques lignes! □

Exercice.

Appliquer ce résultat à la fonction f telle que $f(x) = x$ sur $[-\pi,\pi]$. En déduire la valeur de $\sum_{k=1}^{\infty} \frac{\sin k}{k}$.

Théorème. *Si $f \in \mathcal{C}^{1pm}([-\pi,\pi])$ et continue sur $[-\pi,\pi]$ alors:*

1. *Les séries $\sum |a_k(f)|$ et $\sum |b_k(f)|$ sont convergentes.*
2. *La série $S_f(x)$ de Fourier de f converge normalement (donc uniformément) vers f sur $[-\pi,\pi]$, c'est-à-dire que pour tout $x \in [-\pi,\pi]$,*

$$f(x) = \frac{a_0(f)}{2} + \sum_{k=1}^{\infty}(a_k(f)\cos(kx) + b_k(f)\sin(kx)).$$

Proof. a/ Par IPP, on a $|a_n(f')| = |b_n(f)|/n$ et $|b_n(f')| = |a_n(f)|/n$. Ainsi, $\sum_{k=1}^{N}(|a_k(f)| + |b_k(f)|) = \sum_{k=1}^{N}(|a_k(f')| + |b_k(f')|)/k$. Mais d'après Cauchy-Schwartz, $\sum_{k=1}^{N}(|a_k(f')|+|b_k(f')|)/k \leq \left(\sum_{k=1}^{N} 1/k^2\right)^{1/2}\left(2\sum_{k=1}^{N} |a_k(f')|^2+|b_k(f')|^2)\right)^{1/2}$. Ainsi en utilisant le Théorème de Bessel pour f' et le fait que $\sum 1/k^2$ converge, on a bien montré que $\sum_{k=1}^{\infty}(|a_k(f)| + |b_k(f)|)$ converge.
b/ C'est juste un cas particulier du Théorème de Dirichlet lorsque f est continue (et alors $f(x) = (f(x^+) + f(x^-))/2$). □

Exercice.

Appliquer ce résultat à la fonction f telle que $f(x) = \max(0, \sin x)$ sur $[-\pi,\pi]$.

Conséquence.

Si on appelle E l'ensemble des fonctions continues et de classe $\mathcal{C}^1$ par morceaux sur $[-\pi,\pi]$, alors la famille $e = (e_k)_{k\in\mathbb{N}}$ telle que $e_0 = 1$ et pour $k \in \mathbb{N}^*$, $e_{2k} = \sqrt{2}\cos(kx)$ et $e_{2k-1} = \sqrt{2}\sin(kx)$ est une sorte de base orthonormale de E (la notion de base s'entend ici au sens hilbertien, voir en L3...). On a bien alors $f = \sum_{k\in\mathbb{N}} < f, e_k > e_k$.

Toute cette étude est également aisément transposable à un intervalle $[a,b]$ (au lieu de $[-\pi,\pi]$) en remplaçant $\cos(kx)$ et $\sin(kx)$ par $\cos(\frac{2\pi}{b-a}\,kx)$ et $\sin(\frac{2\pi}{b-a}\,kx)$

Remarque.

La convergence des séries de Fourier est restée longtemps mystérieuse. Kolmogorov (mathématicien russe) avait montré en 1925 que l'on peut construire une fonction intégrable dont la série de Fourier diverge presque partout. Il a fallu attendre Carlesson (mathématicien suédois) en 1966 pour que soit montrée la convergence (presque partout) des séries de Fourier pour des fonctions continues (et plus généralement pour des fonctions de carrés intégrables).

4 Formes linéaires et espace dual

Dans toute la suite on note E un $\mathbb{R}$-espace vectoriel.

4.1 Premières définitions et propriétés

Définition. *On appelle forme linéaire une application linéaire de E dans $\mathbb{R}$. L'ensemble de toutes les formes linéaires de E dans $\mathbb{R}$, encore noté $E^* = \mathcal{L}(E, \mathbb{R})$ est appelé l'espace dual de E.*

Exemple.

1. Dans le cas où $E = \mathbb{R}$ et $E = \mathbb{R}^2$, déterminer explicitement E^*.
2. Si $E = \mathcal{C}^0([a, b])$ alors l'application $f \in E \mapsto \int_a^b f(t)dt$ est une forme linéaire.

Définition. *Un hyperplan de E est un sous-espace vectoriel H de E admettant un supplémentaire de dimension 1, c'est-à-dire que $E = H \oplus D$ avec $dim D = 1$.*

Exemple.

Dans le cas où $E = \mathbb{R}^n$ muni du produit scalaire canonique, montrer que pour tout $a \in E$ non nul alors $a^\perp$ est un hyperplan de E. Réciproquement, montrer que tout hyperplan de E est de la forme $a^\perp$, où $a \in E$ est non nul.

Propriété. *(H hyperplan de E) $\Longleftrightarrow$ (H noyau d'une forme linéaire non nulle de E^*).*

Proof. $\Longrightarrow$ Soit $D = \mathrm{V}ect(a)$. Alors pour tout $x \in E$, il exsite un unique couple $(y, \lambda) \in H \times \mathbb{R}$ tel que $x = y + \lambda\, a$. On pose $\phi(x) = \lambda$. Alors ϕ est bien une forme linéaire. De plus $\ker \phi = H$.
$\Longleftarrow$ On suppose donc que $H = \ker f$, où f est une forme linéaire. Soit $a \notin H$. On note $D = \mathrm{V}ect(a)$. Pour tout $x \in E$, $x = x - f(x)a/f(a) + f(x)a/f(a)$. Mais $f\big(x - f(x)a/f(a)\big) = 0$ donc $x - f(x)a/f(a) \in H$. De plus $f(x)a/f(a) \in D$. Donc $E = H + D$. De plus par définition de D, $H \cap D = \{0\}$. D'où $E = H \oplus D$. □

Remarque.

L'équation d'un hyperplan H est donc celle issue du fait que H est le noyau d'une forme linéaire.

Propriété. *Deux formes linéaires non nulles ayant le même noyau sont proportionnelles.*

Proof. Soit ϕ_1 et ϕ_2 deux formes linéaires ayant le même noyau. Soit $a \notin \ker \phi_1$. Alors comme dans la preuve précédente, pour tout $x \in E$, $x = x - \phi_1(x)a/\phi_1(a) + \phi_1(x)a/\phi_1(a)$. On sait que $x - \phi_1(x)a/\phi_1(a) \in \ker \phi_1$ donc on a également $x - \phi_1(x)a/\phi_1(a) \in \ker \phi_2$. Aussi pour tout $x \in E$, $\phi_2(x) = 0 + \phi_2\big(\phi_1(x)a/\phi_1(a)\big) = \big(\phi_2(a)/\phi_1(a)\big)\phi_1(x)$: les deux formes linéaires sont bien proportionnelles. □

4.2 Notion d'orthogonalité pour les formes linéaires

Définition. *Pour $x \in E$ et $f \in E^*$, on note $< f^*, x >:= f^*(x)$ (Attention ce n'est pas un produit scalaire!). On dit que $x \in E$ et $u^* \in E^*$ sont orthogonaux si $< f^*, x >= f^*(x) = 0$ (donc si x est dans l'hyperplan noyau de f^*).*
Plus généralement, pour tout $A \subset E$ et $A^* \subset E^*$ (pas nécessairement des s.e.v.), on dit que A et A^* sont orthogonaux si $\forall (x, f^*) \in A \times A^*$, $< f^*, x >= f^*(x) = 0$.

Propriété. *Pour tout $A \subset E$ et $A^* \subset E^*$ (pas nécessairement des s.e.v.),*

$$A^\perp := \big\{f^* \in E^*, \forall x \in A, < f^*, x >= 0\big\} \quad \text{est un s.e.v. de } E^*;$$
$$A^{*\perp} := \big\{x \in E, \forall f^* \in A^*, < f^*, x >= 0\big\} \quad \text{est un s.e.v. de } E.$$

Proof. A partir de maintenant, toutes les démonstrations peuvent être calquées sur celles du chapitre précédent concernant le s.e.v. orthogonal. □

Propriété. *Pour tout A, $B \subset E$ et A^*, $B^* \subset E^*$ (pas nécessairement des s.e.v.),*

- *si $A \subset B$, alors $B^\perp \subset A^\perp$, et si $A^* \subset B^*$, alors $B^{*\perp} \subset A^{*\perp}$.*

- $A \subset (A^{\perp})^{\perp}$ *et* $A^* \subset (A^{*\perp})^{\perp}$.

Exemple.

Déterminer $E^{\perp}$ et $E^{*\perp}$.

Propriété. *On suppose ici que E est de dimension finie n.*

1. $dimE^* = n$.
2. *Pour $e = (e_1, \cdots, e_n)$ base de E, si pour tout $i, j = 1, \cdots, n$, $f_j^*(e_i) = \delta_{ij}$, alors $(f_1^*, \cdots, f_n^*)$ est une base de E^* appelée base duale de e.*
3. *Pour F et F^* s.e.v. respectivement de E et E^*, $dimF^{\perp} = dimE - dimF$ et $dimF^{*\perp} = dimE^* - dimF^*$.*
4. *Pour tout $A \subset E$ et $A^* \subset E^*$ (pas nécessairement des s.e.v.), $(A^{\perp})^{\perp} = Vect(A)$ et $(A^{*\perp})^{\perp} = Vect(A^*)$.*
5. *Soit F un s.e.v. de E de dimension $1 \leq p < n$. Alors il existe $n-p$ formes linéaires $f_1^*, \cdots, f_{n-p}^*$ de E^* telles que*
$$F = \bigcap_{i=1}^{n-p} \ker f_i^*.$$

Proof. 1/ il est clair que $E^* = \mathcal{L}(E, \mathbb{R})$. Or en dimension finie, $\dim \mathcal{L}(E,F) = \dim E \, dimF$ donc ici on a bien $\dim E^* = n$.
2/ on voit aisément que si $\sum_{i=1}^n \lambda_i f_i^* = 0$ alors pour tout $j = 1, \cdots, n$, $\sum_{i=1}^n \lambda_i f_i^*(e_j) = 0$, ce qui entraîne que pour tout $j = 1, \cdots, n$, $\lambda_j = 0$. Ce sont donc bien n vecteurs libres et comme la dimension de E^* est n, cela forme une base de E^*.
3/ Si on suppose que $m = \dim F$, $(e_1, \cdots, e_m)$ base de F, alors $F^{\perp} = \{f^*, \ f^*(e_j) = 0 \text{ pour tout } j \in \{1, \cdots, m\}\}$. Donc F est le noyau de toute forme linéaire de $F^{\perp}$ et $F^{\perp}$ est un ensemble de formes linéaires dont l'image est de dimension $\dim E - \dim F$, donc $\dim F^{\perp} = \dim E - \dim F$.
4/ Il est clair d'après 3/ que $\dim (A^{\perp})^{\perp} = \mathrm{Vect}(A)$, et on sait que $A \subset (A^{\perp})^{\perp}$. Donc $(A^{\perp})^{\perp} = \mathrm{Vect}(A)$.
5/ Soit $(e_1, \cdots, e_p)$ base de F, $(e_{p+1}, \cdots, e_n)$ d'autres vecteurs telles que $(e_1, \cdots, e_n)$ base de E. On peut alors noter $(f_1^*, \cdots, f_n^*)$ la base duale de $(e_1, \cdots, e_n)$. Alors pour tout $x \in F$, il existe un unique p-uplet $(\lambda_1, \cdots, \lambda_p)$ telle que $x = \sum_{i=1}^p \lambda_i e_i$. Mais il est clair que pour tout $j = p+1, \cdots, n$, $f_j^*(x) = 0$ car $f_j^*(e_i) = 0$ pour tout $i = 1, \cdots, p$. En conséquence, $x \in \ker f_j^*$. Comme cela est vrai pour tout $j = p+1, \cdots, n$, on a pour tout $x \in F$, $x \in \bigcap_{i=p+1}^n \ker f_i^*$, d'où $F \subset \bigcap_{i=p+1}^n \ker f_i^*$. Mais comme les f_i^* forme une base duale, $\dim \bigcap_{i=p+1}^n \ker f_i^* = p$, donc on a bien $F = \bigcap_{i=p+1}^n \ker f_i^*$. □

Conséquence.

Une conséquence de la dernière propriété est le fait que tout s.e.v. d'un e.v. de dimension finie peut s'écrire comme un système de $n - p$ équations (par exemple, on savait déjà qu'en dimension 3, un plan vectoriel est représenté par une équation et une droite vectorielle par 2 équations...).

Propriété. *Si $(E, < \cdot, \cdot >)$ est un espace préhilbertien, alors pour tout $x \in E$, $\phi_x : y \in E \mapsto < x, y >$ est une forme linéaire de E^*. De plus, l'application linéaire $u : x \in E \mapsto \phi_x \in E^*$ est injective, donc E est isomorphe à $u(E)$, s.e.v. de E^* (en dimension finie, $u(E) = E^*$).*

Proof. Soit $x \in E$. Alors d'après les propriétés du produit scalaire, $\phi_x(0) = 0$ et pour tout $(\lambda_1, \lambda_2) \in \mathbb{R}^2$, pour tout $(y_1, y_2) \in E^2$, $\phi_x(\lambda_1 y_1 + \lambda_2 y_2) = < x, \lambda_1 y_1 + \lambda_2 y_2 > = \lambda_1 < x, y_1 > + \lambda_2 < x, y_2 > = \lambda_1 \phi_x(y_1) + \lambda_2 \phi_x(y_2)$. ϕ_x est bien une forme linéaire.
Pour montrer que u est injective, il suffit de montrer que $\ker u = \{0_{E^*}\}$. Mais $\ker u = \{x \in E, \phi_x = 0\}$. Mais si $\phi_x = 0$ alors nécessairement $< x, y > = 0$ pour tout $y \in E$, donc en particulier pour $y = x$ et on en déduit que $x = 0$. Ainsi u est bien injective.
On a $\dim E = \dim u(E) + \dim(\ker u)$, donc $\dim E = \dim u(E)$: d'où le fait que E est isomorphe à $u(E)$. □

Conséquence.

Une conséquence de la dernière propriété est le fait que dans un espace préhilbertien de dimension finie (et même pour des espaces de dimension infinis appelés espaces de Hilbert) toute forme linéaire s'écrit de manière unique comme un produit scalaire par rapport à un vecteur: c'est le théorème de représentation de Riesz...

5 Application linéaire adjointe et réduction des matrices symétriques et orthogonales

5.1 Adjoint d'une application linéaire

Propriété. *Si $(E, < \cdot, \cdot >)$ est un espace euclidien (de dimension finie) et u une application linéaire de E dans E, alors il existe une unique application linéaire notée u^* et appelée application linéaire adjointe de u telle que*

$$< u(x), y >=< x, u^*(y) > \quad \text{pour tout } x,\, y \in E. \tag{1}$$

De plus, si e est une base orthonormale de E, alors $M_e(u^*) = {}^tM_e(u)$.

Proof. On se place dans une base orthonormale $(e_1, \cdots, e_n)$ de E. Alors $u(e_i) = \sum_{k=1}^n u_{ki} e_k$ et pour tout endomorphisme v de E, $v(e_i) = \sum_{k=1}^n v_{ki} e_k$. Donc si $x = \sum_{k=1}^n x_k e_k$ et $y = \sum_{k=1}^n y_k e_k$, alors $< u(x), y >= \sum_{j=1}^n y_j \big(\sum_{i=1}^n x_i u_{ji} \big)$ et $< x, v(y) >= \sum_{i=1}^n x_i \big(\sum_{j=1}^n y_i v_{ij} \big)$. Par suite, $< u(x), y >=< x, v(y) > \Longleftrightarrow \sum_{i,j} x_i y_j (u_{ji} - v_{ij}) = 0$. Il suffit donc de considérer v tel que $v_{ij} = u_{ji}$ pour que pour tout x et y, $< u(x), y >=< x, v(y) >$. De plus $v = u^*$ est unique. $\square$

Définition. *Plus généralement, si E est préhilbertien, on appellera aussi u^* application linéaire adjointe de u si elle existe une solution au problème (1).*

Remarque.

Pour un espace de Hilbert, c'est-à-dire un espace préhilbertien complet (toutes les suites de Cauchy sont convergentes), alors il existe toujours un adjoint à un endomorphisme (ceci se montre grâce au théorème de représentation de Riesz).

Propriété. *Si $(E, < \cdot, \cdot >)$ est un espace préhilbertien, si u et v sont des applications linéaires de E dans E admettant des adjoints u^* et v^*, alors:*

- $(u^*)^* = u$ *et* $(u_0 v)^* = v_0^* u^*$;
- *si u est un isomorphisme, alors u^* en est également un et $(u^*)^{-1} = (u^{-1})*$.*

Proof. • On a pour tout $(x, y) \in E^2$, $< u(x), y >=< x, u^*(y) >$. Par définition, on a $< x, u^*(y) >=< (u^*)^*(x), y >$. Donc $< u(x), y >=< (u^*)^*(x), y >$, soit pour tout $x, y \in E$, $< u(x) - (u^*)^*(x), y >= 0$. Donc pour tout $x \in E$, $u(x) - (u^*)^*(x)$ est orthogonal à E, donc $u(x) - (u^*)^*(x) = \{0\}$ et ainsi $u(x) = (u^*)^*(x)$.
De même, pour tout $(x, y) \in E^2$, $< (u_0 v)(x), y >=< x, (u_0 v)^*(y) >$. Mais $< (u_0 v)(x), y >=< u(v(x)), y >=< v(x), u^*(y) >=< x, v^*(u^*(y)) >$ par définition de u^* et v^*. Donc $< x, (u_0 v)^*(y) >=< x, v^*(u^*(y)) >$ pour tout $x, y \in E^2$ et on conclut comme précédemment.
• On a $\ker(u) = \{0\}$. Soit $y \in \ker(u^*)$. Alors $< u(x), y >=< x, u^*(y) >= 0$ pour tout $x \in E$. Donc y est orthogonal à $\mathrm{Im}(u) = E$ car u est un isomorphisme, donc $y = 0$. $\square$

Définition. *Si $u = u^*$ on dit que u est symétrique (on dit parfois aussi auto-adjoint). Ceci se dit également en dimension finie pour une matrice carrée: M est symétrique si $M = {}^tM$.*

5.2 Application linéaire orthogonale dans un espace euclidien

Ici nous nous contenterons de travailler dans un espaces euclidien E de dimension n.

Définition. *Soit $(E, < \cdot, \cdot >)$ un espace euclidien. On dit qu'une application linéaire u de E dans E est orthogonale (on dit aussi que u est une isométrie vectorielle) si*

$$\|u(x)\| = \|x\| \quad \text{pour tout } x \in E. \tag{2}$$

Propriété. *Si u est orthogonale alors u est un isomorphisme.*

Proof. On a $u(x) = 0 \Longleftrightarrow \|u(x)\| = 0$ d'après une des propriétés de base vérifiées par les normes. Or, si u est orthogonale, alors $\|u(x)\| = 0 = \|x\|$, donc également $x = 0$. Ainsi $\ker u = \{0\}$: u est bien un endomorphisme injectif donc un isomorphisme. $\square$

Exemple.

Symétrie orthogonale: si F est un s.e.v. de E tel que $E = F \oplus F^\perp$, on définit la symétrie orthogonale sur F par $s_F(x) = x$ pour tout $x \in F$ et $s_F(x) = -x$ pour $x \in F^\perp$. On peut montrer que $s_F = 2p_F - Id$.

Proposition. *Pour une application linéaire u de E dans E, les propriétés suivantes sont équivalentes:*

1. *u est orthogonale.*
2. *Pour tout $(x,y) \in E^2$, $< u(x), u(y) >=< x, y >$.*
3. *$u_0^* u = u_0 u^* = Id_E$.*
4. *u transforme toute base orthonormale en une autre base orthonormale.*

Proof. 1.$\Longrightarrow$2. Pour $x,y \in E$, $\|u(x+y)\|^2 = \|u(x)+u(y)\|^2 = \|u(x)\|^2 + \|u(y)\|^2 + 2 < u(x), u(y) >= \|x\|^2 + \|y\|^2 + 2 < u(x), u(y) >$ car u est linéaire et est une isométrie. Par ailleurs, $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2 < x, y >$. Mais on a également $\|u(x+y)\|^2 = \|x+y\|^2$ car u est une isométrie, donc $< u(x), u(y) >=< x, y >$.
2.$\Longleftrightarrow$3. Par définition u^* est telle que $< x, u(y) >=< y, u^*(x) >$. On peut appliquer ceci à $u(x)$ plutôt qu'à x et on a $< u(x), u(y) >=< y, u^*(u(x)) >$. Mais d'après 2., $< u(x), u(y) >=< x, y >$ donc pour tout $x, y \in E$, $< y, u^*(u(x)) >=< x, y >$, soit $< y, u^*(u(x)) - x >= 0$. On a vu un peu plus haut que cela induit que pour tout $x \in E$, $u^*(u(x)) - x = 0$, donc on a bien $u_0^* u = u_0 u^* = Id_E$.
2.$\Longrightarrow$4. Pour $(e_i)_{i\in I}$ une base orthonormale de E, alors $< u(e_i), u(e_j) >=< e_i, e_j >$, donc si $i \neq j$, $< u(e_i), u(e_j) >= 0$ et si $i = j$, $< u(e_i), u(e_i) >= 1$: la famille $(u(e_i))_{i\in I}$ est bien une base orthonormale.
4.$\Longrightarrow$1. Si $(u(e_i))_{i\in I}$ est une base orthonormale quand $(e_i)_{i\in I}$ est une base orthonormale de E, alors pour $x \in E$, $x = \sum_{i\in I} \alpha_i e_i$ avec $(\alpha_i)_{i\in I}$ une famille de réels, donc $u(x) = \sum_{i\in I} \alpha_i u(e_i)$ car u est linéaire et
$\|u(x)\|^2 = \left\langle \sum_{i\in I} \alpha_i u(e_i), \sum_{i\in I} \alpha_i u(e_i) \right\rangle = \sum_{i,j\in I} \alpha_i \alpha_j < u(e_i), u(e_j) >= \sum_{i\in I} \alpha_i^2$ car $(u(e_i))_{i\in I}$ est une base orthonormale. De la même manière, $\|x\|^2 = \left\langle \sum_{i\in I} \alpha_i e_i, \sum_{i\in I} \alpha_i e_i \right\rangle = \sum_{i,j\in I} \alpha_i \alpha_j < e_i, e_j >= \sum_{i\in I} \alpha_i^2$. Donc on a bien $\|u(x)\| = \|x\|$ pour tout $x \in E$: u est orthogonale. □

Définition. *On appelle matrice orthogonale de $\mathcal{M}_n(\mathbb{R})$ une matrice P telle que $P^{-1} = {}^tP$. Une matrice orthogonale est celle d'une application linéaire orthogonale dans une base orthonormale.*

Propriété. *Soit $(E, < .,. >)$ un espace euclidien de dimension n et de base orthonormale e. Soit e' une autre base orthonormale de E. Alors si P est la matrice de passage de e dans e', on a $P^{-1} = {}^tP$ et $|\det P| = 1$ (car ${}^tP\,P = I_n$): P est une matrice orthogonale.*

Proof. D'après les propriétés précédentes, on peut toujours écrire que P est la matrice dans e d'un endomorphisme orthogonal (car e' est une autre base orthonormale). Soit u^* l'endomorphisme adjoint de u. Alors on sait que sa matrice est ${}^tP$ (voir plus haut). De plus comme u est orthogonal, $u^*\,u = u\,u^* = Id$ donc en se plaçant dans la base e, alors ${}^tP\,P = P\,{}^tP = I_n$: on en déduit que $P^{-1} = {}^tP$. De plus, comme $\det {}^tP\,P = \det({}^tP)\det(P) = \det^2(P) = \det(I_n) = 1$, on a bien $|\det P| = 1$. □

Exercice.

Déterminer l'ensemble des matrices orthogonales en dimension 2 (on distinguera deux cas, suivant le signe du déterminant).

5.3 Diagonalisation des matrices symétriques

Dans tout ce qui suit nous donnerons les résultats pour les matrices symétriques, mais ils sont également valables pour les applications linéaires symétriques.

Propriété. *Soit $M \in \mathcal{M}_n(\mathbb{R})$, matrice symétrique. Alors M n'admet que des valeurs propres réelles et le polynôme caractéristique de M est scindé dans $\mathbb{R}[X]$ (c'est-à-dire que $\chi_M(\lambda) = \prod_{i=1}^{p} (\lambda_i - \lambda)^{m_i}$, où les λ_i sont les valeurs propres, les m_i étant leur multiplicité).*

Proof. On a $M = {}^tM$ car M est symétrique. Si λ est une valeur propre de M, alors il existe X un vecteur colonne non nul tel que $M\,X = \lambda X$. D'où ${}^t\overline{M}\,M\,X = M\,M\,X = \lambda^2 X$ car M est réelle, et ainsi ${}^t\overline{X}\,{}^t\overline{M}\,M\,X = \lambda^2\,{}^t\overline{X}\,X$, soit ${}^t(\overline{M\,X})\,M\,X = \lambda^2\,{}^t\overline{X}\,X$. Or pour $Y = (y_1, \cdots, y_n) \in \mathbb{C}^n$ un vecteur colonne alors ${}^t\overline{Y}\,Y = \sum_{i=1}^n |y_i|^2 \in \mathbb{R}_+$, d'où $\lambda^2 = {}^t(\overline{M\,X})\,(M\,X) / {}^t\overline{X}\,X \in \mathbb{R}_+$ (car ${}^t\overline{X}\,X > 0$). Donc $\lambda \in \mathbb{R}$. De ceci on en déduit également que χ_M est bien scindé dans $\mathbb{R}$ car si ce n'était pas le cas, cela signifierait qu'il existe des valeurs propres complexes non réelles. □

Propriété. *Soit $M \in \mathcal{M}_n(\mathbb{R})$, matrice symétrique. Alors si λ_1 et λ_2 sont deux valeurs propres distinctes de M, leurs sous-espaces propres associés sont orthogonaux dans $\mathbb{R}^n$.*

Proof. Soit $x_1 \in \ker\big(M - \lambda_1\,I_n\big)$ et $x_2 \in \ker\big(M - \lambda_2\,I_n\big)$ deux vecteurs non nuls. Soit également u tel que M soit la matrice de u dans une base de E. Alors $u(x_1) = \lambda_1 x_1$ et $u(x_2) = \lambda_2 x_2$. Mais comme $u = u^*$ alors $< u(x_1), x_2 >= \lambda_1 < x_1, x_2 >=< x_1, u(x_2) >= \lambda_2 < x_1, x_2 >$. Du fait que $\lambda_1 \neq \lambda_2$, on en déduit que $< x_1, x_2 >= 0$, donc x_1 et x_2 sont orthogonaux. □

Théorème. *Soit $M \in \mathcal{M}_n(\mathbb{R})$, matrice symétrique. Alors il existe une base orthonormale de E constituée de vecteurs propres de M, c'est-à-dire que*

$$E = \ker\big(M - \lambda_1 I_n\big) \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} \ker\big(M - \lambda_p I_n\big)$$

Ainsi, il existe une matrice P orthogonale telle que $D = {}^tP.M.P$, où D est la matrice diagonale constituée des valeurs propres de M sur la diagonale.

Proof. On sait déjà (voir le cours général sur la diagonalisation) que $E = \ker\big(M - \lambda_1 I_n\big)^{m_1} \oplus \cdots \oplus \ker\big(M - \lambda_p I_n\big)^{m_p}$ où les m_i sont les différentes multiplicités des λ_i. Si les m_i sont égaux à 1, il n'y a pas de problème. Mais si par exemple $m_i = 2$ alors pour $X \in \ker\big(M - \lambda_i I_n\big)^2$ non nul, alors $\big(M - \lambda_i I_n\big)^2 X = 0$, soit encore ${}^t\big(M - \lambda_i I_n\big)\big(M - \lambda_i I_n\big)X = 0$ du fait que M et I_n soit symétrique, et en multipliant par ${}^tX$ à droite on obtient que ${}^tX{}^t\big(M - \lambda_i I_n\big)\big(M - \lambda_i I_n\big)X = 0$, d'où $\|\big(M - \lambda_i I_n\big)X\|^2 = 0$ (car pour tout vecteur colonne réel Y de taille n, ${}^tY\,Y = \|Y\|^2$ la norme canonique sur $\mathbb{R}^n$) et donc $\big(M - \lambda_i I_n\big)X = 0$: ainsi $X \in \ker\big(M - \lambda_i I_n\big)$. On a donc $\ker\big(M - \lambda_i I_n\big)^2 = \ker\big(M - \lambda_i I_n\big)$ (l'inclusion inverse est toujours vérifiée).

Plus généralement, si m_i est paire, avec la même méthode, lorsque $X \in \ker\big(M - \lambda_1 I_n\big)^{m_i}$ alors $X \in \ker\big(M - \lambda_1 I_n\big)^{m_i/2}$. Si m_i est impaire, on doit multiplier une fois par ${}^t\big(M - \lambda_i I_n\big)$ à gauche avant de multiplier par ${}^tX$ pour qu'ainsi on aboutisse au fait que lorsque $X \in \ker\big(M - \lambda_1 I_n\big)^{m_i}$ alors $X \in \ker\big(M - \lambda_1 I_n\big)^{(m_i+1)/2}$. Donc par itération, d'une manière générale, $\ker\big(M - \lambda_i I_n\big)^{m_i} = \ker\big(M - \lambda_i I_n\big)$. Par conséquent M est diagonalisable dans $\mathbb{R}$: il existe P une matrice de passage réelle telle que $D = P^{-1} M\, P$ soit une matrice diagonale. Enfin, grâce à la propriété précédente on déduit que les sous-espaces propres sont orthogonaux entre eux: la matrice est donc diagonalisable dans une base orthonormale.

Enfin, on sait que M est la matrice de u dans une base orthonormale. On vient de voir que u (ou M) est diagonalisable dans une autre base orthonormale, donc la matrice P de changement de base permet de passer d'une base orthonormale à une autre. D'après la sous-section précédente on sait que P est donc une matrice orthogonale et donc $P^{-1} = {}^tP$. □

Exemple.

Diagonaliser dans une base orthonormale la matrice $A = \begin{pmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix}$.

5.4 Réduction des matrices orthogonales

Propriété. *Si u est une application linéaire orthogonale (une isométrie vectorielle) dans un espace euclidien alors il existe un s.e.v. F de dimension 1 ou 2 stable par u. Plus généralement, on peut écrire que $E = F_1 \overset{\perp}{\oplus} \cdots \overset{\perp}{\oplus} F_p$, où les (F_i) sont des s.e.v. de dimension 1 ou 2 stables par u.*

Proof. Soit a un vecteur propre associé à une valeur propre réelle de u (si elle existe). Alors $F = \mathrm{Vect}(a)$ convient (donc F s.e.v. de dim$=1$).

Sinon, s'il n'existe pas de valeur propre réelle, on considère $v = u + u^*$. Alors v est symétrique donc v admet des valeurs propres réelles. Soit a un vecteur propre réel de v. Alors $u(a) + u^*(a) = \lambda\, a$. Comme u est orthogonal, $u^*u = Id$ donc $a + u^2(a) = \lambda\, u(a)$ soit $u^2(a) = \lambda\, u(a) - a$. De ceci on en déduit que $F = \mathrm{Vect}(a, u(a))$ est stable par u.

Pour montrer la décomposition on utilise une récurrence. En effet la proposition est vraie en dimension 1 et 2. supposons qu'elle est vraie en dimension n et $n+1$ et montrons qu'elle est alors vraie en dimension $n+2$. Soit F_1 un s.e.v. de dimension 1 ou 2 stable par u. Alors $F_1^\perp$ est également stable par u (car si $x \in F_1$ et $y \in F_1^\perp$ alors $<x, y> = 0$ mais $<x, y> = <u(x), u(y)>$ car u orthogonal, donc $<u(x), u(y)> = 0$ et $u(x) \in F_1$). De plus u restreint à $F_1^\perp$ (c'est-à-dire l'application $v_1 : x \in F_1^\perp \mapsto u(x) \in F_1^\perp$) est aussi une isométrie. Donc comme $F_1^\perp$ est de dimension n ou $n+1$ on peut utiliser la formule de récurrence et ainsi décomposer $F_1^\perp$ comme somme directe orthogonale de s.e.v. de dimension 1 ou 2. Enfin, il suffit d'écrire $E = F_1 \overset{\perp}{\oplus} F_1^\perp$ pour conclure. □

Proposition. *Si M est une matrice orthogonale alors il existe une base orthonormale e de E telle que:*

$$M = \begin{pmatrix} A_1 & \cdots 0 \cdots & 0 & 0 & \cdots 0 \cdots & 0 \\ 0 & \ddots & 0 & 0 & \vdots & 0 \\ 0 & \cdots 0 \cdots & A_q & 0 & \cdots 0 \cdots & 0 \\ 0 & \cdots 0 \cdots & 0 & \pm 1 & \cdots 0 \cdots & 0 \\ 0 & \vdots & 0 & 0 & \ddots & 0 \\ 0 & \cdots 0 \cdots & 0 & 0 & \cdots 0 \cdots & \pm 1 \end{pmatrix}$$

où les A_i sont des matrices de la forme $A_i := \begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$.

Proof. Il est facile de voir que si λ est une valeur propre et x un vecteur propre associé non nul alors $\|f(x)\| = |\lambda|\|x\| = \|x\|$ donc $|\lambda| = 1$. Si $\dim F_i = 1$ alors $\lambda_i = \pm 1$. Si $\dim F_i = 2$, alors u restreint à F_i peut avoir pour matrice dans F_i une matrice de type $\begin{pmatrix} \cos\theta_i & -\sin\theta_i \\ \sin\theta_i & \cos\theta_i \end{pmatrix}$, seule type de matrice dont le module des valeurs propres est 1. □

Exemple.

Réduire la matrice $A := \dfrac{1}{9}\begin{pmatrix} -7 & -4 & 4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{pmatrix}$ et en déduire quelle est la transformation géométrique de $\mathbb{R}^3$ dans $\mathbb{R}^3$ associée à cette matrice.

6 Formes bilinéaires symétriques et formes quadratiques

6.1 Définition et premières propriétés

Définition. *Soit E un espace vectoriel. Une application $\psi : (x,y) \in E^2 \mapsto \psi(x,y) \in \mathbb{R}$ est une forme bilinéaire symétrique si :*

1. $\forall x \in E,\ \forall (y_1,y_2) \in E^2,\ \forall (\lambda_1,\lambda_2) \in \mathbb{R}^2,\ \psi(x, \lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 \psi(x,y_1) + \lambda_2 \psi(x,y_2)$;
2. $\forall (x,y) \in E^2,\ \psi(x,y) = \psi(y,x)$.

On peut associer de manière unique à ψ une application $\Phi : x \in E \mapsto \Phi(x) := \psi(x,x) \in \mathbb{R}$ appelée forme quadratique et telle que:

1. $\forall x \in E,\ \forall \lambda \in \mathbb{R},\ \Phi(\lambda \cdot x) = \lambda^2 \Phi(x)$;
2. $\forall (x,y) \in E^2,\ \Phi(x+y) = \Phi(x) + \Phi(y) + 2\,\psi(x,y)$.

Propriété. *Réciproquement également: à toute forme quadratique on peut associer une unique forme bilinéaire symétrique.*

Proof. Ceci est évident du fait de la dernière propriété: $\forall (x,y) \in E^2,\ \Phi(x+y) = \Phi(x) + \Phi(y) + 2\,\psi(x,y)$, donc $\psi(x,y) = \frac{1}{2}\big(\Phi(x+y) - \Phi(x) - \Phi(y)\big)$. □

Exemple.

Pour $E = \mathbb{R}^3$, avec $x = (x_1,x_2,x_3)$ et $y = (y_1,y_2,y_3)$, considérer les formes bilinéaires symétriques $\psi_1(x,y) = x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3$ et $\psi_2(x,y) = x_2 \cdot y_2 - 3x_1 \cdot y_2$.

Définition. *On dit que $(x,y) \in E^2$ sont ψ-orthogonaux si $\psi(x,y) = 0$.*

Exemple.

En reprenant l'exemple précédent, déterminer des vecteurs ψ_1 et ψ_2-orthogonaux au vecteur $v = (1,1,1)$.

Définition. *On dit qu'une forme quadratique Φ sur E est :*

- *positive, si* $\forall x \in E,\ \Phi(x) \geq 0$;
- *définie, si* $\forall x \in E,\ (\Phi(x) = 0)$;
- *dégénérée si* $\ker \Phi := \big\{x \in E,\ \forall y \in E,\ \psi(x,y) = 0\big\}$, *qui est un s.e.v. de E, n'est pas réduit au vecteur nul.* $\Longleftrightarrow$ $(x = 0)$.

Exemple.

En reprenant l'exemple précédent, étudier si ψ_1 et ψ_2 sont définies ou/et positive.

Propriété. *Soit E un espace vectoriel de dimension finie et soit $e = (e_1,\cdots,e_n)$ une base de E.*

- *Pour $(x,y) \in E^2$ de coordonnées $(x_1,\cdots,x_n)$ et $(y_1,\cdots,y_n)$ dans la base E,*

$$\psi(x,y) = \sum_{i=1}^n \sum_{j=1}^n [\psi(e_i,e_j)]\, x_i y_j \ \text{ et } \ \Phi(x) = \sum_{i=1}^n \sum_{j=1}^n [\psi(e_i,e_j)]\, x_i x_j = \sum_{i=1}^n [\Phi(e_i)]\, x_i^2 + 2 \sum_{1 \leq i < j \leq n} [\psi(e_i,e_j)]\, x_i x_j.$$

- *Réciproquement, toute forme quadratique s'écrit :*

$$\Phi(x) = \sum_{1 \le i \le j \le n} \alpha_{ij}\, x_i x_j, \quad \textit{où } \alpha_{ij} \in \mathbb{R} \textit{ pour tout } i, j = 1, \cdots, n, \quad \textit{avec } \alpha_{ij} = \alpha_{ji}.$$

Proof. • On écrit $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{j=1}^n x_j e_j$, d'où $\psi(x,y) = \sum_{i=1}^n \sum_{j=1}^n [\psi(e_i,e_j)]\, x_i y_j$ d'après la propriété de bilinéarité. Ensuite, il suffit d'écrire $\Phi(x) = \psi(x,x)$.
• Ceci est une conséquence de ce qui précède, et de la symétrie. □

Conséquence.

Soit E un espace vectoriel de dimension finie et soit $e = (e_1, \cdots, e_n)$ une base de E. On définit la matrice symétrique $M_e(\Phi) = (\psi(e_i,e_j))_{1 \le i,j \le n} \in \mathcal{M}_(\mathbb{R})$. Alors si $(x,y) \in E^2$ ont pour matrice colonne de coordonnées X et Y dans e, on a :

$$\psi(x,y) = {}^tX \cdot M_e(\Phi) \cdot Y \quad \text{et} \quad \Phi(x) = {}^tX \cdot M_e(\Phi) \cdot X.$$

6.2 Application linéaire symétrique associée

On suppose ici que E est un espace euclidien muni du produit scalaire $< \cdot, \cdot >$.

Propriété. *Pour toute forme quadratique sur E, il existe une unique application linéaire symétrique u telle que pour tout $x \in \mathbb{E}$, $\Phi(x) =< x, u(x) >$. De plus, $\psi(x,y) =< x, u(y) >$ pour tout x, $y \in E$.*

Proof. Pour $y \in E$, $x \in E \to \phi(x,y)$ est une forme linéaire de E (du fait de linéarité de ϕ. On sait donc d'après le Théorème de représentation de Riesz, qu'il existe $z \in E$, tel que $\forall x \in E$, $\phi(x,y) =< x, z >$. Donc pour tout y il existe un tel z; notons $z = u(y)$. D'après la bilinéarité de ϕ, u est une application linéaire (pour tout $x \in E$, tout $y_1, y_2 \in E$ et $\lambda_2 \in \mathbb{R}$, on a $\phi(x,y_1) + \phi(x, \lambda_2 y_2) = \phi(x,y_1) + \lambda_2 \phi(x,y_2) =< x, u(y_1) + \lambda_2 < x, u(y_2) >=< x, u(y_1) + \lambda_2 u(y_2) >$, mais aussi $\phi(x,y_1) + \phi(x, \lambda_2 y_2) = \phi(x, y_1 + \lambda_2 y_2) =< x, u(y_1 + \lambda_2 y_2) >$, donc $< x, u(y_1) + u(y_2) >=< x, u(y_1 + y_2) >$ et comme cela est vrai pour tout $x \in E$, alors $u(y_1) + u(y_2) = u(y_1 + y_2)$). Il ne reste plus qu'à montrer que u est unique. Mais si pour tout $x \in E$, tout $y \in E$, $< x, u(y) >=< x, u'(y) >$ alors $< x, u(y) - u'(y) >= 0$ et comme cela est vrai pour tout $x \in \mathbb{R}$, alors $u(y) = u'(y)$ pour tout $y \in E$. u est bien unique. □

Propriété. *Dans toute base e de E, la matrice de Φ dans e est également la matrice de u dans e.*

Conséquence.

Une conséquence de ceci est que l'espace vectoriel constitué par l'ensemble des formes quadratiques sur E (pourquoi est-ce un e.v.?) est isomorphe à l'ensemble des applications linéaires symétriques de E.

Exercice.

Soit E un espace euclidien et u une application linéaire de E dans E (pas forcément symétrique). Montrer que l'application $q : x \in E \mapsto < x, u(x) >$ est une forme quadratique sur E et déterminer l'application linéaire symétrique associée à q.

6.3 Réduction des formes quadratiques

Proposition. *Dans E un espace vectoriel de dimension n, il existe une base orthonormale e (constitué de vecteurs propres de $M_e(\Phi)$) et des n réels $(\lambda_i)_{1 \le i \le n} \in BBr^n$ (valeurs propres de $M_e(\Phi)$ pouvant être prises égales), tels que pour $x \in E$ de coordonnées $(x_1, \cdots, x_n)$ dans e,*

$$\Phi(x) = \sum_{i=1}^n \lambda_i \cdot x_i^2.$$

Proof. On utilise la représentation $\Phi(X) = {}^tX\, M_e(\Phi)\, X$ et comme $M_e(\Phi)$ est symétrique, on peut la diagonaliser dans une base orthonormale, d'où $M_e(\Phi) = P\,D\,{}^tP$ avec P une matrice orthogonale et D une matrice diagonale réelle, donc $\Phi(X) = {}^tXP\,D\,{}^tPX$ et ainsi avec $Y = {}^tPX$, on obtient $\Phi(X) = {}^tY\,D\,Y$. □

Corollaire. *Avec les mêmes hypothèses et notations que précédemment, Φ est positive si et seulement si toutes les valeurs propres de $M_e(\Phi)$ sont positives et Φ est définie positive si et seulement si toutes les valeurs propres de $M_e(\Phi)$ sont strictement positives.*

Définition. *Soit E un espace vectoriel de dimension n et Φ une forme quadratique sur E. On appelle signature de Φ le couple signature$(\Phi) = (p, q) \in \mathbb{N}^2$ tel que p (respectivement q) est le nombre de valeurs propres de $M_e(\Phi)$ strictement positives (respectivement négatives).*

Propriété. *Si signature$(\Phi) = (p, q)$ est tel que $p + q < n$ alors Φ est dégénérée.*

Exemple.

Reprendre l'exemple des formes bilinéaires symétriques de $\mathbb{R}^3$, $\psi_1(x, y) = x_1 \cdot y_1 + x_2 \cdot y_2 + x_3 \cdot y_3$ et $\psi_2(x, y) = x_2 \cdot y_2 - 3x_1 \cdot y_2$.

Propriété. *Il existe un autre procédé pour réduire une forme quadratique dans un espace euclidien directement à partir de son expression sous la forme $\Phi(x) = \sum_{1 \leq i \leq j \leq n} \alpha_{ij}\, x_i x_j$: le procédé d'orthogonalisation de Gauss. Celui-ci consiste à rassembler tous les termes se rapportant à x_1, puis d'écrire*

$$\alpha_{11}x_1^2 + \sum_{j=2}^{n} \alpha_{1j}x_1x_j = \alpha_{11}\big(x_1 + \sum_{j=2}^{n} \frac{\alpha_{1j}}{2\alpha_{11}}x_j\big)^2 - \big(\sum_{j=2}^{n} \frac{\alpha_{1j}}{2\alpha_{11}}x_j\big)^2.$$

On recommence ensuite avec x_2 et ainsi de suite...

6.4 Application aux coniques

Soit $(0, \vec{i}, \vec{j})$ un repère orthonormé du plan. Soit $(a, b, c, d, e, f) \in \mathbb{R}^6$. On appelle conique $\mathcal{C}$ un l'ensemble des points $M(x, y)$ tels que :

$$a \cdot x^2 + 2b \cdot x \cdot y + c \cdot y^2 + d \cdot x + e \cdot y + f = 0.$$

On associe à $\mathcal{C}$ la forme quadratique :

$$\Phi(x\vec{i} + y\vec{j}) = a \cdot x^2 + 2b \cdot x \cdot y + c \cdot y^2.$$

Après avoir réduit la forme quadratique (on détermine ses valeurs propres λ_1 et λ_2) par un changement de variable lié à un changement de base orthonormale, puis avec un changement de repère, si $\lambda_1 \neq 0$ et $\lambda_2 \neq 0$, on en arrive à :

$$(M(x', y') \in \mathcal{C}) \iff \lambda_1 \cdot (x')^2 + \lambda_2 \cdot (y')^2 + k = 0, \text{ avec } k \in \mathbb{R},$$

où x' et y' sont des fonctions affines de x et y. Alors :

1. si signature$(\Phi) = (2, 0)$ ou signature$(\Phi) = (0, 2)$, $\mathcal{C}$ peut être une ellipse ($k < 0$), un point ($k = 0$) ou $\emptyset$ ($k > 0$).
2. si signature$(\Phi) = (0, 2)$, $\mathcal{C}$ peut être une ellipse ($k > 0$), un point ($k = 0$) ou $\emptyset$ ($k < 0$).
3. si signature$(\Phi) = (1, 1)$, $\mathcal{C}$ peut être une hyperbole ($k \neq 0$) ou deux droites ($k = 0$).
4. si signature$(\Phi) = (1, 0)$, ou signature$(\Phi) = (0, 1)$, on reprend la réduction de la forme quadratique et on montre que $\mathcal{C}$ peut être une parabole, deux droites parallèles, une droite ou $\emptyset$.

6.5 Application à l'optimisation des fonctions à plusieurs variables