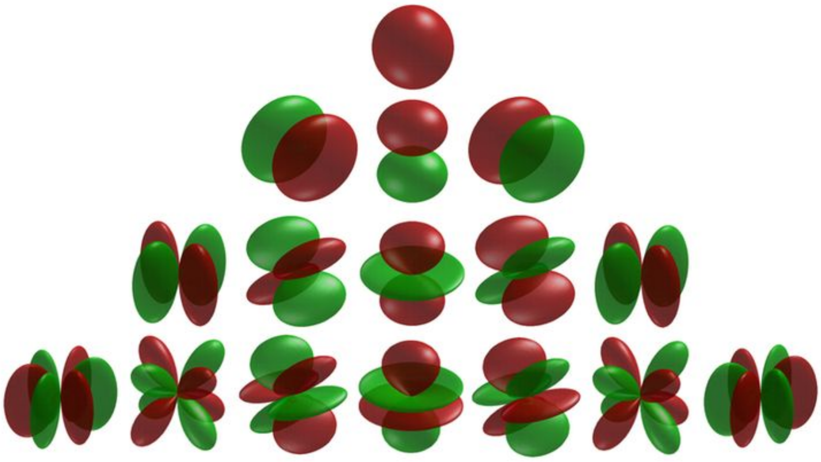
Ab1 : Espaces préhilbertiens réels



Le calcul des harmoniques spheriques utilise les polynômes orthogonaux de Legendre

I Produit scalaire sur un R-espace vectoriel

I.1 Définitions

Soit E un **R**-espace vectoriel. On appelle produit scalaire sur E toute application

$$\begin{array}{rccc} (.|.) : & E \times E & \longrightarrow & \mathbf{R} \\ & (x,y) & \longmapsto & (x|y) \end{array}$$

bilinéaire symétrique définie positive, c'est-à-dire vérifiant les propriétés suivantes :

1. $\forall (x, y_1, y_2) \in E^3 \quad \forall \lambda \in \mathbf{R} \quad (x|y_1 + \lambda y_2) = (x|y_1) + \lambda (x|y_2)$
   (linéarité à droite)
2. $\forall (x_1, x_2, y) \in E^3 \quad \forall \lambda \in \mathbf{R} \quad (x_1 + \lambda x_2|y) = (x_1|y) + \lambda (x_2|y)$
   (linéarité à gauche)
3. $\forall (x, y) \in E^2 \quad (y|x) = (x|y)$
   (symétrie)
4. $\forall x \in E \quad (x|x) \in \mathbf{R}^+$
   (positivité)
5. $(x|x) = 0 \Longrightarrow x = 0_E$
   (définition)

Dans la pratique, on commence par vérifier la symétrie, alors la linéarité à gauche implique la linéarité à droite et réciproquement.

Un **R**-espace vectoriel muni d'un produit scalaire est appelé espace préhilbertien réel. Les autres notations usuelles pour un produit scalaire sont $x.y$, $\langle x, y \rangle$.

I.2 Exemple-avertissement

Montrer que

$$(P,Q) \longmapsto \int_0^{+\infty} e^{-t}\, P(t)\, Q(t)\, \mathrm{d}t$$

définit un produit scalaire sur $\mathbf{R}[X]$ (on confond comme d'habitude polynômes et fonctions polynômes).

On remarquera quelques confusions possibles dans le vocabulaire : ici, montrer que le produit scalaire est « défini » peut être compris comme « bien défini » (l'intégrale existe) ou comme « défini » dans le sens $(P|P) = 0 \Rightarrow P = 0$.

I.3 Exemples fondamentaux

a $\mathbf{R}^n$

A deux éléments quelconques $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ de $\mathbf{R}^n$, on associe

$$(x|y) = \sum_{i=1}^{n} x_i y_i$$

On définit ainsi un produit scalaire, appelé produit scalaire canonique sur $\mathbf{R}^n$. Si X et Y désignent les matrices colonnes des composantes de x et de y dans la base canonique, on remarque que

$$(x|y) = {}^t X Y$$

b $\mathscr{M}_n(\mathbf{R})$

Suivant la démarche du paragraphe précédent, il y a un produit scalaire canonique sur $\mathscr{M}_n(\mathbf{R})$: si A et B sont deux matrices,

$$(A|B) = \sum_{1 \leq i,j \leq n} A_{i,j} B_{i,j}$$

Il est intéressant (on peut même dire important) de savoir que ce produit scalaire s'exprime en utilisant la trace.

c $\mathscr{C}\big([a,b],\mathbf{R}\big)$

où a et b sont deux réels, $a < b$. A deux éléments quelconques f et g de cet espace des fonctions continues sur $[a,b]$ à valeurs réelles, on associe

$$(f|g) = \int_a^b f(t)\, g(t)\, \mathrm{d}t$$

On définit ainsi un produit scalaire.
Mais si on remplace $\mathscr{C}\big([a,b],\mathbf{R}\big)$ par $\mathscr{C}_{\mathrm{pm}}\big([a,b],\mathbf{R}\big)$ (avec des notations « évidentes »), ce n'est plus tout-à-fait un produit scalaire.

d fonctions continues de carré intégrable sur un intervalle quelconque (h.p.)

Soit E l'ensemble des fonctions f continues sur I, intervalle de $\mathbf{R}$, à valeurs réelles, de carré intégrable sur I, c'est-à-dire telles que $|f^2| = f^2$ soit intégrable sur I. De l'inégalité

$$|fg| \leq \frac{1}{2}\big(f^2 + g^2\big)$$

on déduit que le produit de deux fonctions de carré intégrable est intégrable. Donc, si f et g sont de carrés intégrables, si λ est un nombre réel quelconque, $f + \lambda g$ est de carré intégrable, car

$$(f + \lambda g)^2 = f^2 + \lambda^2 g^2 + 2\lambda f g\ .$$

Comme la fonction nulle est de carré intégrable, E est un $\mathbf{R}$-espace vectoriel. On a vu que pour deux éléments f et g de E on pouvait définir

$$(f|g) = \int_I fg\ .$$

On définit ainsi un produit scalaire sur E.

e Suites de carré sommable (h.p.)

On note $\ell^2_{\mathbf{R}}(\mathbf{N})$ les suites réelles de carré sommable, c'est-à-dire les suites réelles $(u_n)_{n\in\mathbf{N}}$ telles que la série $\sum u_n^2$ soit (absolument) convergente. De l'inégalité

$|ab| \leq \dfrac{1}{2}(a^2+b^2)$ on déduit que $\ell^2_{\mathbf{R}}(\mathbf{N})$ est un **R**-espace vectoriel (voir méthodes au paragraphe précédent) et que, en définissant, pour tout couple (u,v) d'éléments de $\ell^2_{\mathbf{R}}(\mathbf{N})$,

$$(u|v) = \sum_{n=0}^{+\infty} u_n v_n$$

on définit un produit scalaire sur $\ell^2_{\mathbf{R}}(\mathbf{N})$.

f Presque un exemple

Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé. Si X et Y sont deux variables aléatoires réelles discrètes de carré sommable sur $(\Omega, \mathscr{A}, P)$, on définit

$$(X|Y) = E(XY)$$

Définit-on ainsi un produit scalaire sur l'espace vectoriel des variables aléatoires réelles discrètes de carré sommable sur $(\Omega, \mathscr{A}, P)$?

I.4 Inégalité de Cauchy-Schwarz

Hermann Schwarz, mathématicien allemand de la deuxième moitié du XIXème siècle, ne doit pas être confondu avec Laurent Schwartz, mathématicien français du XXème siècle. Donc, pas de t au Schwarz de l'inégalité de Cauchy-Schwarz.

Soit x, y deux éléments d'un espace préhilbertien réel E, λ un nombre réel quelconque. On sait qu'alors

$$(x+\lambda y|x+\lambda y) \in \mathbf{R}^+$$

ce qui donne, en développant par bilinéarité et en utilisant la symétrie,

$$\lambda^2(y|y)+2\lambda(x|y)+(x|x) \geq 0$$

Si y n'est pas nul, on obtient un trinôme du second degré qui garde un signe constant sur $\mathbf{R}$. Son discriminant est donc négatif ou nul, ce qui s'écrit

$$(x|y)^2-(x|x)(y|y) \leq 0$$

ou encore

Proposition

$$\left|(x|y)\right| \leq \sqrt{(x|x)}\sqrt{(y|y)}\ .$$

Cette inégalité, appelée inégalité de Cauchy-Schwarz, est encore vraie si y est nul (c'est alors une égalité).
Bien sûr, à plus forte raison,

$$(x|y) \leq \sqrt{(x|x)}\sqrt{(y|y)}\ .$$

On peut aussi écrire :

$$(x|y)^2 = (x|x)\ (y|y)$$

Proposition L'inégalité de Cauchy-Schwarz (avec la valeur absolue !) est une égalité si et seulement si x et y sont liés.

L'égalité $(x|y) = \sqrt{(x|x)}\sqrt{(y|y)}$ est obtenue si et seulement si x et y sont « positivement liés », c'est-à-dire si et seulement $y = 0$ ou $x = \lambda y$ avec $\lambda \geq 0$.

Méthode plus rapide : On développe, si x et y non nuls, pour $\epsilon = \pm 1$,

$$\begin{aligned} \left\| \frac{x}{\|x\|} + \epsilon \frac{y}{\|y\|} \right\|^2 &= 1 + 1 + 2\epsilon \frac{(x|y)}{\|x\| \, \|y\|} \\ &\geq 0 \end{aligned}$$

ce qui donne presque immédiatement l'inégalité et le cas d'égalité.

Approfondissement L'inégalité est encore vraie si $(.|.)$ est une forme bilinéaire symétrique positive, sans que l'on ait besoin de la supposer définie (voir par exemple l'inégalité de Cauchy-Schwarz pour les espérances, chapitre P4). En revanche, les cas d'égalité sont plus nombreux.

I.5 Norme associée à un produit scalaire

De l'inégalité de Cauchy-Schwarz, on déduit que, pour tout couple (x, y) d'éléments de E,

$$\begin{aligned} (x+y|x+y) &= (x|x) + (y|y) + 2(x|y) \\ &\leq (x|x) + (y|y) + 2\sqrt{(x|x)}\sqrt{(y|y)} \\ &\leq \left(\sqrt{(x|x)} + \sqrt{(y|y)} \right)^2 \end{aligned}$$

et finalement :

$$\sqrt{(x+y|x+y)} \leq \sqrt{(x|x)} + \sqrt{(y|y)} \, .$$

L'application $x \mapsto \|x\| = \sqrt{(x|x)}$ est donc une application de E dans $\mathbf{R}^+$ vérifiant les propriétés suivantes (pour tous x et y dans E, λ dans $\mathbf{R}$) :

$\|\lambda x\| = |\lambda| \|x\|$

$\|x + y\| \leq \|x\| + \|y\|$

$\|x\| = 0 \Rightarrow x = 0_E$

C'est donc une norme sur E. Un espace préhilbertien réel est donc muni d'une structure d'espace vectoriel normé, la norme précédente est dite associée au produit scalaire ; on parle de **norme euclidienne**. La distance associée est une distance euclidienne.

I.6 Exemples fondamentaux

a $\mathbf{R}^n$

Pour tous éléments $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$ de $\mathbf{R}^n$, on a :

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \sqrt{\sum_{i=1}^{n} x_i^2}\,\sqrt{\sum_{i=1}^{n} y_i^2}\,.$$

Et aussi, d'ailleurs :

$$\sum_{i=1}^{n} |x_i y_i| \leq \sqrt{\sum_{i=1}^{n} x_i^2}\,\sqrt{\sum_{i=1}^{n} y_i^2}\,.$$

La norme est définie par

$$\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{{}^tX\,X}.$$

b $\mathscr{C}\big([a,b],\mathbf{R}\big)$

où a et b sont deux réels, $a < b$. Pour tous éléments f et g de cet espace, on a

$$\left|\int_a^b f(t)\,g(t)\,\mathrm{d}t\right| \leq \sqrt{\int_a^b f(t)^2\,\mathrm{d}t}\,\sqrt{\int_a^b g(t)^2\,\mathrm{d}t}\,.$$

On peut évidemment mettre des valeurs absolues sur f et g,

$$\int_{[a,b]} |fg| \leq \sqrt{\int_{[a,b]} f^2}\,\sqrt{\int_{[a,b]} g^2}\,.$$

La norme associée au produit scalaire est définie par

$$N_2(f) = \sqrt{\int_a^b f(t)^2\,\mathrm{d}t}\,.$$

c fonctions de carré intégrable (h.p.)

Soit f, g continues sur I, intervalle de **R**, à valeurs réelles, de carré intégrable sur I. On a

$$\left|\int_I fg\right| \leq \sqrt{\int_I f^2}\,\sqrt{\int_I g^2}\,.$$

Et, aussi

$$\int_I |fg| \leq \sqrt{\int_I f^2}\,\sqrt{\int_I g^2}\,.$$

Ce qui s'écrit aussi $N_1(fg) \leq N_2(f)N_2(g)$.
La norme est définie par

$$N_2(f) = \sqrt{\int_I f^2}\,.$$

On peut utilement remarquer que ces inégalités intégrales sont encore valables pour des fonctions continues par morceaux.

Une application célèbre : Montrer que la fonction $x \mapsto \ln(\Gamma(x))$ est convexe (on rappelle que, si $x > 0$, $\Gamma(x) = \displaystyle\int_0^{+\infty} e^{-t}t^{x-1}\,dt$ et que l'on peut dériver indéfiniment cette expression sous le signe $\int$).

d Suites de carré sommable (h.p.)

Soit u, v deux suites réelles, telles que les séries $\sum u_n^2$ et $\sum v_n^2$ convergent. Alors la série de terme général $u_n v_n$ est absolument convergente et

$$\left|\sum_{n=0}^{+\infty} u_n v_n\right| \leq \sqrt{\sum_{n=0}^{+\infty} u_n^2}\,\sqrt{\sum_{n=0}^{+\infty} v_n^2}$$

A rapprocher du paragraphe précédent

e Espérances

Si X et Y sont des variables aléatoires réelles discrètes de carré intégrable sur un espace probabilisé $(\Omega, \mathscr{A}, P)$, alors XY est d'espérance finie, et

$$|E(XY)| \leq \sqrt{E(X^2)}\,\sqrt{E(Y^2)}$$

I.7 « Polarisation »

Soit x et y deux éléments d'un espace préhilbertien réel, α et β deux nombres réels. Alors
$\|\alpha x+\beta y\|^2 =$
On obtient alors les identités dites de polarisation, qui permettent de retrouver le produit scalaire à partir de la norme :
$(x|y) = \dfrac{1}{2}(\|x+y\|^2 - \|x\|^2 - \|y\|^2)$

$(x|y) = \dfrac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$.

Seule la première de ces formules est au programme.

I.8 Identité du parallélogramme

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Très utile dans plusieurs problèmes géométriques, dans la projection sur un convexe compact...A mémoriser géométriquement !

I.9 Continuité du produit scalaire

Soit $\left(E,(.|.)\right)$ un espace préhilbertien réel. Alors l'application

$$\begin{array}{rccc} (.|.): & E\times E & \longrightarrow & \mathbf{R} \\ & (x,y) & \longmapsto & (x|y) \end{array}$$

est continue (E étant muni de la norme associée à $(.|.)$). *Ce résultat n'est au programme qu'en dimension finie.*

II Orthogonalité

II.1 Définitions

On dit que deux vecteurs sont orthogonaux lorsque

$$(x|y) = 0$$

On écrit $x \perp y$.

Une famille orthogonale est une famille de vecteurs deux à deux orthogonaux.

Un vecteur unitaire est un vecteur de norme 1 ($\|x\| = 1$).

Une famille orthonormale est une famille orthogonale de vecteurs unitaires.

II.2 Indépendance linéaire

Proposition Une famille orthogonale de vecteurs **non nuls**, en particulier une famille orthonormale, est libre.

C'est une méthode utile pour démontrer qu'une famille est libre…ne pas oublier « non nuls ».

II.3 Théorème de Pythagore

Théorème Si $(x_i)_{1 \leq i \leq n}$ est une famille orthogonale,

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2$$

La réciproque est vraie quand $n = 2$.

II.4 Sous-espaces orthogonaux

Deux sous-espaces F et G sont orthogonaux lorsque tout vecteur de F est orthogonal à tout vecteur de G. On note $F \perp G$. La somme $F+G$ est alors directe. On la notera $F \bigoplus G$

Plus généralement, si les F_i $(1 \leq i \leq p)$ sont des sous-espaces deux à deux orthogonaux (si $i \neq j$, alors $F_i \perp F_j$), alors leur somme est directe : on peut la noter

II.5 Orthogonal d'un sous-espace vectoriel

Définition L'orthogonal $F^\perp$ d'un sous-espace F est l'ensemble des vecteurs de E orthogonaux à tout vecteur de F.

$$F^\perp = \left\{x \in E\ ;\ \forall y \in F \quad (x|y) = 0\right\}$$

Proposition $F^\perp$ est un sous-espace vectoriel de E, orthogonal à F. Autrement dit :

$$F \perp F^\perp$$

De plus, pour tout sous-espace vectoriel G :

$$(G \perp F) \Longrightarrow \left(G \subset F^\perp\right)$$

On a toujours $F \cap F^\perp = \{0_E\}$, mais en général, $F \bigoplus F^\perp \neq E$.

A retenir En général, l'orthogonal d'un sous-espace vectoriel n'est pas un supplémentaire de ce sous-espace vectoriel : on peut avoir $F \oplus F^\perp \subsetneq E$.

Un exemple classique Soit $E = \mathscr{C}\big([0,1], \mathbf{R}\big)$, muni de son produit scalaire usuel, et F le sev des fonctions polynômes réelles sur $[0,1]$. Le théorème de Weierstrass permet de voir (c'est un exercice classique!) que $F^\perp = \{0_E\}$. Et donc que

$$F \oplus F^\perp = F \neq E$$

On retiendra donc que la somme $F \bigoplus F^\perp$ est toujours directe, mais que ce n'est en général pas E, donc l'orthogonal de F n'est pas nécessairement un supplémentaire de F.

Comme souvent, notre intuition est bonne seulement en dimension finie : il suffit d'ailleurs que F soit de dimension finie pour que $F \oplus F^\perp = E$,voir un peu plus loin la projection sur un sous-espace de dimension finie.

III Espaces euclidiens

III.1 Définition

On appelle espace euclidien un espace préhilbertien réel de dimension finie.

III.2 Existence de bases orthonormales

Théorème Tout espace euclidien a des bases orthonormales.

Démonstration Récurrence sur la dimension de l'espace.

Soit $(e_1,\dots,e_n)$ une base orthonormale de E, x un vecteur de E. Alors

$$x=\sum_{i=1}^{n}(e_i|x)e_i$$

$$\|x\|=\sqrt{\sum_{i=1}^{n}(e_i|x)^2}$$

et, si y est un autre vecteur :

$$(x|y)=\sum_{i=1}^{n}(e_i|x)(e_i|y)\ .$$

ou encore, en notant X et Y les vecteurs colonnes des composantes respectives de x et y dans cette base orthonormale. Alors

$$(x|y)={}^tXY\quad ;\quad \|x\|=\sqrt{{}^tXX}\ .$$

III.3 Isomorphisme canonique avec l'espace dual

Soit E un espace euclidien. Pour tout élément a de E, on définit

$$\begin{array}{rccc}\phi_a\ : & E & \longrightarrow & \mathbf{R}\\ & x & \longmapsto & (a|x)\end{array}$$

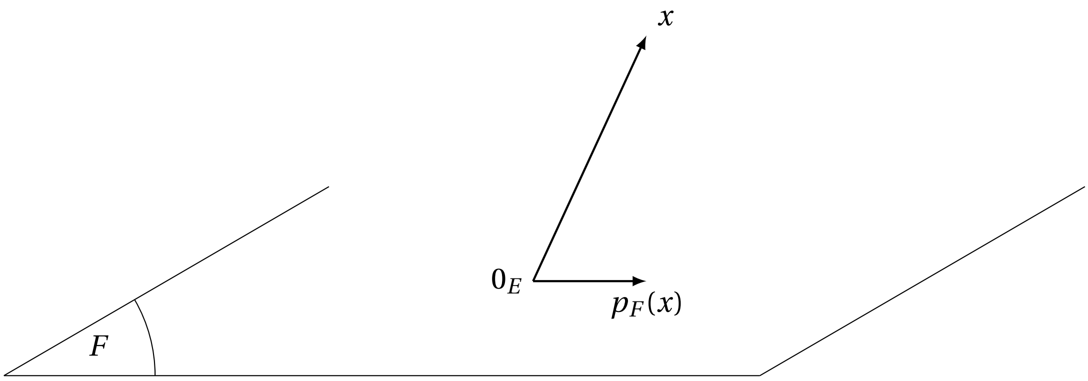
ϕ_a est une application linéaire, c'est donc un élément de $E^*=\mathscr{L}(E,\mathbf{R})$.
L'application $a\longmapsto\phi_a$ est un isomorphisme de E sur E^*, appelé isomorphisme canonique de E sur E^*.
On utilise fréquemment ce résultat sous la forme suivante :

Proposition : Soit ϕ une forme linéaire non nulle sur l'espace euclidien E. Il existe alors un unique élément a de E tel que $\forall x\in E\quad (a|x)=\phi(x)$.

Applications : Définition du produit vectoriel, du gradient.

IV Projection orthogonale sur un s.e.v. de dimension finie

5/2 : Les suites totales ne sont plus au programme.



IV.1 Théorème de projection

Théorème Soit F un sous-espace **de dimension finie** d'un espace préhilbertien réel E (*E n'est pas nécessairement de dimension finie*).

Pour tout $x \in E$ il existe $y \in F$ unique tel que

$$y \in F \qquad \text{et} \qquad x - y \in F^{\perp}\ .$$

On note $y = p_F(x)$; p_F est donc caractérisé par

$$\forall x \in E \qquad p_F(x) \in F \qquad \text{et} \qquad x - p_F(x) \in F^{\perp}\ .$$

p_F est la projection sur F parallèlement à $F^{\perp}$, on l'appelle projection orthogonale sur F.

$F^{\perp}$ est donc un supplémentaire de F (on dit que $F^{\perp}$ est le supplémentaire orthogonal de F).

L'application $z \in F \longmapsto \|x - z\|$ atteint son minimum en un point unique : $p_F(x)$. Autrement dit,

$$\forall z \in F \setminus \{p_F(x)\} \qquad \|x - z\| > \|x - p_F(x)\|\ .$$

On a donc $d(x, F) = \|x - p_F(x)\|$.

Si $(e_1, \ldots, e_p)$ est une base orthonormale de F, alors, pour tout vecteur x de E :

$$p_F(x) = \sum_{i=1}^{p} (e_i|x)e_i \ .$$

Le théorème de Pythagore montre que $\|x\|^2 = \|p_F(x)\|^2 + \|x - p_F(x)\|^2$, on en déduit l'inégalité de Bessel :

$$\sum_{i=1}^{p} (e_i|x)^2 \leq \|x\|^2 \ .$$

Si E est de dimension finie, on a de plus

$$\dim(F^\perp) \ + \ \dim(F) = \dim(E)$$

On peut noter aussi que $\left(F^\perp\right)^\perp = F$, mais cette relation n'est pas très utile, il est surtout intéressant de comprendre qu'elle n'est pas vraie en toute généralité (si F n'est pas de dimension finie...).

Si F n'est pas de dimension finie, il se peut que $F \bigoplus F^\perp$ ne soit pas E tout entier : alors F n'a pas de supplémentaire orthogonal, et il n'y a pas de projection orthogonale sur F. Mais il se peut aussi que $F \bigoplus F^\perp = E$. Alors la projection orthogonale sur F est bien définie.

IV.2 Orthonormalisation de Gram-Schmidt

Théorème Soit $(e_1,\dots,e_n)$ une base d'un espace euclidien. Alors il existe une base orthonormale $(f_1,\dots,f_n)$ de cet espace telle que

$\forall i \in \{1,\dots,n\} \quad \mathrm{Vect}(e_1,\dots,e_i) = \mathrm{Vect}(f_1,\dots,f_i)$.

La base $(f_1,\dots,f_n)$ est dite obtenue à partir de la base $(e_1,\dots,e_n)$ par orthonormalisation de Schmidt.

Construction - Formule

Les conditions $\mathrm{Vect}(e_1) = \mathrm{Vect}(f_1)$ et $\|f_1\| = 1$ équivalent à

$$f_1 = \pm\frac{1}{\|e_1\|}e_1$$

Si $1 \leq j \leq n-1$, notons p_j la projection orthogonale sur $\mathrm{Vect}(e_1,\dots,e_j)$, on définit

$$f_{j+1} =$$

La formule est aussi importante que le théorème.

La décomposition QR Le procédé d'orthonormalisation de Schmidt a une interprétation matricielle, la décomposition QR, qui est un thème d'exercice classique, voir chapitre suivant car on a besoin des matrices orthogonales.

Corollaire (base orthonormale incomplète) Toute famille orthonormale de vecteurs d'un espace euclidien peut être complétée en une base orthonormale de cet espace.

Un exercice On suppose donné un produit scalaire $(.|.)$ sur $E = \mathbf{R}[X]$. Montrer qu'il existe une base orthogonale de $(E,(.|.))$ formée de polynômes dont le coefficient dominant est 1 (on évitera de dire « unitaires »).

IV.3 Projecteurs orthogonaux associés à une décomposition de l'espace en somme directe orthogonale

Soit $(F_i)_{1 \leq i \leq p}$ une famille de p sous-espaces d'un espace euclidien E, deux à deux orthogonaux. On suppose que la somme (directe orthogonale) des F_i est E. On associe à cette décomposition de E en somme directe orthogonale les projecteurs orthogonaux suivants : p_i est la projection orthogonale sur F_i, c'est-à-dire la projection sur F_i parallèlement à $\bigoplus_{j \neq i} F_j$. Les p_i sont des projecteurs orthogonaux vérifiant

$$\sum_{i=1}^{p} p_i = \mathrm{Id}_E \quad \text{et} \quad i \neq j \Rightarrow p_i \circ p_j = \Theta$$

Table des matières