Chapitre 10 : Espaces préhilbertiens

1 Produit scalaire, norme et orthogonalité

1.1 Produit scalaire

Définition 1. Soit E un $\mathbb{R}$-espace vectoriel. On appelle **produit scalaire sur** E toute « forme bilinéaire symétrique définie positive ». C'est-à-dire toute application $(.|.) : E \times E \to \mathbb{R}$ qui vérifie :

1. bilinéaire : $\forall (x, y, z) \in E^3,\ \forall (\lambda, \mu) \in \mathbb{R}^2$,
$$(\lambda x + \mu y|z) = \lambda(x|z) + \mu(y|z) \text{ et } (z|\lambda x + \mu y) = \lambda(z|x) + \mu(z|y)$$
2. symétrique : $\forall (x, y) \in E^2,\ (x|y) = (y|x)$
3. définie : $\forall x \in E,\ (x|x) = 0 \Longleftrightarrow x = 0_E$
4. positive : $\forall x \in E,\ (x|x) \geq 0$

Un espace vectoriel réel muni d'un produit scalaire est appelé un *espace préhilbertien* et si, de plus, il est de dimension finie on dira que c'est un *espace euclidien*.

Définition 2. Produit scalaire canonique dans $\mathbb{R}^n$
L'application $(.|.) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ qui à deux vecteurs $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ et $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ associe :
$$(x|y) = \sum_{k=1}^{n} x_k y_k$$
est un produit scalaire sur $\mathbb{R}^n$ appelé *produit scalaire canonique.*
On peut définir ce produit scalaire matriciellement en utilisant les matrices colonnes qui représentent les vecteurs dans la base canonique :
$$\begin{array}{ccc} \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ (X, Y) & \mapsto & (X^T)Y \end{array}$$

Remarque Ce n'est pas le seul produit scalaire qu'on peut mettre sur $\mathbb{R}^n$ mais c'est le plus usuel. C'est celui sur lequel se focalise le programme officiel.

1.2 Distance et norme associées à un produit scalaire

Définition 3. Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$.

- On appelle **norme sur E associée au produit scalaire** $(.|.)$ l'application $\|.\| : E \to \mathbb{R}_+$ définie par :
$$\forall x \in E,\ \|x\| = \sqrt{(x|x)}$$
- On appelle **distance sur E associée au produit scalaire** $(.|.)$ l'application $\mathrm{d} : E \times E \to \mathbb{R}_+$ définie par :
$$\forall (x, y) \in E^2,\ d(x, y) = \|x - y\|$$

Remarques

1. Dans $\mathbb{R}^n$ muni du produit scalaire canonique, on retrouve la distance « habituelle ».

2. Une norme quelconque vérifie les mêmes propriétés que la norme canonique dans $\mathbb{R}^2$ ou $\mathbb{R}^3$ comme par exemple $\|\lambda x\| = |\lambda|\|x\|$.

Théorème 1. Inégalité de Cauchy-Schwarz
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$. Pour tout $(x, y) \in E^2$, on a

$$|(x|y)| \leq \|x\|\|y\|$$

On a égalité si et seulement si x et y sont colinéaires.

Théorème 2. Inégalités triangulaires
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$. Pour tout $(x, y, z) \in E^3$, on a

$$\big|\|x\| - \|y\|\big| \leq \|x + y\| \leq \|x\| + \|y\|$$

On a égalité dans l'inégalité de droite si et seulement si x et y sont colinéaires de même sens.
On peut donner une version « distance » de ces inégalités :

$$\big|\mathrm{d}(x, y) - \mathrm{d}(x, z)\big| \leq \mathrm{d}(x, z) \leq \mathrm{d}(x, y) + \mathrm{d}(y, z)$$

Proposition 1. Identités du parallélogramme et de polarisation
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$.

1. Identité du parallélogramme : $\forall (x, y) \in E^2,\ \|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$
2. Identités de polarisation : $\forall (x, y) \in E^2,\ (x|y) = \frac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right) = \frac{1}{4}\left(\|x + y\|^2 - \|x - y\|^2\right)$

1.3 Orthogonalité

1.3.1 Vecteurs orthogonaux

Définition 4. Vecteurs unitaires, orthogonaux
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$.

1. On dit qu'un vecteur x de E est **unitaire** si $\|x\| = 1$.
2. Deux vecteurs de E, x et y sont dits **orthogonaux** si $(x|y) = 0$. On notera alors $x \perp y$.
3. Une famille de vecteurs de E, $(x_1, \ldots, x_n)$ est dite **orthogonale** si, pour tout $(i, j) \in [\![1, n]\!]^2$ distincts, x_i et x_j sont orthogonaux.
4. Une famille de vecteurs de E, $(x_1, \ldots, x_n)$ est dite **orthonormale** si elle est orthogonale et si tous ses vecteurs sont unitaires. On peut résumer en disant :

$$\forall (i, j) \in [\![1, n]\!]^2,\ (x_i|x_j) = \delta_{i,j} \text{ où } \delta_{i,j} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}$$

Proposition 2. Dans $\mathbb{R}^n$ muni de son produit scalaire canonique, la base canonique est orthonormale.

Théorème 3. Liberté d'une famille orthogonale
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$.
Si $(x_1, \ldots, x_n)$ est une famille orthogonale de vecteurs <u>non nuls</u> de E alors elle est libre.

Théorème 4. Pythagore
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$.
Si $(x_1, \dots, x_n)$ est une famille orthogonale de vecteurs de E alors

$$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \sum_{k=1}^{n} \|x_k\|^2$$

Remarque La réciproque est vraie pour une famille de 2 vecteurs mais devient fausse pour $n \geq 3$.

1.3.2 Bases orthonormées

Proposition 3. Procédé de Gram-Schmidt (HP)
Dans un espace préhilbertien E muni d'un produit scalaire $(.|.)$, il est toujours possible de transformer une famille libre $(e_1, \dots, e_n)$ en une famille orthonormale $(u_1, \dots, u_n)$ telle que $\mathrm{Vect}\,(e_1, \dots, e_n) = \mathrm{Vect}\,(u_1, \dots, u_n)$.

Remarque Par exemple pour 3 vecteurs libres (e_1, e_2, e_3) les formules sont les suivantes :

$$u_1 = \frac{e_1}{\|e_1\|}\ ,\ u_2 = \frac{e_2 - (e_2|u_1)u_1}{\|e_2 - (e_2|u_1)u_1\|}\ ,\ u_3 = \frac{e_3 - (e_3|u_1)u_1 - (e_3|u_2)u_2}{\|e_3 - (e_3|u_1)u_1 - (e_3|u_2)u_2\|}$$

Proposition 4. Existence d'une base orthonormée
Soit E un espace euclidien (donc de <u>dimension finie</u>) non nul muni d'un produit scalaire $(.|.)$. Alors E possède une base orthonormée.

Exemple On a déjà signalé que la base canonique est une base orthonormée de $\mathbb{R}^n$ muni de son produit scalaire canonique. Ce n'est bien sûr pas la seule : $\mathbb{R}^n$ possède une infinité de bases orthonormées pour son produit scalaire canonique.

Corollaire. Tout sous-espace vectoriel d'un espace euclidien possède une base orthonormée.

Théorème 5. Coordonnées dans une base orthonormée
Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et $(e_1, \dots, e_n)$ <u>une base orthonormée</u> de E et $x \in E$. Alors les coordonnées de x dans la base $(e_1, \dots, e_n)$ sont $((x|e_1), \dots, (x|e_n))$, i.e.

$$x = \sum_{k=1}^{n} (x|e_k)e_k$$

Application Si $u \in \mathcal{L}(E)$ et que $A = (a_{i,j})_{1 \leq i,j \leq n}$ est la matrice représentative de u dans une <u>base orthonormée</u> de E, alors $a_{i,j} = (u(e_j)|e_i)$.

Théorème 6. Expression du produit scalaire et de la norme dans une base orthonormée
Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et $(x, y) \in E^2$ de coordonnées $(x_1, \ldots, x_n)$ et $(y_1, \ldots, y_n)$ dans une base orthonormée de E. Alors on a :

$$(x|y) = \sum_{k=1}^{n} x_k y_k \quad \text{et} \quad \|x\| = \sqrt{\sum_{k=1}^{n} x_k^2}$$

En termes matriciels, si X est la matrice colonne des coordonnées de x dans cette BON et Y celle de y, on a

$$(x|y) = (X^T)Y \quad \text{et} \quad \|x\| = \sqrt{(X^T)X}$$

Remarque En d'autres termes, **dans n'importe quelle base orthonormée**, un produit scalaire quelconque (par exemple le produit scalaire canonique dans $\mathbb{R}^n$) et sa norme s'écrivent comme le produit scalaire canonique de $\mathbb{R}^n$ et sa norme dans la base canonique.

1.3.3 Sous-espaces vectoriels orthogonaux

Définition 5. Sous-espaces orthogonaux (HP)
Soit E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$ et soient F et G deux sous-espaces vectoriels de E. On dit que F et G sont orthogonaux si tout vecteur de F est orthogonal à tout vecteur de G, i.e. :

$$\forall f \in F,\ \forall g \in G,\ (f|g) = 0$$

Remarque Si F et G sont de dimension finie, alors il suffit de vérifier l'orthogonalité des vecteurs d'une base de F avec ceux d'une base de G.

Définition 6. Orthogonal d'un sous-espace
Soient E un espace préhilbertien réel muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E. On appelle **orthogonal** de F dans E l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F, i.e. :

$$F^{\perp} = \{x \in E\ ;\ \forall f \in F,\ (x|f) = 0\}$$

Exemple $\{0_E\}^{\perp} = E$ et $E^{\perp} = \{0_E\}$.

Proposition 5. Avec les mêmes notations que la définition précédente.

1. $F^{\perp}$ est un sous-espace vectoriel de E orthogonal à F.
2. $F \cap F^{\perp} = \{0\}$

Proposition 6. Soient E un espace euclidien (donc de dimension finie) muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E. On a

1. Pour tout $x \in E$, il existe un unique couple $(x_F, x_{F^{\perp}}) \in F \times F^{\perp}$ tel que $x = x_F + x_{F^{\perp}}$.
2. La concaténation d'une base de F et d'une base de $F^{\perp}$, donne une base de E.
3. $\dim E = \dim F + \dim F^{\perp}$.

2 Théorème spectral

2.1 Matrices orthogonales (HP)

Définition 7. Soit $A \in \mathcal{M}_n(\mathbb{R})$. On dit que A est **orthogonale** si $(A^T)A = A(A^T) = I_n$; i.e. si A est inversible d'inverse A^T. L'ensemble des matrices orthogonales de taille n est noté $O(n)$.

Remarque Si A est orthogonale alors A^T l'est aussi. Si A et B appartiennent à $O(n)$ alors $AB \in O(n)$ mais a priori $A + B$ n'est pas orthogonale.

Théorème 7. Si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases orthonormées de $\mathbb{R}^n$ (pour son produit scalaire canonique) alors la matrice de passage P de $\mathcal{B}$ vers $\mathcal{C}$ est une matrice orthogonale et vérifie donc $P^{-1} = P^T$.

Remarque C'est très pratique pour simplifier le calcul d'inverse qui apparaît dans le théorème de changement de bases matriciel (cf $A = PDP^{-1}$).

Proposition 7. Réciproquement : si $P \in O(n)$ alors ses vecteurs colonnes forment une base orthonormée de $\mathbb{R}^n$ muni du produit scalaire canonique. De même pour ses vecteurs lignes.

2.2 Théorème spectral

Définition 8. On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est **symétrique** si $A^T = A$. L'ensemble des matrices symétriques à coefficients dans $\mathbb{K}$ est noté $\mathcal{S}_n(\mathbb{K})$. C'est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{K})$ de dimension $\dfrac{n(n+1)}{2}$.

Proposition 8. Soit $A \in \mathcal{S}_n(\mathbb{R})$ une matrice symétrique réelle. Deux vecteurs propres de A correspondant à des valeurs propres distinctes sont orthogonaux.

Théorème 8. Théorème spectral
Toute matrice symétrique **réelle** est orthogonalement diagonalisable. C'est-à-dire : pour toute matrice réelle symétrique $A \in \mathcal{S}_n(\mathbb{R})$, il existe une matrice diagonale réelle D, il existe une matrice orthogonale $P \in O(n)$ telles que

$$A = PDP^{-1} = PD(P^T)$$

Remarque Ce théorème nous apprend deux choses : A est diagonalisable et de plus il existe une base orthonormée de $\mathcal{M}_{n,1}(\mathbb{R})$ constituée de vecteurs propres pour A.

Attention C'est faux pour les matrices symétriques complexes. Par exemple $A = \begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$ est symétrique mais non diagonalisable !

Exemple Pour toute matrice $A \in \mathcal{M}_n(\mathbb{R})$, $(A^T)A$ est symétrique réelle et donc diagonalisable orthogonalement.

Corollaire. Soit $u \in \mathcal{L}(\mathbb{R}^n)$ tel que sa matrice dans une base orthonormée de $\mathbb{R}^n$ soit symétrique. Alors u est diagonalisable et il existe une base orthonormée de $\mathbb{R}^n$ (pour le produit scalaire canonique) formée de vecteurs de propres pour u.

Remarque **Les espaces propres de u sont deux à deux orthogonaux**. Cela peut être pratique pour les déterminer.

2.3 Un exemple en géologie : le tenseur des contraintes

2.4 Un exemple probabiliste : matrices de covariance

Définition 9. Soit $X = (X_1, \ldots, X_n)$ un vecteur aléatoire réel tel que, pour tout $i \in [\![1, n]\!]$, X_i admet un moment d'ordre 2. On définit **la matrice de covariance de** X, notée Γ_X, par $\Gamma_X = (\mathrm{Cov}\,(X_i, X_j))_{(i,j)\in[\![1,n]\!]^2}$.

Remarque Sur la diagonale de Γ_X, on retrouve les variances des X_i.

Proposition 9. Si $X = (X_1, \ldots, X_n)$ un vecteur aléatoire réel comme précédemment, alors sa matrice de covariance est symétrique réelle et donc orthogonalement diagonalisable.

Remarque Ce fait est le tout début d'une théorie : l'Analyse en Composante Principale (ACP) qui est utilisée en statistiques, mais aussi en traitement d'image...

Proposition 10. **Effet d'une application linéaire**
Soient $X = (X_1, \ldots, X_n)$ un vecteur aléatoire réel, $A \in \mathcal{M}_{m,n}(\mathbb{R})$ et $Y = AX$. Alors Y est un vecteur aléatoire réel, dont on peut noter les coordonnées $(Y_1, \ldots, Y_m)$, (ce sont m variables aléatoires) et on a

$$\Gamma_Y = A\Gamma_X A^T$$

Remarque On sait qu'il existe $P \in O(n)$ telle que $(P^T)\Gamma_X P$ est diagonale. Ce qui signifie que le vecteur aléatoire $Y = (P^T)X$ a ses coordonnées non corrélées, ce qui peut donner l'illusion de l'indépendance... !

3 Projections orthogonales et distance à un sous-espace vectoriel

3.1 Projections orthogonales

Définition 10. Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E (nécessairement de dimension finie). On a vu que pour tout vecteur $x \in E$, il existe un unique couple
$(x_F, x_{F^\perp}) \in F \times F^\perp$ tel que $x = x_F + x_{F^\perp}$.
On appelle projection orthogonale sur F l'application p qui à tout vecteur $x \in E$ associe l'unique vecteur $x_F \in F$ défini ci-dessus.

Théorème 9. Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E. Soit p la projection orthogonale sur F.
Si $(e_1, \dots, e_m)$ est **une base orthonormée de F** alors, l'expression de p est donnée par :

$$\forall x \in E,\ p(x) = \sum_{k=1}^{m} (x|e_k)e_k.$$

Exemple Si $F = \mathrm{Vect}\,(v)$ avec v un vecteur non nul de E, alors la projection orthogonale sur F est définie par :

$$\forall x \in E,\ p(x) = \frac{(x|v)v}{\|v\|^2}$$

Proposition 11. Propriétés des projections orthogonales
Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E et p la projection orthogonale sur F.

1. p est linéaire
2. $p \circ p = p$
3. $\mathrm{Im} p = \ker(p - \mathrm{Id}_E) = F$
4. $\ker p = F^\perp$

Remarque Sauf le cas pathologique $p = \mathrm{Id}_E$, une projection orthogonale n'est jamais bijective. Et donc la matrice d'une projection orthogonale n'est pas orthogonale ! (car pas inversible)

Proposition 12. Réduction des projections orthogonales
Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E et p la projection orthogonale sur F.

1. p est diagonalisable. Si p n'est pas nulle, ni égale à Id_E, son spectre est égal à $\{0, 1\}$ et $E_1(p) = F$, $E_0(p) = F^\perp$.
2. Dans une base orthonormée de E, la matrice de p est symétrique.

Remarque Merci d'utiliser le point 2 pour vérifier vos calculs !

Proposition 13. Soient E un espace euclidien muni d'un produit scalaire $(.|.)$ et F un sous-espace vectoriel de E et p la projection orthogonale sur F. Si on note $p^\perp$ la projection orthogonale sur $F^\perp$ alors

$$p + p^\perp = \mathrm{Id}_E$$

Exemple Si, dans $\mathbb{R}^3$ muni d'un produit scalaire $(.|.)$, F est un sous-espace vectoriel de dimension 2 donné par un un vecteur normal u, alors la projection orthogonal sur F est définie par :

$$\forall x \in \mathbb{R}^3,\ p(x) = x - \frac{(x|u)u}{\|u\|^2}$$

3.2 Distance à un sous-espace vectoriel

Définition 11. Soient E un espace préhilbertien réel, A une partie non vide de E et $x \in E$. On appelle distance de x à A le réel $\mathrm{d}(x, A) = \inf\{\mathrm{d}(x, a)\ ,\ a \in A\}$.

Remarque A priori cette borne inférieure n'est pas un minimum, i.e. elle n'est pas forcément atteinte.

Théorème 10. Distance à un sous-ev de dim finie
Soient E un espace préhilbertien réel, F un sous-espace vectoriel de dimension finie de E et $x \in E$. On note p la projection orthogonale sur F.

1. $d(x, F) = \|x - p(x)\| = \mathrm{d}(x, p(x))$. La distance de x à F est donc un minimum, atteint en un unique point : le projeté de x sur F.
2. Pour tout $f \in F$, $\|x - f\|^2 = \|p(x) - f\|^2 + \mathrm{d}(x, F)^2$. En particulier, $\mathrm{d}(x, F)^2 = \|x\|^2 - \|p(x)\|^2$.

Application importante : droite de régression linéaire par la méthode des moindres carrés.