1 Généralités

1.1 Produit scalaire

[Produit scalaire] On appelle **produit scalaire** sur E toute forme bilinéaire symétrique définie positive, c'est-à-dire toute application $\phi : E \times E \to \mathbb{R}$ telle que :
- ϕ est bilinéaire : pour tous $x_1, x_2, y_1, y_2 \in E$ et $\lambda \in \mathbb{R}$,

$$\phi(\lambda x_1 + x_2, y_1) = \lambda\phi(x_1, y_1) + \phi(x_2, y_1)$$
$$\phi(x_1, \lambda y_1 + y_2) = \lambda\phi(x_1, y_1) + \phi(x_1, y_2)$$

- ϕ est symétrique : pour tous $x, y \in E$, $\phi(x, y) = \phi(y, x)$.
- ϕ est définie positive : pour tout $x \in E$, $\phi(x, x) \geq 0$ et $\phi(x, x) = 0$ si et seulement si $x = 0_E$.

On note généralement le produit scalaire $\langle \cdot|\cdot \rangle$, $(\cdot|\cdot)$ ou $\langle \cdot, \cdot \rangle$.

[Espaces préhilbertiens réels]
- On appelle **espace préhilbertien réel** tout $\mathbb{R}$-espace vectoriel muni d'un produit scalaire. Notation usuelle : $(E, \langle \cdot|\cdot \rangle)$.
- Un espace préhilbertien réel de dimension finie est appelé **espace euclidien**.

1.2 Exemples fondamentaux

[$\mathbb{R}^n$ muni de son produit scalaire canonique] Le produit scalaire canonique est défini par :

$$\forall X, Y \in \mathbb{R}^n, \langle X|Y \rangle = X^\top Y = \sum_{i=1}^{n} x_i y_i$$

en notant $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

[$E = \mathcal{C}([a, b], \mathbb{R})$] Muni de $(f, g) \mapsto \int_a^b f(t)g(t)dt$.

[$E = \mathbb{R}[X]$] Muni de $(P, Q) \mapsto \int_0^1 P(t)Q(t)dt$.

[$E = \mathcal{M}_n(\mathbb{R})$] Muni de $(A, B) \mapsto \mathrm{Tr}(A^\top B)$.

1.3 Norme euclidienne

[Norme euclidienne] Pour tout vecteur x de l'espace préhilbertien réel E, $\langle x|x \rangle > 0$. On pose alors :
- $\|x\| = \sqrt{\langle x|x \rangle}$ pour tout $x \in E$ (norme euclidienne)
- $d(x, y) = \|x - y\|$ pour tous $x, y \in E$ (distance)

Si $x \neq 0_E$, $\frac{x}{\|x\|}$ est de norme 1, il est dit unitaire.

[Identités remarquables] Soit $(E, \langle \cdot|\cdot \rangle)$ un espace préhilbertien réel. Pour tous $x, y \in E$,
1. $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x|y \rangle$
2. $\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x|y \rangle$

3. Identité du parallélogramme : $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$
4. Identité de polarisation : $\langle x|y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$

[Inégalité de Cauchy-Schwarz] Soit $(E, \langle\cdot|\cdot\rangle)$ un espace préhilbertien réel. Alors,

$$\forall x, y \in E, |\langle x|y\rangle| \leq \|x\| \cdot \|y\|$$

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

Soient $x, y \in E$ et $\lambda \in \mathbb{R}$.

— Si $x = 0_E$, le résultat est immédiat. Supposons $x \neq 0_E$.

— $\langle \lambda x + y|\lambda x + y\rangle \geq 0$ et $\langle \lambda x + y|\lambda x + y\rangle = \lambda^2\langle x|x\rangle + 2\lambda\langle x|y\rangle + \langle y|y\rangle$.

C'est un trinôme en λ de signe constant donc $\Delta \leq 0$:

$$\Delta = (2\langle x|y\rangle)^2 - 4\langle x|x\rangle\langle y|y\rangle \leq 0$$

D'où $|\langle x|y\rangle| \leq \sqrt{\langle x|x\rangle}\sqrt{\langle y|y\rangle} = \|x\| \cdot \|y\|$.

Cas d'égalité : $\Delta = 0$ donc il existe λ_0 tel que $\lambda_0 x + y = 0_E$.

[Norme] L'application $\|\cdot\|$ est une norme sur E.

2 Orthogonalité

2.1 Vecteurs orthogonaux

[Vecteurs orthogonaux] Deux vecteurs x et y de E sont dits **orthogonaux** si $\langle x|y\rangle = 0$.

[Pythagore] Soient $x, y \in E$. $\|x+y\|^2 = \|x\|^2 + \|y\|^2 \iff \langle x|y\rangle = 0$.

Le vecteur nul est le seul vecteur orthogonal à tous les autres.

2.2 Familles orthogonales et orthonormales

[Familles orthogonales et orthonormales]

— Une famille $(e_i)_{i\in I}$ est dite **orthogonale** si :

$$\forall (i,j) \in I^2, i \neq j \implies \langle e_i|e_j\rangle = 0$$

— Elle est dite **orthonormale** si de plus : $\forall i \in I, \|e_i\| = 1$.

[Pythagore généralisé] Soit $(e_1, \ldots, e_n)$ une famille orthogonale de vecteurs de E. Alors,

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

Une famille orthogonale constituée de vecteurs non nuls est libre. En particulier, toute famille orthonormale est libre.

[Décomposition dans une base orthonormée] Soient E un espace euclidien de dimension $n \in \mathbb{N}^*$ et $(e_1, \ldots, e_n)$ une base orthonormée de E. Alors,

$$\forall x \in E, x = \sum_{i=1}^{n} \langle x|e_i\rangle e_i$$

2.3 Procédé d'orthonormalisation de Gram-Schmidt

[Gram-Schmidt] Soit $(u_1, \dots, u_p)$ une famille libre de vecteurs de E. Il existe une unique famille orthonormée $(e_1, \dots, e_p)$ telle que :

1. $\forall k \in \{1, \dots, p\},\ \mathrm{Vect}(e_1, \dots, e_k) = \mathrm{Vect}(u_1, \dots, u_k)$
2. $\forall k \in \{1, \dots, p\},\ \langle u_k, e_k \rangle > 0$

Algorithme :

1. Initialisation : $e_1 = \frac{u_1}{\|u_1\|}$
2. Pour k de 2 à p :

$$e'_k = u_k - \sum_{i=1}^{k-1} \langle u_k, e_i \rangle e_i$$
$$e_k = \frac{e'_k}{\|e'_k\|}$$

[Orthonormalisation dans $\mathbb{R}^3$] Soit $\mathcal{B} = (u_1, u_2, u_3)$ avec $u_1 = (1,1,0)$, $u_2 = (1,0,1)$, $u_3 = (0,1,1)$.

Étape 1 : $\|u_1\| = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$
$e_1 = \frac{u_1}{\|u_1\|} = \frac{1}{\sqrt{2}}(1,1,0)$
Étape 2 : $\langle u_2, e_1 \rangle = \frac{1}{\sqrt{2}}(1 \times 1 + 0 \times 1 + 1 \times 0) = \frac{1}{\sqrt{2}}$
$e'_2 = u_2 - \langle u_2, e_1 \rangle e_1 = (1,0,1) - \frac{1}{2}(1,1,0) = (\frac{1}{2}, -\frac{1}{2}, 1)$
$\|e'_2\| = \sqrt{(\frac{1}{2})^2 + (-\frac{1}{2})^2 + 1^2} = \sqrt{\frac{6}{4}} = \frac{\sqrt{6}}{2}$
$e_2 = \frac{2}{\sqrt{6}}(\frac{1}{2}, -\frac{1}{2}, 1) = \frac{1}{\sqrt{6}}(1,-1,2)$
Étape 3 : $\langle u_3, e_1 \rangle = \frac{1}{\sqrt{2}}$, $\langle u_3, e_2 \rangle = \frac{1}{\sqrt{6}}(0 - 1 + 2) = \frac{1}{\sqrt{6}}$
$e'_3 = u_3 - \frac{1}{\sqrt{2}}e_1 - \frac{1}{\sqrt{6}}e_2 = (0,1,1) - \frac{1}{2}(1,1,0) - \frac{1}{6}(1,-1,2) = (-\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$
$\|e'_3\| = \frac{2}{3}\sqrt{3}$
$e_3 = \frac{3}{2\sqrt{3}}(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}) = \frac{1}{\sqrt{3}}(-1,1,1)$
La base orthonormée obtenue est donc :

$$\left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right)$$

2.4 Orthogonal d'un sous-espace vectoriel

[Orthogonal] Soit F un sous-espace vectoriel de E. On appelle **orthogonal** de F l'ensemble :

$$F^\perp = \{x \in E \mid \forall y \in F, \langle x|y \rangle = 0\}$$

Soient F et G deux sous-espaces vectoriels de E.
— $F^\perp$ est un sous-espace vectoriel de E
— Si $F \subset G$, alors $G^\perp \subset F^\perp$
— $F \subset (F^\perp)^\perp$
Soit F un sous-espace vectoriel de dimension finie de E. Alors, $E = F \oplus F^\perp$.

2.5 Projection orthogonale et distance

[Projecteur orthogonal] Soit F un sous-espace vectoriel de E de dimension finie. On appelle **projection orthogonale** sur F la projection sur F parallèlement à $F^{\perp}$.

En notant p la projection orthogonale sur F, sous-espace vectoriel de E de dimension finie n,

— Si $x \in E$, $p(x)$ est caractérisé par : $p(x) \in F$ et $x - p(x) \in F^{\perp}$

— Si $(e_1, \dots, e_n)$ est une base orthonormale de F, alors :

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

[Distance] Soient $x \in E$ et F un sous-espace vectoriel de E de dimension finie. On appelle **distance** de x à F le réel :

$$d(x, F) = \inf_{u \in F} \|x - u\|$$

Soient $x \in E$ et F un sous-espace vectoriel de E de dimension finie. Alors :

$$d(x, F) = \|x - p(x)\|$$

où p est la projection orthogonale sur F.

3 Suites totales (pour votre culture)

[Suite totale] On dit qu'une suite de vecteurs $(e_i)_{i \in \mathbb{N}}$ de E est **totale** si pour tout $x \in E$,

$$\forall \epsilon > 0, \exists y \in \mathrm{Vect}(e_i), \|x - y\| < \epsilon$$

Soient $(e_i)_{i \in \mathbb{N}}$ une suite orthonormale totale d'éléments de E et pour tout $n \in \mathbb{N}$, p_n le projecteur orthogonal sur $\mathrm{Vect}(e_0, \dots, e_n)$. Alors, pour tout x de E, $(p_n(x))_{n \in \mathbb{N}}$ converge vers x.

[Égalité de Parseval] Si $(e_i)_{i \in \mathbb{N}}$ est une suite orthonormale totale d'éléments de E, alors pour tout $x \in E$:

$$\|x\|^2 = \sum_{i=0}^{+\infty} \langle x | e_i \rangle^2$$

4 Méthode des moindres carrés

4.1 Principe général

On dispose de n points $(x_i, y_i)_{1 \leq i \leq n}$ et on cherche à trouver la droite $y = ax + b$ qui minimise la somme des carrés des écarts :

$$\min_{(a,b) \in \mathbb{R}^2} \sum_{i=1}^{n} (y_i - ax_i - b)^2$$

4.2 Formulation géométrique

Dans $\mathbb{R}^n$, on pose :
— $Y = (y_1, \ldots, y_n)^T$ (vecteur des observations)
— $X = (x_1, \ldots, x_n)^T$ (vecteur des variables explicatives)
— $U = (1, \ldots, 1)^T$ (vecteur constant)
Le problème revient à trouver la projection orthogonale de Y sur $\mathrm{Vect}(X, U)$.

4.3 Système normal

Les conditions d'orthogonalité donnent le système :

$$\begin{cases} \langle Y - aX - bU, X \rangle = 0 \\ \langle Y - aX - bU, U \rangle = 0 \end{cases}$$

Ce qui conduit à :

$$\begin{cases} a \sum x_i^2 + b \sum x_i = \sum x_i y_i \\ a \sum x_i + nb = \sum y_i \end{cases}$$

4.4 Exemple numérique

Soient les points : $(1, 2)$, $(2, 3)$, $(3, 5)$, $(4, 7)$, $(5, 8)$.
Calculons les coefficients :

$$\sum x_i = 15, \qquad \sum y_i = 25,$$
$$\sum x_i^2 = 55, \qquad \sum x_i y_i = 91$$

Le système devient :

$$\begin{cases} 55a + 15b = 91 \\ 15a + 5b = 25 \end{cases}$$

Solution : $a = 1.5$, $b = 0.5$. La droite des moindres carrés est donc $y = 1.5x + 0.5$.

4.5 Matriciellement

Le problème peut se reformuler comme :

$$\min_{\beta \in \mathbb{R}^2} \|Y - M\beta\|^2$$

où $M = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$ et $\beta = \begin{pmatrix} a \\ b \end{pmatrix}$.
La solution est donnée par :

$$\beta = (M^T M)^{-1} M^T Y$$

Ce qui généralise à des modèles polynomiaux ou multidimensionnels.