12 Produit scalaire dans $\mathbb{R}^n$

Les objectifs : Mise en place de deux résultats fondamentaux pour les applications : la projection orthogonale sur un sous-espace d'une part et la diagonalisation des matrices symétriques d'autre part.
Dans ce chapitre, les mots « vecteurs » et « points » peuvent être considérés comme interchangeables.

✐ Dans l'ensemble du chapitre, E désigne $\mathbb{R}^n$ qui est un $\mathbb{R}$-espace vectoriel.

1 Produit scalaire dans $\mathbb{R}^n$

1.1 Produit scalaire.

Définition

Définition 1.1 : Produit scalaire usuel de $\mathbb{R}^n$

On appelle `produit scalaire usuel dans` $\mathbb{R}^n$ *l'application ϕ de $E \times E$ dans $\mathbb{R}$ définie par :*

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

où $x = (x_1, \cdots, x_n)$ et $y = (y_1, \cdots, y_n)$

$\mathbb{R}^n$ muni de son produit scalaire usuel est dit `espace euclidien`.

Notations : Selon les cas, on trouvera pour désigner le produit scalaire de x par y les notations :

$$\phi(x, y) =< x, y >= (x|y)$$

Remarque

Remarque 1.1. Écriture matricielle

Si $X = \mathcal{M}_{\mathcal{B}}(x)$ et $Y = \mathcal{M}_{\mathcal{B}}(y)$ où $\mathcal{B}$ désigne la base canonique de $\mathbb{R}^n$, alors $(x|y) = {}^tXY = {}^tYX$

Propriété

prop.1.1.Bilinéarité

$\forall x, y, z \in E$, $\forall \lambda \in \mathbb{R}$,

$$\phi(\lambda x + y, z) = \lambda\phi(x, z) + \phi(y, z) \Leftrightarrow (\lambda x + y|z) = \lambda(x|z) + (y|z)$$

$$\phi(x, \lambda y + z) = \lambda\phi(x, y) + \phi(x, z) \Leftrightarrow (x|\lambda y + z) = \lambda(x|y) + (x|z)$$

Propriété

prop.1.2 Forme symétrique définie positive

$\forall x, y \in E$,

① $\phi(x,y) = \phi(y,x)$ *[ϕ est* **symétrique***]*

② $\phi(x,x) \geq 0$ *[ϕ est* **positive***]*

③ $\phi(x,x) = 0 \Rightarrow x = 0_E$ *[ϕ est dite* **définie***]*

Remarque

Remarque 1.2. Produit scalaire

Si E est un $\mathbb{R}$-e.v., on appelle `produit scalaire sur` E *toute forme bilinéaire ϕ symétrique, positive et définie.*

Remarque

Remarque 1.3

Soit $\mathcal{F} = \{u_1, \cdots, u_p\}$ et $\mathcal{G} = \{v_1, \cdots, v_q\}$ deux familles de vecteurs de E.

Si $u \in$ Vect$\{F\} \Leftrightarrow \exists \lambda_1, \cdots, \lambda_p \in \mathbb{R}/u = \sum_{i=1}^{p} \lambda_i u_i$ et $v \in$ Vect$\{G\} \Leftrightarrow \exists \mu_1, \cdots, \mu_q \in \mathbb{R}/v = \sum_{j=1}^{q} \mu_j v_j$

alors :

$$(u|v) = \left(\sum_{i=1}^{p} \lambda_i u_i \middle| \sum_{j=1}^{q} \mu_j v_j\right) = \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_i \mu_j (u_i|v_j)$$

1.2 Norme euclidienne

Définition

Définition 1.2 : Norme euclidienne

On appelle `norme euclidienne associée au produit scalaire` ϕ *l'application définie sur E par :*

$$x \longmapsto \sqrt{\phi(x,x)} = \sqrt{(x|x)} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

On notera $\|x\|$ la norme de x associée au produit scalaire $< . >$ et on remarque que $\|x\|^2 = {}^t X X$

Propriété

prop.1.3. Propriétés

$\forall x, y \in E$, $\forall \lambda \in \mathbb{R}$:

- $\|x\| \geq 0$ *et* $\|x\| = 0 \Leftrightarrow x = 0$
- $\|\lambda x\| = |\lambda| \, \|x\|$
- $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2(x|y)$

Remarque

Remarque 1.4

On dira qu'un vecteur $x \in E$ est `normé` *ou* `unitaire` *si $\|x\| = 1$.*
Soit $x \in E / x \neq 0_E$ et $\|x\| \neq 1$. On forme un vecteur normé de E colinéaire à x en écrivant :

$$y = \frac{1}{\|x\|}x$$

Propriété

prop.1.4. Inégalité de Cauchy-Schwarz

Soit $x, y \in E$, *alors :* $(x|y)^2 \leq (x|x) \cdot (y|y)$, *ou encore :*

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

Remarque

Remarque 1.5

L'inégalité précédente s'écrit pour le produit scalaire dans $\mathbb{R}^n$:

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

Exemple

Application à l'obtention d'inégalités

Pour tout $(x_1, \cdots, x_n) \in \mathbb{R}^n$: $(x_1 + \cdots + x_n)^2 \leq n(x_1^2 + \cdots + x_n^2)$

Remarque

Remarque 1.6

Si x et y sont deux vecteurs non nuls, alors :

$$\frac{(x|y)}{\|x\|\,\|y\|} \in [-1, 1] \Rightarrow \exists! \alpha \in [0, \pi] / cos(\alpha) = \frac{(x|y)}{\|x\|\,\|y\|}$$

Par définition, cet angle unique est appelé `angle des deux vecteurs non nuls` x `et` y.

Propriété

prop.1.5. Inégalité triangulaire

$\forall x, y \in E,$

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

1.3 Bases orthonormées

Définition

Définition 1.3 : vecteurs orthogonaux

On dit que deux vecteurs x et y de E sont orthogonaux si $(x|y) = 0$. On note alors $x \perp y$.

Remarque

Remarque 1.7

Par symétrie du produit scalaire, si $(x|y) = 0$, alors $(y|x) = 0$, ce qui justifie la symétrie de la définition précédente.

Définition

Définition 1.4 : sous-espaces vectoriels orthogonaux

Soient F et G deux sous-espaces vectoriels de E.
F et G sont dit `orthogonaux` *si tout vecteur de F est orthogonal à chaque vecteur G ou encore :*

$$\forall x \in F, \forall y \in G, (x|y) = 0$$

Remarque

Remarque 1.8

Soit F un sous-espace vectoriel de $\mathbb{R}^n$. On désigne par $F^\perp$ l'orthogonal de F défini par :

$$F^\perp = \{x \in \mathbb{R}^n / \forall y \in F, (x|y) = 0\}$$

On montre facilement que $F^\perp$ est un sous-espace vectoriel de $\mathbb{R}^n$ tel que $F^\perp \cap F = \{0_{\mathbb{R}^n}\}$.
Par ailleurs, si $B_F = (u_1, \cdots, u_q)$ est une base de F, alors :

$$x \in F^\perp \Leftrightarrow \forall i \in [\![1, q]\!],\ (x|u_i) = 0$$

Propriété

Prop. 1.6. Théorème de Pythagore

Deux vecteurs x et y sont orthogonaux si, et seulement si :

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2$$

Définition

Définition 1.5 : Famille orthonormée de $\mathbb{R}^n$ ou d'un ss-e.v. de $\mathbb{R}^n$

Soit F un sous-espace vectoriel de $\mathbb{R}^n$ (éventuellement $F = \mathbb{R}^n$) et soit $\{x_i\}_{i \in I}$ une famille de vecteurs de F.
$\{x_i\}_{i \in I}$ est une famille orthonormée si les éléments de cette famille sont orthogonaux deux à deux et de norme égale à 1.

Propriété

Prop. 1.7.

Une famille orthogonale de vecteurs non nuls de F est libre.
Si le cardinal de cette famille est égale à la dimension de F, il s'agit d'une base `orthogonale` *de F.*
Si la famille est orthonormale, on parlera de base `orthonormale` *de F*

Exemple

Exemple 1.1

Si $\mathbb{R}^n$ est muni de son produit scalaire usuel, alors sa base canonique est orthonormée

Remarque

Remarque 1.9

Le produit scalaire et la norme se calculent de la même manière dans toutes les bases orthonormales.

Propriété

Prop. 1.8. Coordonnées dans une b.o.n

Soit $\mathcal{B} = (u_1, \cdots, u_n)$ une base orthonormé de $\mathbb{R}^n$. Alors pour tout $x \in \mathbb{R}^n$, on a :

$$x = \sum_{k=1}^{n} (x|u_k)u_k \text{ et } \|x\|^2 = \sum_{k=1}^{n} (x|u_k)^2$$

Propriété

Prop. 1.9. Interprétation matricielle

Soit $\mathcal{B}$ la base canonique de $\mathbb{R}^n$ et $\mathcal{B}' = (u_1, \cdots, u_n)$ une base orthonormée de $\mathbb{R}^n$.
Si P désigne la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$, alors : $\boxed{{}^tPP = I_n}$.
On dit que P est une `matrice orthogonale`. Elle est inversible et vérifie $P^{-1} = {}^tP$.

2 Matrices symétriques réelles

Propriété

Prop. 2.1. Spectre des matrices symétriques réelles

Soit A une matrice symétrique réelle de $\mathcal{M}_n(\mathbb{R})$. Alors $Sp(A) \subset \mathbb{R}$

Remarque

Remarque 1.8

Si A est une matrice symétrique réelle, alors les vecteurs propres associés à des valeurs propres distinctes sont orthogonaux

Propriété

Prop. 2.2. Diagonalisation des matrices symétriques

Toute matrice carrée symétrique réelle est diagonalisable dans une base orthonormale de vecteurs propres.
Ou encore :
Soit A une matrice symétrique réelle de $\mathcal{M}_n(\mathbb{R})$. Alors il existe une matrice D diagonale et une matrice P inversible telles que $A = PDP^{-1}$ avec $P^{-1} = {}^tP$

Exemple

Exemple 2.1

Montrer que $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ *est diagonalisable dans une base orthonormale de vecteurs propres*

3 Projection orthogonale

Définition

Définition 2.1 : distance entre vecteurs et d'un vecteur à une partie non vide

Soit x, y *deux vecteurs de* $\mathbb{R}^n$ *et* $\mathcal{A}$ *une partie non vide de* E. *Alors :*

- $d(x,y) = \|y-x\| = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}.$
- $d(x,\mathcal{A}) = \underset{y\in\mathcal{A}}{inf}\, d(x,y) = inf\{\|x-y\|, y \in A\}$

Définition

Définition 2.2 : Projection orthogonale

Soit F *un sous-espace vectoriel de* $\mathbb{R}^n$.
On appelle projection orthogonale sur F *de* $\mathbb{R}^n$ *un endomorphisme* p *de* $\mathbb{R}^n$ *défini par :*

$$\forall x \in \mathbb{R}^n,\ p(x) \in F/(x-p(x)|y) = 0,\ \forall y \in F$$

Propriété

Prop. 2.3. Existence et unicité de la projection sur un ssev

Soit $\mathcal{B}_F = (u_1, \cdots, u_q)$ *une base orthonormée d'un sous-espace vectoriel* F *de* $\mathbb{R}^n$.
Si p *est la projection orthogonale de* $\mathbb{R}^n$ *sur* F, *alors :*

$$\forall x \in \mathbb{R}^n,\ p(x) = \sum_{i=1}^{q}(u_i|x)u_i$$

Propriété

Prop. 2.4.

Soit F *un sous-espace vectoriel de* $\mathbb{R}^n$ *dont* $B_F = (u_1, \cdots, u_q)$ *est une base et soit* p *la projection orthogonale sur* F *de* $\mathbb{R}^n$. *Alors :*

① $p \circ p = p$

② $Im(p) = F$

③ $Ker(p) = F^{\perp}$

Remarque

Remarque 1.9

On vient d'obtenir que pour tout SEV F de $\mathbb{R}^n$, $\dim(F)+\dim(F^\perp)=n=\dim(\mathbb{R}^n)$ et puisque $F\cap F^\perp=\{0\}$, on construit une base orthonormale de $\mathbb{R}^n$ en juxtaposant une base orthonormale de F et de $F^\perp$.

Exemple

Exemple 2.2

Soit $\mathcal{B}$ la base canonique de $\mathbb{R}^n$ et soit p la projection orthogonale sur F, droite vectorielle engendrée par le vecteur a qu'on supposera normé. Alors :

$$\forall x\in\mathbb{R}^n,\ p(x)=(a|x)a$$

Si on suppose que $a=(a_1,\cdots,a_n)_{\mathcal{B}}$, et $\|a\|=1$ alors :

$$M=\mathcal{M}_{\mathcal{B}}(p)=\begin{pmatrix} a_1^2 & a_1a_2 & \cdots & a_1a_n \\ a_1a_2 & a_2^2 & \cdots & a_2a_n \\ \vdots & \vdots & & \vdots \\ a_1a_n & a_2a_n & \vdots & a_n^2 \end{pmatrix}\in\mathcal{S}_n(\mathbb{R})$$

Exemple

Exemple 2.3

Soit $E=\mathbb{R}^3$ et $F=\{(x,y,z)\in\mathbb{R}^3/x+y-z=0\}$.
Écrire les projections orthogonales p et q respectivement sur F et sur $F^\perp$

Propriété

Prop. 2.5. Distance à un sous-espace vectoriel

Soit F un sous-espace vectoriel de $\mathbb{R}^n$. Alors pour tout $x\in\mathbb{R}^n$,

$$\boxed{d(x,F)=inf\{\|x-y\|,y\in F\}=\|x-p(x)\|}$$

⚠ Soit F sous-espace vectoriel de $\mathbb{R}^n$ et p la projection orthogonale sur F de $\mathbb{R}^n$. Alors :

$\forall x\in\mathbb{R}^n$, $\|x-p(x)\|=inf\{\|x-y\|,y\in F\}$ et $p(x)$ est l'unique vecteur de F qui réalise ce minimum

La projection orthogonale de x sur F peut être vue comme la meilleure approximation de x par un élément de F .

Conséquence : L'ajustement affine d'un nuage de points $\{(x_i,y_i),i\in[\![0,n]\!]\}$ par la méthode des moindres carrés peut-être interprété en termes de projection sur un sous-espace de dimension 2.

Géométrie et principales caractéristiques de séries statistiques

① **Introduction :** On considère une série statistique multivariée $S = \{M_x, M_y\}$ de taille n portant sur deux caractères x et y de telle façon que $M_x = \{x_1, \cdots, x_n\}$ et $M_y = \{y_1, \cdots, y_n\}$, les x_i étant supposés deux à deux distincts (On parle parfois de variable « explicative » pour x tandis que y est la variable « expliquée ».
Cette série peut-être représentée par un nuage de points $M_i(M_{x_i}, M_{y_i})$ pour i compris entre 1 et n et $\mathbb{R}^2$ devient l'espace des individus.

On supposera $\mathbb{R}^n$ muni de sa base canonique $\mathcal{B} = (e_1, \cdots, e_n)$ et de son produit scalaire usuel.
Dès lors, M_x et M_y peuvent être considérés comme des vecteurs de $\mathbb{R}^n$ qui devient l'espace des variables.

On notera par la suite $U = (1, \cdots, 1)$, vecteur de $\mathbb{R}^n$ de norme $\|U\|^2 = n$.

② **Moyenne et variance :**

- Moyenne : $\bar{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i = \boxed{\dfrac{1}{n}(M_x|U)}$ et de même : $\bar{y} = \boxed{\dfrac{1}{n}(M_y|U)}$.

  ✐ *Remarque* : $\bar{x}U$ peut se lire comme le projeté orthogonal de M_x sur $\mathrm{Vect}\{U\} = \mathrm{Vect}\{U_1\}$ où $U_1 = \dfrac{U}{\|U\|} = \dfrac{U}{\sqrt{n}}$. En effet $p(M_x) = (M_x|U_1)U_1 = \dfrac{(M_x|U)}{n}U = \bar{x}U$.

  Par la suite, on centrera les séries statistiques M_x et M_y en posant :

$$V_x = (x_1 - \bar{x}, \cdots, x_n - \bar{x}) = M_x - \bar{x}U = M_x - \frac{(M_x|U)}{\|U\|^2}U \text{ et } V_y = M_y - \bar{y}U$$

  V_x et V_y sont deux vecteurs orthogonaux à U.

- Variance : $\boxed{s_x^2 = \displaystyle\sum_{i=1}^{n} \dfrac{1}{n}(x_i - \bar{x})^2 = \dfrac{\|V_x\|^2}{n}}$ et $\boxed{s_y^2 = \dfrac{\|V_y\|^2}{n}}$.

③ **Covariance et coefficient de corrélation :**

$$\boxed{(V_x|V_y)} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{x}) = \boxed{n\mathrm{Cov}(x, y) = ns_{x,y}}$$

Dès lors :

$$\boxed{\cos(V_x, V_y)} = \frac{(V_x|V_y)}{\|V_x\|\,\|V_y\|} = \frac{ns_{x,y}}{ns_x s_y} = \boxed{\frac{s_{x,y}}{s_x s_y} = r_{x,y}}$$

Conséquence :

- Si $r_{x,y} = 0$, alors $\cos(V_x, V_y) = 0$ et donc $V_x \perp V_y$. On dira que x et y sont deux variables `orthogonales` (c'est en particulier le cas lorsqu'elles sont indépendantes).
- Si $|r_{x,y}| = 1$ alors $\cos(V_x, V_y) = \pm 1$ et donc V_x et V_y sont colinéaires. On en déduit que $\exists a \in \mathbb{R}/V_y = aV_x$ ou encore :

$$M_y - \bar{y}U = a(M_x - \bar{x}U) \text{ soit } \boxed{M_y = aM_x + (\bar{y} - a\bar{x})U}$$

- Si $\dfrac{\sqrt{3}}{2} \leq |r_{x,y}| < 1$ soit environ $0,866 \leq |r_{x,y}| < 1$, alors $\alpha \in \left[-\dfrac{\pi}{6}, \dfrac{\pi}{6}\right] \setminus \{0\}$. On dira qu'il existe une bonne corrélation linéaire entre x et y.
- Si $\dfrac{1}{2} \leq |r_{x,y}| < \dfrac{\sqrt{3}}{2}$, soit environ $0.5 \leq |r_{x,y}| < 0.866$ alors $|\alpha| \in \left]-\dfrac{\pi}{6}, \dfrac{\pi}{3}\right]$. On dira que la corrélation linéaire entre x et y est médiocre.
- Si $0 < |r_{x,y}| < \dfrac{1}{2}$ on dira que la corrélation est mauvaise.

④ **Prédicteur linéaire :**

Il s'agit de déterminer l'équation d'une droite qui réalisera un ajustement affine du nuage de points $(x_i, y_i)_{i \in [\![0,n]\!]}$, autrement dit d'une droite qui « approche au mieux » le nuage.

Le critère retenu consiste à minimiser $\displaystyle\sum_{i=1}^{n} M_iH_i^2$ où $H_i(x_i, ax_i + b)$.

Pour ça on pose, a et b étant deux réels, $P_{a,b} = \{ax_i + b = \widetilde{y_i}, i \in [\![1, n]\!]\} = aM_x + bU$.

En chaque point, l'erreur commise par l'ajustement est : $e_i = y_i - (ax_i + b) = y_i - P_{a,b}(x_i)$.

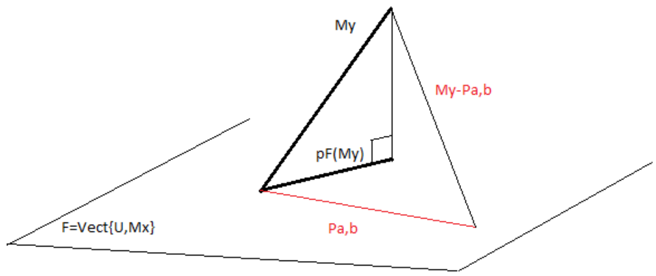
On peut donc constituer le vecteur des erreurs $M_e = \{e_i, 1 \leq i \leq n\} = \{M_iH_i, 1 \leq i \leq n\}$

Conclusion : On cherche à déterminer $a, b \in \mathbb{R} / \|M_e\|^2 = \|M_y - P_{a,b}\|^2$ soit minimale.

Si on désigne par F le plan vectoriel engendré par U et V_x, alors $\|M_y - P_{a,b}\|$ n'est autre que la distance de M_y à l'élément $P_{a,b}$ de F.

Rechercher une droite d'équation $y = ax + b = P_{a,b}(x)$ telle que la somme $\displaystyle\sum_{i=1}^{n} M_iH_i^2$ soit minimale revient donc à chercher $P_{a,b}$ de sorte que la distance $\|M_y - P_{a,b}\|$ soit minimale. Ceci est réalisé si, et seulement si :

$$\boxed{P_{a,b} \text{ est la projection orthogonale de } M_y \text{ sur le plan vectoriel } F}$$



Expression de la projection orthogonale sur F :

- On commence par former une base orthonormée de $F = \text{Vect}\{U, V_x\}$:
  $\{U, V_x\}$ est une famille orthogonale de F car $(U|V_x) = 0$. Il suffit donc de la normer :
  Prenons $U_1 = \dfrac{1}{\sqrt{n}}U$ et $W_x = \dfrac{V_x}{\|V_x\|} = \dfrac{V_x}{\sqrt{n}s_x}$.
  Alors $\mathcal{B}_F = (U_1, W_x)$ est une base orthonormée de F.

- D'après la proposition 2.3, si p_F désigne la projection orthogonale sur F, alors :
  $$p_F(V_y) = (V_y|U_1)U_1 + (V_y|W_x)W_x = (V_y|W_x)W_x \text{ car } V_y \text{ et } U \text{ sont orthogonaux}$$
  Dès lors :
  $$p_F(V_y) = \frac{1}{ns_x^2}(V_y|V_x)V_x = \frac{1}{ns_x^2}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})(M_x - \bar{x}U) = \frac{s_{x,y}}{s_x^2}(M_x - \bar{x}U)$$
  or par linéarité de la projection :
  $$p_F(V_y) = p_F(M_y - \bar{y}U) = p_F(M_y) - \bar{y}p_F(U) = p_F(M_y) - \bar{y}U \text{ car } U \in F$$
  ou encore :
  $$p_F(M_y) = \frac{s_{x,y}}{s_x^2}(M_x - \bar{x}U) + \bar{y}U = aM_x + bU \text{ où } a = \frac{s_{x,y}}{s_x^2} \text{ et } b = (\bar{y} - a\bar{x})U$$

Conclusion : $\boxed{\text{La droite de régression de } y \text{ en } x \text{ est la droite d'équation } y = ax + b \text{ où } a = \dfrac{s_{x,y}}{s_x^2} \text{ et } b = \bar{y} - a\bar{x}}$