Notes de cours 11

$\mathbb{R}^n$ Euclidien

Table des matières

1 Quelques problèmes de minimum

1.1 L'espérance

Théorème 1. *Soit X une v.a. réelle admettant une variance alors la fonction f définie par*

$$\forall a \in \mathbb{R},\, f(a) = \mathbb{E}((X-a)^2)$$

a pour minimum $\mathbb{V}(X)$. Ce minimum est atteint uniquement en $a = \mathbb{E}(X)$.

Démonstration. On a, pour $a \in \mathbb{R}$, $f(a) = a^2 - 2a.\mathbb{E}(X) + \mathbb{E}(X^2)$. Ce trinôme atteint son minimum en son unique point critique vérifiant $f'(a) = 0$, *i.e.* $2(a - \mathbb{E}(X)) = 0$, *i.e.* $a = \mathbb{E}(X)$. Le minimum vaut $\mathbb{E}((X-\mathbb{E}(X))^2) = \mathbb{V}(X)$ □

Reformulation

Soit $(x_1, \ldots, x_n) \in \mathbb{R}^n$, soit N une v.a. uniformément distribuée sur $\{1, \ldots, n\}$ et soit $X = x_N$. On a

$$\mathbb{E}(X) = \frac{1}{n}\sum_{k=1}^{n} x_n,\, \mathbb{V}(X) = \frac{1}{n}\sum_{k=1}^{n} x_n^2 - \left(\frac{1}{n}\sum_{k=1}^{n} x_n\right)^2,$$

La fonction f définie par

$$\forall a \in \mathbb{R},\, f(a) = \frac{1}{n}\sum_{k=1}^{n}(x_k - a)^2$$

atteint son minimum uniquement en $a = \frac{1}{n}\sum_{k=1}^{n} x_k$. Ce minimum vaut $\mathbb{V}(X)$.

1.2 La covariance : moindres carrés

Théorème 2. *Soit X, Y deux v.a. réelles admettant une variance,* centrées, X non constante *alors la fonction f définie par*

$$\forall a, b \in \mathbb{R},\, f(a,b) = \mathbb{E}((Y-(aX+b))^2)$$

a pour minimum $\frac{\mathbb{V}(X)\mathbb{V}(Y) - \mathbb{C}ov(X,Y)^2}{\mathbb{V}(X)}$. Ce minimum est atteint uniquement en $(a_0, b_0) = (\frac{\mathbb{C}ov(X,Y)}{\mathbb{V}(X)}, 0)$.

Démonstration. Remarque : On suppose que X est non constante, car sinon, $X = 0$ et on est ramenés au cas précédent.

On a, pour $a, b \in \mathbb{R}$, $f(a,b) = \mathbb{V}(Y) + a^2\mathbb{V}(X) - 2.a\mathbb{C}ov(X,Y) + b^2$.

Cette fonction de deux variables est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et un couple (a,b) est critique pour f ssi

$$\frac{\partial f}{\partial a}(a,b) = 0 \text{ et } \frac{\partial f}{\partial b}(a,b) = 0 \; i.e. \; a = \frac{\mathbb{C}ov(X,Y)}{\mathbb{V}(X)} = 0 \text{ et } b = 0$$

Ce couple $(a_0, b_0) = (\frac{\mathbb{C}ov(X,Y)}{\mathbb{V}(X)}, 0)$ minimise f.

En effet, si $(a,b) \in \mathbb{R}^2$, en utilisant la technique de la forme canonique,

$$f(a,b) = \mathbb{V}(Y) + b^2 + \mathbb{V}(X)(a-a_0)^2 - \mathbb{V}(X)a_0^2$$

et donc

$$f(a,b) - f(a_0,b_0) = b^2 + \mathbb{V}(X)(a-a_0)^2 \geq 0$$

f atteint donc son minimum en son unique point critique (a_0, b_0) et ce minimum vaut $\mathbb{V}(Y) - \frac{\mathbb{C}ov(X,Y)^2}{\mathbb{V}(X)}$.

□

Après décentrage, on obtient que la variable Z, fonction affine de X, la plus proche de Y au sens des moindres carrés est donnée par

$$Z = \frac{\mathbb{C}ov(X,Y)}{\mathbb{V}(X)}(X - \mathbb{E}(X)) + \mathbb{E}(Y)$$

Calculer Z de la sorte s'appelle effectuer la *régression linéaire* de Y sur X.

Reformulation : droite des moindres carrés

Soit $(x_1,\dots,x_n),(y_1,\dots,y_n) \in \mathbb{R}^n$, soit N une v.a. uniformément distribuée sur $\{1,\dots,n\}$ et soit $X = x_N, Y = y_N$. On a

$$\mathbb{E}(X) = \frac{1}{n}\sum_{k=1}^{n} x_k,\ \mathbb{V}(X) = \frac{1}{n}\sum_{k=1}^{n} x_k^2 - \left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)^2,$$

On a, de même,

$$\mathbb{E}(Y) = \frac{1}{n}\sum_{k=1}^{n} y_k,\ \mathbb{V}(Y) = \frac{1}{n}\sum_{k=1}^{n} y_k^2 - \left(\frac{1}{n}\sum_{k=1}^{n} y_k\right)^2,$$

On a, de même,

$$\mathbb{C}ov(X,Y) = \frac{1}{n}\sum_{k=1}^{n} x_k.y_k - \left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)\left(\frac{1}{n}\sum_{k=1}^{n} y_k\right)$$

La fonction f définie par

$$\forall a,b \in \mathbb{R},\ f(a,b) = \frac{1}{n}\sum_{k=1}^{n}(y_k - (a.x_k + b))^2$$

atteint son minimum uniquement en $(a_0,b_0) = ...$

1.3 La distance d'un point à une droite vectorielle

Distance dans $\mathbb{R}^3$

La distance euclidienne entre deux points A et B de $\mathbb{R}^3$ est définie par

$$AB = d(A,B) = \left((x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2\right)^{\frac{1}{2}}$$

où (x_A,y_A,z_A), resp. (x_B,y_B,z_B) sont les coordonnées de A, resp. B.

Distance à une droite

Etant donné une droite vectorielle $D \subset \mathbb{R}^3$, un point A dans $\mathbb{R}^3$, on cherche le(s) point(s) M du plan D le(s) plus proche(s) de A. En posant $M = (x,y,z)$, cela revient à chercher le(s) minimant(s) de la fonction f définie par

$$f(x,y,z) = (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2$$

avec la contrainte $(x,y,z) \in D$.

La valeur du minimum sera appelée la distance de A à la droite D.

Proposition 3. *Le minimum de $f(x,y,z)$ sous la contrainte $(x,y,z) \in D$ est atteint en un unique point $A' \in D$. Le point A' est caractérisé par*

$$A' \in D \text{ et } \forall w \in D, w \text{ et } \vec{AA'} \text{ sont orthogonaux}$$

A' s'appelle le projeté orthogonal *de A sur D.*

On a besoin, pour préciser l'énoncé et lisser le calcul, de définir les quantités et concept suivants :

1. Si $u = (x_u, y_u, z_u) \in \mathbb{R}^3$, la *norme* de u est définie par
$$\|u\| := \left(x_u^2 + y_u^2 + z_u^2\right)^{\frac{1}{2}}$$
2. Si $u = (x_u, y_u, z_u), v = (x_v, y_v, z_v) \in \mathbb{R}^3$, leur *produit scalaire* est défini par
$$\langle u, v\rangle = x_u.x_v + y_u.y_v + z_u.z_v$$
3. $u, v \in \mathbb{R}^3$ sont dits orthogonaux si $\langle u, v\rangle = 0$.

Démonstration. La difficulté est de passer d'un problème avec contrainte à un problème sans contrainte. On commence donc par décrire la droite D. Celle-ci est donnée par un vecteur directeur $u \neq 0$ et on a

$$(x,y,z) \in D \Leftrightarrow \exists \lambda \in \mathbb{R}, (x,y,z) = \lambda.u$$

Minimiser la fonction f définie par

$$f(x,y,z) = (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2$$

avec la contrainte $(x,y,z) \in D$ revient donc à minimiser la fonction g définie par

$$\forall \lambda \in \mathbb{R}, g(\lambda) = (\lambda.x_u - x_A)^2 + (\lambda.y_u - y_A)^2 + (\lambda.z_u - z_A)^2$$

La fonction g est un trinôme en λ, qui, après développement, se réécrit

$$\forall \lambda \in \mathbb{R}, g(\lambda) = \lambda^2.\|u\|^2 - 2\lambda.\langle u, A\rangle + \|A\|^2$$

Le minimum est atteint en l'unique point critique de g déterminé par $g'(\lambda_0) = 0$, *i.e.*

$$\lambda_0 = \frac{\langle u, A\rangle}{\|u\|^2}$$

et on a

$$\forall \lambda \in \mathbb{R}, g(\lambda) = \|u\|^2.(\lambda - \lambda_0)^2 + \|A\|^2 - \frac{(\langle u, A\rangle)^2}{\|u\|^2}$$

En repassant à f, on obtient que le minimum de $f(x,y,z)$ sous la contrainte $(x,y,z) \in D$ est atteint uniquement en

$$A' = \lambda_0.u = \frac{\langle u, A\rangle}{\|u\|^2}.u$$

et vaut

$$d(A,D)^2 = d(A,A')^2 = \|A\|^2 - \frac{(\langle u, A\rangle)^2}{\|u\|^2}$$

Il nous reste à vérifier que le seul point A' satisfaisant à la caractérisation donnée est le point A'.

1. Le point A' satisfait à la caractérisation : en effet, $A' \in D$ par construction et, si $w \in D$, $w = \lambda.u$, on vérifie assez aisémment que
$$\left\langle w, \vec{AA'} \right\rangle = 0$$
2. Il n'y a qu'un point M (c'est donc A' !) satisfaisant à la caractérisation : si $M \in D$, $M = \lambda.u$ pour un certain $\lambda \in \mathbb{R}$ et, comme on doit avoir
$$\left\langle u, \vec{AM} \right\rangle = 0$$
alors, forcément,
$$\lambda = \frac{\langle u, A \rangle}{\|u\|^2}$$

Tout ceci se dessine ! □

1.4 La distance d'un point à un plan vectoriel

Distance à un plan

Etant donné un plan vectoriel $P \subset \mathbb{R}^3$, un point A dans $\mathbb{R}^3$, on cherche le(s) point(s) M du plan P le plus proche de A. En posant $M = (x, y, z)$, cela revient à chercher le(s) minimant(s) de la fonction f définie par
$$f(x, y, z) = (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2$$
avec la contrainte $(x, y, z) \in P$.

La valeur du minimum sera appelée la distance de A au plan P.

Proposition 4. *Le minimum de $f(x, y, z)$ sous la contrainte $(x, y, z) \in P$. est atteint en un unique point $A' \in P$. Le point A' est caractérisé par*
$$A' \in P \text{ et } \forall t \in P, t \text{ et } \vec{AA'} \text{ sont orthogonaux}$$
A' s'appelle le projeté orthogonal *de A sur P.*

Démonstration. La difficulté est de passer d'un problème avec contrainte à un problème sans contrainte. On commence donc par décrire le plan P. Celui-ci est donné par une base (v, w) et on a
$$(x, y, z) \in P \Leftrightarrow \exists (\lambda, \mu) \in \mathbb{R}^2, (x, y, z) = \lambda.v + \mu.w$$
Minimiser la fonction f définie par
$$f(x, y, z) = (x - x_A)^2 + (y - y_A)^2 + (z - z_A)^2$$
avec la contrainte $(x, y, z) \in P$ revient donc à minimiser la fonction g définie par
$$\forall (\lambda, \mu) \in \mathbb{R}^2, g(\lambda, \mu) = (\lambda.x_v + \mu.x_w - x_A)^2 + (\lambda.y_v + \mu.y_w - y_A)^2 + (\lambda.z_v + \mu.z_w - z_A)^2$$
Les points critiques de cette fonction sont déterminés par le système
$$\left\{ \begin{array}{rcl} x_v.(\lambda.x_v + \mu.x_w - x_A) + y_v.(\lambda.y_v + \mu.y_w - y_A) + z_v.(\lambda.z_v + \mu.z_w - z_A) & = & 0 \\ x_w.(\lambda.x_v + \mu.x_w - x_A) + y_w.(\lambda.y_v + \mu.y_w - y_A) + z_w.(\lambda.z_v + \mu.z_w - z_A) & = & 0 \end{array} \right.$$
i.e. après simplification de notations
$$\left\{ \begin{array}{rcl} \lambda.\|v\|^2 + \mu.\langle v, w \rangle & = & \langle v, A \rangle \\ \lambda.\langle v, w \rangle + \mu.\|w\|^2 & = & \langle w, A \rangle \end{array} \right.$$
Le déterminant de ce système est $\Delta = \|v\|^2.\|w\|^2 - \langle v, w \rangle^2$ et une ligne (difficile) de raisonnement est la suivante

1. Comme v, w sont non colinéaires, $\Delta \neq 0$,
2. le système en considération admet une unique solution (λ_0, μ_0) donnée par

$$\lambda_0 = \frac{\langle v,A\rangle . \|w\|^2 - \langle w,A\rangle \langle v,w\rangle}{\Delta} \text{ et } \mu_0 = \frac{\langle w,A\rangle . \|v\|^2 - \langle v,A\rangle \langle v,w\rangle}{\Delta}$$

3. Démontrer que pour tout couple (λ, μ), $g(\lambda, \mu) - g(\lambda_0, \mu_0) \geq 0$ en faisant apparaitre une somme de carrés [1]
4. Repasser ensuite en $A' = \lambda_0.v + \mu_0.w$ etc...

La difficulté du programme esquissé vient de la trop grande généralité laissée aux vecteurs v et w.

En effet, admettons un moment que l'on puisse supposer $\|v\|^2 = \|w\|^2 = 1$ et $\langle v, w\rangle = 0$. Le système donnant λ_0, μ_0 devient alors

$$\begin{cases} \lambda_0 & = & \langle v,A\rangle \\ \mu_0 & = & \langle w,A\rangle \end{cases}$$

et on a, en développant les carrés

$$\forall(\lambda,\mu) \in \mathbb{R}^2, g(\lambda,\mu) = \lambda^2 + \mu^2 - 2\lambda. \langle v,A\rangle - 2\lambda. \langle w,A\rangle + \|A\|^2 = (\lambda-\lambda_0)^2 + (\mu-\mu_0)^2 + \|A\|^2 - \langle v,A\rangle^2 - \langle w,A\rangle^2$$

Il est alors clair que

$$\forall(\lambda,\mu) \in \mathbb{R}^2, g(\lambda,\mu) \geq g(\lambda_0,\mu_0) = \|A\|^2 - \langle v,A\rangle^2 - \langle w,A\rangle^2$$

Le minimum de g sur $\mathbb{R}^2$ est atteint uniquement en (λ_0, μ_0) et vaut $\|A\|^2 - \langle v,A\rangle^2 - \langle w,A\rangle^2$.

En repassant à f, on obtient que le minimum de $f(x,y,z)$ sous la contrainte $(x,y,z) \in P$ est atteint uniquement en

$$A' = \lambda_0.v + \mu_0.w = \langle v,A\rangle .v + \langle w,A\rangle .w$$

et vaut

$$d(A,P)^2 = \|A\|^2 - \langle v,A\rangle^2 - \langle w,A\rangle^2$$

Modulo les conditions imposées sur v et w, il nous reste à vérifier que le seul point A' satisfaisant à la caractérisation donnée est le point A'.

1. Le point A' satisfait à la caractérisation : en effet, $A' \in P$ par construction et, si $t \in P$, $t = v$, $t = w$, puis $t = \lambda.v + \mu.w$, on vérifie assez aisément que

$$\left\langle t, \vec{AA'} \right\rangle = 0$$

2. Il n'y a qu'un point M (c'est donc A' !) satisfaisant à la caractérisation : si $M \in P$, $M = \lambda.v + \mu.w$ pour un certain couple (λ, μ) et, comme on doit avoir

$$\left\langle v, \vec{AM} \right\rangle = \left\langle w, \vec{AM} \right\rangle = 0$$

alors, forcément,

$$\lambda = \langle v,A\rangle \text{ et } \mu = \langle w,A\rangle$$

Pour finir, il nous reste à montrer que l'on peut trouver une base (v, w) de P vérifiant les conditions imposées. On part d'une base (v_0, w_0) de P.

1. Ca demande beaucoup de dextérité !

1. (Etape de normalisation de v_0) On a $v_0 \neq 0$ et donc $\|v_0\| \neq 0$ et, en posant
$$v = \frac{1}{\|v_0\|} v_0$$
on obtient un vecteur $v \in P$ vérifiant $\|v\|^2 = 1$.
2. (Etape d'orthogonalisation de (v, w_0)) On peut choisir $\lambda \in \mathbb{R}$ de sorte que $w_1 := w_0 - \lambda . v \in P$ vérifie $\langle w_1, v\rangle = 0$. Il suffit en effet de prendre
$$\lambda = \langle w_0, v\rangle$$
(v, w_1) est une base de P car c'est une famille libre comportant deux éléments.
3. (Etape de normalisation de w_1) On a $w_1 \neq 0$ et donc $\|w_1\| \neq 0$ et, en posant
$$w = \frac{1}{\|w_1\|} w_1$$
on obtient un vecteur $w \in P$ vérifiant $\|w\|^2 = 1$ et $\langle v, w\rangle = 0$.

Tout ceci se dessine ! □

1.5 Le lien entre les deux calculs

D'un point de vue équation cartésienne, le plan P peut s'écrire, pour un vecteur $u \in \mathbb{R}^3$, non nul, comme
$$P = \{(x,y,z) \in \mathbb{R}^3,\, x_u.x + y_u.y + z_u.z = 0\} = \{X \in \mathbb{R}^3,\, \langle u, X\rangle = 0\}$$
On peut supposer, en normalisant u, que $\|u\|^2 = 1$.

Si (v, w) est une base de P vérifiant
$$\|v\|^2 = 1,\, \|v\|^2 = 1,\, \langle v, w\rangle = 0$$
On a obtenu une famille $\mathscr{U} = (u, v, w)$ de vecteurs de $\mathbb{R}^3$ vérifiant
$$\|u\|^2 = 1,\, \|v\|^2 = 1,\, \|v\|^2 = 1 \text{ et } \langle u, v\rangle = 0,\, \langle u, w\rangle = 0,\, \langle v, w\rangle = 0$$
Cette famille est une base de $\mathbb{R}^3$ (vérifier qu'elle est libre). On dit que c'est une base orthonormée de $\mathbb{R}^3$.

Les coordonnées d'un vecteur X sur la base $\mathscr{U}$ sont
$$_{\mathscr{U}}[X] = \begin{pmatrix} \langle u, X\rangle \\ \langle v, X\rangle \\ \langle w, X\rangle \end{pmatrix}$$
On vérifie (et le calcul est un peu pénible à ce stade) que pour $X, Y \in \mathbb{R}^3$, on a
$$\|X\|^2 = \langle u, X\rangle^2 + \langle v, X\rangle^2 + \langle w, X\rangle^2$$
et
$$\langle X, Y\rangle = \langle u, X\rangle \langle u, Y\rangle + \langle v, X\rangle \langle v, Y\rangle + \langle w, X\rangle \langle w, Y\rangle$$
Si $A \in \mathbb{R}^3$, $A' = \langle u, A\rangle . u$ son projeté orthogonal sur $D = \mathrm{Vect}\langle u\rangle$, $A'' = \langle v, A\rangle . v + \langle w, A\rangle . w$ son projeté orthogonal sur P, on a
$$\|A\|^2 = \|A'\|^2 + \|A''\|^2 = d(A, P)^2 + d(A, D)^2$$
Ce qui se dessine, là encore. On reconnait le dessin typique du théorème de PYTHAGORE.

En conclusion

Le but de ce chapitre est de donner un cadre permettant de systématiser et de simplifier ce type de calculs et d'en donner une interprétation géométrique.

2 Distances et orthogonalité dans $\mathbb{R}^n$

Norme et produit scalaire

Définition 5. *1. On appelle* produit scalaire *de deux vecteurs $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$ de $\mathbb{R}^n$ le nombre réel*

$$\langle X, Y \rangle = \sum_{i=1}^{n} x_i . y_i$$

2. On appelle norme *d'un vecteur $X = (x_1, \dots, x_n)$ le nombre réel* positif*, noté $\|X\|$ défini par*

$$\|X\|^2 = \sum_{i=1}^{n} x_i^2 = \langle X, X \rangle$$

Distance

Définition 6. *Si A, B sont deux points de $\mathbb{R}^n$, de coordonnées respectives $(x_{A,1}, \dots, x_{A,n})$ et $(x_{B,1}, \dots, x_{B,n})$, le vecteur $\vec{AB}$ est le vecteur (de $\mathbb{R}^n$) de coordonnées $(x_{B,1} - x_{A,1}, \dots, x_{B,n} - x_{A,n})$. La distance entre A et B est (deux notations),*

$$d(A, B) = AB := \|\vec{AB}\| = \left(\sum_{i=1}^{n} (x_{B,i} - x_{A,i})^2 \right)^{\frac{1}{2}}$$

Une remarque

Dans $\mathbb{R}^n$, points et vecteurs sont les mêmes objets, seule la façon de les utiliser diffère. Si A et B sont deux points, on peut définir le vecteur $\vec{AB}$ dont les composantes sont les différences des composan,tes des vecteurs A et B.

Premières Propriétés

Si $X \in \mathbb{R}^n$, $\langle X, X \rangle \geq 0$ et $\|X\|$ est donc bien défini.

On a $\|X\| = 0 \Leftrightarrow X = 0$. Un vecteur $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ sera systématiquement écrit comme la colonne de ses coordonnées relativement à la base canonique

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Avec cette convention, on a

Proposition 7 (Ecriture matricielle du produit scalaire.)**.** *Pour X, $Y \in \mathbb{R}^n$, on a*

$$\langle X, Y \rangle = X^\top . Y$$

Proposition 8 (Symétrie.)**.** *Pour $X, Y \in \mathbb{R}^n$, $\langle X, Y \rangle = \langle Y, X \rangle$.*

Proposition 9 (Linéarité)**.** *Si $Y \in \mathbb{R}^n$ est fixé, l'application $\mathscr{L}_Y : \mathbb{R}^n \to \mathbb{R}$ définie par*

$$\forall X \in \mathbb{R}^n, \mathscr{L}_Y(X) = \langle Y, X \rangle$$

est une forme linéaire

Proposition 10 (Formes linéaires)**.** *Réciproquement, soit $\mathscr{L} : \mathbb{R}^n \to \mathbb{R}$ une forme linéaire, il existe un vecteur $Y \in \mathbb{R}^n$ tel que $\mathscr{L} = \mathscr{L}_Y$*

Exemple : On considère $n = 3$ et

$$H = \{(x_1, x_2, x_3) \in \mathbb{R}^3,\, 2x_1 + 3x_2 - x_3 = 0\}$$

Exhiber un vecteur Y de sorte que $H = \operatorname{Ker} \mathscr{L}_Y$. Quelle est la dimension de H ?

D'une façon générale, si $Y \neq 0$ est un vecteur fixé de $\mathbb{R}^n$, la dimension de $H = \operatorname{Ker} \mathscr{L}_Y$ est $n-1$. On dit que H est un *hyperplan* (vectoriel) de $\mathbb{R}^n$. L'hyperplan orthogonal à Y

Démonstration. Ecrivons L, la matrice de l'application linéaire $\mathscr{L}$ relativement à la base canonique (au départ) de $\mathbb{R}^n$ et la base (1) de $\mathbb{R}$. L est une matrice comportant une ligne. En posant $Y = L^\top$, on obtient le vecteur cherché. □

Exercice 1.— L'espace $\mathbb{R}^4$ est muni du produit scalaire canonique. Pour chacun des sev E_i suivants, donner une base de $E_i^\perp = \{y \in \mathbb{R}^4, \forall x \in E_i,\, \langle x, y\rangle = 0\}$.

On commencera par démontrer le principe suivant, valable dans $\mathbb{R}^n$: Si $E = \operatorname{Vect}\langle u_1, \dots, u_k\rangle \subset \mathbb{R}^n$ et

$$F = \{y \in \mathbb{R}^n, \langle u_1, y\rangle = 0, \dots, \langle u_k, y\rangle = 0\}$$

alors

$$E^\perp = F$$

1. $E_1 = \operatorname{Vect}\left\langle \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \\ 3 \\ 2 \end{pmatrix} \right\rangle$

2. $E_3 = \operatorname{Vect}\left\langle \begin{pmatrix} 3 \\ 2 \\ 0 \\ 4 \end{pmatrix} ; \begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \end{pmatrix} ; \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} \right\rangle$

3. $E_2 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4,\, 2x_1 + 3x_2 - x_4 = 0\}$

Norme d'une somme

Proposition 11. *Si $X, Y \in \mathbb{R}^n$ sont des vecteurs quelconques, on a*

$$\|X+Y\|^2 = \|X\|^2 + \|Y\|^2 + 2\langle X, Y\rangle$$

et

$$\|X-Y\|^2 = \|X\|^2 + \|Y\|^2 - 2\langle X, Y\rangle$$

Démonstration. Il suffit d'écrire, par linéarité,

$$\begin{aligned} \|X+Y\|^2 &= \langle X+Y, X+Y\rangle = \langle X+Y, X\rangle + \langle X+Y, Y\rangle \\ &= \langle X, X\rangle + \langle Y, X\rangle + \langle X, Y\rangle + \langle Y, Y\rangle \end{aligned}$$

et d'utiliser la symétrie du produit scalaire. □

L'inégalité de CAUCHY–SCHWARZ

Théorème 12 (Inégalité de CAUCHY–SCHWARZ pour le produit scalaire.)**.**

$$\forall X,Y \in \mathbb{R}^n,\ |\langle X,Y\rangle| \leq \|X\|.\|Y\|$$

Démonstration. Soit X et Y deux vecteurs quelconques dans $\mathbb{R}^n$ et f la fonction d'une variable réelle définie par

$$\forall t \in \mathbb{R},\ f(t) = \|X+t.Y\|^2$$

alors f est polynômiale de dégré ≤ 2.

1. Si $Y=0$, l'inégalité voulue est trivialement vraie
2. Si $Y\neq 0$, f est de degré exactement 2, de coefficient dominant strictement positif et de discriminant $\Delta = 4(\langle X,Y\rangle^2 - \|X\|^2\|Y\|^2)$. Comme f est postive, ce discriminant est ≤ 0, ce qui donne l'égalité voulue

On peut commenter sur le cas d'égalité : Si on a égalité dans l'inégalité de CAUCHY–SCHWARZ c'est soit que $Y=0$, soit que $Y\neq 0$ et que le discriminant Δ est nul, *i.e.* , il existe $t\in\mathbb{R}$ tel que $X=-t.Y$. Ces deux cas équivalent au fait que X et Y sont colinéaires. □

Exercice 2.— Soit $a_1,a_2,\ldots,a_n$ des nombres réels. Montrer que :

$$\left(\frac{a_1+\ldots+a_n}{n}\right)^2 \leqslant \frac{a_1^2+\ldots+a_n^2}{n}$$

Exercice 3.—
1. Soit $x_1,\ldots,x_n$ des nombres réels strictement positifs. Montrer que :

$$\left(\sum_{k=1}^{n}\frac{1}{x_k}\right)\left(\sum_{k=1}^{n}x_k\right) \geqslant n^2$$

2. Trouver le minimum de $\left(\sum_{k=1}^{n}\frac{1}{x_k}\right)\left(\sum_{k=1}^{n}x_k\right)$ quand $(x_1,\ldots,x_n)$ décrit $(\mathbb{R}^{+*})^n$.

Exercice 4.—(Facultatif/Spécial 5/2) Une application aux séries. Soient $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ deux suites à termes réels. On suppose que les séries $\sum_{n\geq 0}x_n^2$ et $\sum_{n\geq 0}y_n^2$ sont convergentes.
1. On suppose les séries $\sum x_n$ et $\sum y_n$ à termes positifs. Montrer, en majorant les sommes partielles par une constante, que la série $\sum_{n\geq 0}(x_n.y_n)$ est convergente.
2. Montrer, en général, que la série $\sum_{n\geq 0}(x_n.y_n)$ est absolument convergente.
3. Montrer l'inégalité

$$\left|\sum_{n=0}^{+\infty}(x_n.y_n)\right| \leq \sum_{n=0}^{+\infty}|x_n|.|y_n| \leq \sqrt{\sum_{n=0}^{+\infty}x_n^2}.\sqrt{\sum_{n=0}^{+\infty}y_n^2}$$

Inégalités triangulaires

Théorème 13 (Inégalité triangulaire I)**.**

$$\forall X,Y\in\mathbb{R}^n,\ \|X+Y\| \leq \|X\|+\|Y\|$$

Théorème 14 (Inégalité triangulaire II)**.**

$$\forall A,B,C \in \mathbb{R}^n,\ d(A,B) \leq d(A,C) + d(C,B)$$

Rappelons que d vérifie aussi

$$\forall A,B \in \mathbb{R}^n,\ d(A,B) = d(B,A)$$

et

$$\forall A,B \in \mathbb{R}^n,\ d(A,B) = 0 \Leftrightarrow A = B$$

Ce sont les trois propriétés caractéristiques d'une *distance*.

Théorème 15 (Inégalité triangulaire III)**.**

$$\forall X,Y \in \mathbb{R}^n,\ |\|X\| - \|Y\|| \leq \|X - Y\|$$

Démonstration.

1. Inégalité triangulaire I. Il suffit de développer $\|X+Y\|^2$ et d'utiliser l'inégalité de CAUCHY–SCHWARZ.
2. Inégalité triangulaire II. On applique l'inégalité triangulaire I à $X = \vec{AC}$, $Y = \vec{CB}$ et $X + Y = \vec{AB}$
3. Inégalité triangulaire III. On applique l'inégalité triangulaire I à $X' = X - Y$, $Y' = Y$ pour obtenir que

   $$\|X\| \leq \|X - Y\| + \|Y\| \text{ } i.e. \text{ } \|X\| - \|Y\| \leq \|X - Y\|$$

   On symétrise ensuite les rôles de X et Y.

$\square$

Orthogonalité

Définition 16. *Si X et Y sont deux vecteurs de $\mathbb{R}^n$, on dit qu'ils sont* orthogonaux *si* $\langle X,Y\rangle = 0$.

Le théorème de PYTHAGORE

Théorème 17. *Pour tous X, Y vecteurs de $\mathbb{R}^n$,*

$$X \text{ et } Y \text{ sont orthogonaux} \Leftrightarrow \|X - Y\|^2 = \|X\|^2 + \|Y\|^2 \Leftrightarrow \|X + Y\|^2 = \|X\|^2 + \|Y\|^2$$

Rq : quel est le lien entre ce résultat et le classique théorème de PYTHAGORE ?
Indication:Dans $\mathbb{R}^2$ représenté par le plan, via un repère *orthonormé* usuel, la longueur d'un vecteur X est sa norme $\|X\|$. Considérer pour deux vecteurs X et Y de $\mathbb{R}^2$ le triangle de points du plan formé par X, Y et $O = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Démonstration. Il suffit de développer en partant de $\|X-Y\|^2 = \langle X-Y, X-Y\rangle$ et de $\|X+Y\|^2 = \langle X+Y, X+Y\rangle$.

$\square$

Cosinus de l'angle de deux vecteurs

Exercice 5.—Dans le plan (muni d'un repère orthonormé), représenter les vecteurs $X = \begin{pmatrix} 2\cos\alpha \\ 2\sin\alpha \end{pmatrix}$ et $Y = \begin{pmatrix} 3\cos\beta \\ 3\sin\beta \end{pmatrix}$ pour des valeurs α et β que vous jugerez représentatives d'une situation générique. Exprimer la quantité $\frac{\langle X,Y\rangle}{\|X\|.\|Y\|}$ en fonction de α et β. Que représente géométriquement cette quantité ?

Dans $\mathbb{R}^n$ en général, pour $X,Y \neq 0$, la quantité $\frac{\langle X,Y\rangle}{\|X\|.\|Y\|}$ est le cosinus d'un nombre réel. C'est le cosinus de l'angle formé par les deux vecteurs X et Y.

Pour le mesurer pratiquement, il suffit de se placer dans le plan engendré par X et Y. (Dessin)

3 Familles orthogonales, Bases orthonormales

3.1 Définitions

Familles orthogonales

Définition 18. *Soit* $(V_1,V_2,\dots,V_p)$ *une famille de vecteurs de* $\mathbb{R}^n$ *; On dit que c'est une* famille orthogonale *si pour toute paire* V_i,V_j, $i \neq j$ *de vecteurs de cette famille, on a*

$$\langle V_i,V_j\rangle = 0$$

Exercice 6.—A une famille $(V_1,V_2,\dots,V_p)$ de vecteurs de $\mathbb{R}^n$, on associe la matrice P, $P = (V_1|\dots|V_p)$ comportant p colonnes et n lignes.

Montrer que La famille $(V_1,V_2,\dots,V_p)$ est une famille orthogonale si et seulement si $P^\top.P$ est une matrice diagonale.

Proposition 19 (liberté des familles orthogonales)**.** *Soit* $(V_1,V_2,\dots,V_p)$ *une famille orthogonale de vecteurs* non nuls *de* $\mathbb{R}^n$*. Cette famille est libre.*

Démonstration. Partant de la relation $\lambda_1.V_1 + \dots + \lambda_p.V_p = 0$, on isole le coefficient λ_i en effectuant le produit scalaire avec le vecteur V_i. Le fait que $V_i \neq 0$ implique que $\langle V_i,V_i\rangle \neq 0$ et que $\lambda_i = 0$. □

Familles orthonormales

Définition 20. *Soit* $(U_1,U_2,\dots,U_p)$ *une famille de vecteurs de* $\mathbb{R}^n$ *; On dit que c'est une* famille orthonormale *si*

1. *pour toute paire* U_i,U_j, $i \neq j$ *de vecteurs de cette famille, on a*

$$\langle U_i,U_j\rangle = 0$$

2. *pour tout vecteur* U_i *de cette famille,* $\|U_i\| = 1$ *(On dit que chaque vecteur* U_i *est* unitaire*.)*

Rq : Une famille orthonormale de vecteurs de $\mathbb{R}^n$ est libre. Elle comporte au maximum n éléments.

Exercice 7.—Comme précédemment, on pose $P = (U_1|\dots|U_p)$ la matrice dont les colonnes sont les U_i. Montrer que $(U_1,U_2,\dots,U_p)$ est une famille orthonormale si et seulement si $P^\top.P$ est la matrice I_p. P est-elle automatiquement inversible ?

Dans le cas $p = n$,

Proposition 21. *Si la famille $\mathscr{U} = (U_1,U_2,\dots,U_n)$ de vecteurs de $\mathbb{R}^n$ est orthonormale alors*

1. *$P = (U_1|\dots|U_n)$ construite précédemment est inversible, d'inverse $P^\top$,*
2. *$\mathscr{U}$ est une base de $\mathbb{R}^n$,*
3. *P est la matrice de passage de $\mathscr{C}_n$, la base canonique de $\mathbb{R}^n$ à la base $\mathscr{U}$.*

Définition 22. *Dans les conditions énoncées ci-dessus, la famille $\mathscr{U}$ est appelée une* base orthonormale *ou* base orthonormée *de* $\mathbb{R}^n$.

Exemples :

1. La base canonique de $\mathbb{R}^n$ est une base orthonormée de $\mathbb{R}^n$,
2. Pour $\theta \in \mathbb{R}$, posons $U_\theta = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, $V_\theta = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$.

   $\mathscr{U}_\theta := (U_\theta, V_\theta)$ est une base orthonormée de $\mathbb{R}^2$.

Exercice 8.—Dans $\mathbb{R}^3$, soit $\theta, \phi \in \mathbb{R}$ et

$$U = \begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{pmatrix}$$

On pose, à la physicienne,

$$V_\theta = \frac{\partial U}{\partial \theta}, V_\phi = \frac{\partial U}{\partial \phi}$$

où les dérivées partielles sont effectuées sur U composante par composante.
1. Donner des formules pour V_θ et V_ϕ.
2. A quelle condition sur θ, ϕ, (U,V_θ,V_ϕ) forme une base orthogonale de $\mathbb{R}^3$. En déduire une base orthonormée de $\mathbb{R}^3$.

Exercice 9.—(Question ouverte). En s'inpirant des deux exemples précédents, pourriez vous proposer un procédé de construction pour fabriquer une famille de bases orthonormales de $\mathbb{R}^4$ paramétrée par trois angles θ, ϕ, ψ ?

3.2 Calcul de la norme et du produit scalaire dans une B.O.N

Soit $\mathscr{C}_n = (e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$ et $\mathscr{U} = (U_1, \ldots, U_n)$ une base orthonormée, P la matrice de passage de $\mathscr{C}_n$ à $\mathscr{U}$. Pour un vecteur X, on note $X^{\mathscr{U}}$ les coordonnées de ce vecteur dans la base $\mathscr{U}$. On a $X = P.X^{\mathscr{U}}$.

Proposition 23. *1. Pour tout vecteur $X \in \mathbb{R}^n$, $\|X\|^2 = \|X^{\mathscr{U}}\|^2$,*

2. Pour tous vecteurs $X, Y \in \mathbb{R}^n$, $\langle X, Y\rangle = \left\langle X^{\mathscr{U}}, Y^{\mathscr{U}}\right\rangle$.

En d'autre termes, on peut calculer le produit scalaire (ou la norme) à partir du moment où l'on connait les coordonnées des vecteurs relativement à une base orthonormée, sans revenir forcément à la base canonique. Exemple : Soient $\theta, \phi \in \mathbb{R}$, $\cos\phi \neq 0$, on a alors, si

$$U = \begin{pmatrix} \cos\theta\cos\phi \\ \sin\theta\cos\phi \\ \sin\phi \end{pmatrix}, V_\theta = \begin{pmatrix} -\sin\theta\cos\phi \\ \cos\theta\cos\phi \\ 0 \end{pmatrix}, V_\phi = \begin{pmatrix} -\cos\theta\sin\phi \\ -\sin\theta\sin\phi \\ \cos\phi \end{pmatrix}$$

et

$$X = a.U + b.V_\theta + c.V_\phi$$

alors

$$\|X\|^2 = a^2 + b^2\cos^2\phi + c^2$$

3.3 Construction de bases orthonormées dans $\mathbb{R}^n$.

Exercice 10.—Dans $\mathbb{R}^3$, on se donne trois vecteurs

$$u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, w = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

1. Déterminer $\lambda \in \mathbb{R}$ de sorte que $u' = \lambda.u$ soit normé. Vérifier que $\mathrm{Vect}\langle u\rangle = \mathrm{Vect}\langle u'\rangle$.
2. On pose $v' = v - \langle v, u'\rangle u'$. Montrer que (u', v') est une famille orthogonale. Est-elle orthonormale ? Vérifier que $\mathrm{Vect}\langle u, v\rangle = \mathrm{Vect}\langle u', v'\rangle$. Déterminer $mu \in \mathbb{R}$ de sorte que $v'' = \mu.v'$ soit normé.
3. On pose $w' = w - \langle w, u'\rangle u' - \langle w, v''\rangle v''$. Montrer que (u', v'', w') est une famille orthogonale. Quelle modification apporter à w' pour que la famille devienne orthonormale ? Vérifier que $\mathrm{Vect}\langle u, v, w\rangle = \mathrm{Vect}\langle u', v'', w'\rangle$.

4. (Facultatif). Ce procédé se généralise à $\mathbb{R}^n$: étant donnée une famille *libre* $(u_1, u_2, \ldots, u_p)$ de vecteurs de $\mathbb{R}^n$, montrer comment construire par récurrence une famille $(u'_1, \ldots, u'_p)$ orthonormale et telle que pour tout $1 \leq k \leq p$,

$$\mathrm{Vect}\langle u_1, \ldots, u_k\rangle = \mathrm{Vect}\left\langle u'_1, \ldots, u'_k\right\rangle$$

Indication:A défaut, on peut aller voir sur le web à "Procédé d'orthonormalisation de GRAM–SCHMIDT". Ce procédé est hors-programme On retiendra de cet exercice les énoncés suivants.

Théorème 24. *Soit F un sev de $\mathbb{R}^n$. Il existe une base de F orthonormale.*

Théorème 25. *Soit F un sev de $\mathbb{R}^n$ de dimension p, $\mathscr{U}$ une B.O.N. de F. Il existe une base de $\mathbb{R}^n$ orthonormale dont les p premiers vecteurs sont les vecteurs de $\mathscr{U}$.*

3.4 Diagonalisation des matrices symétriques réelles.

Théorème 26 (Le théorème spectral). *Soit $M \in \mathscr{M}_n(\mathbb{R})$ une matrice symétrique. Il existe une base orthonormale de $\mathbb{R}^n$ diagonalisant M.*

Exercice 11.——Un lemme utile pour ce théorème.– Soit M une matrice symétrique réelle, λ, μ deux valeurs propres réelles distinctes de M associées respectivement aux vecteurs propres X et Y. En calculant $\langle M.X, Y \rangle$ de deux façons distinctes, montrer qu'alors X et Y sont orthogonaux.

Exercice 12.— Soit $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$

Trouver une matrice P inversible telle que $P^\top.A.P$ soit diagonale.

Exercice 13.—Soit M une matrice anti-symétrique réelle. Montrer que la seule valeur propre réelle possible de M est $\lambda = 0$

Exercice 14.— Diagonaliser dans une base orthonormée de $\mathbb{R}^4$ la matrice $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$.

4 Projections orthogonales

4.1 Distance à un ensemble

On rappelle que la distance entre deux points A et B ou entre deux vecteurs X et Y de $\mathbb{R}^n$ est :

$$AB = d(A,B) = \|\vec{AB}\|,\, d(X,Y) = \|Y - X\|$$

Définition 27. *Soit $F \subset \mathbb{R}^n$ une partie non vide, $A \in \mathbb{R}^n$. On appelle* distance *de A à F la quantité*

$$d(A,F) := \inf\{AM,\, M \in F\}$$

Rq : Si $A \in F$ alors $d(A,F) = 0$. On s'intéresse ici exclusivement au cas où F est un sev de $\mathbb{R}^n$.

4.2 Le cas de la projection sur une droite ou un hyperplan

Le cas F droite vectorielle

Proposition-Définition 28. *Soit $F = Vect\langle U \rangle \in \mathbb{R}^n$ une droite vectorielle. On suppose que $U \neq 0$ et que $\|U\| = 1$. Soit $A \in \mathbb{R}^n$ alors il existe un unique point $M_A \in F$ tel que $d(A,F) = AM_A$.*

On a $M_A = \langle A, U \rangle .U$.

Le point M_A s'appelle le projeté orthogonal *de A sur la droite F. On définit l'application $p_F : \mathbb{R}^n \to \mathbb{R}^n$ par*

$$\forall A \in \mathbb{R}_n,\, p_F(A) = \langle A, U \rangle .U$$

Cette application, linéaire*, s'appelle la* projection orthogonale *sur F.*

Démonstration. Soit $A \in \mathbb{R}^n$, fixé. Posons, pour $t \in \mathbb{R}$, $f(t) = \|A - t.U\|^2$. On a

$$\forall t \in \mathbb{R},\, f(t) = \|A\|^2 - 2.t\,\langle A, U \rangle + t^2.$$

La fonction f est un trinôme et admet un minimum atteint uniquement en $t_A = \langle A, U \rangle$.

Le point $M_A = t_A.U$ ets donc l'unique point réalisant le minimum de $d(A,M)$ lorsque M parcourt F □

Propriétés de la projection sur une droite

On conserve les notations de la proposition précédente.

1. p_F est un endomorphisme de $\mathbb{R}^n$ et $p_F \circ p_F = p_F$.
2. Si $U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$ avec $\|U\|^2 = \sum_{i=1}^n u_i^2 = 1$ alors la matrice de p_F dans la base canonique de $\mathbb{R}^n$ est la matrice $P_F = U.U^\top$.
3. P_F est symétrique et $P_F^2 = P_F$.
4. $\operatorname{Ker} p_F = \{X \in \mathbb{R}^n, \langle U, X\rangle = 0\}$, $\operatorname{Im} p_F = \operatorname{Vect}\langle U\rangle = F$.
5. p_F est diagonalisable dans une base orthonormée de $\mathbb{R}^n$. Elle est semblable à... (Quelle matrice ?)

Le cas F hyperplan de $\mathbb{R}^n$

Proposition-Définition 29. *Soit F un hyperplan vectoriel de $\mathbb{R}^n$.* $\dim F = n-1$ *et il existe $U \in \mathbb{R}^n$, $\|U\| = 1$ tel que $F = \{X \in \mathbb{R}^n, \langle U, X\rangle = 0\}$.*

Soit $A \in \mathbb{R}^n$. Il existe un unique point $M_A \in F$ tel que $d(A,F) = AM_A$.

On a $M_A = A - \langle A, U\rangle .U$.

Le point M_A s'appelle le projeté orthogonal *de A sur l'hyperplan F. On définit l'application $p_F : \mathbb{R}^n \to \mathbb{R}^n$ par*

$$\forall A \in \mathbb{R}_n,\ p_F(A) = A - \langle A, U\rangle .U$$

Cette application, linéaire*, s'appelle la* projection orthogonale *sur F.*

Démonstration. Soit $A \in \mathbb{R}^n$, $G = \operatorname{Vect}\langle U\rangle$ et p_G la projection orthogonale sur G et $M_A = A - p_G(A)$.

On a $M_A \in F$ et pour tout $M \in F$,

$$\|AM\|^2 = \|AM_A\|^2 + \|M_A M\|^2$$

Il suffit d'écrire $M - A = (M - M_A) + (M_A - A)$ et se servir du fait, à vérifier, que $\langle M - M_A, M_A - A\rangle = 0$.

On a donc, pour tout $M \in F$, $\|AM\| = d(A,M) \geq \|AM_A\| = d(A, M_A)$ avec égalité si et seulement si $M = M_A$.

En reformulant, on a donc, comme $M_A \in F$ que

$$d(A,F) = AM_A$$

et que M_A est le seul point M de F pour lequel cette égalité est vraie.

Remarquons qu'on a de plus

$$\forall A \in \mathbb{R}^n,\ d(A,F)^2 = (\langle A, U\rangle)^2$$

□

Proposition-Définition 30. *L'application $p_F : \mathbb{R}^n \to \mathbb{R}^n$ est un endormorphisme. On a*

1. *$p_F + p_G = Id_{\mathbb{R}^n}$ où $G = Vect\langle U\rangle$*
2. *$p_F \circ p_F = p_F$,*
3. *$Im\, p_F = F$, $Ker\, p_F = Vect\langle U\rangle = G$. Soit P_F la matrice de p_F dans la base canonique.*
1. *$P_F = I_n - U^\top .U$*
2. *P_F est symétrique et vérifie $P_F^2 = P_F$.*
3. *P_G est diagonalisable, elle est semblable à ...*

4.3 Projection orthogonale sur un sev F de $\mathbb{R}^n$

Définition, existence et unicité

Le programme officiel de BCPST2 donne la définition suivante :

Définition 31. *On appelle* projection orthogonale *sur un sous espace F de $\mathbb{R}^n$ un endomorphisme p de $\mathbb{R}^n$ tel que*

(i) $\forall x \in \mathbb{R}^n, p(x) \in F$

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

Cette définition doit être travaillée pour garantir l'existence d'(au moins) une telle projection orthogonale sur un sev F de $\mathbb{R}^n$ donné. Cependant, sous réserve d'existence, on a unicité : si p et p' vérifient (i) et (ii) alors, pour tout $x \in \mathbb{R}^n$, $p(x) = p'(x)$. En effet, si $x \in \mathbb{R}^n$ alors comme $p'(x), p(x) \in F$, d'après (ii) pour p et pour p'

$$\left\langle x - p(x), p'(x) - p(x) \right\rangle = 0 = \left\langle x - p'(x), p(x) - p'(x) \right\rangle$$

En ajoutant ces deux identités, on obtient $\|p'(x) - p(x)\|^2 = 0$, c'est à dire $p(x) = p'(x)$. On peut reformuler la définition 31 en termes matriciels.

Définition 32. *Une matrice $P \in \mathscr{M}_n(\mathbb{R})$ est matrice (relativement à la base canonique de $\mathbb{R}^n$) de projection orthogonale sur un sous-espace F de $\mathbb{R}^n$ si*

(i) $\operatorname{Im} P \subset F$ *;*

(ii) $\forall X \in \mathbb{R}^n, \forall Y \in F, (X - P.X)^\top.Y = 0$

Matrices de projections orthogonales

Ayant réglé la question de l'unicité éventuelle, on peut maintenant jouer sur cette définition pour obtenir caractérisation, puis existence de (la) projection orthogonale sur un sous-espace F.

Théorème 33. *Soit une matrice $P \in \mathscr{M}_n(\mathbb{R})$. On a équivalence entre P est matrice (relativement à la base canonique) de projection orthogonale sur $F = \operatorname{Im} P$ et*

1. $P^2 = P$
2. $P^\top = P$

Démonstration. Il y a deux cas très particuliers, si $P = 0$ ou $P = I_n$ auxquels cas l'équivalence est évidente. Soit $P \in \mathscr{M}_n(\mathbb{R})$. On procède par double équivalence.

$\Leftarrow$ Si P vérifie $P^2 = P$ et $P^\top = P$ alors :
Posons $E_1 = \operatorname{Ker}(P - I_n)$. On a $E_1 = \operatorname{Im} P = F$. En effet si $X \in \operatorname{Im} P$, il existe X' tel $X = P.X'$ et donc $P.X = P^2.X' = P.X' = X$ et donc $X \in E_1$ et réciproquement si $P.X = X$, il est clair que $X \in \operatorname{Im} P = F$. Si $X \in \mathbb{R}^n$, $Y \in F$, il vient $Y = P.Y$ et on a alors, en utilisant la règle $(P.X)^\top = X^\top.P^\top$, puis $P^\top = P$,

$$(X - P.X)^\top.Y = X^\top.Y - X^\top.P^\top.Y = X^\top.(Y - P^\top.Y) = X^\top.(Y - P.Y) = 0$$

La matrice P vérifie donc la définition 32.

$\Rightarrow$ Réciproquement, si la matrice P vérifie la définition 32.

1. Si $X \in F$ alors $P.X = X$; en effet, soit $X \in \mathbb{R}^n$, dans le point (ii) de la définition 32, on peut prendre comme vecteur $Y = X - P.X$ sachant que ce vecteur est dans F vu que $X \in F$ et $P.X \in \operatorname{Im} P \subset F$ par (i) de la définition 32. On obtient alors $X - P.X^\top.(X - P.X) = 0$, *i.e.* $\|X - P.X\|^2 = 0$, *i.e.* $X = P.X$. On en déduit que $\operatorname{Im} P = F$ et que $\forall X \in F, P.X = X$.
2. On déduit de (i) et du point précédent que $P^2 = P$. En effet si $X \in \mathbb{R}^n$ est quelconque, $P.X \in F$ et donc $P.(P.X) = (P.X)$, *i.e.* $P^2.X = P.X$.
3. De (ii) et du point précédent, on déduit que

$$\forall X, Y \in \mathbb{R}^n, (P.X)^\top.Y = (P.X)^\top.P.Y = X^\top.P.Y$$

En effet : Soit $X, Y \in \mathbb{R}^n$ quelconques. Par (ii), comme $P.X \in F$, $(Y - P.Y)^\top.P.X = 0$ et donc, en transposant, $(P.X)^\top.(Y - P.Y) = 0$ et donc

$$(P.X)^\top.Y = (P.X)^\top.P.Y$$

Par symétrie, on a aussi

$$(P.Y)^\top.X = (P.Y)^\top.P.X$$

et en tranposant cette identité

$$(P.X)^\top.Y = (P.X)^\top.P.Y = (P.Y)^\top.P.X = X^\top.P.Y$$

4. Du point précédent appliqué aux couples de vecteurs (X, Y) où X et Y décrivent indépendamment les n vecteurs de la base canonique de $\mathbb{R}^n$, on déduit (entrée par entrée) que $P^\top = P$.

$\square$

Théorème 34. *Soit F un sev de $\mathbb{R}^n$, p_F, la projection orthogonale sur F existe et, pour calculer sa matrice P (relativement à la base canonique), on a les formules :*

(BCPST2) si $(U_1, \ldots, U_r)$ est une base orthonormale de F, alors

$$P = U_1.U_1{}^\top + \cdots + U_r.U_r{}^\top;$$

(HP) si $(V_1, \ldots, V_r)$ est une base quelconque de F, en formant $V = (V_1|\ldots|V_r) \in \mathscr{M}_{n,r}(\mathbb{R})$, la matrice $V^\top.V \in \mathscr{M}_r(\mathbb{R})$ est inversible[2] *et*

$$P = V.(V^\top.V)^{-1}.V^\top.$$

Démonstration. 1. Soit $(U_1, \ldots, U_r)$ une base orthonormée de F. Nommons temporairement

$$P = U_1.U_1{}^\top + \cdots + U_r.U_r{}^\top$$

(a) On a $\operatorname{Im} P = F$. L'inclusion $\operatorname{Im} P \subset F$ étant due au fait que

$$\forall X \in \mathbb{R}^n, P.X = \sum_{k=1}^{r} \langle x, U_k \rangle .U_k$$

L'inclusion réciproque étant due au fait que pour $1 \leq k \leq r$, $P.U_k = U_k$

2. Attention ! ! ! Si $r < n$, V n'est pas inversible, on ne peut simplifier cette écriture ! ! Si $r = n$, V est inversible et la simplification donne $P = I_n$.

(b) On a $P^\top = P$ par les régles sur les transposées de sommes et de produits.

(c) Finalement

$$\begin{aligned} P^2 &= \left(\sum_{k=1}^{r} U_k.U_k{}^\top\right)^2 = \left(\sum_{k=1}^{r} U_k.U_k{}^\top\right).\left(\sum_{\ell=1}^{r} U_\ell.U_\ell{}^\top\right) \\ &= \sum_{1\leq k,\ell\leq r} U_k.U_k{}^\top.U_\ell.U_\ell{}^\top \end{aligned}$$

Comme, pour tout $\ell,k \in \{0,\dots,n\}$, $U_k{}^\top.U_\ell = \begin{cases} 1 & \text{si } k=\ell \\ 0 & \text{si } k\neq\ell \end{cases}$, il vient, la somme se simplifiant,

$$P^2 = \sum_{1\leq k\leq r} U_k.U_k{}^\top = P$$

D'après le théorème 33, P est matrice de projection orthogonale sur F. On peut donc définir p_F, l'endomorphisme projection orthogonale sur F.

2. Montrons d'abord que $V^\top.V \in \mathscr{M}_r(\mathbb{R})$ est inversible. Soit $X \in \mathbb{R}^r$, $X \in \operatorname{Ker} V^\top.V$. On a alors $V^\top.V.X = 0$ et donc

$$0 = \left\langle V^\top.V.X, X\right\rangle = \|V.X\|^2$$

D'où $X \in \operatorname{Ker} V$ mais, par le théorème du rang $\operatorname{Ker} V = \{0\}$ et donc $X = 0$. On a donc $\operatorname{Ker} V^\top.V = \{0\}$, ce qui, pour une matrice carrée, garantit l'inversibilité. Posons $Q = V.(V^\top.V)^{-1}.V^\top$.

(a) On a $Q^2 = Q$ et $Q^\top = Q$.

(b) On a clairement $\operatorname{Im} Q \subset \operatorname{Im} V = F$. Si $X \in \operatorname{Im} V$, il existe $Y \in \mathbb{R}^r$ tel que $X = V.Y$ et on a

$$Q.X = V.\underbrace{(V^\top.V)^{-1}.V^\top.V}_{I_r}.Y = V.Y = X$$

et donc $X \in \operatorname{Im} Q$.

(c) En conclusion, on a $\operatorname{Im} Q = F$ et, par le théorème tout juste démontré, $Q = P_F$, la matrice de p_F.

□

La première formule est officiellement au programme sous une forme « endormorphisme ».

Théorème 35 (Théorème officiel BCPST2). *Soit F un sous-espace de $\mathbb{R}^n$ de dimension $r > 0$, $(u_1,\dots,u_r)$ une B.O.N. de F. La projection orthogonale sur F, p_f se calcule par*

$$\forall x \in \mathbb{R}^n,\ p_F(x) = \sum_{k=1}^{r} \langle x, u_k\rangle .u_k = \sum_{k=1}^{r} u_k.\langle u_k, x\rangle$$

La deuxième formule proposée n'est pas officiellement au programme. La démontrer peut par contre faire l'objet d'une belle question d'écrit ou d'oral. Dans la même veine de façon « duale »

Exercice 15.— Soit $E \in \mathscr{M}_{r,n}(\mathbb{R})$ une matrice de rang r. Soit

$$F = \{X \in \mathbb{R}^n,\ E.X = 0\}$$

L'espace F est donc décrit par un système de r équations cartésiennes, indépendantes, chaque ligne de E représentant une ligne d'équation.

1. Quelle est la dimension de F ?
2. Montrer que la matrice $E.E^\top \in \mathscr{M}_r(\mathbb{R})$ est inversible.
3. Montrer que $Q = I_n - .E^\top.(E.E^\top)^{-1}.E$ est la matrice de la projection orthogonale sur F relativement à la base canonique.
4. Montrer que pour tout $x \in \mathbb{R}^n$,

$$d(x,F)^2 = \left\langle (E.E^\top)^{-1}.E.x, E.x \right\rangle$$

et simplifier cette formule dans le cas où les lignes de E forment une famille orthonormée dans $\mathbb{R}^n$.

Exercice 16.— Soit F un sev de $\mathbb{R}^n$. p_F la projection orthogonale sur F et

$$G = \operatorname{Ker} p_F$$

1. $G = \{x \in \mathbb{R}^n, \forall y \in F, \langle x, y\rangle = 0\}$
2. La projection orthogonale sur G est donnée par $p_G = \mathrm{id}_{\mathbb{R}^n} - p_F$ et $F = \operatorname{Ker} p_G$.
3. $F = \{x \in \mathbb{R}^n, \forall y \in G, \langle x, y\rangle = 0\}$
4. (PYTHAGORE). On a

$$\forall x \in \mathbb{R}^n, x = p_F(x) + p_G(x) \text{ et } \|x\|^2 = \|p_F(x)\|^2 + \|p_G(x)\|^2$$

Exercice 17.—Dans $\mathbb{R}^3$, on se donne trois vecteurs

$$u = \begin{pmatrix}1\\1\\1\end{pmatrix}, v = \begin{pmatrix}1\\2\\1\end{pmatrix}, w = \begin{pmatrix}-1\\1\\1\end{pmatrix}$$

et $F = \operatorname{Vect}\langle u, v\rangle$. Donner quelques moyens de calculer $d(A,F)$ lorsque A a pour coordonnées (x_A, y_A, z_A).

Exercice 18.— Soit $a = (a_1, a_2, a_3) \in \mathbb{R}^3$ avec $\|a\|^2 = 1$.
1. Donner la matrice P_a, relativement à la base canonique de $\mathbb{R}^3$, de p_a, la projection orthogonale sur $\operatorname{Vect}\langle a\rangle$.
2. Donner la matrice V_a, relativement à la base canonique de $\mathbb{R}^3$, de l'endomorphisme v_a défini [3] par

$$\forall x = (x_1, x_2, x_3) \in \mathbb{R}^3,\ v_a(x) = \begin{pmatrix} \begin{vmatrix} a_2 & x_2 \\ a_3 & x_3 \end{vmatrix} \\ \\ \begin{vmatrix} a_3 & x_3 \\ a_1 & x_1 \end{vmatrix} \\ \\ \begin{vmatrix} a_1 & x_1 \\ a_2 & x_2 \end{vmatrix} \end{pmatrix} = a \wedge x$$

V_a est-elle symétrique ?
3. Montrer que V_a^2 est symétrique (sans gros calcul).
4. Calculer V_a^2 et montrer que $V_a^2 = P_a - I_3$.
5. Quelles sont les valeurs propres de V_a^2 ? V_a est-elle diagonalisable sur $\mathbb{R}$?

3. il s'agit dans chaque composante du déterminant de la matrice 2×2

6. V_a est-elle diagonalisable sur $\mathbb{C}$?

Exercice 19.— Dans $\mathbb{R}^n$ muni de son produit scalaire canonique :
1. Ecrire la matrice (relativement à la base canonique) de la projection orthogonale p sur $F = \text{Vect}\langle u_1, u_2\rangle$ avec

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ et } u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

2. Déterminer

$$F^{\perp} = \{v \in \mathbb{R}^4, \forall u \in F, \langle u, v\rangle = 0\}$$

On pourra commencer par prouver que $v \in F^{\perp} \Leftrightarrow \langle v, u_1\rangle = \langle v, u_2\rangle = 0$.

Ecrire la matrice (relativement à la base canonique) de la projection orthogonale q sur $F^{\perp}$.

Quelques commentaires sur les valeurs propres d'une projection, le théorème spectral

Si p_F est la projection orthogonale sur un certain s.e.v F de $\mathbb{R}^n$ alors $p_F^2 = p_F$ et donc les seules valeurs propres possibles de p_F sont 0 et 1. Par ailleurs, les espaces (propres de p_F si non nuls)

$$E_1 := \text{Ker}\,(p_F - \text{id}) = F = \text{Im}\, p_F \text{ et } E_0 := \text{Ker}\, p_F$$

vérifient, par le théorème du rang $\dim E_0 + \dim E_1 = n$, et donc p_F est diagonalisable. Par aillleurs, de l'orthogonalité des vecteurs propres d'une atrice symétrique associée à des valeurs propres distinctes, on a

$$\forall x \in E_0, \forall y \in E_1, \langle x, y\rangle = 0$$

Si $(u_1, \ldots u_r)$ est une base (orthonormée ou pas) de E_1, $u_{r+1}, \ldots, u_n$ est une base de E_0 alors $\mathscr{U} = (u_1, \ldots, u_n)$ est ne base de $\mathbb{R}^n$ composée de vecteurs de p_F. Dans cette base, la matrice de p_F est

$$^{\mathscr{U}}\,[p_F]_{\mathscr{U}} = \begin{pmatrix} I_r & 0_{r,n-r} \\ 0_{n-r,r} & 0_{n-r,n-r} \end{pmatrix}$$

L'endormorphisme $p_G = \text{id} - p_F$ est la projection orthogonale sur $G = E_0$. Son noyau est F. Il a pour matrice dans a base $\mathscr{U}$

$$^{\mathscr{U}}\,[p_G]_{\mathscr{U}} = \begin{pmatrix} 0_{r,r} & 0_{r,n-r} \\ 0_{n-r,r} & I_{n-r} \end{pmatrix}$$

Le théorème spectral affirme l'existence, pour un endomorphisme $f : E \to E$ (avec $E = \mathbb{R}^n$) dont la matrice relativement à une base orthonormée est symétrique d'une famille $\lambda_1, \ldots, \lambda_p$ de nombres réels distincts tels que, en posant

$$E_k = \text{Ker}\,(f - \lambda_k.\text{id})$$

on a

$$f = \sum_{k=1}^{p} \lambda_k . p_{E_k}$$

La solution du problème de minimisation

Théorème 36. *Soit F un sev de $\mathbb{R}^n$, p_F la projection orthogonale sur F. On a*

$$\forall x \in \mathbb{R}^n, \forall y \in F, \|x-y\|^2 = \|x-p_F(x)\|^2 + \|p_F(x)-y\|^2$$

En conséquence $p_F(x)$ vérifie

1. $p_F(x) \in F$
2. $d(x, p_F(x)) = \inf_{y\in F} d(x,y) = d(x,F)$
3. $p_F(x)$ *est l'unique minimant de* $y \mapsto d(x,y)$ *sur* F.

Démonstration. Soit $x \in \mathbb{R}^n$. Pour $y \in F$, on a

$$\|x-y\|^2 = \|(x-p_F(x))-(p_F(x)-y)\|^2 = \|(x-p_F(x))\|^2 + \|(p_F(x)-y)\|^2 - 2\underbrace{\langle x-p_F(x), p_F(x)-y\rangle}_{=0 \text{ car } p_F(x)-y\in F}$$

On a donc, pour $y \in F$,

$$\|x-y\| \geq \|x-p_F(x)\|$$

avec égalité si et seulement si $y = p_F(x)$ □

5 Retour aux statistiques

On peut reformuler les résultats des paragraphes 1.1 et 1.2 en termes de projections orthogonales.

On considère que l'on a a disposition des séries de données réelles, indiquant les valeurs de caractères sur une population de N individus,

$$X = (x_k)_{k\in\{1,\dots,N\}}, Y = (y_k)_{k\in\{1,\dots,N\}}$$

On note $\bar{x} = \frac{1}{N}\sum_{k=1}^{N} x_k$, la moyenne de la série X, σ_x l'écart-type de cette série, avec des notations analogues pour la série Y et enfin, on note

$$c_{xy} = \frac{1}{N}\sum_{k=1}^{N}(x_k-\bar{x}).(y_k-\bar{y}) = \frac{1}{N}\sum_{k=1}^{N} x_k.y_k - \bar{x}.\bar{y}$$

la covariance de la série bivariée (X,Y).

On note $\mathbb{1}$ le vecteur de $\mathbb{R}^N$ ne comportant que des 1. On remarque les liens suivant avec le produit scalaire usuel sur $\mathbb{R}^N$:

$$\bar{x} = \frac{1}{N}\langle X, \mathbb{1}\rangle, \sigma_x^2 = \frac{1}{N}\|X-\bar{x}\mathbb{1}\|^2, c_{xy} = \frac{1}{N}\langle X-\bar{x}\mathbb{1}, Y-\bar{y}\mathbb{1}\rangle,$$

relations que l'on peut aussi écrire matriciellement.

5.1 Moyenne

Si $a \in \mathbb{R}$, on pose[4] $f(a) := \frac{1}{N}\sum_{k=1}^{N}(x_k - a)^2 = \frac{1}{N}\|X - a.\mathbb{1}\|^2$.

Rendre $f(a)$ minimal, revient donc à rendre minimale la quantité $\|X - u\|$ lorque u varie dans $F = \mathrm{Vect}\langle \mathbb{1}\rangle$, *i.e.* à calculer $d(X,F)$. Ce minimum est atteint (uniquement) en $u = p_F(X)$ où p_F désigne la projection orthogonale sur F.

Comme $(u_0) := (\frac{1}{\sqrt{N}}.\mathbb{1})$ est une base orthonormale de F, la matrice de p_F relativement à la base canonique est

$$P_F = u_0.u_0^{\top} = \frac{1}{N}\mathbb{1}.\mathbb{1}^{\top}$$

et on a matriciellement[5],

$$u = p_F(X) = \frac{1}{N}\mathbb{1}.\mathbb{1}^{\top}.X = \bar{x}.\mathbb{1}$$

La quantité $f(a)$ est donc minimale (uniquement) en $a = \bar{x}$ et ce minimum vaut $\sigma_x^2 = \frac{1}{N}.d(X, \mathrm{Vect}\langle\mathbb{1}\rangle)^2$.

5.2 Droite des moindres carrés

Si $a,b \in \mathbb{R}$, on pose[6] $f(a,b) := \frac{1}{N}\sum_{k=1}^{N}(y_k - (a.x_k + b))^2 = \frac{1}{N}\|Y - a.X - b.\mathbb{1}\|^2$.

Rendre $f(a,b)$ minimal, revient donc à rendre minimale la quantité $\|Y - u\|$ lorque u varie dans $F = \mathrm{Vect}\langle \mathbb{1}, X\rangle$, *i.e.* à calculer $d(Y,F)$. Ce minimum est atteint (uniquement) en $u = p_F(Y)$ où p_F désigne la projection orthogonale sur F. Supposons X et $\mathbb{1}$ non colinéaires de sorte que $\dim F = 2$.

Posons $u_0 = \frac{1}{\sqrt{N}}.\mathbb{1}$, $v_x = X - \langle u_0, X\rangle.u_0 = X - \bar{x}.\mathbb{1}$ et[7] $u_x = \frac{1}{\|v_x\|}v_x = \frac{1}{\sqrt{N}.\sigma_X}.v_x$.

La famille (u_0, u_x) est une base orthonormale de F, la matrice de p_F relativement à la base canonique est

$$P_F = u_0.u_0^{\top} + u_x.u_x^{\top} = \frac{1}{N}\mathbb{1}.\mathbb{1}^{\top} + \frac{1}{N.\sigma_x^2}.(X - \bar{x}.\mathbb{1}).(X - \bar{x}.\mathbb{1})^{\top}$$

et on a matriciellement,

$$u = p_F(Y) = \frac{1}{N}\mathbb{1}.\mathbb{1}^{\top}.Y + \frac{1}{N.\sigma_x^2}.(X - \bar{x}.\mathbb{1}).(X - \bar{x}.\mathbb{1})^{\top}.Y$$

En remarquant que v_x et $\bar{y}\mathbb{1}$ sont orthogonaux, on a

$$\begin{aligned} \frac{1}{N.\sigma_x^2}.(X - \bar{x}.\mathbb{1}).(X - \bar{x}.\mathbb{1})^{\top}.Y &= \frac{1}{N.\sigma_x^2}.(X - \bar{x}.\mathbb{1}).(X - \bar{x}.\mathbb{1})^{\top}.(Y - \bar{y}.\mathbb{1}) \\ &= \frac{1}{N.\sigma_x^2}.(X - \bar{x}.\mathbb{1}).N.c_{xy} = \frac{c_{xy}}{\sigma_x^2}.(X - \bar{x}.\mathbb{1}) \end{aligned}$$

et finalement

$$u = p_F(Y) = \bar{y}.\mathbb{1}. + \frac{c_{xy}}{\sigma_x^2}.(X - \bar{x}.\mathbb{1})$$

La quantité $f(a,b)$ est donc minimale (uniquement) en $a = \frac{c_{xy}}{\sigma_x^2}$ et $b = \bar{y} - \frac{c_{xy}}{\sigma_x^2}.\bar{x}$.

4. Il s'agit de la fonction f définie en 1.1. On a remplacé n par N.
5. remarquer que $\mathbb{1}.\mathbb{1}^{\top}$ est la matrice $N \times N$ composée uniquement de 1
6. Il s'agit de la fonction f définie en 1.2.
7. On a $\|v_x\|^2 = N.\sigma_x^2$

6 Un exemple concret

Des systèmes sans solution exacte

Revenons au problème, traité dans le chapitre de bases de l'algèbre linéaire de déterminer les coefficients d'une formule à partir de données expérimentales. Rappelons le problème en considération

On sait (d'expérience !) que l'indice de réfraction n [8] de certains matériaux dépend de la longueur d'onde λ de l'onde incidente suivant une loi [9] de la forme

$$n = p_0 + p_1 . \frac{1}{\lambda^2} + p_2 . \frac{1}{\lambda^4}$$

où les coefficients réels p_0, p_1, p_2 sont caractéristiques du matériau. On a effectué des mesures, à l'aide d'un goniomètre, pour un certain type de verre, regroupées dans le tableau :

$\lambda\,(nm)$	404,7	434,7	546,1	576,9	579,1	614,9	671,6
n	1,540	1,536	1,526	1,525	1,524	1,523	1,521

Peut-on, à partir de ces mesures, déterminer les coefficients p_0, p_1, p_2 ? C'est un problème plus compliqué que celui de l'interpolation déjà traité car, en exprimant naïvement les égalités attendues, on tombe sur un système linéaire comportant beaucoup plus d'équations que d'inconnues.

On dispose ici de *sept* nombres réels distincts z_0, z_1,... z_6 (les carrés des inverses des longueurs d'onde), des sept valeurs de n correspondantes, n_0, n_1, ..., n_6 et il s'agit de trouver un polynôme P de degré ≤ 2 tel que

$$\forall i = 0, 1, .., 6,\ P(z_i) = n_i$$

En écrivant les choses telles qu'elles viennent, on voit que ce problème se réduit à résoudre le système linéaire d'inconnue (p_0, p_1, p_2),

$$\left\{ \begin{array}{cccc} p_0 & +p_1.z_0 & +p_2.z_0^2 & = & n_0 \\ p_0 & +p_1.z_1 & +p_2.z_1^2 & = & n_1 \\ \vdots & & & = \vdots \\ p_0 & +p_1.z_6 & +p_2.z_6^2 & = & n_6 \end{array} \right.$$

On laisse le lecteur remplacer les paramètres z_i, n_i par leurs valeurs obtenues à partir du tableau, fourrer le système linéaire dans un système de calcul afin d'obtenir... un système qui n'admet pas de solution.

La meilleure solution approchante

L'absence de solution est due, en partie, aux erreurs de mesures que l'on peut faire. Si l'on croit que la loi de CAUCHY est exacte (ce qui est déjà faux, c'est une loi empirique) et que l'on fait des mesures absolument sans erreurs, on devrait trouver une unique solution. Comme ces erreurs existent quoiqu'il arrive, que notre loi n'est pas exacte, il ne faut pas espérer avoir de solution... On relaxe donc le problème de la façon suivante : on cherche les(des) paramètres (p_0, p_1, p_2) tels que la quantité

$$SC(p_0, p_1, p_2) = \sum_{k=0}^{6} \left(n_k - (p_0 + p_1.z_k + p_2.z_k^2)\right)^2$$

soit minimale. Le/Un triplet minimisant sera considéré comme une bonne solution à notre problème de détermination des paramètres : Ce sera une solution « aux moindres carrés ».

8. rapport entre célérité de la lumière dans le vide et vitesse de phase de l'onde lumineuse, grandeur sans dimension

9. dite loi de CAUCHY, *cf.* `https://fr.wikipedia.org/wiki/Loi_de_Cauchy_(optique)`, liée à la séparation de la lumière blanche par un prisme

Le lien avec les projections orthogonales

Reformulons cette question à l'aide des outils de géométrie développés auparavant. Soit

$$V = \begin{pmatrix} 1 & z_0 & z_0^2 \\ 1 & z_1 & z_1^2 \\ \vdots & & \vdots \\ 1 & z_6 & z_6^2 \end{pmatrix}, N = \begin{pmatrix} n_0 \\ n_1 \\ \vdots \\ n_6 \end{pmatrix}, X = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix}$$

On a

$$SC(p_0, p_1, p_2) = \|N - V.X\|^2$$

et lorsque X parcourt $\mathbb{R}^3$, $V.X$ parcourt $\operatorname{Im} V$. L'application $X \in \mathbb{R}^3 \mapsto V.X \in \mathbb{R}^7$ est de rang 3, tout comme la matrice V. Par restriction, c'est un isomorphisme de $\mathbb{R}^3$ sur $\operatorname{Im} V$.

SC est minimale pour $V.X = P_{\operatorname{Im} V}(N)$ où $P_{\operatorname{Im} V}$ est la matrice (base canonique) de la projection orthogonale sur $\operatorname{Im} V$. On a donc, par la formule (ii) du théorème 34,

$$V.X = V.(V^\top.V)^{-1}.V^\top.N$$

i.e. , par injectivité de V,

$$\begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} = (V^\top.V)^{-1}.V^\top.N$$

Implémentation

On peut implémenter cette formule en Python en y injectant nos données et tracer, conjointement, à fin de comparaison : les données, la courbe indice $= f(\lambda)$ obtenue ainsi que l'une de celle obtenue par interpolation avec trois valeurs de λ choisies parmi les 7.

Utiliser le script `indice-verre.py`. Les paramètres obtenus par moindres carrés et par interpolation

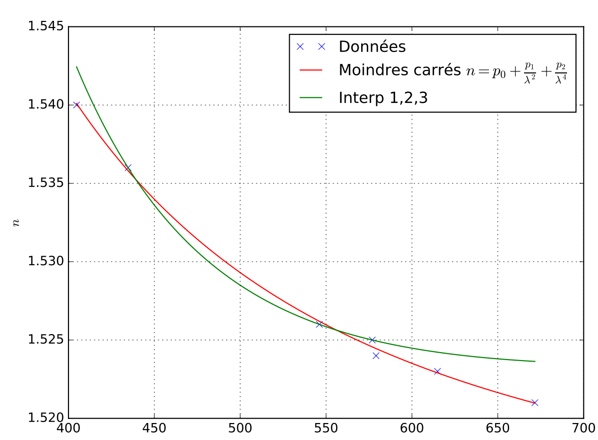

FIGURE 1 – Comparaison entre courbes obtenues par moindre carrés et courbes obtenues par interpolation

de trois valeurs diffèrent grandement

1. `> Paramètres moindres carrés= [ 369325.69 -1751726.94 2079437.34]`
2. `> Paramètres interp= [ 1020661.61 -4802523.90 5651709.20]`

7 Exercices Supplémentaires

Exercice 20.— Vrai - Faux Dans $\mathbb{R}^n$ muni de son produit scalaire canonique :
1. $\forall x,y,z,t \in \mathbb{R}^n$, $\forall \alpha,\beta \in \mathbb{R}$, $\alpha \langle x,y\rangle + \beta \langle z,y\rangle = \langle \alpha x + \beta z, y\rangle$
2. $\forall x,y,z,t \in \mathbb{R}^n$, $\langle x,y\rangle + \langle z,t\rangle = \langle x+z, y+t\rangle$
3. $\forall x,y \in \mathbb{R}^n$, $\langle x,y\rangle = 0 \Rightarrow x = 0$ ou $y = 0$
4. $\forall x \in \mathbb{R}^n$, $\langle x,x\rangle = 0 \Rightarrow x = 0$
5. $\forall x \in \mathbb{R}^n$, $(\forall z \in \mathbb{R}^n,\ \langle x,z\rangle = 0) \Rightarrow x = 0$
6. $\forall (x,y) \in (\mathbb{R}^n)^2,\ \|x+y\| = \|x\| + \|y\|$
7. Toute famille de vecteurs orthogonaux deux à deux est libre.
8. Si A est symétrique réelle et si $\mathscr{C}$ est une base de vecteurs propres de A, alors $\mathscr{C}$ est une base orthonormée de $\mathbb{R}^n$.

Exercice 21.— Soit f un endomorphisme non nul de $\mathbb{R}^n$ tel que $\forall x \in \mathbb{R}^n, \quad \langle f(x), x\rangle = 0$.
1. Montrer que : $\forall x,y \in (\mathbb{R}^n)^2, \quad \langle f(x), y\rangle = -\langle x, f(y)\rangle$
2. Montrer que

$$\operatorname{Ker} f = (\operatorname{Im} f)^{\perp} := \{y \in \mathbb{R}^n, \forall x \in \mathbb{R}^n, \langle y, f(x)\rangle = 0\}$$

3. Montrer que si λ est valeur propre de f, alors $\lambda = 0$. f est-il diagonalisable sur $\mathbb{R}$?
4. Soit $\mathscr{C}$ la réunion d'une base orthonormale de $\operatorname{Ker} f$ et d'une base orthonormale de $\operatorname{Im} f$. Donner la forme de la matrice de f dans cette base.

Exercice 22.— Endomorphismes orthogonaux. Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une application (pas forcément linéaire), on dit que f est une isométrie, si

$$\forall x,y \in \mathbb{R}^n,\ \|f(x) - f(y)\| = \|x - y\|$$

Si, de plus, f est endomorphisme de $\mathbb{R}^n$, on dit aussi que f est un endomorphisme *orthogonal*
1. Montrer que si f est une isométrie, alors f est injective. Montrer que si f est bijective alors f^{-1} est aussi une isométrie de $\mathbb{R}^n$.
2. Une projection orthogonale est-il une isométrie ? (ie : est-ce un « endomorphisme orthogonal » au sens que l'on vient de définir)
3. Soit $f : \mathbb{R}^n \to \mathbb{R}^n$ une isométrie. Montrer que f_0 définie par $\forall x \in \mathbb{R}^n$, $f_0(x) = f(x) - f(0)$ est aussi une isométrie. Réciproque ?
4. Soit f une application de $\mathbb{R}^n$ dans $\mathbb{R}^n$ *surjective* (attention : f n'est pas supposée linéaire !) telle que $f(0) = 0$
Montrer, en calculant $\|f(x) - f(y)\|^2$ de deux façons différentes, que :

$$f \text{ est une isométrie } \iff \forall (x,y) \in (\mathbb{R}^n)^2,\ \langle f(x), f(y)\rangle = \langle x,y\rangle .$$

En déduire que f est linéaire [10]. On pourra utiliser le principe d'identification suivant : Soient x, $x' \in \mathbb{R}^n$ alors

$$x = x' \Leftrightarrow \forall z \in \mathbb{R}^n,\ \langle x,z\rangle = \langle x',z\rangle$$

5. Soit $f \in \mathscr{L}(\mathbb{R}^n)$. Montrer que : f est une isométrie $\iff$ la matrice de f dans une B.O.N. est orthogonale.

10. Important : on a défini une isométrie uniquement en termes de distances, il s'avère *a posteriori* que les isométries surjectives de $\mathbb{R}^n$ sont linéaires, à translation près