INU Champollion

Licence Info 3ème Année

Maths pour l'I.A.

Thierry Montaut

2021

Plan du cours

- Compléments d'algèbre linéaire
  - Rappels de diagonalisation
  - Application au calcul des puissances de matrice. Chaînes de Markov
  - **Espace euclidien $\mathbb{R}^n$: Produits scalaires, orthogonalité**
  - **Bases orthogonales. Projections orthogonales,**
  - **Diagonalisation des matrices symétriques**
- Compléments d'analyse
- Compléments de probabilité et de statistiques

Produit scalaire

Définition

De manière générale, si E est un $\mathbb{R}$ espace vectoriel, une application $f : E \times E \to \mathbb{R}$ est appelée **produit scalaire** *sur E ssi f vérifie les propriétés suivantes :* $\forall (X, Y, Z) \in E^3, \ \ \forall \lambda \in \mathbb{R}$

1. $f(X, Y + Z) = f(X, Y) + f(X, Z)$
   $f(X, \lambda Y) = \lambda f(X, Y)$.
   On dit que f est linéaire à droite.
2. $\forall (X, Y, Z) \in E^3, \ \ \forall \lambda \in \mathbb{R}$
   $f(X + Y, Z) = f(X, Z) + f(Y, Z)$
   $f(\lambda X, Y) = \lambda f(X, Y)$.
   On dit que f est linéaire à gauche.
   Une application linéaire à gauche et à droite est dite **bilinéaire**.

Définition

3. *f est* **symétrique** *:*

$$f(X,Y) = f(Y,X)$$

4. *f est* **"définie positive"** *:*

$$\forall X \in E, f(X,X) \geq 0 \text{ (positive)}$$

$$f(X,X) = 0 \iff X = 0 \text{ (définie).}$$

Un produit scalaire est donc une **application bilinéaire, symétrique, définie positive***.*
Dans toute la suite, on notera $(X|Y)$ le produit scalaire de X et de Y.

l'espace vectoriel $\mathbb{R}^n$.

1. Dans tout ce chapitre, on ne s'intéressera qu'à l'espace vectoriel de dimension finie $E = \mathbb{R}^n$ que l'on cherchera à enrichir à l'aide de produits scalaires et de normes.
2. On considère $\mathbb{R}^n$ muni de sa base canonique $(e_1, \dots, e_n)$ et on représente $X \in \mathbb{R}^n$ par son vecteur composante Xdans cette base. Donc pour $X = (x_1, \dots, x_n) = x_1 e_1 + \dots + x_n e_n$, on considérera

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

Dit autrement, nous considérerons toujours nos vecteurs de $\mathbb{R}^n$ comme des vecteurs colonnes.

Produit scalaire de $\mathbb{R}^n$.

Définition

On définit un produit scalaire sur $\mathbb{R}^n$ de la manière suivante :
Si $X =^t (x_1, \ldots, x_n)$ et $Y =^t (y_1, \ldots, y_n)$ alors

$$(X|Y) = \sum_{i=1}^{n} x_i y_i =^t X.Y$$

Ce produit scalaire est dit produit scalaire canonique de $\mathbb{R}^n$. Nous ne considérerons que lui dans toute la suite même si ...

Ce n'est pas l'unique produit scalaire sur $\mathbb{R}^n$. Par exemple, on pourra vérifier que

$$(X|Y) = x_1y_1 + 2x_2y_2 + 3x_3y_3 + x_1y_2 + x_2y_1$$

définit bien un produit scalaire sur $\mathbb{R}^3$ mais pas

$$(X|Y) = 2x_1^2 + x_3^2 + 3x_1x_2 + 5x_2x_3$$

ni

$$(X|Y) = x_1^2 - x_2^2 + 2x_3^2$$

Espace euclidien

Définition

On appelle **espace euclidien** *un $\mathbb{R}$ espace vectoriel E* **de dimension finie** *muni d'un produit scalaire.*
Quand nous parlerons de l'espace euclidien $\mathbb{R}^n$ canonique, cela voudra dire que nous munissons $\mathbb{R}^n$ de son produit scalaire canonique.

Notons (sans dire encore pourquoi...)

$$||X|| = \sqrt{(X|X)} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Alors

Théorème

Inégalité de Cauchy-Schwarz

$\forall (X, Y) \in E^2$:

$$|(X|Y)| \leq ||X||.||Y||$$

Proposition

Inégalité Triangulaire

$\forall (X, Y) \in E^2$:

$$||X + Y|| \leq ||X|| + ||Y||$$

Normes sur un espace vectoriel

Définition

Soit E un $\mathbb{R}$ espace vectoriel, $N : E \to \mathbb{R}$ est une norme sur E ssi elle vérifie les 4 propriétés suivantes :

1. $\forall X \in E, \quad N(X) \geq 0.$
2. $\forall X \in E, \quad N(X) = 0 \iff X = 0.$
3. $\forall X \in E, \quad \forall \lambda \in \mathbb{R}, \quad N(\lambda X) = |\lambda| N(X).$
4. $\forall (X, Y) \in E^2, \quad N(X + Y) \leq N(X) + N(Y).$

Définition

L'application

$$||.|| : \left(\begin{array}{ccc} E & \rightarrow & \mathbb{R} \\ X & \mapsto & ||X|| = \sqrt{(X|X)} \end{array} \right.$$

est une norme sur E dite **norme euclidienne**
(i.e. associée à un produit scalaire).

On a donc

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

Normes usuelles sur $\mathbb{R}^n$

La norme euclidienne n'est la seule norme de $\mathbb{R}^n$. On utilise souvent les deux autres normes classiques suivantes :

- La norme 1 :

$$\|X\|_1 = \sum_{i=1}^{n} |x_i|$$

- La norme infinie ou norme sup :

$$\|X\|_\infty = \max_{i \in [\![1,n]\!]} |x_i|$$

Distances

Les normes nous permettent de calculer des distances. Plus précisément, si on fixe une norme $||.||$ sur E, on définit :

Définition

- *La distance entre deux vecteurs X et Y de $\mathbb{R}^n$ par*

$$d(X,Y) = ||X - Y||,$$

- *Si $A = (x_1, \ldots, x_n)$ et $B = (y_1, \ldots, y_n)$ sont deux point de $\mathbb{R}^n$ alors*

$$d(A,B) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2},$$

- *La distance entre X et une partie F de E par*

$$d(X,F) = min\{d(X,Y), \ \ Y \in A\}.$$

Normes usuelles sur $\mathbb{R}^n$

Pour les trois normes classiques de $\mathbb{R}^2$:

- Calculer la distance entre les vecteurs $(1,2)$ et $(4,-1)$.
- Dessiner l'ensemble des éléments de $\mathbb{R}^2$ à une distance 1 de $(0,0)$.
- Calculer la distance entre les vecteurs $(1,2)$ et la droite d'équation $y=0$.

Orthogonalité dans $\mathbb{R}^n$.

Définition

X et Y sont dits orthogonaux ssi $(X|Y) = 0$. *On note alors* $X \perp Y$.

Proposition

Théorème de Pythagore dans $\mathbb{R}^n$

$$X \perp Y \;\Leftrightarrow\; ||X+Y||^2 = ||X||^2 + ||Y||^2.$$

Familles et bases orthogonales

Définition

- *Si* $(X_1, \dots, X_p)$ *est une famille de vecteurs non nuls 2 à 2 orthogonaux, alors elle est libre.*
  Une telle famille est dite **famille orthogonale**.
- *Si de plus* $\forall i \in \llbracket 1, n \rrbracket, ||X_i|| = 1$, *la famille est dite* **orthonormée**.
- *Si* $p = n$, *cette famille (dont on rappelle qu'elle est toujours libre) constitue donc une base de* E.
- *On parle alors de* **base orthogonale (BOG)** *et de* **bases orthonormées (BON)**.

Orthogonal d'un sous-espace F de E.

Définition

Soit F un sous-espace vectoriel de E de base $(f_1, \ldots, f_p)$,
On appelle **orthogonal de** F **dans** E *l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F i.e.*

$$F^{\perp} = \{X \in E, \ \forall Y \in F, \ X \perp Y\}.$$

Mais être orthogonal à tout vecteur de F équivaut à être orthogonal à tous les vecteurs d'une base de F, donc

$$F^{\perp} = \{X \in E, \ \forall i \in [\![1, p]\!], \ X \perp f_i\}.$$

Exemple : Soit $E = \mathbb{R}^4$ et $F = vect\{(1,1,0,1),(0,1,2,1)\}$
Donner les équations cartésiennes puis une base de $A^{\perp}$.

Projection orthogonale

Théorème

Théorème fondamental

Soit F un s.e.v. de E. On a

1. $E = F \oplus F^{\perp}$.
2. $(F^{\perp})^{\perp} = F$.

C'est pourquoi $F^{\perp}$ est appelé "le supplémentaire orthogonal" de F.

Hyperplan

Définition

Rappel : Si E est un espace vectoriel de dimension finie n, on appelle hyperplan de E tout sev de E de dimension $n-1$.

Propriété

- *D'après le théorème fondamental, le supplémentaire orthogonal d'un hyperplan H est une droite vectorielle.*
- *Tout vecteur non nul de son orthogonal en est donc une base. Un tel vecteur $n = (a_1, \ldots, a_n)$ est appelé vecteur normal à H.*
- *Une équation de H est alors $a_1 x_1 + \ldots + a_n x_n = 0$.*

Projection orthogonale

Définition

Soit F un s.e.v. de E.
F et $F^{\perp}$ étant supplémentaires, on peut définir la projection sur F parallèlement à $F^{\perp}$.

$$p_F : X = x_F + x_{F^{\perp}} \mapsto x_F.$$

On l'appelle la **projection orthogonale** *sur F.*

Propriété

1. $\forall X \in E, \;\; X - p_F(X) \in F^{\perp}$. *Cela permet même de caractériser le projeté orthogonal puisque* $p_F(X)$ *est l'unique vecteur* Y *de* F *tel que* $X - Y \in F^{\perp}$.
2. *Le projeté orthogonal est très utile en calcul numérique et en particulier dans les problèmes de "meilleure approximation" puisqu'il minimise la distance entre* X *et les vecteurs de* F.
3. *Plus précisément, la distance de* X *au sev* F *est la distance de* X *à son projeté orthogonal sur* F *:*

$$d(X,F) = min_{Y\in F}(||X - Y||) = ||X - p_F(X)||$$

Calcul du projeté orthogonal, cas d'une base quelconque

Soit $\mathcal{B} = (e_1, \ldots, e_p)$ une base quelconque de F. On sait que $p_F(X)$ est l'unique vecteur Y de F tel que $X - Y$ soit orthogonal à F.
Notons $p_F(X) = \sum_{i=1}^p y_i e_i$ son écriture dans la base de F.
$p_F(X)$ est caractérisé de manière unique par :

$$\forall j \in \llbracket 1, p \rrbracket, \;\; (X - p_F(X) | e_j) = 0,$$

i.e.

$$\forall j \in \llbracket 1, p \rrbracket, \;\; (p_F(X) | e_j) = (X | e_j),$$

i.e.

$$\forall j \in \llbracket 1, p \rrbracket, \;\; \sum_{i=1}^{p} y_i (e_i | e_j) = (X | e_j),$$

i.e. $p_F(X)$ est caractérisé de manière unique par le fait que son vecteur coordonnée $\begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ est l'unique solution du système linéaire :

$$\begin{cases} (e_1|e_1)y_1 + \ldots + (e_p|e_1)y_p = (X|e_1) \\ \vdots \\ (e_1|e_i)y_1 + \ldots + (e_p|e_i)y_p = (X|e_i) \\ \vdots \\ (e_1|e_p)y_1 + \ldots + (e_p|e_p)y_p = (X|e_p) \end{cases}$$

Proposition

Si on ne dispose que d'une base $b = (e_1, \ldots, e_p)$ quelconque de F, on détermine en pratique le projeté orthogonal d'un vecteur X par $p_F(X) = \sum_{i=1}^{p} y_i e_i$ où le vecteur $(y_1, \ldots, y_p)$ est l'unique solution du système linéaire :

$$\begin{cases} (e_1|e_1)y_1 + \ldots + (e_p|e_1)y_p = (X|e_1) \\ \vdots \\ (e_1|e_i)y_1 + \ldots + (e_p|e_i)y_p = (X|e_i) \\ \vdots \\ (e_1|e_p)y_1 + \ldots + (e_p|e_p)y_p = (X|e_p) \end{cases}$$

Ce système est appelé le **système des équations normales**.

Calcul du projeté orthogonal, cas d'une base orthogonale

Dans le cas où la base b est orthogonale, seuls les coefficients diagonaux du système précédent sont non nuls. Ce système se résume donc au système diagonal

$$\begin{cases} (e_1|e_1)y_1 = (X|e_1) \\ \vdots \\ (e_i|e_i)y_i = (X|e_i) \\ \vdots \\ (e_p|e_p)y_p = (X|e_p) \end{cases}$$

de solution : $y_i = \frac{(X|e_i)}{(e_i|e_i)}, \ \ \forall i \in \ [\![1, p]\!]$

Proposition

Si on dispose d'une base orthogonale $b = (e_1, \ldots, e_p)$ de F, on détermine en pratique le projeté orthogonal d'un vecteur X par la formule

$$p_F(X) = \sum_{i=1}^{p} \frac{(X|e_i)}{(e_i|e_i)} e_i$$

Si la base est orthonormée, on peut encore réduire la formule en :

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

On voit dans ces calculs tout l'intérêt de disposer de BOG, donc toute l'importance de savoir les construire...

Exemple : Soit $E = \mathbb{R}^2$
Calculer le projeté orthogonal de $A = (1, 4)$ sur la droite d'équation $y = x$. Quelle est la distance de A à cette droite ?

Bases orthogonales, procédé de Gram-Schmidt

Théorème

Procédé d'orthogonalisation de Gram-Schmidt
Soit $(X_1, \ldots, X_p)$ *une famille libre de vecteurs de* E. $\forall i \in [\![1, p]\!]$, *on note* $H_i = Vect(X_1, \ldots, X_i)$.
On définit une famille $(e_1, \ldots, e_p)$ *par récurrence en posant*

$$e_1 = X_1$$

et $\forall i \in [\![1, p]\!]$,

$$e_i = X_i - p_{H_{i-1}}(X_i) = X_i - \sum_{j=1}^{i-1} \frac{(X_i|e_j)}{(e_j|e_j)} e_j.$$

Alors $\forall i \in [\![1, p]\!]$, *la famille* $(e_1, \ldots, e_i)$ *est une base orthogonale de* H_i.

Exemple : Soit $E = \mathbb{R}^3$
Orthogonaliser la famille $\{(1,-2,2),(-1,0,-1),(5,-3,-7)\}$.

Matrices remarquables

Soit $A \in \mathcal{M}_n(\mathbb{R})$.

Proposition

1. *A est dite symétrique ssi* $A = {}^t A$,
2. *A est dite orthogonal ssi A est inversible et que* ${}^t A = A^{-1}$. *On note* O_n *l'ensemble des matrices orthogonales de taille n.*

Matrices orthogonales

Théorème

On a équivalence entre :

1. *A est une matrice orthogonale,*
2. ${}^tA.A = I_n$,
3. *A conserve le produit scalaire :* $\forall X, Y \in E,\ \ (AX)|AY) = (X|Y)$,
4. *A conserve la norme :* $\forall X \in E,\ \ ||AX|| = ||X||$.
5. *Les colonnes de A constituent une BON de* $\mathbb{R}^n$

C'est pourquoi les géomètres parlent aussi de matrices d'isométrie.

Matrices symétriques

Proposition

1. *Si A est une matrice symétrique :* $\forall X, Y \in E,\ \ (AX|Y) = (X|AY)$,
2. *Les sous-espaces propres d'une matrice symétrique sont 2 à 2 orthogonaux.*

Diagonalisation d'une matrice symétrique

Théorème

Si A est une matrice symétrique réelle, A est diagonalisable dans une base orthonormée de vecteurs propres. La matrice de passage P est alors orthogonale, d'où :

$$A = PDP^{-1} = PD^tP.$$

En pratique : On diagonalise "classiquement" la matrice. On sait que les différents sous-espaces propres sont deux à deux orthogonaux. Il n'Y a donc qu'à orthogonaliser les bases des sous-espaces propres de dimension supérieure à 1 puis à normaliser tout ce beau monde. On obtient ainsi une BON de vecteurs propres

Exemple : Diagonaliser dans une base orthonormée la matrice : $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.