Harmonisation mathématique - Algèbre 2
M1 SID

Marcello BERNARDARA [1]

15 octobre 2014

1. Institut de Mathématiques, Bâtiment 1R2, Bureau 138 - marcello.bernardarama@math.univ-toulouse.fr - 05 61 55 87 59

Ce texte reprend largement le poly de cours de Sophie JAN écrit au profit des précédentes promotions du Master 1 SID

Chapitre 1

Quelques rappels

1.1 Quelques bases

Définition 1.1 *Soient A une matrice réelle $m \times n$. On appelle* **matrice transposée de** A *et on note $A^\top$ la matrice réelle $n \times m$ telle que $(A^\top)_{ij} = A_{ji}$: les colonnes de $A^\top$ sont les lignes de A.*

Exemple : Si $A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & 6 \end{pmatrix}$, alors $A^\top = \begin{pmatrix} 1 & -4 \\ -2 & 5 \\ 3 & 6 \end{pmatrix}$.

Proposition 1.2 $(A+B)^\top = A^\top + B^\top$; $(Ax)^\top = x^\top A^\top$; $(AB)^\top = B^\top A^\top$; $(A^{-1})^\top = (A^\top)^{-1}$.

Définition 1.3 *Une matrice A est* **symétrique** *si $A^\top = A$.*

1.2 Calcul de déterminant

Soit A une matrice carrée
- $\det(A^\top) = \det(A)$;
- si A est triangulaire, son déterminant est le produit des éléments diagonaux ;
- si A a une ligne (colonne) nulle, alors $\det(A) = 0$;
- si A a deux lignes (colonnes) identiques, alors $\det(A) = 0$.

Si B est construite à partir de A
- en échangeant 2 lignes (colonnes), alors $\det(B) = -\det(A)$;
- en multipliant une ligne (colonne) de A par un réel λ, alors $\det(B) = \lambda\det(A)$;
- en ajoutant un multiple d'une ligne (resp. colonne) de A à une autre ligne (resp. colonne), alors $\det(B) = \det(A)$.

Calcul pratique d'un déterminant :
- Cas particulier des matrices 3×3 ;
- Développement par rapport à une ligne ou une colonne après avoir utilisé les propriétés ci-dessus.

Le déterminant du produit de 2 matrices est égal au produit des déterminants (lorsque tout a un sens).

1.3 Inverse de matrice carrée

Théorème 1.4 *Soit A une matrice carrée. Les trois propositions suivantes sont équivalentes :*

1. *A est inversible ;*
2. *le système linéaire $Ax = 0$ n'a que la solution nulle $x = 0$;*
3. $\det(A) \neq 0$.

Calcul pratique d'un inverse de matrice :

- Si on note M_{ij} la matrice obtenue à partir de A en enlevant la ligne i et la colonne j et $A_{ij} = (-1)^{i+j}\det(M_{ij})$ le cofacteur de a_{ij}, on a $A^{-1} = \dfrac{adj A}{\det(A)}$ avec $adj A$ la transposée de la matrice des cofacteurs.On appelle mineur de a_{ij} le déterminant de M_{ij}.
  Peu utile sauf pour des matrices 2×2.
- algorithme de Gauss pour transformer $(A|I)$ en $(I|A^{-1})$.

1.4 Eléments propres et diagonalisation d'une matrice carrée

Définition 1.5 *Soit A une matrice carrée de taille n. On dit que $\lambda \in \mathbb{R}$ est une valeur propre de A s'il existe un vecteur non nul $u \in \mathbb{R}^n$ tel que $Au = \lambda u$. Un tel vecteur u est appelé vecteur propre de A correspondant à la valeur propre λ.*

Calcul pratique des valeurs propres d'une matrice carrée A de taille $n \times n$ en résolvant l'équation $\det(A - \lambda I) = 0$.

Théorème 1.6 *Une matrice carrée A de taille $n \times n$ est diagonalisable (c'est à dire représentable par une matrice diagonale D) si et seulement si elle possède n vecteurs propres linéairement indépendants. La matrice diagonale D ainsi définie a pour éléments les valeurs propres de A et la matrice P telle que $A = PDP^{-1}$ a pour colonnes les vecteurs propres correspondants.*

1.5 Matrices (semi-)définies positives

Définition 1.7 *On dit qu'une matrice réelle symétrique A est*

- **semi définie-positive** *si pour tout $x \in \mathbb{R}^n$, $x^\top Ax \geq 0$;*
- **définie-positive** *si pour tout $x \in \mathbb{R}^n$, $x^\top Ax \geq 0$ et $x^\top Ax = 0 \Leftrightarrow x = 0_{\mathbb{R}^n}$.*

Proposition 1.8 *Si A est une matrice réelle symétrique,*

- *A est semi-définie positive si et seulement si ses valeurs propres sont toutes ≥ 0 ;*
- *A est définie positive si et seulement si ses valeurs propres sont toutes > 0 ;*

Proposition 1.9 *(Critère de Sylvester) Si A est une matrice réelle symétrique,*

- *A est semi-définie positive si et seulement si tous ses mineurs principaux sont ≥ 0 ;*
- *A est définie positive si et seulement si tous ses mineurs principaux sont > 0.*

Chapitre 2

Formes bilinéaires et produit scalaire

2.1 Formes bilinéaires

Définition 2.1 *Soit $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ une application. On dit que β est une* **forme bilinéaire** *si elle est linéaire par rapport à chacune de ses variables :*

- *pour tous $x, x', y \in \mathbb{R}^n$ et tout $\lambda, \lambda' \in \mathbb{R}$, on a $\beta(\lambda x + \lambda' x', y) = \lambda\beta(x, y) + \lambda'\beta(x', y)$;*
- *pour tous $x, y, y' \in \mathbb{R}^n$ et tout $\mu, \mu' \in \mathbb{R}$, on a $\beta(x, \mu y + \mu' y') = \mu\beta(x, y) + \mu'\beta(x, y')$.*

Exemple : L'application $\beta_2 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ qui au couple (x, y) associe leur déterminant dans la base canonique de $\mathbb{R}^2$ est une forme bilinéaire.

Par définition, $\beta_2(x, y) = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - x_2 y_1$.

On a donc

$$\begin{aligned} \beta_2(\lambda x + \lambda' x', y) &= (\lambda x_1 + \lambda' x'_1) y_2 - (\lambda x_2 + \lambda' x'_2) y_1 \\ &= \lambda(x_1 y_2 - x_2 y_1) + \lambda'(x'_1 y_2 - x'_2 y_1) \\ &= \lambda\beta_2(x, y) + \lambda'\beta_2(x', y) \end{aligned}$$

et de même pour la linéarité par rapport à la seconde variable. Ainsi, β_2 est bien une forme bilinéaire.

2.2 Matrice d'une forme bilinéaire et changement de base

Soient β une forme bilinéaire sur $\mathbb{R}^n$ et $\mathcal{B} := \{e_1, e_2, ..., e_n\}$ une base de $\mathbb{R}^n$. Alors si $x, y \in \mathbb{R}^n$, on a $x = \sum_{i=1}^n x_i e_i$ et $y = \sum_{j=1}^n y_j e_j$ et

$$\beta(x, y) = \beta\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n x_i \beta\left(e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \beta(e_i, e_j).$$

Ainsi, β est entièrement déterminée par la donnée des n^2 valeurs de $\beta(e_i, e_j)$, $1 \leq i, j \leq n$.

Définition 2.2 *La matrice de la forme bilinéaire β sur $\mathbb{R}^n$ dans la base $\mathcal{B}$ est*

$$M = (\beta(e_i, e_j))_{1 \leq i \leq n, 1 \leq j \leq n}.$$

De plus, $\beta(x, y) = X^\top M Y$ avec X (resp. Y) le vecteur colonne des composantes de x (resp. y) dans la base $\mathcal{B}$.

D'un autre côté, si on peut écrire $\beta(x, y) = X^\top MY$ avec X (resp. Y) les coordonnées de x (resp. y) dans la base canonique de $\mathbb{R}^n$, alors β est une forme bilinéaire.

Exemples :

- La matrice de β_2 dans la base canonique de $\mathbb{R}^2$ est $A_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.
- La matrice de β_2 dans la base $\mathcal{B}' = \{(1,2), (1,-1)\}$ est $A_2' = \begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix}$.

Théorème 2.3 *Soient $\mathcal{B}$ et $\mathcal{B}'$ deux bases de $\mathbb{R}^n$ et P la matrice de changement de base de $\mathcal{B}$ à $\mathcal{B}'$. Si une forme bilinéaire β est représentée par la matrice A dans la base $\mathcal{B}$, alors elle est représentée dans la base $\mathcal{B}'$ par la matrice $B = P^\top AP$.*

Exemple : Calculons, en utilisant ce théorème, la matrice A_2' de β_2 dans la base $\mathcal{B}'$.

La matrice de passage de la base canonique à $\mathcal{B}'$ est $P = \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix}$. On a donc immédiatement

$$A_2' = P^\top A_2 P = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} A_2 \begin{pmatrix} 1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & -3 \\ 3 & 0 \end{pmatrix}.$$

Rappel : si on note $e_1 = (1,2)$ et $e_2 = (1,-1)$ et c_1, c_2 les vecteurs de la base canonique, on a $e_1 = c_1 + 2c_2$ et $e_2 = c_1 - c_2$. Ainsi si x_1' et x_2' sont les coordonnées de x dans la base $\mathcal{B}'$, on a

$$x_1' e_1 + x_2' e_2 = x_1'(c_1 + 2c_2) + x_2'(c_1 - c_2) = (x_1' + x_2')c_1 + (2x_1' - x_2')c_2.$$

On obtient donc les coordonnées de x dans la base canonique en calculant PX'. En notant X (resp. Y) les coordonnées de x (resp. y) dans la base canonique, on retrouve bien $X^\top A_2 Y = (PX')^\top A_2 (PY') = (X')^\top P^\top A_2 P Y'$, c'est à dire $A_2' = P^\top A_2 P$.

2.3 Formes bilinéaires symétriques

Définition 2.4 *Soit $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ une forme bilinéaire. On dit que β est* **symétrique** *si $\beta(x, y) = \beta(y, x)$ pour tous $x, y \in \mathbb{R}^n$.*

Exemple : L'application β_2 n'est pas symétrique.

En effet, avec $x = (1, 0)$ et $y = (0, 1)$, on a $\beta_2(x, y) = 1 \neq -1 = \beta_2(y, x)$. Pour que β_2 soit symétrique, il faudrait que le coefficient de $x_1 y_2$ soit égal à celui de $x_2 y_1$.

Théorème 2.5 *Soit β une forme bilinéaire sur $\mathbb{R}^n$. Alors β est symétrique si et seulement si toute représentation matricielle de β est une matrice symétrique.*

Exemple : La forme bilinéaire β_2 n'est pas symétrique, puisque sa matrice dans la base canonique de $\mathbb{R}^2$ n'est pas symétrique.

Théorème 2.6 *Soit β une forme bilinéaire symétrique sur $\mathbb{R}^n$. Alors il existe une base de $\mathbb{R}^n$ sur laquelle β est représentée par une matrice diagonale.*

Preuve : Ce sera vu dans le cours sur la diagonalisation des matrices symétriques. $\square$

2.4 Produit scalaire

Définition 2.7 *On dit que l'application $\beta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ est un* **produit scalaire** *si β est une forme bilinéaire symétrique et si*

- *pour tout vecteur $x \in \mathbb{R}^n$, $\beta(x,x) \geq 0$;*
- *et $\beta(x,x) = 0 \Rightarrow x = 0_{\mathbb{R}^n}$.*

Exemple : L'application $\beta_3 : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ définie par $\beta(x,y) = 2x_1y_1 + 3x_2y_2 + 2x_1y_2 + 2y_1x_2$ si $x = (x_1, x_2)$ et $y = (y_1, y_2)$ est un produit scalaire. En effet, comme $\beta(x,y) = X^\top A_3 Y$, avec

$$A_3 = \begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix}$$

matrice symétrique, β_3 est une forme bilinéaire symétrique. De plus, $\beta_3(x,x) = 2x_1^2 + 3x_2^2 + 4x_1x_2 = 2(x_1^2 + 2x_1x_2) + 3x_2^2 = 2(x_1 + x_2)^2 + x_2^2 \geq 0$ et $\beta_3(x,x) = 0$ si et seulement si $x = 0$.

Remarque 2.8 *Pour tout $n \geq 2$, l'application $\tilde{\beta}_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ définie par $\tilde{\beta}_0(x,y) = x_1y_1 + x_2y_2 + ... + x_ny_n$ si $x = (x_1, x_2, ..., x_n)$ et $y = (y_1, y_2, ..., y_n)$ est un produit scalaire, appelé* **produit scalaire usuel ou canonique** *et $\tilde{\beta}_0(x,y) = X^\top Y$.*

En pratique, pour montrer que β est un produit scalaire, on met $\beta(x,y)$ sous la forme $X^\top MY$ et on montre que M est symétrique définie-positive.

A partir de maintenant, pour toute matrice M symétrique, on notera $\langle x, y\rangle_M := X^\top MY$ le produit scalaire de matrice M dans la base canonique de $\mathbb{R}^n$. Naturellement, si $M = I$, on notera simplement $\langle x, y\rangle := X^\top Y$.

2.5 Norme

Définition 2.9 (Rappel) *Une application $N : \mathbb{R}^n \to \mathbb{R}$ est une* **norme** *si pour tous $x, y \in \mathbb{R}^n$ et tout $\lambda \in \mathbb{R}$,*

1. *$N(x) \geq 0$ et $N(x) = 0 \Leftrightarrow x = 0$;*
2. *$N(\lambda x) = |\lambda| N(x)$;*
3. *$N(x + y) \leq N(x) + N(y)$.*

Soit M une matrice symétrique définie-positive.

Proposition 2.10 (Inégalité de Cauchy-Schwarz) *Soient x et y des vecteurs de $\mathbb{R}^n$. Alors on a*

$$|\langle x, y\rangle_M| \leq \sqrt{\langle x, x\rangle_M}\sqrt{\langle y, y\rangle_M}$$

et l'égalité a lieu si et seulement si x et y sont colinéaires.

Proposition 2.11 *Soit M une matrice symétrique définie-positive et $\|.\|_M : \mathbb{R}^n \to \mathbb{R}$ l'application qui à tout $x \in \mathbb{R}^n$ associe $\sqrt{\langle x, x\rangle_M}$. Alors $\|.\|_M$ est une norme sur $\mathbb{R}^n$, appelée norme associée au produit scalaire $\langle ., .\rangle_M$ de $\mathbb{R}^n$.*

Preuve : On a trivialement $\|0_{\mathbb{R}^n}\|_M = 0$ et $\|x\|_M > 0$ pour tout $x \neq 0_{\mathbb{R}^n}$.

De plus $\|\lambda x\|_M = \sqrt{\langle \lambda x, \lambda x\rangle_M} = \sqrt{\lambda^2 \langle x, x\rangle_M} = |\lambda| \|x\|_M$.

Il reste à démontrer l'inégalité triangulaire. Soient $x, y \in \mathbb{R}^n$. Alors $\langle x + y, x + y\rangle_M = (X+Y)^\top M(X+Y) = \langle x, x\rangle_M + 2\langle x, y\rangle_M + \langle y, y\rangle_M$ et, par Cauchy-Schwarz, $\langle x+y, x+y\rangle_M \leq \langle x, x\rangle_M + 2\sqrt{\langle x, x\rangle_M}\sqrt{\langle y, y\rangle_M} + \langle y, y\rangle_M = (\sqrt{\langle x, x\rangle_M} + \sqrt{\langle y, y\rangle_M})^2$. Par croissance de la racine carrée, on conclue que $\|x + y\|_M =\leq \|x\|_M + \|y\|_M$. $\square$

On dira que le vecteur x est M**-unitaire** ou M**-normé** ou M**-normalisé** s'il est de norme 1 pour la norme $\|.\|_M$, c'est-à-dire si $\|x\|_M = 1$. Tout vecteur non nul x peut être divisé par sa norme pour construire un vecteur unitaire $\tilde{x} = \dfrac{x}{\|x\|_M}$. On appelle cette opération la **normalisation** de x.

Chapitre 3

M-orthogonalité

Soit M une matrice carrée symétrique définie-positive de taille $n \times n$, $\langle ., . \rangle_M$ le produit scalaire associé et $\|.\|_M$ la norme associée.

3.1 M-orthogonalité de deux vecteurs

Définition 3.1 *Soient $x, y \in \mathbb{R}^n$. On dit que x et y sont M-**orthogonaux** si le produit scalaire $\langle x, y \rangle_M = 0$.*

Remarque 3.2
- *Si $M = I$, on parle d'orthogonalité.*
- *Tout vecteur est M-orthogonal au vecteur nul.*

Exemples :
- Dans $\mathbb{R}^3$, on considère les vecteurs $x = (1, 1, 1)$, $y = (1, 2, -3)$ et $z = (1, -4, 3)$. On a

$$\begin{array}{rcl}\langle x, y \rangle &=& 1 \times 1 + 1 \times 2 + 1 \times (-3) = 0, \\ \langle x, z \rangle &=& 1 \times 1 + 1 \times (-4) + 1 \times 3 = 0, \\ \langle y, z \rangle &=& 1 \times 1 + 2 \times (-4) + (-3) \times 3 = -16.\end{array}$$

  On en déduit donc que x est orthogonal à y et z, mais y n'est pas orthogonal à z.
- Trouvons un vecteur non nul $z \in \mathbb{R}^3$ orthogonal à $x = (1, 2, 1)$ et $y = (2, 5, 4)$.

$$\begin{array}{rcl}z \text{ est orthogonal à } x \text{ et } y &\Leftrightarrow& \langle z, x \rangle = 0 \text{ et } \langle z, y \rangle = 0 \\ &\Leftrightarrow& z_1 + 2z_2 + z_3 = 0 \text{ et } 2z_1 + 5z_2 + 4z_3 = 0 \\ &\Leftrightarrow& z_1 = -2z_2 - z_3 \text{ et } z_2 + 2z_3 = 0 \\ &\Leftrightarrow& z_2 = -2z_3 \text{ et } z_1 = 3z_3.\end{array}$$

  Quel que soit le choix de $z_3 \in \mathbb{R}^*$, le vecteur $z = z_3(3, -2, 1)^\top$ est orthogonal à x et y. Il n'y a pas unicité.
  Calculons la norme de z : $\|z\|^2 = 14z_3^2$. Pour avoir z unitaire, on peut donc choisir $z_3 = \pm 1/\sqrt{14}$.

3.2 Orthogonal d'un sous-espace vectoriel

Définition 3.3 *Soit F un sous-espace vectoriel de $\mathbb{R}^n$ et soit $x \in \mathbb{R}^n$. On dit que x est M-orthogonal à F si pour tous $y \in F$, $\langle x, y \rangle_M = 0$.*

Proposition 3.4 *Soit F le s.e.v. de $\mathbb{R}^n$ engendré par des vecteurs u_1, u_2, ..., u_p de $\mathbb{R}^n$ ($p \leq n$). Pour qu'un vecteur $x \in \mathbb{R}^n$ soit M-orthogonal à F, il faut et il suffit que x soit M-orthogonal à u_i pour tout $i = 1, ..., p$.*

Preuve : Soit $x \in \mathbb{R}^n$.

- Si $x \in \mathbb{R}^n$ est M-orthogonal à F, alors on a bien x M-orthogonal à u_i pour tout i, puisque tous les u_i sont dans F.
- Réciproquement, on suppose que x est M-orthogonal à u_i pour tout i. Soit $y \in F$. Par définition de vecteurs générateurs, il existe λ_1, λ_2, ..., λ_p dans $\mathbb{R}$ tels que $y = \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_p u_p$. Donc

$$\begin{aligned}
\langle x, y \rangle_M &= \langle x, \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_p u_p \rangle_M \\
&= \lambda_1 \langle x, u_1 \rangle_M + \lambda_2 \langle x, u_2 \rangle_M + ... + \lambda_p \langle x, u_p \rangle_M = 0.
\end{aligned}$$

  Le vecteur x est donc bien M-orthogonal à F.

$\square$

Exemple : Soit $F_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. Alors $F_1 = \{(x_1, x_2, -x_1 - x_2) \in \mathbb{R}^3, x_1, x_2 \in \mathbb{R}\} = \text{Vect}((1, 0, -1), (0, 1, -1))$. Cela signifie que F_1 est engendré par les vecteurs $u_1 = (1, 0, -1)$ et $u_2 = (0, 1, -1)$.

- Soit $\langle ., . \rangle$ le produit scalaire usuel. Cherchons les vecteurs orthogonaux à F_1 :

$$\begin{aligned}
v = (v_1, v_2, v_3) \in \mathbb{R}^3 \text{ est orthogonal à } F_1 &\Leftrightarrow \langle v, u_1 \rangle = 0 \text{ et } \langle v, u_2 \rangle = 0 \\
&\Leftrightarrow v_1 - v_3 = 0 \text{ et } v_2 - v_3 = 0 \\
&\Leftrightarrow v_1 = v_3 \text{ et } v_2 = v_3.
\end{aligned}$$

  Les vecteurs orthogonaux à F_1 sont donc de la forme $v_3(1, 1, 1)$, $v_3 \in \mathbb{R}$.
- Considérons le produit scalaire défini pour $x = (x_1, x_2, x_3)$ et $y = (y_1, y_2, y_3)$ par $\langle x, y \rangle_M$ avec $M = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ et cherchons les vecteurs M-orthogonaux à F_1 :

$$\begin{aligned}
v = (v_1, v_2, v_3) \in \mathbb{R}^3 \text{ est } M\text{-orthogonal à } F_1 &\Leftrightarrow \langle v, u_1 \rangle_M = 0 \text{ et } \langle v, u_2 \rangle_M = 0 \\
&\Leftrightarrow 2v_1 + v_2 = 0 \text{ et } v_1 + v_2 = 0 \\
&\Leftrightarrow v_1 = v_2 = 0.
\end{aligned}$$

  Les vecteurs M-orthogonaux à F_1 sont donc de la forme $(0, 0, v_3)$, $v_3 \in \mathbb{R}$.

Proposition 3.5 *Soit F un s.e.v. de $\mathbb{R}^n$. L'ensemble des vecteurs de $\mathbb{R}^n$ qui sont M-orthogonaux à F est un s.e.v. de $\mathbb{R}^n$. Ce s.e.v. s'appelle* **l'orthogonal** *de F pour le produit scalaire associé à M et se note $F^{\perp_M}$.*

Remarque 3.6 *Si $M = I$, on notera simplement $F^{\perp}$.*

Exemples :

- Dans l'exemple précédent, on a montré que
  - $F_1^{\perp} = \mathrm{Vect}((1,1,1))$;
  - $F_1^{\perp_M} = \mathrm{Vect}((0,0,1))$.
- Soit $u = (1,3,-4) \in \mathbb{R}^3$. Cherchons $u^{\perp}$.

$$\begin{array}{rcl} v = (v_1, v_2, v_3) \in \mathbb{R}^3 \text{ est orthogonal à } u & \Leftrightarrow & \langle v, u \rangle = 0 \\ & \Leftrightarrow & v_1 + 3v_2 - 4v_3 = 0 \\ & \Leftrightarrow & v_1 = -3v_2 + 4v_3. \end{array}$$

  Ainsi, $u^{\perp} = \{(-3v_2 + 4v_3, v_2, v_3), v_2, v_3 \in \mathbb{R}\} = \mathrm{Vect}((-3,1,0),(4,0,1))$.
- Soient $\mathbb{R}^4$ muni du produit scalaire usuel et P le plan d'équations $x + 2y - z + 3t = 0$ et $x + y + 5z = 0$. Cherchons des équations de l'orthogonal $P^{\perp}$ du plan P.
  Tout d'abord,

$$\begin{array}{rcl} P & = & \{(x,y,z,t) \in \mathbb{R}^4 : x = -y - 5z \text{ et } y - 6z + 3t = 0\} \\ & = & \{(x,y,z,t) \in \mathbb{R}^4 : x = -y - 5z \text{ et } t = -\frac{1}{3}y + 2z\} \\ & = & \{(-y - 5z, y, z, -\frac{1}{3}y + 2z), y, z \in \mathbb{R}\} \\ & = & \mathrm{Vect}((-1,1,0,-\frac{1}{3}),(-5,0,1,2)) \\ & = & \mathrm{Vect}((-3,3,0,-1),(-5,0,1,2)). \end{array}$$

  Le plan P est donc engendré par les 2 vecteurs $u_1 = (-3,3,0,-1)$ et $u_2 = (-5,0,1,2)$. On a donc

$$\begin{array}{rcl} v = (a,b,c,d) \in \mathbb{R}^4 \text{ est orthogonal à } P & \Leftrightarrow & \langle v, u_1 \rangle = 0 \text{ et } \langle v, u_2 \rangle = 0 \\ & \Leftrightarrow & -3a + 3b - d = 0 \text{ et } -5a + c + 2d = 0. \end{array}$$

  Ainsi, $P^{\perp}$ est le plan d'équations $-3a + 3b - d = 0$ et $-5a + c + 2d = 0$.

Chapitre 4

Bases M-orthonormées et procédé d'orthogonalisation de Gram-Schmidt

Dans tout ce chapitre, on considère M une matrice réelle carrée symétrique définie-positive.

4.1 Bases M orthogonales

Proposition 4.1 *Soient u_1, u_2, ..., u_p des vecteurs de $\mathbb{R}^n$ non nuls et 2 à 2 M-orthogonaux. Alors les vecteurs u_1, u_2, ..., u_p sont linéairement indépendants.*

Preuve : Supposons $\lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_i u_i + ... + \lambda_p u_p = 0$. Soit i entier entre 1 et p. Alors

$$\begin{array}{rcl} 0 & = & \langle u_i, \lambda_1 u_1 + \lambda_2 u_2 + ... + \lambda_i u_i + ... + \lambda_p u_p \rangle_M \\ & = & \lambda_1 \langle u_i, u_1 \rangle_M + \lambda_2 \langle u_i, u_2 \rangle_M + ... + \lambda_i \langle u_i, u_i \rangle_M + ... + \lambda_p \langle u_i, u_p \rangle_M \\ & = & \lambda_i \|u_i\|_M^2. \end{array}$$

Comme le vecteur u_i est non nul, on en déduit que $\lambda_i = 0$. Comme ce raisonnement est valable pour tout i entre 1 et p, on a montré l'indépendance linéaire des vecteurs u_1, u_2, ..., u_p. □

Remarque 4.2
- *Attention, la proposition précédente n'est pas une équivalence. En effet, les vecteurs $(1,0)$ et $(1,1)$ sont linéairement indépendants, mais ils ne sont pas orthogonaux.*
- *Une matrice Q $n \times p$ ($p \leq n$) formée de vecteurs colonnes M-orthogonaux vérifie $Q^\top M Q$ diagonale.*

Corollaire 4.3 *Si u_1, u_2, ..., u_n sont des vecteurs de $\mathbb{R}^n$ non nuls et deux à deux M-orthogonaux, alors ils forment une base, dite M-**orthogonale**, de $\mathbb{R}^n$.*

Théorème 4.4 (Pythagore) *Soient u_1, u_2, ..., u_p des vecteurs de $\mathbb{R}^n$ non nuls et 2 à 2 M-orthogonaux. Alors $\|u_1 + u_2 + ... + u_p\|_M^2 = \|u_1\|_M^2 + \|u_2\|_M^2 + ... + \|u_p\|_M^2$.*

Preuve :

$$\begin{array}{rcl} \|u_1 + u_2 + ... + u_p\|_M^2 & = & \langle u_1 + u_2 + ... + u_p, u_1 + u_2 + ... + u_p \rangle_M \\ & = & \langle \sum_{i=1}^p u_i, \sum_{j=1}^p u_j \rangle_M = \sum_{i=1}^p \sum_{j=1}^p \langle u_i, u_j \rangle_M = \sum_{i=1}^p \langle u_i, u_i \rangle_M = \sum_{i=1}^p \|u_i\|_M^2. \end{array}$$

□

4.1.1 Base orthogonale et combinaisons linéaires

Exemple : On considère les vecteurs suivants : $u_1 = (1, 2, 1)$, $u_2 = (2, 1, -4)$ et $u_3 = (3, -2, 1)$. Il est aisé de vérifier que ces 3 vecteurs sont deux à deux orthogonaux. Ils forment donc une base orthogonale de $\mathbb{R}^3$. Essayons d'écrire le vecteur $v = (7, 1, 9)$ comme combinaison linéaire de u_1, u_2 et u_3 : on cherche x_1, x_2, $x_3 \in \mathbb{R}$ tels que $v = x_1u_1 + x_2u_2 + x_3u_3$.

- Méthode 1 : on procède par identification et on aboutit au système linéaire suivant :

$$\left\{\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 7 \\ 2x_1 & + & x_2 & - & 2x_3 & = & 1 \\ x_1 & - & 4x_2 & + & x_3 & = & 9 \end{array}\right.$$

  dont la solution est $x_1 = 3$, $x_2 = -1$ et $x_3 = 2$.
- Méthode 2 : on utilise le produit scalaire. Pour tout $i = 1, 2, 3$, on a

$$\langle v, u_i\rangle = \langle x_1u_1 + x_2u_2 + x_3u_3, u_i\rangle = x_i\langle u_i, u_i\rangle.$$

  On obtient $x_i = \dfrac{\langle v, u_i\rangle}{\langle u_i, u_i\rangle}$ et donc $x_1 = \dfrac{18}{6} = 3$, $x_2 = \dfrac{-21}{21} = -1$ et $x_3 = \dfrac{28}{14} = 2$.

La procédure décrite dans la seconde méthode est générale :

Théorème 4.5 *Soit $\{u_1, u_2, ..., u_n\}$ une base M-orthogonale de $\mathbb{R}^n$. Alors $\forall v \in \mathbb{R}^n$, on a*

$$v = \frac{\langle v, u_1\rangle_M}{\langle u_1, u_1\rangle_M}u_1 + \frac{\langle v, u_2\rangle_M}{\langle u_2, u_2\rangle_M}u_2 + ... + \frac{\langle v, u_n\rangle_M}{\langle u_n, u_n\rangle_M}u_n.$$

Preuve : On cherche les $c_i \in \mathbb{R}$, $i = 1, ..., n$ tels que $v = \sum_{i=1}^n c_iu_i$.

Soit $j \in \{1, 2, ..., n\}$. On a

$$\begin{array}{rcll} \langle v, u_j\rangle_M & = & \langle \sum_{i=1}^n c_iu_i, u_j\rangle_M & \\ & = & \sum_{i=1}^n c_i\langle u_i, u_j\rangle_M & \\ & = & c_j\langle u_j, u_j\rangle_M & \text{par } M\text{-orthogonalité des vecteurs } u_j \end{array}$$

et donc $c_j = \dfrac{\langle v, u_j\rangle_M}{\langle u_j, u_j\rangle_M}$. □

On donne une interprétation géométrique du coefficient c_j dans la section suivante.

4.1.2 Projections

Soient v et u deux vecteurs de $\mathbb{R}^n$. On appelle projection M-orthogonale de v sur u et on note $proj_M(v, u)$ un vecteur cu où c est un réel tel que $v' = v - cu$ est M-orthogonal à u :

$$0 = \langle v - cu, u\rangle_M = \langle v, u\rangle_M - c\langle u, u\rangle_M \Leftrightarrow c = \frac{\langle v, u\rangle_M}{\langle u, u\rangle_M}.$$

Le scalaire c est unique et est appelé **coefficient de Fourier** ou **composante** de v par rapport à u.

Cette notion se généralise de la manière suivante :

Théorème 4.6 *Soient $u_1, u_2, ..., u_r$ des vecteurs non nuls de $\mathbb{R}^n$, deux à deux M-orthogonaux, et $v \in \mathbb{R}^n$. On pose $v' = v - (c_1u_1 + c_2u_2 + ... + c_ru_r)$ avec $c_i = \dfrac{\langle v, u_i\rangle_M}{\langle u_i, u_i\rangle_M}$ pour tout $i = 1, 2, ..., r$. Alors v' est M-orthogonal à $u_1, u_2, ..., u_r$.*

Les c_i du théorème sont les composantes de v par rapport aux u_i.
Preuve : Soit $i \in \{1, 2, ..., r\}$.

$$\begin{array}{rcll}
\langle v', u_i\rangle_M & = & \langle v - (c_1u_1 + c_2u_2 + ... + c_ru_r), u_i\rangle_M & \\
& = & \langle v, u_i\rangle_M - \sum_{j=1}^{r} c_j\langle u_j, u_i\rangle_M & \\
& = & \langle v, u_i\rangle_M - c_i\langle u_i, u_i\rangle_M & \text{par orthogonalité des vecteurs } u_j \\
& = & 0 & \text{par définition de } c_i
\end{array}$$

□

Remarque 4.7 *Dans le théorème précédent, si on note $W := Vect(u_1, u_2, ..., u_r)$, comme les vecteurs u_i sont 2 à 2 M-orthogoanux, ils forment une base de W. En fait, $c_1u_1+c_2u_2+...+c_ru_r$ est la projection M-orthogonale de v sur W, notée $proj_M(v, W)$.*

4.2 Procédé d'orthogonalisation de Gram-Schmidt

Soit $\mathcal{V} = \{v_1, v_2, ..., v_n\}$ une base quelconque de $\mathbb{R}^n$. On peut construire à partir de $\mathcal{V}$ une base M-orthogonale $\mathcal{U} = \{u_1, u_2, ..., u_n\}$ de $\mathbb{R}^n$ de la manière suivante :

$$\begin{array}{rcl}
u_1 & = & v_1 \\
u_2 & = & v_2 - c_{21}u_1 \\
\dots & & \\
u_k & = & v_k - c_{k1}u_1 - c_{k2}u_2 - ... - c_{k,k-1}u_{k-1} \\
\dots & & \\
u_n & = & v_n - c_{n1}u_1 - c_{n2}u_2 - ... - c_{n,k-1}u_{k-1} - ... - c_{n,n-1}u_{n-1}
\end{array}$$

avec $c_{ki} = \dfrac{\langle v_k, u_i\rangle_M}{\langle u_i, u_i\rangle_M}$.

D'après le théorème précédent, chaque u_k est M-orthogonal aux u_j qui le précèdent. On obtient donc bien une base M-orthogonale. On peut en déduire aisément une base M-**orthonormale** en normalisant chacun des vecteurs u_i.

Remarque 4.8
- *Chacun des u_k est CL de v_k et des u_j qui le précèdent. On en déduit par récurrence que tout u_k est CL de v_1, v_2, ..., v_k.*
- *Puisque la M-orthogonalité des vecteurs n'est pas affectée lorsqu'on change la longueur des vecteurs, on peut lors des calculs se débarrasser des dénominateurs en multipliant chaque u_k par un facteur approprié avant de calculer u_{k+1}.*

Exemple : On se place dans $\mathbb{R}^4$ muni du produit scalaire usuel. Soient u_1, u_2 et u_3 les vecteurs de $\mathbb{R}^4$ définis par

$$v_1 = (2, 1, 0, 2), v_2 = (-4, 1, 0, -1), v_3 = (1, 3, -4, -1)$$

et W le s.e.v. engendré par ces 3 vecteurs.

1. On cherche à construire une base orthogonale de W en utilisant le procédé de Gram-Schmidt.
   - On pose $u_1 = v_1$.
   - On a $u_2 = v_2 - c_{21}u_1$ avec $c_{21} = \dfrac{\langle v_2, u_1\rangle}{\langle u_1, u_1\rangle}$.
     Or $\langle v_2, u_1\rangle = -9$ et $\langle u_1, u_1\rangle = \|u_1\|^2 = 9$. Donc $c_{21} = -1$ et $u_2 = v_2 + u_1 = (-2, 2, 0, 1)$.
   - On a $u_3 = v_3 - c_{31}u_1 - c_{32}u_2$ avec $c_{31} = \dfrac{\langle v_3, u_1\rangle}{\langle u_1, u_1\rangle}$ et $c_{32} = \dfrac{\langle v_3, u_2\rangle}{\langle u_2, u_2\rangle}$.
     Or $\langle v_3, u_1\rangle = 3$, $\langle v_3, u_2\rangle = 3$ et $\langle u_2, u_2\rangle = 9$. Donc $c_{31} = 1/3$, $c_{32} = 1/3$ et $u_3 = v_3 - (u_1 + u_2)/3 = (1, 3, -4, -1) - (0, 3, 0, 3)/3 = (1, 2, -4, -2)$.

   Par construction, (u_1, u_2, u_3) forme une B.O.G. de W.
2. On en déduit facilement une base orthonormée de W en construisant $w_i = \dfrac{v_i}{\|v_i\|}$. Ici, ça donne $w_1 = \dfrac{1}{3}(2, 1, 0, 2)$, $w_2 = \dfrac{1}{3}(-2, 2, 0, 1)$, $w_3 = \dfrac{1}{5}(1, 2, -4, -2)$.

Théorème 4.9 *Soit $\mathcal{V} = \{v_1, v_2, ..., v_n\}$ une base quelconque de $\mathbb{R}^n$. Alors il existe une base M-orthonormée $\mathcal{U} = \{u_1, u_2, ..., u_n\}$ de $\mathbb{R}^n$ telle que la matrice de changement de base de $\mathcal{V}$ à $\mathcal{U}$ soit triangulaire. En d'autres termes,*

$$u_k = a_{k1}v_1 + a_{k2}v_2 + ... + a_{kk}v_k, \forall k = 1, 2, ..., n.$$

Théorème 4.10 *Soit $\mathcal{U} = \{u_1, u_2, ..., u_r\}$ une base M-orthogonale d'un sous-espace F de $\mathbb{R}^n$. Parce qu'on est en dimension finie, on peut compléter la base $\mathcal{U}$ pour obtenir une base M-orthogonale de $\mathbb{R}^n$: il existe des vecteurs u_{r+1}, ..., u_n tels que u_1,u_2,...,u_r,u_{r+1}, ..., u_n forment une base M-orthogonale de $\mathbb{R}^n$.*

4.3 Quelques propriétés des bases M-orthonormées

Remarque 4.11
- *Une matrice Q de taille $n \times p$, formée de vecteurs colonnes M-orthonormés, vérifie $Q^\top M Q = I_p$.*
- *Si $p = n$ et si on considère le produit scalaire usuel ($M = I_n$) de $\mathbb{R}^n$, alors Q est inversible et $Q^{-1} = Q^\top$.*

Théorème 4.12 *Soit $(e_1, e_2, ..., e_n)$ une base M-orthonormée de $\mathbb{R}^n$.*
- *Pour tout vecteur $x \in \mathbb{R}^n$, on a $x = \displaystyle\sum_{i=1}^{n} \langle x, e_i\rangle_M e_i$.*
- *Pour tous vecteurs $x, y \in \mathbb{R}^n$, on a $\langle x, y\rangle_M = \displaystyle\sum_{i=1}^{n} \langle x, e_i\rangle_M \langle y, e_i\rangle_M$.*

Preuve :
- Soit $x \in \mathbb{R}^n$. On peut écrire $x = \displaystyle\sum_{i=1}^{n} x_i e_i$. On a alors pour tout j

$$\langle x, e_j\rangle_M = \langle\sum_{i=1}^{n} x_i e_i, e_j\rangle_M = \sum_{i=1}^{n} x_i \langle e_i, e_j\rangle_M = x_j.$$

- Soient $x, y \in \mathbb{R}^n$. On peut écrire $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{j=1}^{n} y_j e_j$. On a alors

$$\langle x, y\rangle_M = \langle\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{n} y_j e_j\rangle_M = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j \langle e_i, e_j\rangle_M = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n}\langle x, e_i\rangle_M \langle y, e_i\rangle_M$$

□

Exemple : On se place dans $\mathbb{R}^2$ muni du produit scalaire usuel. Les vecteurs $e_1 = \frac{1}{\sqrt{2}}(1,1)$ et $e_2 = \frac{1}{\sqrt{2}}(-1,1)$ forment une B.O.N. $\mathcal{E}$ de $\mathbb{R}^2$.

Calculons les coordonnées dans cette base $\mathcal{E}$ des vecteurs de la base canonique de $\mathbb{R}^2$: $c_1 = (1,0)$ et $c_2 = (0,1)$:

$$\begin{array}{rcccl} c_1 & = & \langle c_1, e_1\rangle e_1 + \langle c_1, e_2\rangle e_2 & = & \frac{1}{\sqrt{2}} e_1 - \frac{1}{\sqrt{2}} e_2 \\ c_2 & = & \langle c_2, e_1\rangle e_1 + \langle c_2, e_2\rangle e_2 & = & \frac{1}{\sqrt{2}} e_1 + \frac{1}{\sqrt{2}} e_2 \end{array}$$

Maintenant, en utilisant le théorème ci-dessus, on retrouve $\langle c_1, c_1\rangle = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 = 1$, $\langle c_1, c_2\rangle = \left(\frac{1}{\sqrt{2}}\right)^2 - \left(\frac{1}{\sqrt{2}}\right)^2 = 0$, $\langle c_2, c_2\rangle = \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = 1$.

4.4 Matrices orthogonales

Définition 4.13 *Une matrice P est dite orthogonale si elle est régulière et vérifie $P^{-1} = P^\top$.*

Théorème 4.14 *Soit P une matrice réelle $n \times n$. Les trois propositions suivantes sont équivalentes :*

1. *P est orthogonale ;*
2. *les lignes de P forment une BON de $\mathbb{R}^n$;*
3. *les colonnes de P forment une BON de $\mathbb{R}^n$.*

Remarque 4.15 *Attention, ce théorème n'est vrai que pour le produit scalaire usuel de $\mathbb{R}^n$!*

Théorème 4.16 *Soient $\mathcal{B} = e_1, e_2, ..., e_n$ et $\mathcal{B}' = e'_1, e'_2, ..., e'_n$ deux BON de $\mathbb{R}^n$. Alors la matrice de changement de base de $\mathcal{B}$ à $\mathcal{B}'$ est une matrice orthogonale.*

Preuve : Posons $e'_j = b_{1j}e_1 + b_{2j}e_2 + ... + b_{nj}e_n$ pour $j = 1, 2, ..., n$. On doit montrer que la matrice B ainsi construite est orthogonale.

Comme $\mathcal{B}'$ est une BON, on a

$$\begin{array}{rcl} \delta_{ij} & = & \langle e'_i, e'_j\rangle \\ & = & \langle\sum_{k=1}^n b_{ki}e_k, \sum_{l=1}^n b_{lj}e_l\rangle \\ & = & \sum_{k=1}^n b_{ki} \sum_{l=1}^n b_{lj}\langle e_k, e_l\rangle \\ & = & \sum_{k=1}^n b_{ki}b_{kj} \\ & = & (B^\top B)_{ij} \end{array}$$

On en déduit bien que B est régulière et que $B^{-1} = B^\top$, c'est à dire que B est orthogonale. □

Théorème 4.17 *Soit $e_1, e_2, ..., e_n$ une BON de $\mathbb{R}^n$ et P une matrice orthogonale. Alors les n vecteurs suivants forment une BON de $\mathbb{R}^n$:*

$$e'_i = p_{1i}e_1 + p_{2i}e_2 + ... + p_{ni}e_n, i = 1, 2, ..., n.$$

Preuve : Il suffit de calculer

$$\begin{array}{rcl} \langle e'_i, e'_j \rangle & = & \langle \sum_{k=1}^n p_{ki}e_k, \sum_{l=1}^n p_{lj}e_l \rangle \\ & = & \sum_{k=1}^n p_{ki} \sum_{l=1}^n p_{lj} \langle e_k, e_l \rangle \\ & = & \sum_{k=1}^n p_{ki}p_{kj} \\ & = & (P^\top P)_{ij} \\ & = & \delta_{ij}. \end{array}$$

□

4.5 Projection orthogonale

Soit M une matrice carrée symétrique définie-positive. Soit F un s.e.v. de de $\mathbb{R}^n$. Notons $G := F^{\perp_M}$.

Théorème 4.18 *Les s.e.v. F et G sont supplémentaires :*

$$\forall x \in \mathbb{R}^n, \exists! f \in F, \exists! g \in G \text{ tels que } x = f + g.$$

On a aussi $(F^{\perp_M})^{\perp_M} = F$.

On remarque que le Théorème précédent n'est pas vrai lorsque F est un s.e.v. d'un espace E de dimension infinie.

Définition 4.19 *Soit $x \in \mathbb{R}^n$ et $f \in F$, $g \in G$ tels que $x = f + g$. L'application de $\mathbb{R}^n$ dans F qui à tout $x \in \mathbb{R}^n$ associe f est une application linéaire appelé projection M-orthogonale de $\mathbb{R}^n$ sur F. On la note p_F^M.*

Proposition 4.20
- *L'application linéaire p_F^M a pour noyau G et pour image F.*
- *Pour tout $x \in \mathbb{R}^n$, $p_F^M(p_F^M(x)) = p_F^M(x)$.*
- *Pour tout $f \in F$, $p_F^M(f) = f$.*
- *Pour tout $x \in \mathbb{R}^n$, $\|x - p_F^M(x)\|_M = \min_{f \in F} \|x - f\|_M$.*

Preuve : du point 4 Soit $x \in \mathbb{R}^n$. Notons $x_F = p_F^M(x)$. Alors $x - x_F = x_{F^{\perp_M}} \in F^{\perp_M}$.

Soit $f \in F$. On a

$$\begin{array}{rcl} \|x - f\|_M^2 & = & \| \underbrace{(x - x_F)}_{\in F^{\perp_M}} + \underbrace{(x_F - f)}_{\in F} \|_M^2 \\ & = & \|x - x_F\|_M^2 + \|x_F - f\|_M^2 \quad \text{par Pythagore} \\ & \geq & \|x - x_F\|_M^2. \end{array}$$

□

Proposition 4.21 *Soit $\{e_1, ..., e_p\}$ une base M-orthonormée de F et p_F^M la projection M-orthogonale de $\mathbb{R}^n$ sur F. Alors pour tout $x \in \mathbb{R}^n$, $p_F^M(x) = \langle x, e_1\rangle_M e_1 + ... + \langle x, e_p\rangle_M e_p$.*

Preuve : Soit $x \in \mathbb{R}^n$. On a $x = p_F(x) + x_{F^{\perp_M}}$. Comme $p_F(x) \in F$, il s'écrit $p_F(x) = \lambda_1 e_1 + ... + \lambda_j e_j + ... + \lambda_p e_p$.

Alors pour tout $1 \leq j \leq p$,

$$\begin{aligned}
\langle x, e_j\rangle_M &= \langle \lambda_1 e_1 + ... + \lambda_j e_j + ... + \lambda_p e_p + x_{F^{\perp_M}}, e_j\rangle_M \\
&= \lambda_1 \underbrace{\langle e_1, e_j\rangle_M}_{=0} + ... + \lambda_j \underbrace{\langle e_j, e_j\rangle_M}_{=1} + ... + \lambda_p \underbrace{\langle e_p, e_j\rangle_M}_{=0} + \underbrace{\langle x_{F^{\perp_M}}, e_j\rangle}_{=0} \\
&= \lambda_j.
\end{aligned}$$

$\square$

Chapitre 5

Diagonalisation des matrices réelles symétriques et décomposition en valeurs singulières (DVS)

5.1 Matrices réelles symétriques

Soit M une matrice réelle carrée symétrique de taille n.

Théorème 5.1 *1. Toutes les valeurs propres de M sont réelles ;*

2. si x et y sont deux vecteurs propres de M associés à deux valeurs propres distinctes λ et μ, alors ils sont orthogonaux pour le produit scalaire usuel de $\mathbb{R}^n$.

Preuve : Montrons le second point : tout d'abord, on a

$$\lambda < u, v >=< Mu, v >= (Mu)^\top v = u^\top M^\top v = u^\top Mv =< u, Mv >= \mu < u, v > .$$

On en déduit que $(\lambda - \mu) < u, v >= 0$. Comme $\lambda \neq \mu$, on a bien $< u, v >= 0$. □

Exemple : On considère $M = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$.

Les valeurs propres de M sont les racines de $(5 - \lambda)^2 - 9 = (2 - \lambda)(8 - \lambda) = 0$, c'est-à-dire ici 2 et 8.

En résolvant $\begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 2 \begin{pmatrix} x \\ y \end{pmatrix}$, on montre que qu'un vecteur propre de M associé à la valeur propre 2 est par exemple $u = (1, -1)$.

De même, on montre qu'un vecteur propre de M associé à la valeur propre 8 est par exemple $v = (1, 1)$.

Il est aisé de vérifier que u et v sont orthogonaux.

Si on pose $Q = (u\ v) = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, par définition des vecteurs propres, on a aussi $MQ = Q \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$.

Corollaire 5.2 *Il existe une matrice orthogonale P telle que la matrice $D := P^\top MP$ soit diagonale.*

Exemple : Reprenons l'exemple précédent. Partant de u et v, on construit la matrice orthogonale P à partir d'une base de vecteurs propres orthogonaux $u' = \dfrac{u}{\|u\|}$ et $v' = \dfrac{v}{\|v\|}$: $P = \dfrac{1}{\sqrt{2}}Q$.
Alors $MP = \dfrac{1}{\sqrt{2}}MQ = \dfrac{1}{\sqrt{2}}QD = PD$ avec $D = \begin{pmatrix} 2 & 0 \\ 0 & 8 \end{pmatrix}$.

Comme P est orthogonale, on a bien $P^\top MP = D$ diagonale.

Proposition 5.3

- *M est semi-définie positive si et seulement si ses valeurs propres sont toutes ≥ 0 ;*
- *M est définie positive si et seulement si ses valeurs propres sont toutes > 0 ;*

Preuve : Montrons le premier point.

Notons $P = (u_1, ..., u_i, ..., u_n)$. On sait que les vecteurs colonnes de P forment une B.O.N. de $\mathbb{R}^n$.

$$\begin{aligned} M \text{ est semi-définie positive} &\Leftrightarrow \text{ pour tout } x \in \mathbb{R}^n,\ x^\top Mx \geq 0 \\ &\Leftrightarrow \text{ pour tout } i = 1, ..., n,\ u_i^\top Mu_i \geq 0 \\ &\Leftrightarrow \text{ pour tout } i = 1, ..., n,\ u_i^\top \lambda_i u_i \geq 0 \\ &\Leftrightarrow \text{ pour tout } i = 1, ..., n,\ \lambda_i \geq 0. \end{aligned}$$

□

Exemple : Reprenons encore la matrice de l'exemple précédent. Comme ses valeurs propres sont 2 et 8, elle est définie-positive.

Corollaire 5.4 *Une forme bilinéaire β est un produit scalaire si sa matrice représentative est symétrique définie positive.*

Exemple : La forme bilinéaire définie par $\beta(x, y) = 5x_1y_1 + 3x_1y_2 + 3x_2y_1 + 5x_2y_2$ est un produit scalaire, parce que sa matrice représentative $M = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$ est symétrique définie positive.

Image de la sphère unité par une matrice carrée réelle symétrique

Considérons $S = \{x \in \mathbb{R}^n : \|x\| = 1\}$ la sphère unité de $\mathbb{R}^n$.

Notons f l'application linéaire associée à la matrice M.

Cherchons l'image $f(S)$ de S par f. Pour ce faire, prenons $x \in S$ et posons $t = f(x)$. Alors $t = Mx = PDP^\top x$.

Posons $y = P^\top x$. Alors $\|y\|^2 = y^\top y = x^\top PP^\top x = x^\top x = \|x\|^2 = 1 \Rightarrow y \in S$, y contient simplement les coordonnées de x dans la B.O.N. des vecteurs propres.

Posons maintenant $z = Dy$. Comme D est diagonale, on a $z_i = \lambda_i y_i$ et donc

$$\left(\frac{z_1}{\lambda_1}\right)^2 + ... + \left(\frac{z_i}{\lambda_i}\right)^2 + ... + \left(\frac{z_n}{\lambda_n}\right)^2 = 1.$$

On obtient donc un ellipsoïde dont les longueurs des demi-axes sont données par les valeurs propres λ_i.

Maintenant, avec $t = Pz$, on revient simplement dans la base canonique.

Exemple : Revenons à l'exemple précédent.

On constate sur la que l'opération $y = P^\top x$ consiste simplement en une rotation du cercle unité (noir) d'angle $\pi/4$ (vert), $z = Dy$ transforme le cercle unité en une ellipse de demi-axes de longueurs 2 et 8 (bleu foncé) et enfin $t = Pz$ faire tourner l'ellipse d'un angle $-\pi/4$ (cyan).

On voit également que l'image du cercle unité par f est une ellipse de demi-axes de longueurs 2 et 8 dans la base des vecteurs propres (rouge).

Intérêt de la diagonalisation de la matrice d'une forme bilinéaire symétrique

M peut être considérée comme la matrice, dans la base canonique, d'une forme bilinéaire symétrique β. La diagonalisation permet alors d'avoir une expression plus simple de la forme bilinéaire dans la base des vecteurs propres.

$$\beta(x,y) = X^\top MY = X^\top PDP^\top Y = \tilde{X}^\top D\tilde{Y} = \sum_{i=1}^{n} \lambda_i \tilde{x}_i \tilde{y}_i$$

avec $\tilde{X} = P^\top X$ et $\tilde{Y} = P^\top Y$.

Exemple : Toujours avec l'exemple précédent, si β est la forme bilinéaire associée à M, l'expression de β est $\beta(x,y) = 5x_1y_1 + 3x_1y_2 + 3x_2y_1 + 5x_2y_2$.

Dans la base $\{u', v'\}$, l'expression est $2\tilde{x}_1\tilde{y}_1 + 8\tilde{x}_2\tilde{y}_2$.

Le lien entre x (y) et $\tilde{x}$ ($\tilde{y}$) est $\tilde{x}_1 = \dfrac{1}{\sqrt{2}}(x_1 - x_2)$ et $\tilde{x}_2 = \dfrac{1}{\sqrt{2}}(x_1 + x_2)$.

5.2 Décomposition en valeurs singulières (DVS)

On considère $\mathbb{R}^n$ et $\mathbb{R}^p$ munis du produit scalaire usuel noté $< .,. >$.

Soit A une matrice à n lignes et p colonnes. On suppose que $n \geq p$ et que le rang de A est r. On a alors $r \leq p$.

Comme $A^\top A$, de taille $p \times p$, est réelle symétrique, il existe deux matrices de taille $p \times p$, $V = (v_1, ..., v_p)$ orthogonale et $D = diag(\lambda_1, ..., \lambda_p)$ telles que $A^\top AV = VD$.

Comme $AA^\top$, de taille $n \times n$, est réelle symétrique, il existe deux matrices de taille $n \times n$, $U = (u_1, ..., u_n)$ orthogonale et $\Delta = diag(\mu_1, ..., \mu_n)$ telles que $AA^\top U = U\Delta$.

Grâce à la propriété $A(A^\top A) = (AA^\top)A$,

- Si $\lambda \neq 0$ est une valeur propre de $A^\top A$ et v un vecteur propre associé, alors λ est une valeur propre de $AA^\top$ associée à Av. Remarquez que v est un vecteur de $\mathbb{R}^p$, donc le produit Av a du sens, et donne un vecteur de $\mathbb{R}^n$.
  En effet, rappelons que $v \neq 0_{\mathbb{R}^p}$ (vecteur propre), d'où on déduit $\lambda v = A^\top Av \neq 0_{\mathbb{R}^p}$, puisque $\lambda \neq 0$, et donc $Av \neq O_{\mathbb{R}^n}$. On conclut en caclulant :
  $$A^\top Av = \lambda v \Rightarrow AA^\top Av = \lambda Av.$$
- De la même manière, si $\mu \neq 0$ est une valeur propre de $AA^\top$ et u un vecteur propre associé, alors μ est une valeur propre de $A^\top A$ associée à $A^\top u$.

On en conclut que les valeurs propres non nulles de $A^\top A$ et de $AA^\top$ sont identiques et que, pour les indices i correspondant, Av_i et u_i sont colinéaires.

De plus,

$$A^\top Av = \lambda v \Rightarrow \|Av\|^2 = v^\top A^\top Av = \lambda v^\top v = \lambda\|v\|^2,$$

donc les valeurs propres de $A^\top A$ sont ≥ 0.

Pour les $i = 1, ..., r$ tels que $\lambda_i \neq 0$, notons $\sigma_i = \sqrt{\lambda_i}$. Les σ_i s'appellent les **valeurs singulières** de A.

Comme Av_i et u_i sont colinéaires, il existe $t \in \mathbb{R}$ tel que $Av_i = tu_i$. Alors on a

$$t^2 = (tu_i)^\top(tu_i) = (Av_i)^\top(Av_i) = v_i^\top A^\top Av_i = v_i^\top \sigma_i^2 v_i = \sigma_i^2.$$

Ainsi on a la relation $Av_i = \sigma_i u_i$.

De cette relation, on tire

$$\underbrace{A}_{n\times p}\underbrace{(v_1, ..., v_r)}_{p\times r} = \underbrace{(u_1, ..., u_r)}_{n\times r}\underbrace{diag(\sigma_1, ..., \sigma_r)}_{r\times r}$$

qui s'écrit encore, parce que les vecteurs v_i sont orthonormés,

$$A = U_r\Sigma_r V_r^\top = \sigma_1 u_1 v_1^\top + ... + \sigma_r u_r v_r^\top$$

avec

$$U_r = (u_1, ..., u_r), \quad \Sigma = diag(\sigma_1, ..., \sigma_r) \text{ et } V_r = (v_1, ..., v_r).$$

En complétant
- la base orthonormale $(v_1, ..., v_r)$ par une base orthonormale du sous-espace propre de $A^\top A$ associé à la valeur propre 0
- et la base orthonormale $(u_1, ..., u_r)$ par une base orthonormale du sous-espace propre de $AA^\top$ associé à la valeur propre 0,

on obtient la formule

$$A = \underbrace{U}_{n\times n}\underbrace{\Sigma}_{n\times p}\underbrace{V^\top}_{p\times p}.$$

Ainsi, pour calculer la décomposition en valeurs singulières d'une matrice A de taille $n \times p$, on procède comme suit :
- on forme $M = A^\top A$;
- on calcule les valeurs propres $\lambda_1 \geq ... \geq \lambda_p$ de M ;
- on note r le plus grand indice i pour lequel $\lambda_i > 0$ et on pose $\sigma_i = \sqrt{\lambda_i}$ pour tout $1 \leq i \leq r$;
- on cherche une B.O.N. de chaque sous-espace propre et on en déduit une B.O.N. v_1, ..., v_p de vecteurs propres de M ;
- pour tout $1 \leq i \leq r$, on construit $u_i = Av_i/\sigma_i$;
- si $r < n$, on cherche une B.O.N. $(u_{r+1}, ..., u_n)$ du sous-espace propre de $AA^\top$ associé à la valeur propres 0 ;
- on forme $V = (v_1, ..., v_p)$, Σ de la taille de A avec les σ_i sur la diagonale principale et $U = (u_1, ..., u_n)$;

- on a finalement $\mathrm{A} = U\Sigma V^\top$.

Remarque 5.5 *Si B est une matrice de taille $n \times p$ avec $n < p$, alors on applique la méthode ci-dessus avec $A = B^\top$, $B^\top = U\Sigma V^\top$, puis on transpose : $A = B^\top = V\Sigma^\top U^\top$.*

Exemple : Prenons $A = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$. Cette matrice n'est pas diagonalisable. Essayons d'en trouver une DVS.

- On forme $M = A^\top A = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$.
- On a déjà calculé ses valeurs propres : $\lambda_1 = 8$ et $\lambda_2 = 2$.
- Ici, $r = 2$ et on pose $\sigma_1 = \sqrt{8} = 2\sqrt{2}$ et $\sigma_2 = \sqrt{2}$.
- On normalise les vecteurs propres déjà calculés plus haut pour construire une B.O.N. de vecteurs propres de M : On choisit donc $v_1 = (1,1)^\top/\sqrt{2}$ et $v_2 = (-1,1)^\top/\sqrt{2}$.
- On pose $u_1 = \dfrac{Av_1}{2\sqrt{2}} = (1,0)^\top$ et $u_2 = \dfrac{Av_2}{\sqrt{2}} = (0,1)^\top$.
- Posons $V = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$, $\Sigma = \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$ et $U = I_2$, alors $A = U\Sigma V^\top$.

Appendices

Annexe A

Preuves

Preuve : de la proposition 1.2

- Prouvons le point 2 : si on suppose que $A = (c_1, c_2, ..., c_n)$, on a $Ax = c_1x_1+c_2x_2+...+c_nx_n$ et donc

$$(Ax)^\top = c_1^\top x_1 + c_2^\top x_2 + ... + c_n^\top x_n = (x_1, x_2, ..., x_n)\begin{pmatrix} c_1^\top \\ c_2^\top \\ \vdots \\ c_n^\top \end{pmatrix} = x^\top A^\top.$$

- Prouvons le point 3 : soient A de taille $m \times n$ et B de taille $n \times p$. Supposons que les colonnes de B soient notées b_1, b_2, ..., b_p. Alors $AB = (Ab_1, Ab_2, ..., Ab_p)$ et par le point 2,

$$(AB)^\top = \begin{pmatrix} (Ab_1)^\top \\ (Ab_2)^\top \\ \vdots \\ (Ab_p)^\top \end{pmatrix} = \begin{pmatrix} b_1^\top A^\top \\ b_2^\top A^\top \\ \vdots \\ b_p^\top A^\top \end{pmatrix} = \begin{pmatrix} b_1^\top \\ b_2^\top \\ \vdots \\ b_p^\top \end{pmatrix} A^\top = B^\top A^\top.$$

- Prouvons le point 4 : partant de $AA^{-1} = A^{-1}A = I$, en prenant la transposée, on a $(A^{-1})^\top A^\top = A^\top (A^{-1})^\top = I^\top = I$ et donc $(A^\top)^{-1} = (A^{-1})^\top$.

$\square$

Preuve : de la proposition 2.10 Si $x = 0_E$, les 2 membres de l'inégalité sont nuls.

Supposons donc maintenant $x \neq 0_E$.

Si x et y sont colinéaires, alors il existe $t \in \mathbb{R}$ tel que $y = tx$. On a donc $| < x, y > | = |t < x, x > | = |t| < x, x >$ et $< y, y >= t^2 < x, x >$. Finalement, on a bien $| < x, y > | = \sqrt{< x, x >< y, y >}$.

Si maintenant x et y sont linéairement indépendants, pour tout $t \in \mathbb{R}$ on a $y + tx \neq 0$ et donc $(y + tx, y + tx) > 0$. Or

$$< y + tx, y + tx >=< x, x > t^2 + 2 < x, y > t + (y, y).$$

Il s'agit d'un polynôme du second degré en t dont le coefficient de plus au degré est > 0. Un tel polynôme ne peut-être > 0 que si son discriminant est < 0. Or $\Delta = 4 < x, y >^2 -4 < x, x >< y, y >$. Donc $< x, y >^2 << x, x >< y, y >$ ou encore $| < x, y > | < \sqrt{< x, x >< y, y >}$. $\square$

Preuve : du théorème 4.4 En exploitant l'orthogonalité des vecteurs, on a aisément

$$\begin{aligned}
\|u_1 + u_2 + ... + u_p\|^2 &= (u_1 + u_2 + ... + u_p, u_1 + u_2 + ... + u_p) \\
&= (u_1, u_1 + u_2 + ... + u_p) + (u_2, u_1 + u_2 + ... + u_p) + ... \\
&\quad +(u_p, u_1 + u_2 + ... + u_p) \\
&= \|u_1\|^2 + \|u_2\|^2 + ... + \|u_p\|^2.
\end{aligned}$$

□

Preuve : du théorème 4.12

- Soit $x \in E$ et x_1, x_2, ..., x_n ses coordonnées dans la base $(e_1, e_2, ..., e_n)$: $x = \sum_{i=1}^{n} x_i e_i$.

  Alors on a $< x, e_j >= \left\langle \sum_{i=1}^{n} x_i e_i, e_j \right\rangle = \sum_{i=1}^{n} x_i < e_i, e_j >= x_j$. On en déduit donc que $x = \sum_{i=1}^{n} < x, e_i > e_i$.
- Soient $x, y \in E$.

  D'après ce que l'on vient de voir, on a $x = \sum_{i=1}^{n} < x, e_i > e_i$ et $y = \sum_{j=1}^{n} < y, e_j > e_j$. On a donc

$$\begin{aligned}
< x, y > &= \left\langle \sum_{i=1}^{n} (x, e_i) e_i, \sum_{j=1}^{n} (y, e_j) e_j \right\rangle \\
&= \sum_{i=1}^{n} < x, e_i > \left\langle e_i, \sum_{j=1}^{n} (y, e_j) e_j \right\rangle \\
&= \sum_{i=1}^{n} < x, e_i > \sum_{j=1}^{n} < y, e_j >< e_i, e_j > \\
&= \sum_{i=1}^{n} < x, e_i >< y, e_i > .
\end{aligned}$$

□