ALGÈBRE LINEAIRE Module 2
Structure Euclidienne
PAD - Notes de cours

S. Rigal, D. Ruiz, et J. C. Satgé

December 5, 2008

Table des Matières

Chapitre 1

Espaces euclidiens – Orthogonalité

Ce premier chapitre met en place les notions de base essentielles pour la suite du cours. Vous découvrirez :

- La structure d'espace vectoriel normé sur l'ensemble des nombres réels et sur l'ensemble des nombres complexes.
- La notion de produit scalaire qui permet de définir l'orthogonalité dans un espace vectoriel dit "euclidien".
- Une des applications fondamentales cette structure euclidienne : les bases orthonormées et les matrices orthogonales.
- Une technique de calcul (il s'agira pour vous de savoir reproduire les méthodes développées dans le cours et illustrées par des exemples) : le procédé d'orthogonalisation de SCHMIDT qui permet de construire, à partir d'une base donnée, une base orthonormée.

1-1 Espaces euclidiens

1-1.1 Espaces vectoriels normés – Généralités

La structure d'espace vectoriel n'est que la structure minimale permettant de traiter des problèmes d'algèbre linéaire. On peut enrichir cette structure en la complétant, en particulier, par la structure métrique. En effet, le concept d'espace vectoriel est uniquement algébrique et, afin de pouvoir étudier, dans ces espaces, des problèmes tels que la convergence, il faut ajouter à la structure d'espaces vectoriel des concepts topologiques, et en particulier la notion de de norme et de distance.

Définition 1-1.1 ***Espaces vectoriels normés***
Un espace vectoriel ***normé*** *sur un corps K ($\mathbb{R}$ ou $\mathbb{C}$) est un espace vectoriel E muni d'une application de E dans $\mathbb{R}$ qui à tout élément $\mathbf{x} \in E$ associe le réel $\|\mathbf{x}\|_E$* ***appelé norme de*** $\mathbf{x}$ *et vérifiant les propriétés suivantes :*

1. $\forall \mathbf{x} \in E \,, \quad \|\mathbf{x}\|_E \geq 0$
2. $\|\mathbf{x}\|_E = 0 \Longleftrightarrow \mathbf{x} = 0$
3. $\forall \mathbf{x} \in E, \ \forall \mathbf{y} \in E \,, \quad \|\mathbf{x}+\mathbf{y}\|_E \leqslant \|\mathbf{x}\|_E + \|\mathbf{y}\|_E$ *(inégalité triangulaire)*
4. $\forall \lambda \in K \quad \forall \mathbf{x} \in E \,, \quad \|\lambda \mathbf{x}\|_E = |\lambda| . \|\mathbf{x}\|_E$

Définition 1-1.2 ***Distance***
Dans un espace vectoriel normé E, la ***distance*** *entre deux éléments $\mathbf{x}$ et $\mathbf{y}$ est définie par la norme $\|\mathbf{x}-\mathbf{y}\|_E$.*

Exemple : Dans $\mathbb{R}^n$, on peut considérer par exemple les normes suivantes :

- $\|\mathbf{x}\|_\infty = \max\limits_{1 \leqslant i \leqslant n} |x_i|$
- $\|\mathbf{x}\|_1 = \sum\limits_{i=1}^{n} |x_i|$
- $\|\mathbf{x}\|_2 = \sqrt{\sum\limits_{i=1}^{n} x_i^2}$
- $\|\mathbf{x}\|_p = \left(\sum\limits_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$, pour $p > 2$.

1-1.2 Produit scalaire canonique dans $\mathbb{R}^n$ – Norme euclidienne

La norme sur $\mathbb{R}^n$ introduite précédemment, et définie par

$$\forall \mathbf{x} \in \mathbb{R}^n, \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}\,,$$

est appelée NORME 2 ou encore NORME EUCLIDIENNE CLASSIQUE sur $\mathbb{R}^n$. Cette norme est associée au PRODUIT SCALAIRE CANONIQUE sur $\mathbb{R}^n$ défini par :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \langle \mathbf{x}\,, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i\, y_i\,.$$

Le produit scalaire canonique de deux vecteurs de $\mathbb{R}^n$ est une application de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$ qui a les propriétés suivantes :

1. Le produit scalaire est une application **bilinéaire**, c'est à dire que, pour tous vecteurs $\mathbf{u}$, $\widetilde{\mathbf{u}}$, $\mathbf{v}$, et $\widetilde{\mathbf{v}}$ de $\mathbb{R}^n$ et tous scalaire λ de $\mathbb{R}$, on a :

$$\begin{aligned} \langle \mathbf{u} + \widetilde{\mathbf{u}}\,,\, \mathbf{v} \rangle &= \langle \mathbf{u}\,,\, \mathbf{v} \rangle + \langle \widetilde{\mathbf{u}}\,,\, \mathbf{v} \rangle & \qquad \langle \lambda \mathbf{u}\,,\, \mathbf{v} \rangle &= \lambda \langle \mathbf{u}\,,\, \mathbf{v} \rangle \\ \langle \mathbf{u}\,,\, \mathbf{v} + \widetilde{\mathbf{v}} \rangle &= \langle \mathbf{u}\,,\, \mathbf{v} \rangle + \langle \mathbf{u}\,,\, \widetilde{\mathbf{v}} \rangle & \qquad \langle \mathbf{u}\,,\, \lambda \mathbf{v} \rangle &= \lambda \langle \mathbf{u}\,,\, \mathbf{v} \rangle \end{aligned}$$

2. Le produit scalaire est une application **symétrique**, c'est à dire que, pour tous vecteurs $\mathbf{u}$ et $\mathbf{v}$ de $\mathbb{R}^n$, on a :

$$\langle \mathbf{u}\,, \mathbf{v} \rangle = \langle \mathbf{v}\,, \mathbf{u} \rangle\,.$$

3. Le produit scalaire est une application **définie positive**, c'est à dire que, pour tout vecteur **$\mathbf{u}$ non nul** de $\mathbb{R}^n$, on a :

$$\langle \mathbf{u}\,, \mathbf{u} \rangle > 0\,.$$

Définition 1-1.3 *$\mathbb{R}^n$ muni de la norme 2 ci-dessus est appelé espace* EUCLIDIEN*. Ceci signifie simplement que la norme 2 considérée* "dérive" *d'un produit scalaire par la relation :*

$$\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x}\,,\, \mathbf{x} \rangle}\,,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire canonique sur $\mathbb{R}^n$.

Propriétés :

- Le produit scalaire canonique de deux vecteurs dans $\mathbb{R}^n$

$$\langle \mathbf{x}\,, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i\, y_i\,,$$

peut aussi s'écrire à l'aide de la **transposition** sous la forme du produit matriciel suivant :

$$\langle \mathbf{x} \, , \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} \, .$$

C'est d'ailleurs cette relation qui nous a permis de décrire le produit matriciel à l'aide de produits scalaires entre des lignes et des colonnes des matrices intervenant dans ce produit.

- Notons l'inégalité de Cauchy-Schwarz :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad |\langle \mathbf{x} \, , \mathbf{y} \rangle| \ \leqslant \ \|\mathbf{x}\|_2 \, \|\mathbf{y}\|_2 \, ,$$

l'égalité n'ayant lieu que si $\mathbf{x} = \mathbf{0}$ ou $\mathbf{y} = \mathbf{0}$ ou $\mathbf{x} = \lambda \mathbf{y}$.

- Notons aussi **l'dentité du parallélogramme**

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \, , \quad \|\mathbf{x} + \mathbf{y}\|_2^2 + \|\mathbf{x} - \mathbf{y}\|_2^2 \ = \ 2\|\mathbf{x}\|_2^2 + 2\|\mathbf{y}\|_2^2 \, .$$

- Enfin, on a les deux égalités suivantes :

$$\begin{aligned} \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \, , \quad \langle \mathbf{x} \, , \mathbf{y} \rangle \ &= \ \frac{1}{4} \left(\|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{y}\|_2^2 \right) \, , \\ \text{et} \quad \langle \mathbf{x} \, , \mathbf{y} \rangle \ &= \ \frac{1}{2} \left(\|\mathbf{x} + \mathbf{y}\|_2^2 - \|\mathbf{x}\|_2^2 - \|\mathbf{y}\|_2^2 \right) \, . \end{aligned}$$

1-1.3 Produit scalaire canonique dans $\mathbb{C}^n$

Le produit scalaire canonique de deux vecteurs de $\mathbb{C}^n$ est défini par :

$$\forall \mathbf{x}, \mathbf{y} \in \mathbb{C}^n, \quad \langle \mathbf{x} \, , \mathbf{y} \rangle = \sum_{i=1}^{n} \overline{x_i} \, y_i \, ,$$

où, pour un nombre complexe $z = a + ib$ donné, $\overline{z} = a - ib$ désigne le complexe conjugué de z. Ce produit scalaire généralise celui défini dans $\mathbb{R}^n$ en ce sens que si les composantes des deux vecteurs $\mathbf{x}$ et $\mathbf{y}$ sont réelles, alors la conjugaison est inopérante et on retrouve le produit scalaire sur $\mathbb{R}^n$.

Propriété : Le produit scalaire canonique sur $\mathbb{C}^n$ peut aussi s'écrire à l'aide de la **transposition** et de la **conjugaison** sous la forme du produit matriciel suivant :

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\mathbf{x}^T} \mathbf{y} \, .$$

La norme euclidienne sur $\mathbb{C}^n$ associée est donné par :

$$\forall \mathbf{x} \in \mathbb{C}^n, \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \, ,$$

où, pour un nombre complexe $z = a + ib$ donné, $|z|^2 = \overline{z} z = a^2 + b^2$, désigne le module de z au carré. Dans $\mathbb{C}^n$ muni de cette norme, on retrouve les mêmes propriétés que celles énoncées précédemment dans le cas de la norme 2 sur $\mathbb{R}^n$. La seule petite différence réside dans les propriétés du produit scalaire, à savoir :

$\forall \mathbf{u}, \mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^n$ et $\forall \lambda, \mu \in \mathbb{C}$:

$$\left.\langle \mathbf{u}\,,\, \lambda\mathbf{v}_1 + \mu\mathbf{v}_2\rangle = \lambda\langle \mathbf{u}\,,\, \mathbf{v}_1\rangle + \mu\langle \mathbf{u}\,,\, \mathbf{v}_2\rangle\right\} \quad \text{(linéarité à droite),}$$

$$\left.\langle \mathbf{u}\,,\, \mathbf{v}\rangle = \overline{\langle \mathbf{v}\,,\, \mathbf{u}\rangle}\right\}$$ **symétrie hermitienne** (qui dans le cas réel se réduit à la symétrie classique),

$$\left.\begin{array}{l}\langle \mathbf{v}\,,\, \mathbf{v}\rangle \geq 0 \\ \langle \mathbf{v}\,,\, \mathbf{v}\rangle = 0 \;\Rightarrow\; \mathbf{v} = 0\end{array}\right\}$$ **définie positivité**.

1-1.4 Produit scalaire sur un espace vectoriel – Espaces euclidiens

Définition 1-1.4 *Soit E un $\mathbb{R}$-espace vectoriel, on appelle **produit scalaire** sur E une application $\varphi : E \times E \to \mathbb{R}$ qui est :*

- ***bilinéaire**, c'est à dire que, pour tous vecteurs $\mathbf{u}$, $\widetilde{\mathbf{u}}$, $\mathbf{v}$, et $\widetilde{\mathbf{v}}$ de E et tout scalaire λ de $\mathbb{R}$, on a :*

$$\begin{array}{ll}\varphi(\mathbf{u}+\widetilde{\mathbf{u}}\,,\,\mathbf{v}) = \varphi(\mathbf{u}\,,\,\mathbf{v}) + \varphi(\widetilde{\mathbf{u}}\,,\,\mathbf{v}) & \varphi(\lambda\mathbf{u}\,,\,\mathbf{v}) = \lambda\varphi(\mathbf{u}\,,\,\mathbf{v}) \\ \varphi(\mathbf{u}\,,\,\mathbf{v}+\widetilde{\mathbf{v}}) = \varphi(\mathbf{u}\,,\,\mathbf{v}) + \varphi(\mathbf{u}\,,\,\widetilde{\mathbf{v}}) & \varphi(\mathbf{u}\,,\,\lambda\mathbf{v}) = \lambda\varphi(\mathbf{u}\,,\,\mathbf{v})\end{array}$$

- ***symétrique**, c'est à dire que, pour tous vecteurs $\mathbf{u}$ et $\mathbf{v}$ de E, on a :*

$$\varphi(\mathbf{u}\,,\mathbf{v}) = \varphi(\mathbf{v}\,,\mathbf{u})$$

- ***définie positive**, c'est à dire que pour tout vecteur $\mathbf{u}$ de E, on a : :*

$$\varphi(\mathbf{u}\,,\mathbf{u}) \geq 0 \quad \underline{\textbf{et}} \quad \varphi(\mathbf{u}\,,\mathbf{u}) = 0 \Leftrightarrow \mathbf{u} = \mathbf{0}$$

Définition 1-1.5 *Un $\mathbb{R}$-espace vectoriel de dimension finie, muni d'un produit scalaire est appelé **espace euclidien**.*

1-1.5 Exemples

1. Soit E l'ensemble $\mathcal{M}_2(\mathbb{R})$ des matrices carrées 2×2. Soit l'application ϕ de E^2 dans $\mathbb{R}$ définie par :

$$\phi\,(\mathbf{A}, \mathbf{B}) = \mathit{trace}\left(\mathbf{A}\mathbf{B}^T\right) .$$

La trace et la transposition des matrices étant des applications linéaires, l'application ϕ est bilinéaire.

De plus, on sait que pour toute matrice carrée, $\mathbf{C}$, $\textit{trace}\left(\mathbf{C}^T\right) = \textit{trace}\left(\mathbf{C}\right)$, et comme $\left(\mathbf{B}\mathbf{A}^T\right)^T = \mathbf{A}\mathbf{B}^T$, on en déduit que $\phi\left(\mathbf{A},\mathbf{B}\right) = \phi\left(\mathbf{B},\mathbf{A}\right)$, ce qui prouve que ϕ est bien symétrique.

Enfin, $\phi\left(\mathbf{A},\mathbf{A}\right) = \textit{trace}\left(\mathbf{A}\left(\mathbf{A}^T\right)\right) = \sum_{(i,j)} a_{i,j}^2 = a_{1,1}^2 + a_{1,2}^2 + a_{2,1}^2 + a_{2,2}^2$ est positif, et pour finir, $\phi\left(\mathbf{A},\mathbf{A}\right) = 0$ entraîne, pour tout i et j, $a_{i,j}^2 = 0$, et donc $\mathbf{A} = \mathbf{0}$.

Nous avons vérifié tous les critères qui font que ϕ est un produit scalaire sur$\mathcal{M}_2\left(\mathbb{R}\right)$ Ceci s'étend sans difficultés à $M_n\left(\mathbb{R}\right)$, $n \geq 2$.

2. Soit E l'ensemble des applications continues de $[0,1]$ dans $\mathbb{R}$ (cet ensemble se note $C_0\left([0,1],\mathbb{R}\right)$). Soit l'application ϕ de E^2 dans $\mathbb{R}$ définie par :

$$\phi\left(f,g\right) = \int_0^1 f(x)g(x)\,dx\ .$$

Du fait de la commutativité de la multiplication dans $\mathbb{R}$ et de la linéarité de l'intégrale, ϕ est trivialement bilinéaire et symétrique.

De plus $\phi\left(f,f\right) = \int_0^1 f(x)^2\,dx$. f^2 étant une fonction continue positive sur $[0,1]$, on a donc $\phi\left(f,f\right) \geq 0$. Pour finir, $\phi\left(f,f\right) = 0 \Rightarrow f = 0$.

Nous avons vérifié tous les critères qui font que ϕ est un produit scalaire sur $C_0\left([0,1],\mathbb{R}\right)$. Ceci s'étend sans difficultés à $C_0\left([a,b],\mathbb{R}\right)$, a et b quelconques.

1-2 Bases orthonormées – Matrices orthogonales

1-2.1 Orthogonalité

On rappelle qu'un espace vectoriel E est dit EUCLIDIEN si E est un espace vectoriel normé de dimension finie et si sa norme est associée à un produit scalaire par la relation

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}\,,\,\mathbf{x}\rangle}\,,$$

où $\langle \cdot, \cdot \rangle$ désigne le produit scalaire sur E.

Sur un espace euclidien, le produit scalaire définit en fait une **mesure d'angle** entre deux vecteurs, et on peut alors introduire la notion d'orthogonalité entre vecteurs :

Définition 1-2.1 *Soit E un espace* EUCLIDIEN *muni du produit scalaire $\langle \cdot, \cdot \rangle$. On dit que deux vecteurs $\mathbf{u}$ et $\mathbf{v}$ non nuls de E sont orthogonaux si et seulement si :*

$$\langle \mathbf{u}\,,\,\mathbf{v}\rangle = 0\,.$$

On notera alors $\mathbf{x} \perp \mathbf{y}$.

Un vecteur $\mathbf{x}$ de E est orthogonal à un ensemble $S \subset E$ si $\langle \mathbf{x}, \mathbf{y}\rangle = 0$ pour tout $\mathbf{y} \in S$.

Propriété : On rappelle le théorème de Pythagore, dans le cas de deux vecteurs orthogonaux :

$$\text{Si } \mathbf{x} \perp \mathbf{y}\ , \quad \text{alors} \quad \|\mathbf{x}+\mathbf{y}\|^2 \ = \ \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\ .$$

Définition 1-2.2 ***Orthogonal d'un ensemble***

Soit E un espace Euclidien, et A un sous ensemble de E. On appelle orthogonal de A l'ensemble, noté $A^\perp$, de tous les vecteurs orthogonaux à A :

$$A^\perp = \{\mathbf{x} \in E \mid \ \forall \mathbf{y} \in A\ , \ \ \langle \mathbf{x}, \mathbf{y}\rangle = 0\}\ .$$

Proposition 1-2.1 *Soient A et B deux sous ensembles d'un espace euclidien E. On a alors :*

1. *$A^\perp$ est un sous espace vectoriel de E*
2. *$A \subset A^{\perp\perp}$*
3. *Si $A \subset B$, alors $B^\perp \subset A^\perp$*
4. *$A^\perp = A^{\perp\perp\perp}$*
5. *$A^{\perp\perp} = \mathsf{Vect}(A)$, c'est à dire le plus petit sous-espace vectoriel contenant A.*

Théorème 1-2.1 ***Supplémentaire orthogonal***
Soit E un espace euclidien, et soit V un sous-espace vectoriel de E. Tout vecteur $\mathbf{x} \in E$ se décompose de manière unique sous la forme :

$$\mathbf{x} = \mathbf{v} + \mathbf{w}\ ,\quad \mathbf{v} \in V\ ,\quad \mathbf{w} \in V^{\perp}\ .$$

On note alors

$$E = V \oplus V^{\perp}\ ,$$

et on a

$$V = V^{\perp\perp}\ ,\quad E^{\perp} = \{0\}\ ,\quad \{0\}^{\perp} = E\ .$$

1-2.2 Bases orthonormées

Définition 1-2.3 ***Famille orthogonale – Famille orthonormée***
*Soit E un espace euclidien muni du produit scalaire $\langle\cdot,\cdot\rangle$. Une famille de vecteurs $(\mathbf{u}_k)_{1\leqslant k\leqslant m}$ de E est dite **orthogonale** si et seulement si*

$$\forall i \neq j \in \{1,\dots,m\}\ ,\quad \langle \mathbf{u}_i, \mathbf{u}_j\rangle = 0\,.$$

*Cette famille est dite **orthonormée** si et seulement si elle est orthogonale et si*

$$\forall k \in \{1,\dots,m\}\ ,\quad \|\mathbf{u}_k\| = 1\,.$$

Proposition 1-2.2 *Soit E un espace euclidien de dimension n, muni du produit scalaire $\langle\cdot,\cdot\rangle$. Toute famille **orthogonale** de vecteurs non nuls de E est nécessairement **libre** dans E (les vecteurs de cette famille sont linéairement indépendants).*

*En particulier, une famille **orthogonale** de n vecteurs non nuls de E forme une **base** de E.*

Définition 1-2.4 ***Base orthonormée***
Soit E un espace euclidien de dimension n, muni du produit scalaire $\langle\cdot,\cdot\rangle$. On appelle base orthonormée de E, toute famille orthonormée de n vecteurs de E.

Rappelons enfin les propriétés suivantes :

Proposition 1-2.3 *Soit E un espace euclidien de dimension n, muni du produit scalaire $\langle\cdot,\cdot\rangle$ et F un sous espace vectoriel de E. La réunion d'une base orthonormée de F et d'une base orthonormée de $F^{\perp}$ est une base orthonormée de E*

Proposition 1-2.4 *Soit E un espace euclidien de dimension n, muni du produit scalaire $\langle\cdot,\cdot\rangle$ et soit $(\mathbf{u}_k)_{1\leqslant k\leqslant n}$ une base orthonormée de E.*

Alors, pour tout vecteur $\mathbf{x} \in E$ se décomposant dans la base $(\mathbf{u}_k)_{1\leqslant k\leqslant n}$ sous la forme

$$\mathbf{x} = \sum_{k=1}^{n} \alpha_k \mathbf{u}_k$$

on a les égalités suivantes

$$\forall 1 \leqslant k \leqslant n\,,\quad \alpha_k \;=\; \langle \mathbf{x}, \mathbf{u}_k \rangle \qquad \textit{(coefficients de Fourier)}$$

$$\|\mathbf{x}\|^2 \;=\; \sum_{k=1}^{n} |\alpha_k|^2 \qquad \textit{(égalité de Parseval).}$$

1-2.3 Matrices orthogonales dans $\mathcal{M}_n(\mathbb{R})$

Définition 1-2.5 ***Matrices orthogonales***
Une matrice de $\mathcal{M}_n(\mathbb{R})$ est dite ORTHOGONALE *si ses colonnes, ou, de manière équivalente, ses lignes, forment une base orthonormée de $\mathbb{R}^n$ pour le produit scalaire canonique dans $\mathbb{R}^n$.*

Elles représentent donc les opérateurs de changement de base de la base canonique de $\mathbb{R}^n$ (qui est orthonormée) vers une base orthonormée de $\mathbb{R}^n$.

Proposition 1-2.5 *Une matrice $\mathbf{Q} \in \mathcal{M}_n(\mathbb{R})$ orthogonale vérifie les propriétés suivantes :*

1. *Le fait que les colonnes de $\mathbf{Q}$ forment un ensemble orthonormé de vecteurs s'écrit aussi $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_n$, la matrice $\mathbf{Q}^T\mathbf{Q}$ étant en effet constituée des produits scalaires des colonnes de $\mathbf{Q}$ les unes avec les autres.*

2. *Le fait que les lignes de $\mathbf{Q}$ forment un ensemble orthonormé de vecteurs s'écrit aussi $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}_n$, la matrice $\mathbf{Q}\mathbf{Q}^T$ étant en effet constituée des produits scalaires des lignes de $\mathbf{Q}$ les unes avec les autres.*

3. *Des deux égalités précédentes, on peut donner la caractérisation suivante d'une matrice orthogonale dans $\mathcal{M}_n(\mathbb{R})$, à savoir qu'une matrice $\mathbf{Q} \in \mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si l'inverse de $\mathbf{Q}$ est égale à sa transposée :*

$$\mathbf{Q}^{-1} = \mathbf{Q}^T\,.$$

Proposition 1-2.6 *Soit $\mathbf{A} \in \mathcal{M}_n(\mathbb{R})$ et $\mathbf{B} \in \mathcal{M}_n(\mathbb{R})$ deux matrices semblables exprimées dans deux bases orthonormées $\mathcal{B}$ et $\mathcal{B}'$. La matrice de passsage $\mathbf{P}$ de la base $\mathcal{B}$ à la base $\mathcal{B}'$ est orthogonale, la formule de changement de base devient donc :*

$$\mathbf{B} = \mathbf{P}^T\mathbf{A}\mathbf{P}$$

1-2.4 Matrices unitaires dans $\mathcal{M}_n(\mathbb{C})$

Dans le cas de $\mathbb{C}^n$ muni de la norme 2 et du produit scalaire canonique, on retrouve des propriétés très similaires aux précédentes dans le cas des matrices orthogonales de $\mathcal{M}_n(\mathbb{C})$.

Les seules différences résident dans le fait que le produit scalaire canonique sur $\mathbb{C}^n$ est à symétrie hermitienne (voir § 1-1.3), et on a en particulier :

- Une matrice de $\mathcal{M}_n(\mathbb{C})$ dont les colonnes ou les lignes forment une base orthonormée de $\mathbb{C}^n$ pour le produit scalaire canonique dans $\mathbb{C}^n$, est dite UNITAIRE (simplement pour faire la différence avec le cas réel).

  Ces matrices représentent là encore les opérateurs de changement de base de la base canonique de $\mathbb{C}^n$ vers une base orthonormée de $\mathbb{C}^n$.

- La définition du produit scalaire canonique de deux vecteurs de $\mathbb{C}^n$, qui fait intervenir le complexe conjugué du vecteur de gauche ($\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\mathbf{x}^T}\mathbf{y}$), conduit à la caractérisation suivante des matrices unitaires dans $\mathcal{M}_n(\mathbb{C})$, à savoir qu'une matrice $\mathbf{Q} \in \mathcal{M}_n(\mathbb{C}^n)$ est unitaire si et seulement si l'inverse de $\mathbf{Q}$ est égale au complexe conjugué de sa transposée :

$$\mathbf{Q}^{-1} = \overline{\mathbf{Q}^T}\,, \qquad (\text{noté aussi } \overline{\mathbf{Q}^T} = \mathbf{Q}^*)\,.$$

1-3 Procédé d'orthogonalisation de Schmidt

Ayant à notre disposition, dans les espaces EUCLIDIENS, les notions d'orthogonalité et de norme, il est souvent utile de travailler dans des bases orthonormées. Le procédé d'orthogonalisation de SCHMIDT permet de construire, à partir d'une base donnée, une base orthonormée pour le produit scalaire de l'espace euclidien dans lequel on travaille.

1-3.1 Introduction par un exemple

Illustrons cette méthode dans $\mathbb{R}^3$ muni du produit scalaire canonique. Soit donc une base de $\mathbb{R}^3$:

$$\left\{ \mathbf{u}_1 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_2 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \mathbf{u}_3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

On vérifie que l'on a bien une base, le déterminant de cette famille de vecteurs étant non nul (il vaut 3).

1$^{\text{ère}}$ étape : Posons $\mathbf{w}_1 = \mathbf{u}_1$ et déterminons r_{12} tel que $\mathbf{w}_2 = (\mathbf{u}_2 - r_{12}\mathbf{w}_1)$ soit orthogonal à $\mathbf{w}_1$. Nous écrivons donc :

$$\begin{aligned} \langle \mathbf{w}_1 \,,\, \mathbf{w}_2 \rangle = 0 &\Leftrightarrow \langle \mathbf{w}_1 \,,\, \mathbf{u}_2 - r_{12}\mathbf{w}_1 \rangle = 0 \\ &\Leftrightarrow \langle \mathbf{w}_1 \,,\, \mathbf{u}_2 \rangle - r_{12} \langle \mathbf{w}_1 \,,\, \mathbf{w}_1 \rangle = 0 \\ &\Leftrightarrow r_{12} = \frac{\langle \mathbf{w}_1 \,,\, \mathbf{u}_2 \rangle}{\langle \mathbf{w}_1 \,,\, \mathbf{w}_1 \rangle}. \end{aligned}$$

Dans le cas présent, on obtient $r_{12} = 1$ et $\mathbf{w}_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

2$^{\text{ème}}$ étape : Gardons les deux vecteurs $\mathbf{w}_1$ et $\mathbf{w}_2$, et déterminons r_{13} et r_{23} tels que $\mathbf{w}_3 = (\mathbf{u}_3 - r_{13}\mathbf{w}_1 - r_{23}\mathbf{w}_2)$ soit orthogonal aux deux vecteurs précédemment construits, $\mathbf{w}_1$ et $\mathbf{w}_2$. On est donc conduit à résoudre le système suivant :

$$\begin{aligned} \begin{cases} \langle \mathbf{w}_1 \,,\, \mathbf{w}_3 \rangle = 0 \\ \langle \mathbf{w}_2 \,,\, \mathbf{w}_3 \rangle = 0 \end{cases} &\Leftrightarrow \begin{cases} \langle \mathbf{w}_1 \,,\, \mathbf{u}_3 - r_{13}\mathbf{w}_1 - r_{23}\mathbf{w}_2 \rangle = 0 \\ \langle \mathbf{w}_2 \,,\, \mathbf{u}_2 - r_{12}\mathbf{w}_1 - r_{23}\mathbf{w}_2 \rangle = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} \langle \mathbf{w}_1 \,,\, \mathbf{u}_3 \rangle - r_{13} \langle \mathbf{w}_1 \,,\, \mathbf{w}_1 \rangle = 0 \\ \langle \mathbf{w}_2 \,,\, \mathbf{u}_3 \rangle - r_{23} \langle \mathbf{w}_2 \,,\, \mathbf{w}_2 \rangle = 0 \end{cases} \quad \text{(sachant que } \langle \mathbf{w}_1 \,,\, \mathbf{w}_2 \rangle = 0\text{)} \\ &\Leftrightarrow \begin{cases} r_{13} = \dfrac{\langle \mathbf{w}_1 \,,\, \mathbf{u}_3 \rangle}{\langle \mathbf{w}_1 \,,\, \mathbf{w}_1 \rangle} \\ r_{23} = \dfrac{\langle \mathbf{w}_2 \,,\, \mathbf{u}_3 \rangle}{\langle \mathbf{w}_2 \,,\, \mathbf{w}_2 \rangle} \end{cases}. \end{aligned}$$

On obtient $r_{13} = \dfrac{1}{2}$ et $r_{23} = 0$, et une base **orthogonale** de $\mathbb{R}^3$ est donc :

$$\left\{ \mathbf{w}_1 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{w}_2 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \mathbf{w}_3 \begin{pmatrix} 1 \\ 1/2 \\ -1/2 \end{pmatrix} \right\}.$$

3ème étape : Il ne reste plus qu'à normaliser cette base en divisant chacun des vecteurs par sa norme, $\|\mathbf{w}_1\| = \sqrt{2}$, $\|\mathbf{w}_2\| = \sqrt{3}$, et $\|\mathbf{w}_3\| = \sqrt{\dfrac{3}{2}}$. On obtient enfin une base **orthonormée** de $\mathbb{R}^3$:

$$\left\{ \mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \mathbf{v}_3 = \sqrt{\frac{2}{3}} \begin{pmatrix} 1 \\ 1/2 \\ -1/2 \end{pmatrix} \right\}.$$

1-3.2 Généralisation

Généralisons en nous plaçant dans $\mathbb{R}^n$. On procède par récurrence, à partir d'une base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, on construit une **base orthogonale** $\mathcal{O} = \{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_n\}$.

Supposons construits $(p-1)$ vecteurs 2 à 2 orthogonaux, $\{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_{p-1}\}$, on pose $\boldsymbol{\varepsilon}_p = \mathbf{e}_p -$ (*combinaison linéaire des vecteurs déjà trouvés*) de telle sorte que $\boldsymbol{\varepsilon}_p$ soit orthogonal à chacun des vecteurs $\mathbf{e}_1, \ldots, \mathbf{e}_{p-1}$. Soit donc

$$\boldsymbol{\varepsilon}_p = \mathbf{e}_p - \sum_{k=1}^{p-1} r_{kp}\, \boldsymbol{\varepsilon}_k \,.$$

La condition $\boldsymbol{\varepsilon}_p \perp \boldsymbol{\varepsilon}_k$ équivaut à $r_{kp} = \dfrac{\langle \boldsymbol{\varepsilon}_k \,,\, \mathbf{e}_p \rangle}{\langle \boldsymbol{\varepsilon}_k \,,\, \boldsymbol{\varepsilon}_k \rangle}$ pour k variant de 1 à $(p-1)$.

Il ne reste plus qu'à normaliser cette base en divisant chacun des vecteurs par sa norme. On posera enfin

$$\widetilde{\mathcal{O}} = \{\widetilde{\boldsymbol{\varepsilon}_1}, \widetilde{\boldsymbol{\varepsilon}_2}, \ldots, \widetilde{\boldsymbol{\varepsilon}_n}\} \;,$$

avec $\widetilde{\boldsymbol{\varepsilon}_k} = \dfrac{\boldsymbol{\varepsilon}_k}{\|\boldsymbol{\varepsilon}_k\|}$ pour $k = 1, \ldots, n$.

1-4 Factorisation QR d'une matrice carrée inversible

1-4.1 Définition – Propriétés

Dans le procédé d'orthonormalisation de SCHMIDT, par exemple, qui permet de construire une base orthonormée $\mathcal{O} = \{\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \ldots, \boldsymbol{\varepsilon}_n\}$ à partir d'une base $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ donnée, on peut exprimer matriciellement l'ensemble des transformations qui sont mises en œuvre, et aboutir ainsi à une relation matricielle très particulière que nous allons détailler dans ce qui suit.

En effet, l'ensemble des transformations du procédé de SCHMIDT sont des transformations linéaires sur un ensemble de vecteurs. Si on reprend les calculs développés au § 1-3.2, on avait abouti à une relation du type :

$$\widetilde{\boldsymbol{\varepsilon}_p} = \frac{\boldsymbol{\varepsilon}_p}{\|\boldsymbol{\varepsilon}_p\|} = \frac{1}{\|\boldsymbol{\varepsilon}_p\|}\left(\mathbf{e}_p - \sum_{k=1}^{p-1} \frac{\langle \boldsymbol{\varepsilon}_k \,,\, \mathbf{e}_p\rangle}{\langle \boldsymbol{\varepsilon}_k \,,\, \boldsymbol{\varepsilon}_k\rangle}\boldsymbol{\varepsilon}_k\right) ,$$

que l'on peut aussi écrire, en exploitant directement les vecteurs normés $\widetilde{\boldsymbol{\varepsilon}_k}$:

$$\widetilde{\boldsymbol{\varepsilon}_p} = \frac{1}{\|\boldsymbol{\varepsilon}_p\|}\left(\mathbf{e}_p - \sum_{k=1}^{p-1} \langle \widetilde{\boldsymbol{\varepsilon}_k} \,,\, \mathbf{e}_p\rangle\, \widetilde{\boldsymbol{\varepsilon}_k}\right) .$$

En posant cette fois $r_{kp} = \langle \widetilde{\boldsymbol{\varepsilon}_k} \,,\, \mathbf{e}_p\rangle$, pour $k = 1, \ldots, (p-1)$, et $r_{pp} = \|\boldsymbol{\varepsilon}_p\|$, on a alors

$$\mathbf{e}_p = \sum_{k=1}^{p} r_{kp}\, \widetilde{\boldsymbol{\varepsilon}_k} .$$

Si on note enfin

- $\mathbf{A} \in \mathcal{M}_n(\mathbb{R})$ la matrice carrée dont les colonnes sont formées des vecteurs de base $\mathbf{e}_1$, $\mathbf{e}_2$, ..., $\mathbf{e}_n$,

- $\mathbf{Q} \in \mathcal{M}_n(\mathbb{R})$ la matrice carrée dont les colonnes sont formées des vecteurs orthonormés $\widetilde{\boldsymbol{\varepsilon}}_1$, $\widetilde{\boldsymbol{\varepsilon}}_2$, ..., $\widetilde{\boldsymbol{\varepsilon}}_n$,

- et $\mathbf{R} \in \mathcal{M}_n(\mathbb{R})$ la matrice carrée triangulaire supérieure formée des coefficients r_{ij}, $i = 1, \ldots, j$, $j = 1, \ldots, n$:

$$\begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1n} \\ 0 & r_{22} & \ldots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & r_{nn} \end{pmatrix} ,$$

on a alors la relation suivante :

$$\mathbf{A} = \mathbf{Q}\mathbf{R}\,.$$

La matrice $\mathbf{Q}$, dont les colonnes forment une base orthonormée de $\mathbb{R}^n$, est *orthonormale.*

La décomposition d'une matrice carrée inversible sous la forme $\mathbf{A} = \mathbf{QR}$ est aussi appelée **factorisation QR de la matrice A**. Le procédé de Schmidt est l'une des façons de construire une telle factorisation.

En repartant du procédé de Schmidt par exemple, qui est récurrent, on peut aussi vérifier que $\mathsf{Vect}\,\{\mathbf{e}_1, \ldots, \mathbf{e}_k\} = \mathsf{Vect}\,\{\widetilde{\varepsilon}_1, \ldots, \widetilde{\varepsilon}_k\}$, pour tout $k \in \{1, \ldots, n\}$.

Proposition 1-4.1 *Soit $\mathbf{A} = \mathbf{QR}$ la factorisation $\mathbf{QR}$ d'une matrice carrée inversible. On a alors,*

$$\forall k \in \{1, \ldots, n\}\ , \quad \mathsf{Vect}\,\{\mathbf{a}_1, \ldots, \mathbf{a}_k\} = \mathsf{Vect}\,\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}\ ,$$

où les vecteurs $\mathbf{a}_j$ et $\mathbf{q}_j$, $j = 1, \ldots, n$, constituent les colonnes des matrices $\mathbf{A}$ et $\mathbf{Q}$ respectivement.

Ceci signifie que l'image des k premières colonnes de la matrice $\mathbf{A}$ est égale à l'image des k premières colonnes de la matrice $\mathbf{Q}$, ceci pour tout $1 \leq k \leq n$.

Cette structure emboitée du point de vue des espaces engendrés par les vecteurs colonnes des matrices $\mathbf{A}$ et $\mathbf{Q}$ se résume, dans la factorisation $\mathbf{QR}$, par le fait que la structure de $\mathbf{R}$ est triangulaire supérieure. En effet, les "0" dans la partie triangulaire inférieure de $\mathbf{R}$ traduisent, dans l'égalité $\mathbf{Q}^T\mathbf{A} = \mathbf{R}$, l'indépendance entre les vecteurs $\mathbf{q}_i$ et $\mathbf{a}_j$ pour tout $i > j$.

Cette propriété est très utile pour décomposer géométriquement l'espace $\mathbb{R}^n$ en sous-espaces orthogonaux, ou pour projeter un vecteur sur le sous-espace engendré par un sous-ensemble de vecteurs linéairement indépendants.

1-4.2 Application

L'intérêt de la factorisation $\mathbf{QR}$ est qu'elle permet de résoudre de manière plus aisée un système linéaire. Par exemple, dans $\mathcal{M}_n(\mathbb{R})$, on a :

$$\mathbf{A}^{-1}\mathbf{b} = (\mathbf{QR})^{-1}\,\mathbf{b} = \mathbf{R}^{-1}\mathbf{Q}^{-1}\mathbf{b} = \mathbf{R}^{-1}(\mathbf{Q}^T\mathbf{b})\,,$$

la multiplication par $\mathbf{Q}^T$ étant triviale à mettre en œuvre, et la résolution d'un système triangulaire supérieur étant facile à obtenir en partant de la variable d'indice n et par substitutions successives en remontant dans les indices.

Exemple : Soit la matrice : $\mathbf{A} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$, $\mathbf{q}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$, $\mathbf{q}_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}$, et $\mathbf{q}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Posons :

$$\mathbf{Q} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{R} = \begin{pmatrix} \mathbf{q}_1^T\mathbf{a}_1 & \mathbf{q}_1^T\mathbf{a}_2 & \mathbf{q}_1^T\mathbf{a}_3 \\ 0 & \mathbf{q}_2^T\mathbf{a}_2 & \mathbf{q}_2^T\mathbf{a}_3 \\ 0 & 0 & \mathbf{q}_3^T\mathbf{a}_3 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 3/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 2 \end{pmatrix}.$$

On vérifie bien

$$\mathbf{QR} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 3/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} = \mathbf{A}\,.$$

Il est alors facile de calculer l'inverse de **A**. On a en effet

$$\mathbf{Q}^{-1} = \mathbf{Q}^T = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}^T \quad \text{et} \quad \mathbf{R}^{-1} = \begin{pmatrix} 1/\sqrt{2} & -3/\sqrt{2} & 1/2 \\ 0 & \sqrt{2} & -1/2 \\ 0 & 0 & 1/2 \end{pmatrix}.$$

D'où

$$\begin{aligned} \mathbf{A}^{-1} = \mathbf{R}^{-1}\mathbf{Q}^{-1} &= \begin{pmatrix} 1/\sqrt{2} & -3/\sqrt{2} & 1/2 \\ 0 & \sqrt{2} & -1/2 \\ 0 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -1 & 1/2 \\ -1 & 1 & -1/2 \\ 0 & 0 & 1/2 \end{pmatrix}. \end{aligned}$$