Notes de cours Algèbre 2

Alexandre Janon

21 décembre 2010

Table des matières

1 Formes bilinéaires

1.1 Définition

Soit E un espace vectoriel sur $\mathbb{R}$.

Définition Une **forme bilinéaire** b sur E est une fonction $b : E \times E \to \mathbb{R}$ telle que

$$\begin{aligned} b(u+w,v) &= b(u,v)+b(w,v) && (1)\\ b(\lambda u,v) &= \lambda b(u,v) && (2)\\ b(u,v+w) &= b(u,v)+b(u,w) && (3)\\ b(u,\lambda v) &= \lambda b(u,v) && (4) \end{aligned}$$

quels que soient $u, v, w \in E$ et $\lambda \in \mathbb{R}$.

Autrement dit, une forme bilinéaire prend en entrée deux vecteurs, renvoie en sortie un nombre réel, et est linéaire par rapport à chacune de ses deux variables d'entrée.

1.2 Exemples et contre-exemples en dimension 2

Plaçons nous dans le cas où $E = \mathbb{R}^2$.

1. Soit :
$$b(u,v) = 2u_1v_1 + 3u_1v_2 - u_2v_1 + 4u_2v_2 \quad \forall u, v \in E$$
où (u_1, u_2) et (v_1, v_2) désignent respectivement les coordonnées de u (resp. v) dans la base canonique de $\mathbb{R}^2$.
Vérifions que c'est bien une forme bilinéaire : soient $u, v, w \in E$ et $\lambda \in \mathbb{R}$. On a :
   (a)
$$\begin{aligned} b(u+w,v) &= 2(u_1+w_1)v_1 + 3(u_1+w_1)v_2 - (u_2+w_2)v_1 + 4(u_2+w_2)v_2\\ &= 2u_1v_1 + 2w_1v_1 + 3u_1v_2 + 3w_1v_2 - u_2v_1 - w_2v_1 + 4u_2v_2 + 4w_2v_2\\ &= 2u_1v_1 + 3u_1v_2 - u_2v_1 + 4u_2v_2 + 2w_1v_1 + 3w_1v_2 - w_2v_1 + 4w_2v_2\\ &= b(u,v) + b(w,v) \end{aligned}$$
   (b)
$$\begin{aligned} b(\lambda u,v) &= 2\lambda u_1v_1 + 3\lambda u_1v_2 - \lambda u_2v_1 + 4\lambda u_2v_2\\ &= \lambda(2u_1v_1 + 3u_1v_2 - u_2v_1 + 4u_2v_2)\\ &= \lambda b(u,v) \end{aligned}$$
   (c)
$$\begin{aligned} b(u,v+w) &= 2u_1(v_1+w_1) + 3u_1(v_2+w_2) - u_2(v_1+w_1) + 4u_2(v_2+w_2)\\ &= 2u_1v_1 + 2u_1w_1 + 3u_1v_2 + 3u_1w_2 - u_2v_1 - u_2w_1 + 4u_2v_2 + 4u_2w_2\\ &= 2u_1v_1 + 3u_1v_2 - u_2v_1 + 4u_2v_2 + 2u_1w_1 + 3u_1w_2 - u_2w_1 + 4u_2w_2\\ &= b(u,v) + b(u,w) \end{aligned}$$
   (d)
$$\begin{aligned} b(u,\lambda v) &= 2u_1(\lambda v_1) + 3u_1(\lambda v_2) - u_2(\lambda v_1) + 4u_2(\lambda v_2)\\ &= \lambda(2u_1v_1 + 3u_1v_2 - u_2v_1 + 4u_2v_2)\\ &= \lambda b(u,v) \end{aligned}$$

2. De manière générale, toute expression du type :
$$b(u,v) = a_{11}u_1v_1 + a_{12}u_1v_2 + a_{21}u_2v_1 + a_{22}u_2v_2,$$
où a_{11}, a_{12}, a_{21} et a_{22} sont des réels (ne dépendant pas de u et de v) définit une forme bilinéaire sur $\mathbb{R}^2$.
En fait, toute forme bilinéaire sur $\mathbb{R}^2$ a une expression de ce type.
3. Soit
$$b(u,v) = u_1v_1 + 4 \quad \forall u, v \in E$$
b n'est pas une forme bilinéaire. En effet :
$$\begin{aligned} b(u+w,v) &= (u_1+w_1)v_1 + 4 \\ &= u_1v_1 + w_1v_1 + 4, \end{aligned}$$
mais
$$\begin{aligned} b(u,v) + b(w,v) &= u_1v_1 + 4 + w_1v_1 + 4 \\ &= u_1v_1 + w_1v_1 + 8, \end{aligned}$$
donc la propriété (1) n'est pas vérifiée.
4. Soit
$$b(u,v) = u_1u_2 + v_1v_2$$
b n'est pas une forme bilinéaire. En effet :
$$b(\lambda u, v) = \lambda^2 u_1u_2 + v_1v_2,$$
alors que :
$$\lambda b(u,v) = \lambda u_1u_2 + \lambda v_1v_2.$$
La propriété (2) n'est donc pas vérifiée.
5. Pour $E = \mathbb{R}^2$, soit
$$b(u,v) = u_1 + v_1$$
b n'est pas bilinéaire, puisque :
$$b(\lambda u, v) = \lambda u_1 + v_1,$$
alors que :
$$\lambda b(u,v) = \lambda(u_1+v_1) = \lambda u_1 + \lambda v_1.$$
La propriété (2) est donc mise en défaut.

1.3 Expression d'une forme bilinéaire dans une base.

Plaçons nous dans le cas d'un espace E de dimension finie, $E = \mathbb{R}^n$.
Nous généralisons la propriété donnée à l'exemple 2. du paragraphe précédent de la façon suivante :

Proposition Soit B une base de E. Toute forme bilinéaire b sur E peut s'écrire de la façon suivante :
$$\forall u, v \in E \qquad b(u,v) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}u_iv_j \qquad (*)$$
où les a_{ij} sont des réels, et $u_1, \ldots, u_n$ (resp. $v_1, \ldots, v_n$) sont les coordonnées de u (resp. v) dans la base B.

Cette propriété nous dit, par exemple, que toute forme bilinéaire sur $\mathbb{R}^3$ s'écrit :
$$b(u,v) \;=\; a_{11}u_1v_1 + a_{12}u_1v_2 + a_{13}u_1v_3 + a_{21}u_2v_1 + a_{22}u_2v_2 + a_{23}u_2v_3 + a_{31}u_3v_1 + a_{32}u_3v_2 + a_{33}u_3v_3$$
où les coefficients a_{ij} sont des réels.

Remarque Les coefficients a_{ij} dépendent de la base B choisie.

1.4 Matrice d'une forme bilinéaire

Définition La relation (*) de la propriété précédente peut s'écrire à l'aide de produits matriciels :

$$b(u,v) = u^T A v$$

où

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$$

La matrice A est appelée **matrice de la forme bilinéaire** b dans la base B.

Exemple.

Dans $\mathbb{R}^2$ muni de sa base canonique, soit la forme bilinéaire :

$$b(u,v) = 2u_1v_1 + 3u_1v_2 - u_2v_1$$

Ecrivons la matrice de b dans la base canonique. Pour cela on place le coefficient multipliant u_iv_j à la i-ème ligne et j-ème colonne de A :

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 0 \end{pmatrix}$$

Dans cet exemple, il n'y a pas de terme en u_2v_2 dans l'expression de $b(u,v)$, on a donc mis un zéro en ligne 2, colonne 2.

Vérifions la relation $b(u,v) = u^T Av$. On a :

$$u^T A = \begin{pmatrix} u_1 & u_2 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 2u_1 - u_2 & 3u_1 \end{pmatrix}$$

donc :

$$u^T Av = \begin{pmatrix} 2u_1 - u_2 & 3u_1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = (2u_1 - u_2)v_1 + 3u_1v_2 = 2u_1v_1 - u_2v_1 + 3u_1v_2.$$

1.5 Formes bilinéaires symétriques

Définition On dit qu'un forme bilinéaire b est **symétrique** si

$$\forall u, v \in E \qquad b(u,v) = b(v,u)$$

En pratique, pour déterminer si une forme bilinéaire est symétrique ou non, nous utilisons :

Proposition b est symétrique si et seulement si la matrice de b (dans au moins une base – si elle est symétrique dans une base, alors elle l'est dans toutes) est symétrique.

Exemple et contre-exemple

Plaçons nous dans $E = \mathbb{R}^3$, muni de la base canonique.

1. Soit :
$$b(u, v) = 2u_1v_1 + 3u_2v_2 + 4u_3v_1 + 4u_1v_3$$
Matrice de b :
$$A = \begin{pmatrix} 2 & 0 & 4 \\ 0 & 3 & 0 \\ 4 & 0 & 0 \end{pmatrix}$$
On $A^T = A$ donc A est symétrique donc b aussi.
2. Soit :
$$b(u, v) = 2u_1v_1 + 3u_2v_2 + 4u_3v_1 - u_1v_3$$
Matrice de b :
$$A = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 3 & 0 \\ 4 & 0 & 0 \end{pmatrix}$$
A n'est pas symétrique donc b non plus.

Une forme bilinéaire est donc symétrique lorsque pour chaque terme en u_iv_j, le terme en u_jv_i apparaît avec le même coefficient.

2 Formes quadratiques

2.1 Définition et exemples.

Définition Soit b une forme bilinéaire symétrique sur E. Soit q l'application $E \to \mathbb{R}$ définie par
$$\forall u \in E \qquad q(u) = b(u, u)$$
On dit que q est la **forme quadratique** associée à b.

Remarque On n'associe pas de forme quadratique à une forme bilinéaire non symétrique.

Définition On appelle **matrice d'une forme quadratique** q dans une base B la matrice de la forme bilinéaire symétrique associée (dans la base B).

Exemples dans $E = \mathbb{R}^2$

1. Soit
$$b(u, v) = 3u_1v_1 + 2u_1v_2 + 2u_2v_1 + 5u_2v_2.$$
Pour écrire q on fait $u = v$ dans l'expression de $b(u, v)$:
$$q(u) = 3u_1^2 + 2u_1u_2 + 2u_2u_1 + 5u_2^2 = 3u_1^2 + 4u_1u_2 + 5u_2^2.$$
On remarque que les termes en u_iv_i deviennent u_i^2, et que les termes en u_iv_j (pour $i < j$) sont doublés et remplacés par u_iu_j.

2. Dans le sens inverse, partant d'une forme quadratique :

$$q(u) = 3u_1^2 + 12u_1u_2 - u_2^2$$

retrouvons la forme bilinéaire symétrique b dont q est la forme quadratique associée.
Pour cela : on remplace les termes u_i^2 par u_iv_i, et on « dédouble » les termes en u_iu_j (pour $i \neq j$) :

$$b(u,v) = 3u_1v_1 + 6u_1v_2 + 6u_2v_1 - u_2v_2$$

La matrice de q est donc :

$$\begin{pmatrix} 3 & 6 \\ 6 & -1 \end{pmatrix}$$

2.2 Définie-positivité

Définition On dit qu'une *forme quadratique* q est **définie positive** si $q(u) > 0$ quelque soit $u \in E$, $u \neq 0$.

Définition On dit qu'une *matrice symétrique* est **définie positive** si c'est la matrice (dans une certaine base) d'une forme quadratique définie positive.

Exemples et contre-exemples dans $\mathbb{R}^2$.

1. Soit

$$q(u) = u_1^2 + u_2^2$$

Le carré d'un réel étant toujours positif ou nul, $q(u) \geq 0$ quelque soit $u \in E$. De plus, si $u \neq 0$ alors au moins une des deux coordonnées de u est non nulle, et alors $q(u)$ est non nul (donc > 0).

2. Soit

$$q(u) = -u_1^2 - u_2^2$$

q n'est pas définie positive : par exemple $q(1,0) = -1 < 0$.

3. Soit

$$q(u) = u_1^2 - u_2^2$$

q n'est pas définie positive : par exemple $q(1,2) = 1 - 4 = -3 < 0$.

4. Soit

$$q(u) = u_1u_2$$

q n'est pas définie positive : par exemple $q(1,-1) = -1 < 0$.

L'étude de la définie-positivité d'une forme quadratique donnée n'est pas toujours si « simple » que ci-dessus. Par exemple, en dimension 2, la forme

$$q(u) = 2u_1^2 - 2u_1u_2 + u_2^2$$

est-elle définie positive ?

Pour répondre à cette question, nous allons voir un critère « calculatoire » qui permet de décider si une forme quadratique est positive ou non. C'est ce critère que nous utiliserons en pratique.

2.3 Critère pratique de définie-positivité

Proposition Une forme quadratique est définie positive si et seulement si les valeurs propres de sa matrice (dans une base quelconque) sont toutes strictement positives.

Exemple d'utilisation.

Répondons donc à la question posée au-dessus : la matrice de $q(u) = 2u_1^2 - 2u_1u_2 + u_2^2$ est

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Son polynôme caractéristique est :

$$\chi_A(x) = \det(A - xI_2) = \begin{vmatrix} 2-x & -1 \\ -1 & 1-x \end{vmatrix} = (2-x)(1-x) - 1 = x^2 - 3x + 1$$

dont les racines sont $\dfrac{3+\sqrt{5}}{2}$ et $\dfrac{3-\sqrt{5}}{2}$, toutes deux strictement positives, donc q est définie positive.
Au passage, remarquons que A est définie positive (puisque c'est la matrice d'une forme quadratique définie positive), malgré qu'elle possède des coefficients négatifs. En fait il n'y a aucun rapport (c'est-à-dire aucune implication, ni directe ni réciproque) entre être définie positive et avoir des coefficients positifs !

2.4 Produits scalaires

Définition Un **produit scalaire** sur E est une forme bilinéaire symétrique b, dont la forme quadratique associée est définie positive. On dit aussi que b **définit un produit scalaire**.

Dans ce cas on note généralement :

$$< u, v >= b(u, v)$$

Exemple : le produit scalaire usuel dans $\mathbb{R}^n$

On a défini le produit scalaire *usuel* de deux vecteurs de $\mathbb{R}^n$:

$$u \cdot v = u_1v_1 + u_2v_2 + \ldots + u_nv_n$$

(où les coordonnées $u_1, \ldots, u_n, v_1, \ldots, v_n$ sont prises relativement à la base canonique de $\mathbb{R}^n$).
C'est bien une forme bilinéaire, dont la matrice dans la base canonique est la matrice identité I_n.
La forme quadratique associée est :

$$q(u) = u_1^2 + u_2^2 + \ldots + u_n^2,$$

qui est bien définie positive (une somme de carrés est toujours ≥ 0 et si $u \neq 0$ alors au moins des u_i est non nul et $q(u)$ est alors non nul). Le produit scalaire usuel est donc bien un produit scalaire !

Autre exemple.

Il y a d'autres produits scalaires que le produit scalaire usuel. Par exemple, au paragraphe précédent, nous avons vu que, dans $\mathbb{R}^2$, la forme quadratique

$$q(u) = 2u_1^2 - 2u_1u_2 + u_2^2$$

est définie positive.
La forme bilinéaire symétrique associée définit donc un produit scalaire sur $\mathbb{R}^2$:

$$< u, v >= 2u_1v_1 - u_1v_2 - u_2v_1 + u_2v_2$$

Dans la suite, nous notons $<,>$ un produit scalaire « quelconque », et réservons la notation $u \cdot v$ au produit scalaire usuel.

2.5 Normes

Définition Soit $<,>$ un produit scalaire sur E. On appelle **norme associée à ce produit scalaire** l'application $||.|| : E \to \mathbb{R}$ définie par :

$$\forall u \in E \qquad ||u|| = \sqrt{< u, u >}.$$

Remarque Le fait que $<,>$ soit défini positif assure le fait que la racine carrée ci-dessus a toujours un sens.

Exemples.

1. La norme associée au produit scalaire usuel est :

$$||u|| = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

   C'est la norme euclidienne usuelle.

2. Pour un produit scalaire autre que l'usuel, on obtient une norme différente de la norme usuelle ! L'intérêt de faire cela est de changer la manière dont on mesure les distances dans E.
   En effet, une norme induit une distance entre deux vecteurs u et v de E par :

$$\mathrm{dist}(u, v) = ||u - v||.$$

   Avec la norme usuelle, on obtient la distance « géométrique », celle mesurée avec une règle graduée (au moins en dimension inférieure à 3...). Dans de nombreuses applications (par exemple en statistique... ou en physique relativiste !) cette distance usuelle ne reflète pas bien les propriétés de la situation à modéliser, d'où la nécessité de considérer d'autres distances.
   Soit par exemple un relevé statistique issu des notes obtenus par des étudiants pour l'obtention d'un diplôme ne comportant (pour simplifier) que deux épreuves, la première affectée du coefficient 10, l'autre du coefficient 1. Pour chaque étudiant, ces notes sont rangées dans un vecteur de $\mathbb{R}^2$:

$$\begin{pmatrix} \mathrm{note}_{\mathrm{epreuve1}} \\ \mathrm{note}_{\mathrm{epreuve2}} \end{pmatrix}$$

   Si l'on mesure la différence entre deux étudiants A et B par la distance géométrique classique :

$$\sqrt{(\mathrm{note}_{\mathrm{epreuve1}}(\mathrm{A}) - \mathrm{note}_{\mathrm{epreuve1}}(\mathrm{B}))^2 + (\mathrm{note}_{\mathrm{epreuve2}}(\mathrm{A}) - \mathrm{note}_{\mathrm{epreuve2}}(\mathrm{B}))^2},$$

   on ne tiendra pas compte du fait que la première épreuve compte dix fois plus que la seconde. Il serait plus judicieux de mesurer cette différence par une autre distance donnant plus de poids à une différence sur la première épreuve :

$$\sqrt{10\,(\mathrm{note}_{\mathrm{epreuve1}}(\mathrm{A}) - \mathrm{note}_{\mathrm{epreuve1}}(\mathrm{B}))^2 + (\mathrm{note}_{\mathrm{epreuve2}}(\mathrm{A}) - \mathrm{note}_{\mathrm{epreuve2}}(\mathrm{B}))^2},$$

   Cette distance est associée au produit scalaire :

$$< u, v >= 10u_1v_1 + u_2v_2.$$

Proposition

1. Si $||u|| = 0$ alors $u = 0$.
2. Quelque soit $u \in E$, $\lambda \in \mathbb{R}$, on a : $||\lambda u|| = |\lambda|\,||u||$.
3. Inégalité triangulaire : quelque soient $u, v \in E$, $||u + v|| \leq ||u|| + ||v||$.
4. Inégalité de Cauchy-Schwarz : quelque soient $u, v \in E$, $| < u, v > | \leq ||u||||v||$.
5. $||u + v||^2 = ||u||^2 + ||v||^2 + 2 < u, v >$

Démonstrations.

1. Soit u tel que $||u|| = 0$, et, par conséquent : $||u||^2 =< u,u >= q(u) = 0$. Si $u \neq 0$ on a nécessairement $q(u) > 0$ (car q est définie positive), contradiction. Donc $u = 0$.
2. $||\lambda u|| = \sqrt{< \lambda u, \lambda u >} = \sqrt{\lambda^2 < u,u >} = |\lambda|\sqrt{< u,u >} = |\lambda|||u||$.
3. Admise.
4. Admise.
5.

$$\begin{aligned}||u+v||^2 &=< u+v, u+v > \\ &=< u, u+v > + < v, u+v > \\ &=< u,u > + < u,v > + < v,u > + < v,v > \\ &= ||u||^2 + 2 < u,v > + ||v||^2\end{aligned}$$

Remarque. On appelle en général **norme** toute application $E \to \mathbb{R}$ qui vérifie les propriétés 1., 2. et 3..
Il existe des normes qui ne sont pas associées à un produit scalaire, par exemple en dimension 2 :

$$||u||_1 = |u_1| + |u_2|$$

ou :

$$||u||_\infty = \max\left(|u_1|, |u_2|\right).$$

Pour de telles normes les propriétés 4. et 5. ne sont pas vérifiées et n'auraient d'ailleurs pas de sens (que choisir pour $< u,v >$?)

3 Orthogonalité

Dans cette section, on se place dans un espace vectoriel E de dimension finie, muni d'un produit scalaire $<,>$ et de la norme associée $||.||$.

3.1 Vecteurs orthogonaux

Définition Deux vecteurs u et v sont dits **orthogonaux** si $< u,v >= 0$.

La notion d'orthogonalité dépend du produit scalaire considéré. Lorsque le produit scalaire est le produit scalaire usuel, on retrouve l'orthogonalité « géométrique ». Pour un autre produit scalaire (et donc une autre norme) la notion d'orthogonalité n'est pas forcément la même : on remarque donc que changer la manière dont on mesure les distances change la notion d'orthogonalité [1] !

3.2 Bases orthogonales, orthonormales

Définition On dit qu'une base $(f_1, \dots, f_n)$ de E est **orthogonale** si les vecteurs la composant sont orthogonaux deux à deux, c'est-à-dire si

$$< f_i, f_j >= 0$$

dès que $i \neq j$.

1. Cette remarque n'est pas forcément très surprenante si l'on pense au théorème de Pythagore (et sa réciproque), qui relie orthogonalité et longueurs.

Définition On dit que cette base est **orthonormale** (ou **orthonormée**) si elle est orthogonale, et que chaque vecteur la composant a pour norme 1.

Exemples.

1. En dimension 2 (mais cela est vrai en toute dimension), pour le produit scalaire usuel, la base canonique
$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$
est orthonormée.
2. La base canonique n'est pas la seule base orthonormée pour le produit scalaire usuel. Par exemple, soit (f_1, f_2) la base composée de
$$f_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad f_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
est orthogonale, puisque $f_1 \cdot f_2 = 1 - 1 = 0$.
Cependant $||f_1|| = \sqrt{1+1} = \sqrt{2}$ et $||f_2|| = \sqrt{2}$ donc (f_1, f_2) n'est pas orthonormée. Pour passer d'une base orthogonale à une base orthonormée, il suffit de diviser chaque vecteur par sa norme, donc
$$\left\{ \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \right\}$$
est une base orthonormée de $\mathbb{R}^2$ pour le produit scalaire usuel.

3.3 Orthogonal d'un sous-espace

Définition Soit F un sous-espace vectoriel de E. On note $F^\perp$, l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de F. $F^\perp$ est appelé l'**orthogonal** de F dans E.

Proposition $F^\perp$ est un sous-espace vectoriel de E, et
$$\dim F^\perp = \dim E - \dim F.$$

Pour déterminer de manière pratique l'orthogonal d'un sous-espace, nous utilisons la proposition suivante.

Proposition Soit $\{e_1, \ldots, e_k\}$ une base de F.
Un vecteur u appartient à $F^\perp$ si et seulement si il est orthogonal à tous les e_i, c'est-à-dire si :
$$\left\{ \begin{array}{rcl} < u, e_1 > & = & 0 \\ < u, e_2 > & = & 0 \\ \vdots & & \vdots \\ < u, e_k > & = & 0 \end{array} \right.$$

Exemple 1.

On prend $E = \mathbb{R}^3$, muni du produit scalaire usuel. Soit F le sous-espace vectoriel de E engendré par :

$$\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

Les vecteurs e_1 et e_2 sont linéairement indépendants (exercice), ils forment donc une base de F.
Soit le vecteur

$$u = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$$

On a :

$$u \cdot e_1 = 0 \times 1 + 1 \times 0 + (-2) \times 0 = 0 \qquad u \cdot e_2 = 0 \times 1 + 1 \times 2 + (-2) \times 1 = 0$$

Donc $u \in F^{\perp}$.
Comme $\dim F^{\perp} = \dim E - \dim F = 3 - 2 = 1$ et que ce vecteur u est non nul, la famille $\{u\}$ est libre et engendre $F^{\perp}$, et forme donc une base de $F^{\perp}$.

Comment aurait-on pu « deviner » un vecteur $u \in F^{\perp}$ si celui-ci ne nous avait pas été fourni ? Une façon de faire est d'écrire les coordonnées inconnues du vecteur u :

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix},$$

puis d'utiliser :

$$u \in F^{\perp} \Leftrightarrow \left\{ \begin{array}{l} u \cdot e_1 = 0 \\ u \cdot e_2 = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} u_1 = 0 \\ u_1 + 2u_2 + u_3 = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} u_1 = 0 \\ u_3 = -2u_2 \end{array} \right. \Leftrightarrow u = \begin{pmatrix} 0 \\ t \\ -2t \end{pmatrix}, t \in \mathbb{R}.$$

On trouve alors un vecteur de $F^{\perp}$ en donnant une valeur à t. Par exemple, en prenant $t = 1$ on retrouve le u donné au début.

Exemple 2.

Prenons toujours $E = \mathbb{R}^3$, muni du produit scalaire usuel, et cherchons une base de l'orthogonal du sous-espace F engendré par le vecteur

$$e_1 = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}.$$

En notant u le vecteur de coordonnées inconnues u_1, u_2, u_3, on a :

$$u \in F^{\perp} \Leftrightarrow \qquad u \cdot e_1 = 0 \Leftrightarrow \qquad 2u_1 + 3u_2 - u_3 = 0 \Leftrightarrow \qquad u = \begin{pmatrix} t \\ t' \\ 2t + 3t' \end{pmatrix}, t, t' \in \mathbb{R}.$$

Pour trouver une base de $F^{\perp}$, il faut donc $\dim E - \dim F = 3 - 1 = 2$ vecteurs linéairement indépendants de la forme ci-dessus, obtenus en donnant des valeurs à t et t'.
En choisissant $(t, t') = (1, 0)$ puis $(t, t') = (0, 1)$, on trouve une famille

$$\mathcal{G} = \left\{ e'_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, e'_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \right\}$$

de vecteurs de $F^{\perp}$.

Il est important de vérifier[2] que cette famille est bien linéairement indépendante : en effet, soient $\lambda_1, \lambda_2 \in \mathbb{R}$ tels que

$$\lambda_1 e'_1 + \lambda_2 e'_2 = 0$$

alors :

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 2\lambda_1 + 3\lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

soit $\lambda_1 = \lambda_2 = 0$.
Donc la famille $\mathcal{G}$ est bien une base de $F^{\perp}$.

3.4 Orthogonalisation de Gram-Schmidt

Partant d'une base quelconque $\{e_1, e_2, \ldots, e_k\}$ d'un sous-espace vectoriel F de E, on veut obtenir une base **orthogonale** $\{f_1, f_2, \ldots, f_k\}$ du même sous-espace F.

Nous procédons par étapes : à chaque étape nous construisons un vecteur de la « nouvelle base » $\{f_1, f_2, \ldots, f_k\}$, en fonction de ceux précédemment construits et d'un vecteur de la base de départ $\{e_1, e_2, \ldots, e_k\}$.
Etape 1. On pose

$$f_1 = e_1.$$

Etape 2. On cherche f_2 sous la forme :

$$f_2 = e_2 + \lambda_1 f_1,$$

où $\lambda_1 \in \mathbb{R}$ est un scalaire que nous allons déterminer.
Pour que notre nouvelle base $\{f_1, f_2, \ldots, f_k\}$ soit une base orthogonale, il faut en particulier que f_1 et f_2 soient orthogonaux, c'est-à-dire que

$$< f_1, f_2 >= 0.$$

En remplaçant f_2 par son expression en fonction de λ_1, on doit donc avoir :

$$< f_1, e_2 + \lambda_1 f_1 >= 0$$

Soit par bilinéarité :

$$< f_1, e_2 > + \lambda_1 < f_1, f_1 >= 0$$

D'où :

$$\lambda_1 = -\frac{< f_1, e_2 >}{< f_1, f_1 >}.$$

On pose donc :

$$f_2 = e_2 - \frac{< f_1, e_2 >}{< f_1, f_1 >} f_1.$$

Etape 3. On cherche f_3 sous la forme :

$$f_3 = e_3 + \lambda_1 f_1 + \lambda_2 f_2$$

où $\lambda_1, \lambda_2 \in \mathbb{R}$ sont des scalaires que nous devons déterminer, en écrivant que l'on doit avoir :

$$\begin{cases} < f_3, f_1 >= 0 \\ < f_3, f_2 >= 0 \end{cases}$$

soit :

$$\begin{cases} < e_3, f_1 > + \lambda_1 < f_1, f_1 > + \lambda_2 < f_2, f_1 >= 0 \\ < e_3, f_2 > + \lambda_1 < f_1, f_2 > + \lambda_2 < f_2, f_2 >= 0 \end{cases}$$

d'où, car $< f_1, f_2 >=< f_2, f_1 >= 0$:

$$\begin{cases} < e_3, f_1 > + \lambda_1 < f_1, f_1 >= 0 \\ < e_3, f_2 > + \lambda_2 < f_2, f_2 >= 0 \end{cases}$$

2. de mauvais choix pour t et t' (p. ex. $(t, t') = (1, 1)$ et $(t, t') = (2, 2)$) peuvent conduire à une famille liée

Soit finalement :

$$\begin{cases} \lambda_1 = -\dfrac{< e_3, f_1 >}{< f_1, f_1 >} \\ \lambda_2 = -\dfrac{< e_3, f_2 >}{< f_2, f_2 >} \end{cases}$$

On pose donc :

$$f_3 = e_3 - \frac{< f_1, e_3 >}{< f_1, f_1 >} f_1 - \frac{< f_2, e_3 >}{< f_2, f_2 >} f_2.$$

Etape i, pour $i \leq k$. On continue ainsi : à l'étape i on pose

$$f_i = e_i - \frac{< f_1, e_i >}{< f_1, f_1 >} f_1 - \ldots - \frac{< f_{i-1}, e_i >}{< f_{i-1}, f_{i-1} >} f_{i-1}.$$

Et l'on s'arrête lorsque l'on a construit le dernier vecteur f_k.

En procédant ainsi, on obtient *toujours* une base $\{f_1, \ldots, f_k\}$ de F orthogonale (il est donc inutile – en théorie – de le vérifier, cependant il est conseillé, afin de détecter les erreurs de calcul, que les produits scalaires deux à deux des vecteurs obtenus sont nuls).

Si on souhaite avoir une famille orthonormale $\{g_1, \ldots, g_k\}$ de F, on construit une base $\{f_1, \ldots, f_k\}$ orthogonale et on pose :

$$g_1 = \frac{1}{||f_1||} f_1 \quad ; \quad g_2 = \frac{1}{||f_2||} f_2 \quad ; \quad \ldots \quad ; \quad g_k = \frac{1}{||f_k||} f_k \quad ;$$

qui est toujours une base orthonormale de F.

Exemple.

Dans $\mathbb{R}^4$, muni du produit scalaire usuel, on considère le sous-espace vectoriel F de base :

$$\left\{ e_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} ; e_2 = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 1 \end{pmatrix} ; e_3 = \begin{pmatrix} 1 \\ 10 \\ 0 \\ -7 \end{pmatrix} \right\}$$

Nous souhaitons, à partir de cette base, construire une base de F orthonormée, en appliquant le procédé de Gram-Schmidt.
Dans la base de F il y a $k = 3$ vecteurs, nous allons donc construire une base $\{f_1, f_2, f_3\}$ orthogonale en 3 étapes.
Etape 1.

$$f_1 = e_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix}.$$

Etape 2.

$$f_2 = e_2 - \frac{< e_2, f_1 >}{< f_1, f_1 >} f_1$$

Or :

$$< e_2, f_1 > = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 1 \\ 1 \end{pmatrix} = 1 + 2 + 3 + 1 = 7$$

$$< f_1, f_1 > = 1 + 4 + 1 + 1 = 7$$

Donc :

$$f_2 = e_2 - f_1 = \begin{pmatrix} 0 \\ -1 \\ 2 \\ 0 \end{pmatrix}$$

Etape 3.

$$f_3 = e_3 - \frac{< e_3, f_1 >}{< f_1, f_1 >} f_1 - \frac{< e_3, f_2 >}{< f_2, f_2 >} f_2$$

Or :

$$< e_3, f_1 >= 1 + 20 - 7 = 14$$

$$< f_1, f_1 >= 7 \text{ (déjà calculé)}$$

$$< e_3, f_2 >= 0 - 10 + 0 + 0 = -10$$

$$< f_2, f_2 >= 1 + 4 = 5$$

Donc :

$$f_3 = e_3 - \frac{14}{7} f_1 + \frac{10}{5} f_2 = e_3 - 2f_1 + 2f_2 = \begin{pmatrix} 1-2+0 \\ 10-4-2 \\ 0-2+4 \\ -7-2+0 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ 2 \\ -9 \end{pmatrix}$$

Une base orthogonale de F est donc formée par $\{f_1, f_2, f_3\}$.
Pour avoir une base orthonormée, on calcule les normes de chacun de ces vecteurs :

$$||f_1|| = \sqrt{< f_1, f_1 >} = \sqrt{7} \quad ; \quad ||f_2|| = \sqrt{< f_2, f_2 >} = \sqrt{5} \quad ; \quad ||f_3|| = \sqrt{102}$$

et on a donc que :

$$\left\{ \frac{1}{\sqrt{7}} f_1 ; \frac{1}{\sqrt{5}} f_2 ; \frac{1}{\sqrt{102}} f_3 \right\}$$

est une base orthonormée de F.

Remarque Dans le procédé de Gram-Schmidt, on obtient à la fin une base de F **à condition de partir d'une base de** F.
Si la famille de départ $\{e_1, \dots, e_k\}$ est liée, on peut tout de même mener le procédé d'orthogonalisation de Gram-Schmidt, et l'on obtiendra en sortie une famille $\{f_1, \dots, f_k\}$ comportant un ou plusieurs vecteurs nuls. On peut démontrer qu'en retirant ces vecteurs nuls de la famille obtenue on obtient bien une base orthogonale de F en partant d'un système seulement générateur de F. En pratique, il est donc inutile de vérifier au préalable si l'on part d'une famille libre, quitte à retirer les éventuels vecteurs nuls de la base construite.

3.5 Projection orthogonale

Définition Soit F un sous-espace vectoriel de E, de base $\{f_1, \dots, f_k\}$ *orthogonale*. Pour $u \in E$, on note $P_F(u)$ le vecteur :

$$P_F(u) = \frac{< f_1, u >}{||f_1||^2} f_1 + \frac{< f_2, u >}{||f_2||^2} f_2 + \ldots + \frac{< f_k, u >}{||f_k||^2} f_k,$$

appelé **projeté orthogonal** de u sur F.

On définit ainsi une application linéaire $P_F : E \to E$, appelée **projection orthogonale** sur F.

Proposition P_F est une application linéaire, à valeurs dans F.

Exemple de calcul de projection orthogonale.

Nous poursuivons l'exemple de la partie précédente (Gram-Schmidt).

On veut calculer le projeté orthogonal sur F du vecteur

$$u = \begin{pmatrix} -52 \\ 24 \\ 7 \\ 18 \end{pmatrix}.$$

On a :

$$P_F(u) = \frac{< u, f_1 >}{< f_1, f_1 >} f_1 + \frac{< u, f_2 >}{< f_2, f_2 >} f_2 + \frac{< u, f_3 >}{< f_3, f_3 >} f_3$$

Or :

$$< f_1, f_1 >= 7 \qquad < f_2, f_2 >= 5 \qquad < f_3, f_3 >= 102$$

et :

$< u, f_1 >= -52+48+7+18 = 21$ $\qquad$ $< u, f_2 >= -24+14 = -10$ $\qquad$ $< u, f_3 >= 52+96+14-162 = 0$

donc :

$$P_F(u) = \frac{21}{7} f_1 + \frac{-10}{5} f_2 + 0 f_3 = 3f_1 - 2f_2 = \begin{pmatrix} 3 \\ 6+2 \\ 3-4 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 8 \\ -1 \\ 3 \end{pmatrix}$$

3.5.1 Matrice d'une projection orthogonale.

Proposition Soit B une base de E orthonormée pour le produit scalare considéré (par exemple, la base canonique si l'on travaille avec le produit scalaire usuel).
Soit A la matrice dont les colonnes sont les coordonnées (exprimées dans la base B) des vecteurs d'une *base orthonormée* de F.
Alors la matrice M de l'application linéaire P_F dans la base B est :

$$M = AA^T.$$

M est toujours une matrice symétrique.

Exemple. On reprend l'exemple précédent. Comme nous travaillons avec le produit scalaire usuel, nous nous plaçons dans la base canonique de $\mathbb{R}^4$. Une matrice A qui convient est la suivante :

$$A = \begin{pmatrix} 1/\sqrt{7} & 0 & -1/\sqrt{102} \\ 2/\sqrt{7} & -1/\sqrt{5} & 4/\sqrt{102} \\ 1/\sqrt{7} & 2/\sqrt{5} & 2/\sqrt{102} \\ 1/\sqrt{7} & 0 & -9/\sqrt{102} \end{pmatrix}$$

Ce qui donne après quelques calculs fastidieux (!) :

$$M = AA^T = \frac{1}{3570} \begin{pmatrix} 545 & 880 & 440 & 825 \\ 880 & 3314 & -128 & -240 \\ 440 & -128 & 3506 & -120 \\ 825 & -240 & -120 & 3345 \end{pmatrix}$$

Remarque : Lors du calcul de ce produit, on peut ne calculer que les termes au-dessus de la diagonale (diagonale comprise), puis compléter par symétrie (la matrice d'une projection orthogonale est toujours symétrique).
Et on peut vérifier le résultat obtenu précédemment :

$$M \begin{pmatrix} -52 \\ 24 \\ 7 \\ 18 \end{pmatrix} = \begin{pmatrix} 3 \\ 8 \\ -1 \\ 3 \end{pmatrix}.$$

3.5.2 Diagonalisation d'une matrice de projection orthogonale.

Proposition Quelque soit $u \in E$, on a :

$$P_F(u) = u \Leftrightarrow u \in F$$

$$P_F(u) = 0 \Leftrightarrow u \in F^{\perp}$$

Corollaires :

1. Une matrice de projection orthogonale (sur un espace différent de E ou de $\{0\}$) a deux valeurs propres : 0 et 1.
2. Pour la matrice associée à P_F, l'espace propre associé à 1 est F, celui associé à 0 est $F^{\perp}$.
3. Une matrice de projection orthogonale est diagonalisable dans une base orthonormée, formée en regroupant une base orthonormée de F et une base orthonormée de $F^{\perp}$.

Exemple. La matrice M écrite plus haut vérifie :

$$M = PDP^{-1}$$

avec

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

et

$$P = \begin{pmatrix} 1/\sqrt{7} & 0 & -1/\sqrt{102} & 136/\sqrt{18697} \\ 2/\sqrt{7} & -1/\sqrt{5} & 4/\sqrt{102} & 2/\sqrt{18697} \\ 1/\sqrt{7} & 2/\sqrt{5} & 2/\sqrt{102} & 1/\sqrt{18697} \\ 1/\sqrt{7} & 0 & -9/\sqrt{102} & -14/\sqrt{18697} \end{pmatrix}$$

où la dernière colonne a été trouvée comme suit : on cherche un vecteur u de norme 1, orthogonal à $\{f_1, f_2, f_3\}$ (de façon à ce que $\{u\}$ forme une base orthonormée de $F^{\perp}$), donc un vecteur $u = (u_1, u_2, u_3, u_4)^T$, de norme 1, solution du système :

$$\begin{cases} u_1 + 2u_2 + u_3 + u_4 = 0 \\ -u_2 + 2u_3 = 0 \\ -u_1 + 4u_2 + 2u_3 - 9u_4 = 0 \end{cases}$$

on trouve un vecteur u' solution de ce système : $u' = (-55, 16, 8, 15)^T$, et on prend :

$$u = \frac{1}{||u'||}u' = \frac{1}{\sqrt{3570}} \begin{pmatrix} -55 \\ 16 \\ 8 \\ 15 \end{pmatrix}$$

3.5.3 Application : base de l'intersection de deux sous-espaces vectoriels.

Soient F et G deux sous-espaces vectoriels de E. On désire trouver une base de l'intersection $F \cap G$. Pour cela nous utilisons le fait suivant, que nous admettrons :

$$u \in F \cap G \Leftrightarrow p_F(p_G(u)) = u$$

où p_F et p_G désignent respectivement la projection orthogonale sur F et sur G.
En pratique, on dispose de bases de F et de G. Après orthonormalisation de ces bases, on peut écrire la matrice M de la projection orthogonale sur F et N la matrice de la projection orthogonale sur G.

On a alors, par le fait écrit plus haut :

$$u \in F \cap G \Leftrightarrow MNu = u$$

Ainsi, en résolvant l'équation $MNu = u$ (c'est à dire en trouvant l'espace propre associé à la valeur propre 1 de MN), on pourra déterminer une base de $F \cap G$.

Exemple. Trouver une base de $F \cap G$, où

$$F = \text{Vect}\left\{\begin{pmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\} \quad G = \text{Vect}\left\{\begin{pmatrix}1/\sqrt{3}\\1/\sqrt{3}\\-1/\sqrt{3}\end{pmatrix}, \begin{pmatrix}0\\1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}\right\}$$

On vérifie que les bases données de F et G sont orthonormales (pour le produit scalaire usuel).
On pose :

$$A = \begin{pmatrix}1/\sqrt{2} & 0\\0 & 1\\1/\sqrt{2} & 0\end{pmatrix} \quad B = \begin{pmatrix}1/\sqrt{3} & 0\\1/\sqrt{3} & 1/\sqrt{2}\\-1/\sqrt{3} & 1/\sqrt{2}\end{pmatrix}$$

d'où les matrices, respectivement de p_F et p_G :

$$M = \frac{1}{2}\begin{pmatrix}1 & 0 & 1\\0 & 2 & 0\\1 & 0 & 1\end{pmatrix} \quad N = \frac{1}{6}\begin{pmatrix}2 & 2 & -2\\2 & 5 & 1\\-2 & 1 & 5\end{pmatrix}$$

d'où :

$$MN = \frac{1}{12}\begin{pmatrix}0 & 3 & 3\\4 & 10 & 2\\0 & 3 & 3\end{pmatrix}$$

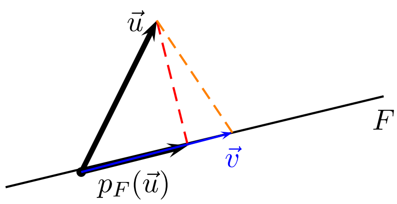
Et ainsi :

$$\begin{aligned} u \in F \cap G &\Leftrightarrow MNu = u \\ &\Leftrightarrow \begin{cases} 3u_2 + 3u_3 = 12u_1 \\ 4u_1 + 10u_2 + 2u_3 = 12u_2 \\ 3u_2 + 3u_3 = 12u_3 \end{cases} \\ &\Leftrightarrow \begin{cases} u_2 = 3u_3 \\ u_1 = u_3 \\ u_3 \in \mathbb{R} \end{cases} \end{aligned}$$

D'où $F \cap G = \text{Vect}\left\{\begin{pmatrix}1\\3\\1\end{pmatrix}\right\}$.

3.5.4 Propriété d'optimalité du projeté orthogonal.

Pour tout vecteur $u \in E$, on démontre que la quantité $||u - v||$, où $v \in F$, est minimale lorsque v est le projeté orthogonal de u sur F.
Autrement dit, le projeté orthogonal de u sur F est, parmi tous les vecteurs de F, le plus proche de u (au sens de la norme que l'on a choisie).



Le vecteur $p_F(\vec{u})$ est le vecteur de F le plus proche de $\vec{u}$. Tout autre vecteur de F (par exemple $\vec{v}$) est situé à une distance plus grande de $\vec{u}$: ici le segment pointillé orange est plus grand que le rouge.

Ce fait est d'une grande importance en pratique. Par exemple, supposons que E soit un espace de « très grande » dimension n. La manipulation sur ordinateur d'un vecteur $u \in E$ (qui peut par exemple représenter un relevé statistique) nécessite le stockage et le traitement des n coordonnées de ce vecteur. Pour des raisons pratiques (mémoire disponible, temps de calcul, ...) on peut vouloir « approcher » u par un vecteur v d'un sous-espace F qui sera de dimension k beaucoup plus petite que n, ce qui permettra de travailler avec des vecteurs moins « grands », et ainsi économiser en coût de stockage ou gagner en temps de calcul. Supposons que la perte d'information liée au remplacement de u par v soit quantifiée par $||u-v||$; nous souhaitons bien entendu minimiser cette perte d'information. Le choix optimal à faire est donc de prendre pour v le projeté orthogonal de u sur F. La difficulté est alors ramenée dans le choix du sous-espace F, et de la norme $||\cdot||$; ce choix dépend du problème que l'on veut résoudre et des compromis que l'on est prêt à faire.

4 Diagonalisation des matrices symétriques

Dans toute cette section, l'espace E est muni du produit scalaire **usuel**, et de la norme usuelle associée à ce produit scalaire.

4.1 Généralités

4.1.1 Matrices orthogonales

Définition Une matrice est dite **orthogonale** si ses vecteurs colonnes forment une base orthonormée (pour le produit scalaire usuel) de E.

Proposition Une matrice P est orthogonale si et seulement si $P^T P = I_n$.
Une matrice P est orthogonale si et seulement si $PP^T = I_n$.
Une matrice P est orthogonale si et seulement si P est inversible, d'inverse P^T.

Exemple. En pratique, pour savoir si une matrice est orthogonale, on calcule le produit de la matrice avec sa transposée (peu importe l'ordre dans lequel on fait ce produit) et on vérifie si l'on obtient la matrice identité. Ainsi, la matrice

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2/3 & -\sqrt{5}/3 \\ 0 & \sqrt{5}/3 & -2/3 \end{pmatrix}$$

est orthogonale (exercice).

4.1.2 Diagonalisation d'une matrice symétrique en base orthonormée

Proposition Toute matrice symétrique est diagonalisable dans une base orthonormée formée de vecteurs propres.

Autrement dit, si une matrice A est symétrique, elle est diagonalisable, donc on peut trouver une matrice diagonale D et une matrice P, dont les colonnes représentent une base de E, formée de vecteurs propres de A, telle que :

$$A = PDP^{-1}$$

ou de manière équivalente :

$$D = P^{-1}AP$$

De plus, la propriété dit que l'on peut choisir une base orthonormée de vecteur propres, ce qui signifie que l'on peut choisir P orthogonale (puisqu'alors les colonnes de P forment une base de E orthonormée). On a donc $P^{-1} = P^T$, et on peut donc écrire :

$$A = PDP^T$$

ou :

$$D = P^T AP$$

En pratique pour trouver P :

1. On recherche les valeurs propres de A, en déterminant les racines de son polynôme caractéristique.
2. On recherche une base de l'espace propre associé à chaque valeur propre.
3. On orthonormalise, pour le produit scalaire usuel, chacune des bases de chacun des espaces propres. On peut démontrer qu'en prenant la réunion de ces différentes bases orthonormées, on obtient une base orthonormée de E.
4. La matrice P recherchée a alors ses colonnes formées des vecteurs de cette base orthonormée, c'est bien une matrice orthogonale. La matrice D contient sur sa diagonale, rangées dans le même ordre que les vecteurs propres associés dans P, les valeurs propres de A, chaque valeur propre étant répétée autant de fois que la dimension du sous-espace propre associé.

Exemple. Diagonalisons

$$A = \begin{pmatrix} 3/2 & 0 & 3/2 \\ 0 & 3 & 0 \\ 3/2 & 0 & 3/2 \end{pmatrix}$$

en base orthonormée.

Valeurs propres de A. Le polynôme caractéristique de A s'écrit :

$$\begin{aligned}
\chi_A(x) = \det(A - xI_2) &= \begin{vmatrix} 3/2 - x & 0 & 3/2 \\ 0 & 3 - x & 0 \\ 3/2 & 0 & 3/2 - x \end{vmatrix} \\
&= (3/2 - x) \begin{vmatrix} 3 - x & 0 \\ 0 & 3/2 - x \end{vmatrix} + 3/2 \begin{vmatrix} 0 & 3/2 \\ 3 - x & 0 \end{vmatrix} \\
&= (3/2 - x)(3 - x)(3/2 - x) + 3/2(-3/2(3 - x)) \\
&= (3 - x)\left[(3/2 - x)^2 - (3/2)^2\right] \\
&= (3 - x)(-x)(3 - x)
\end{aligned}$$

Donc les valeurs propres de A sont 0 et 3.

Espace propre associé à 0. Soit $u = (x, y, z)$ un vecteur propre de A associé à 0 ; il vérifie $Au = 0$ soit :

$$\begin{cases} (3/2)x + (3/2)z = 0 \\ 3y = 0 \\ (3/2)x + (3/2)z = 0 \end{cases}$$

Ce système est équivalent à

$$\begin{cases} x = -z \\ y = 0 \end{cases}$$

Donc une base de l'espace propre de A associé à 0 est : $\{(1, 0, -1)\}$.

Il n'y a qu'un seul vecteur dans cette base, donc elle est orthogonale ; il ne reste qu'à l'orthonormer, en divisant le vecteur par sa norme. On obtient la base : $\left\{(1/\sqrt{2}, 0, -1/\sqrt{2})\right\}$.

Espace propre associé à 3. Soit $u = (x, y, z)$ un vecteur propre de A associé à 3 ; il vérifie $Au = 3u$ soit :

$$\begin{cases} (3/2)x + (3/2)z = 3x \\ 3y = 3y \\ (3/2)x + (3/2)z = 3z \end{cases}$$

Ce système est équivalent à

$$\begin{cases} x = z \\ y \in \mathbb{R} \end{cases}$$

Donc l'espace propre de A associé à 3 est :

$$\{e_1 = (1, 0, 1), e_2 = (1, 1, 1)\}$$

Il faut orthogonaliser cette base, avec un Gram-Schmidt à deux étapes :

$$\text{Etape 1 : } f_1 = e_1 = (1, 0, 1).$$
$$\text{Etape 2 : } f_2 = e_2 - \frac{e_2 \cdot f_1}{f_1 \cdot f_1} f_1 = e_2 - \frac{2}{2} e_1 = (0, 1, 0).$$

En divisant chaque vecteur de la base f_1, f_2 par sa norme, on obtient la base $\left\{(1/\sqrt{2}, 0, 1/\sqrt{2}), (0, 1, 0)\right\}$.
Ecriture de P. On réunit maintenant dans les colonnes de P les bases orthonormées de chacun des espaces propres. On choisit de mettre d'abord le vecteur de la base associée à la valeur propre 0, puis ceux associés à la valeur propre 3 :

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix}$$

Ecriture de D. Compte tenu du choix de l'ordre fait précédemment, on a :

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

La valeur propre 3 est présente deux fois puisqu'il y a deux vecteurs propres associés à cette valeur propre.
A titre d'exercice, on pourra vérifier que P est orthogonale, et que l'on a les relations : $A = PDP^T$ et $D = P^T AP$.

4.2 Ecriture réduite d'une forme quadratique

Soit q une forme quadratique, de matrice A. Comme A est symétrique, on peut diagonaliser A dans une base orthonormée et écrire

$$A = PDP^T$$

avec P orthogonale et D diagonale.
On a, quelque soit $u \in E$:

$$q(u) = u^T Au = u^T PDP^T u = (P^T u)^T D(P^T u) = v^T Dv$$

en posant $v = P^T u$.
Notons $\lambda_1, \dots, \lambda_n$ les coefficients apparaissant sur la diagonale de D. On peut écrire :

$$q(u) = v^T Dv = \sum_{i=1}^{n} \lambda_i v_i^2$$

On dit que l'on a mis q sous **écriture réduite**. Les coefficients $\lambda_1, \dots, \lambda_n$ sont appelés les **valeurs propres** de q.

Exemple.

Dans $\mathbb{R}^2$, mettons $q(u) = u_1^2 - 6u_1u_2 + u_2^2$ sous écriture réduite.
La matrice de q dans la base canonique vaut :

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix}$$

Le polynôme caractéristique de cette matrice est :

$$\chi_A(x) = (1-x)(1-x) - 9 = x^2 - 2x - 8$$

ayant comme racines 4 et -2.
Le vecteur $u = (x, y)$ est vecteur propre de A associé à 4 si et seulement si :

$$Au = 4u$$

soit

$$\begin{pmatrix} x - 3y \\ -3x + y \end{pmatrix} = \begin{pmatrix} 4x \\ 4y \end{pmatrix}$$

soit encore :

$$x = -y$$

Le sous-espace propre de A associé à 4 est donc engendré par $(1, -1)$. Soit en divisant par la norme : $(1/\sqrt{2}, -1/\sqrt{2})$.
De manière analogue, le sous-espace associé à -2 est engendré par $(1, 1)$. Soit en divisant par la norme : $(1/\sqrt{2}, 1/\sqrt{2})$.
Nous avons donc notre base orthonormée formée de vecteurs propres :

$$\left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} ; \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \right\}$$

Ainsi

$$P = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \qquad D = \begin{pmatrix} -2 & 0 \\ 0 & 4 \end{pmatrix}$$

et on peut vérifier que P est orthogonale, et que $A = PDP^T$.
On a :

$$v = P^T u = \begin{pmatrix} \dfrac{u_1 + u_2}{\sqrt{2}} \\ \dfrac{u_1 - u_2}{\sqrt{2}} \end{pmatrix}$$

et l'écriture réduite de q est :

$$q(u) = -2\left(\frac{u_1 + u_2}{\sqrt{2}}\right)^2 + 4\left(\frac{u_1 - u_2}{\sqrt{2}}\right)^2$$

En développant les deux carrés, on peut constater que cette égalité est bien vérifiée pour tout u.

4.3 Application 1 : extrema d'une forme quadratique

Introduction : extrema de combinaisons linéaires de carrés

1. Considérons la forme quadratique de $\mathbb{R}^2$ suivante :

$$q(u) = u_1^2 + 2u_2^2$$

En augmentant u_1 et u_2, on peut rendre q aussi grande que l'on veut, q n'a donc pas de maximum sur $\mathbb{R}^2$. Par contre $q(u)$ est toujours positive (c'est une somme de deux nombres positifs), et vaut 0 lorsque $u_1 = u_2 = 0$; le minimum de q sur $\mathbb{R}^2$ est donc 0, et ce minimum est atteint en $(0, 0)$.

2. Considérons maintenant :
$$q(u) = -u_1^2 - 2u_2^2$$
Cette fois-ci, l'augmentation de u_1 et u_2 permet de diminuer q autant que l'on veut, q n'a donc pas de minimum. Par contre $q(u) \leq 0$ quelque soit u, et vaut 0 lorsque $u_1 = u_2 = 0$. Son maximum sur $\mathbb{R}^2$ est donc 0, et ce maximum est atteint en $(0,0)$.
3. Enfin, soit :
$$q(u) = u_1^2 - 2u_2^2$$
Fixons u_2 et augmentons u_1 : q augmente, et peut être aussi rendue grande que l'on veut (donc q n'a pas de maximum). En fixant u_1 et en faisant tendre u_2 vers l'infini, on fait tendre $q(u)$ vers $-\infty$ donc q n'a pas de minimum.

Pour savoir si une forme quadratique a un minimum ou un maximum (ou ni l'un ni l'autre), il faut donc la mettre sous la forme d'une combinaison linéaire de carrés (forme réduite), puis considérer les signes des coefficients apparaissant devant les carrés (valeurs propres).

Généralisation

Partant de l'écriture réduite de q :
$$q(u) = \lambda_1 v_1^2 + \lambda_2 v_2^2 + \ldots + \lambda_n v_n^2$$
nous pouvons distinguer plusieurs cas suivant le signe des valeurs propres λ_i :

1. **Si toutes les valeurs propres sont positives ou nulles,** alors q n'a pas de maximum (q peut être arbitrairement proche de $+\infty$), mais q a pour minimum 0, atteint en un v tel que $v_i = 0$ pour tout i tel que $\lambda_i \neq 0$.
2. **Si toutes les valeurs propres sont négatives ou nulles,** alors q n'a pas de minimum (q peut être arbitrairement proche de $-\infty$), mais q a pour maximum 0, atteint en un v tel que $v_i = 0$ pour tout i tel que $\lambda_i \neq 0$.
3. **S'il y a des valeurs propres strictement positives, et d'autres strictement négatives,** alors q n'a ni minimum ni maximum (elle peut être arbitrairement proche de $+\infty$ ou de $-\infty$).

Exemples.

– Dans $\mathbb{R}^2$, la forme $q(u) = u_1^2 - 6u_1u_2 + u_2^2$ considérée précédemment, dont l'écriture réduite est
$$q(u) = -2\left(\frac{u_1+u_2}{\sqrt{2}}\right)^2 + 4\left(\frac{u_1-u_2}{\sqrt{2}}\right)^2$$
n'a ni minimum, ni maximum dans $\mathbb{R}^2$.
– Dans $\mathbb{R}^3$, considérons q dont l'écriture réduite est :
$$q(u) = \left(\frac{2u_1+u_2}{\sqrt{3}}\right)^2 + 3\left(\frac{-u_1+2u_2+3u_3}{\sqrt{14}}\right)^2$$
Les valeurs propres sont 1, 3 et 0, donc q n'a pas de maximum, et a pour minimum 0. Ce minimum est atteint pour les u vérifiant le système d'équations :
$$\begin{cases} \dfrac{2u_1+u_2}{\sqrt{2}} = 0 \\ \dfrac{2u_2+u_3}{\sqrt{14}} = 0 \end{cases}$$
soit :
$$u = \begin{pmatrix} t \\ -2t \\ \dfrac{5}{3}t \end{pmatrix}, \quad t \in \mathbb{R}.$$

4.4 Application 2 : extrema d'une forme quadratique, sous contrainte de norme

Introduction : exemple de problème d'optimisation sous contrainte

Un cambrioleur a réussi à s'introduire dans une maison où sont présents plusieurs objets, de poids et de valeur différents. Il souhaite évidemment maximiser la somme des valeurs des objets qu'il va dérober. Pour cela il est clair que la meilleure solution consiste à tout emporter. Cependant notre voleur, peu expérimenté, est venu à pied avec un simple sac à dos (!), et il ne peut donc pas transporter plus qu'un certain poids. Il va donc chercher à maximiser la somme des valeurs des objets emportés (c'est son *objectif*), mais sans que la somme des poids dépasse la capacité maximale de son sac (c'est sa *contrainte*). – Ce problème (connu sous le nom de problème du sac à dos) est un exemple de problème d'optimisation sous contrainte.
Nous allons nous intéresser à un autre exemple de problème d'optimisation sous contrainte : notre objectif (à maximiser ou à minimiser) est $q(u)$, où q est une forme quadratique, et la contrainte est que le vecteur u doit être de norme 1 : $||u|| = 1$.

Optimisation d'une forme quadratique sous contrainte de norme

Nous voulons maximiser q sur l'ensemble S des vecteurs u de norme 1 : $S = \{u \in \mathbb{R}^n / ||u|| = 1\}$ (où la norme est celle associée au produit scalaire usuel).
Commençons par écrire q sous forme réduite :

$$q(u) = \lambda_1 v_1^2 + \lambda_2 v_2^2 + \ldots + \lambda_n v_n^2$$

et supposons que les valeurs propres sont rangées dans l'ordre croissant, i.e.

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$$

On a la majoration :

$$\begin{aligned} q(u) &= \lambda_1 v_1^2 + \lambda_2 v_2^2 + \ldots + \lambda_n v_n^2 \\ &\leq \lambda_n v_1^2 + \lambda_n v_2^2 + \ldots + \lambda_n v_n^2 \\ &= \lambda_n (v_1^2 + v_2^2 + \ldots + v_n^2) \end{aligned}$$

Soit maintenant $u \in S$. On a :

$$\begin{aligned} v_1^2 + v_2^2 + \ldots + v_n^2 &= v^T v \\ &= (P^T u)^T (P^T u) \text{ car } v = P^T u \\ &= u^T P P^T u \\ &= u^T u \text{ car } P P^T = I \\ &= ||u||^2 = 1 \text{ car } u \in S \end{aligned}$$

Ainsi, $q(u) \leq \lambda_n$ quelque soit u de norme 1. Ainsi la valeur λ_n est un candidat pour être le maximum de q sur S. Montrons maintenant que cette valeur est atteinte par q sur S, cherchons donc $u \in S$ tel que $q(u) = \lambda_n$.
On voit, à partir de l'écriture

$$q(u) = \lambda_1 v_1^2 + \lambda_2 v_2^2 + \ldots + \lambda_{n-1} v_{n-1}^2 + \lambda_n v_n^2,$$

que si u est tel que :

$$\left\{ \begin{array}{ccc} v_1 & = & 0 \\ v_2 & = & 0 \\ \vdots & = & \vdots \\ v_{n-1} & = & 0 \\ v_n & = & 1 \end{array} \right.$$

alors $q(u) = \lambda_n$. Ce système se réécrit :

$$v = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Soit encore :

$$P^T u = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

D'où

$$u = (P^T)^{-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Soit

$$u = P \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

car, P étant orthogonale, son inverse est sa transposée.
Ainsi le vecteur u cherché est la dernière colonne de P, autrement dit un vecteur propre (de la matrice de q) de norme 1 associé à la plus grande valeur propre λ_n.

Remarque : le vecteur en lequel ce maximum est atteint n'est pas unique : si u est un vecteur de norme 1 pour lequel $q(u) = \lambda_n$ alors $-u$ est un autre vecteur de norme 1 pour lequel $q(-u) = \lambda_n$. Ces deux vecteurs sont les seuls vecteurs de norme 1 pour lesquels le maximum est atteint, si λ_n est une valeur propre simple (de multiplicité 1). Si λ_n est une valeur propre ayant une multiplicité > 1, il y aura plusieurs vecteurs, linéairement indépendants, pour lesquels le maximum de q sera atteint.

Pour trouver le minimum de q sur S, on peut procéder de la même façon : on trouve que le minimum de q est sa plus petite valeur propre λ_1, et que ce minimum est atteint en un vecteur propre de norme 1 associé à cette plus petite valeur propre.

Exemple.

Pour $q(u) = u_1^2 - 6u_1u_2 + u_2^2$, on a vu en 4.2. que :
– la plus petite valeur propre de la matrice de q est $\lambda_1 = -2$
– la plus grande est $\lambda_2 = 4$
– un vecteur propre de norme 1 associé à λ_1 est $w_1 = (1/\sqrt{2}, 1/\sqrt{2})$
– un vecteur propre de norme 1 associé à λ_2 est $w_2 = (1/\sqrt{2}, -1/\sqrt{2})$
Le minimum de q sur S est donc -2, atteint (par exemple) en w_1.
Le maximum de q sur S est donc 4, atteint par ex. en w_2.

4.5 Racine carrée d'une matrice

4.5.1 Généralités

Définition Soit A une matrice (carrée, à coefficients réels), nous dirons qu'une matrice (carrée, à coefficients réels) X est une *racine carrée* de A si $A = MM^T$.

Remarques :

1. Une matrice n'a pas nécessairement de racine carrée. Par exemple la matrice 1×1 : $A = (-1)$ n'en a pas.
2. On peut démontrer qu'une matrice A a une racine carrée si et seulement si A est symétrique et elle n'a que des valeurs propres positives ou nulles (on dit que A est une matrice *positive*).
3. *En général*, pour calculer une racine carrée d'une matrice, il ne suffit pas de prendre la racine carrée de chacun des coefficients de la matrice ! A titre d'exemple, on pourra vérifier que
$$X = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$
n'est pas racine carrée de
$$A = \begin{pmatrix} 1 & 4 \\ 4 & 9 \end{pmatrix}$$
puisque $MM^T \neq A$.
D'ailleurs, la matrice A est bien symétrique mais elle a une valeur propre strictement négative, elle n'a donc pas de racine carrée.
4. Si une matrice a une racine carrée, alors elle en a au moins une autre : si M est racine de A alors $-M$ est également racine de A.
En fait il y en a beaucoup d'autres, parmi lesquelles figurent plusieurs matrices *triangulaires.* Lorsque l'on a écrit $A = MM^T$ avec M triangulaire, on dit que l'on a effectué une *décomposition de Cholesky* de A.

Nous allons voir une méthode pour calculer *une*racine carrée d'une matrice symétrique A n'ayant que des valeurs propres positives ou nulles.
En général la racine carrée ainsi trouvée ne sera pas triangulaire, il ne s'agira donc pas d'une décomposition de Cholesky – en fait, il existe un algorithme *ad hoc* pour le calcul de telles décompositions.

4.5.2 Calcul d'une racine carrée de matrice

Soit A une matrice symétrique positive (i.e. n'ayant que des valeurs propres positives ou nulles).
Comme A est symétrique, on peut la diagonaliser dans une base orthonormée de E, autrement dit trouver D diagonale et P orthogonale telle que
$$A = PDP^T$$
Ecrivons
$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$
où les λ_i sont les valeurs propres de A ; comme elles sont toutes positives ou nulles, on peut former la matrice :
$$S = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\lambda_n} \end{pmatrix}$$
On remarque que $SS^T = D$, donc S est une racine carrée de D. Pour une matrice diagonale, on peut prendre la racine carrée de chacun des coefficients et obtenir une racine carrée de la matrice, même si ce n'est pas le cas en général, cf. remarque 3.
On a donc :
$$A = PDP^T = PSS^TP^T = (PS)(PS)^T$$
donc :
$$M = PS$$
est une racine carrée de A.

Exemple

Cherchons une racine carrée de :

$$A = \begin{pmatrix} 7/2 & -1/2 \\ -1/2 & 7/2 \end{pmatrix}$$

Commençons par diagonaliser A en base orthonormée. Son polynôme caractéristique est :

$$\chi_A(x) = \left(\frac{7}{2} - x\right)^2 - \left(\left(-\frac{1}{2}\right) \times \left(-\frac{1}{2}\right)\right) = (3-x)(4-x)$$

Donc les valeurs propres de A sont 3 et 4. Toutes les valeurs propres de A sont positives ou nulles, donc A est une matrice symétrique positive, donc elle admet (au moins) une racine carrée.
Cherchons les espaces propres de A : soit $u = (x, y)$; l'équation $Au = 3u$ équivaut à $x = y$ donc E_3 est engendré par $(1, 1)$; l'équation $Au = 4u$ équivaut à $x = -y$ donc E_4 est engendré par $(-1, 1)$. En divisant par les normes de ces vecteurs on trouve

$$P = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

Et par ailleurs :

$$D = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$$

La matrice S du cours vaut donc dans notre cas :

$$S = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 2 \end{pmatrix}$$

Une racine carrée de A est donc donnée par :

$$M = PS = \begin{pmatrix} \sqrt{3/2} & -2/\sqrt{2} \\ \sqrt{3/2} & 2/\sqrt{2} \end{pmatrix} = \begin{pmatrix} \sqrt{3/2} & -\sqrt{2} \\ \sqrt{3/2} & \sqrt{2} \end{pmatrix}$$

(car $\dfrac{2}{\sqrt{2}} = \sqrt{2}$).
Vérification : on peut calculer MM^T et s'assurer que ce produit vaut A.

4.5.3 Application : simulation de gaussiennes corrélées

On veut simuler un couple de variables aléatoires (X, Y) de loi normale centrée (de moyenne 0), et de matrice de variance-covariance une matrice A (symétrique positive – une matrice de variance covariance l'est toujours) donnée.
Problème : on ne dispose que d'un générateur de variables aléatoires centrées, réduites (de variance 1) et indépendantes. Comment effectuer la simulation demandée en utilisant ce générateur ?
Soit (U, V) un couple de variables aléatoires gaussiennes centrées, réduites et indépendantes. La matrice de covariance de (U, V) est :

$$\begin{pmatrix} \mathrm{Var}(U) & \mathrm{Cov}(U,V) \\ \mathrm{Cov}(U,V) & \mathrm{Var}(V) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

Nous allons chercher (X, Y) sous la forme :

$$\begin{pmatrix} X \\ Y \end{pmatrix} = M \begin{pmatrix} U \\ V \end{pmatrix},$$

où M est une matrice 2×2 à déterminer pour que la matrice de covariance de (X, Y) soit A.

La matrice de covariance de (X,Y) est alors donnée par :

$$\begin{aligned}
\mathbf{E}\left(\begin{pmatrix}X\\Y\end{pmatrix}\begin{pmatrix}X\\Y\end{pmatrix}^T\right) &= \mathbf{E}\left(M\begin{pmatrix}U\\V\end{pmatrix}\left(M\begin{pmatrix}U\\V\end{pmatrix}\right)^T\right)\\
&= \mathbf{E}\left(M\begin{pmatrix}U\\V\end{pmatrix}\begin{pmatrix}U\\V\end{pmatrix}^T M^T\right)\\
&= M\mathbf{E}\left(\begin{pmatrix}U\\V\end{pmatrix}\begin{pmatrix}U\\V\end{pmatrix}^T\right)M^T\\
&= MM^T
\end{aligned}$$

la dernière égalité est justifiée par le fait que l'espérance qui apparaît est la matrice de covariance de (U,V), qui est la matrice identité.
Ainsi on veut avoir $MM^T = A$, donc comme matrice M on doit prendre une racine carrée de A.
Pour résoudre notre problème on peut donc :

1. calculer une racine carrée M de la matrice de covariance A à simuler ;
2. pour obtenir une réalisation du couple (X,Y) :
   (a) on simule deux v.a. normales centrées réduites U et V ;
   (b) on calcule le vecteur
   $$\begin{pmatrix}X\\Y\end{pmatrix} = M\begin{pmatrix}U\\V\end{pmatrix},$$
   dont les deux composantes sont respectivement des réalisations de la variable X (resp. Y).

Exemple

Simulons un couple (X,Y) de v.a. gaussiennes centrées de matrice de covariance :

$$A = \begin{pmatrix}7/2 & -1/2\\ -1/2 & 7/2\end{pmatrix},$$

dont on a trouvé une racine carrée

$$M = \begin{pmatrix}\sqrt{3/2} & -\sqrt{2}\\ \sqrt{3/2} & \sqrt{2}\end{pmatrix}.$$

Si (U,V) est une réalisation d'un couple de v.a. gaussiennes centrées réduites, indépendantes, alors le couple (X,Y) donné par :

$$\left\{\begin{array}{lcl} X &=& \sqrt{\frac{3}{2}}U - \sqrt{2}V\\ Y &=& \sqrt{\frac{3}{2}}U + \sqrt{2}V\end{array}\right.$$

est une réalisation du couple (X,Y) cherché.

Avec `R`, pour générer un échantillon de taille 10000 :

```
> u=rnorm(10000, mean=0, sd=1)
> v=rnorm(10000, mean=0, sd=1)
> x=sqrt(3/2)*u-sqrt(2)*v
> y=sqrt(3/2)*u+sqrt(2)*v
```

Vérification :

```
> cov(x,y)
[1] -0.5220036
> var(x)
[1] 3.525983
> var(y)
[1] 3.458432
```