ALGÈBRE BILINÉAIRE

DAVID DOS SANTOS FERREIRA — L2

Table des matières

1.	Espaces euclidiens	1
2.	Groupe orthogonal	18
3.	Formes quadratiques	34
4.	Dualité	50

Dans ces notes, E désigne un espace vectoriel de dimension finie dont le corps des scalaires est le corps des nombres réels $\mathbf{R}$.

1. Espaces euclidiens

1.1. **Premières définitions.** Un espace euclidien est une espace vectoriel réel de dimension finie muni d'une forme bilinéaire symétrique définie positive. Tâchons de définir chacun de ces termes.

Définition 1.1. *Une forme bilinéaire est une application*

$$\begin{aligned} b : E \times E &\to \mathbf{R} \\ (x,y) &\mapsto b(x,y) \end{aligned}$$

à valeurs dans le corps des scalaires, qui est linéaire par rapport à chacune des variables

$$\begin{aligned} b(\lambda x + \mu y, z) &= \lambda b(x,z) + \mu b(y,z) \\ b(x, \lambda y + \mu z) &= \lambda b(x,y) + \mu b(x,z). \end{aligned}$$

La forme quadratique associée est l'application $q : E \to \mathbf{R}$ définie par

$$q(x) = b(x,x), \quad x \in E.$$

Remarque 1.2. La linéarité implique

$$b(0,y) = b(x,0) = 0, \quad x,y \in E.$$

Une forme quadratique est homogène d'ordre 2 : pour tout $\lambda \in \mathbf{R}$ et pour tout $x \in E$

$$q(\lambda x) = \lambda^2 q(x).$$

En particulier, elle est paire : $q(-x) = q(x)$.

Si l'on choisit une base $\mathcal{B} = (e_1, \ldots, e_n)$ de E alors en décomposant les vecteurs dans la base

$$x = \sum_{j=1}^{n} x_j e_j, \quad y = \sum_{j=1}^{n} y_j e_j$$

et en utilisant la linéarité, on obtient

$$b(x, y) = \sum_{j=1}^{n} \sum_{k=1}^{n} b(e_j, e_k) x_j y_k.$$

La matrice

$$B = (b(e_j, e_k))_{1 \leq j,k \leq n}$$

est appelée la matrice de la forme bilinéaire b dans la base $\mathcal{B}$ et est notée

$$B = \mathrm{Mat}_{\mathcal{B}}(b)$$

pour lever toute ambiguité. On a alors la formule matricielle

$$b(x, y) = {}^{\mathrm{t}}XBY$$

avec

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

La linéarité implique la formule de développement suivante

$$q(x + y) = q(x) + q(y) + b(x, y) + b(y, x).$$

Définition 1.3. *Un produit scalaire est une forme bilinéaire*

(*i*) *symétrique : $b(y, x) = b(x, y)$ pour tous $x, y \in E$,*

(*ii*) *positive : la forme quadratique associée est à valeurs positives,*

(*iii*) *définie : la forme quadratique associée ne s'annule qu'en $x = 0$.*

On utilise la notation

$$\langle x, y \rangle = b(x, y)$$

lorsque b est un produit scalaire. Un espace euclidien est un espace vectoriel réel de dimension finie muni d'un produit scalaire.

Exemple 1.4. Voici des exemples classiques de produit scalaire.

1. Sur $E = \mathbf{R}^n$, le produit $\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j = {}^{\mathrm{t}}YX$ est un produit scalaire. Plus généralement, si E est un espace vectoriel et si $(e_1, \ldots, e_n)$ est une base de E alors

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j, \quad x = \sum_{j=1}^{n} x_j e_j, \quad y = \sum_{j=1}^{n} y_j e_j$$

   définit un produit scalaire sur E.

2. Sur $E = \mathrm{M}(n, \mathbf{R})$, le produit

$$\langle A, B \rangle = \mathrm{tr}({}^{\mathrm{t}}AB)$$

   est un produit scalaire.

3. Sur $E = \mathbf{R}_d[X]$, le produit
$$\langle P, Q\rangle = \int_0^1 P(x)Q(x)\,\mathrm{d}x$$
est un produit scalaire.

Remarque 1.5. La matrice d'une forme bilinéaire symétrique b est symétrique
$${}^{\mathrm{t}}B = B$$
car $b(e_j, e_k) = b(e_k, e_j)$. Réciproquement, si la matrice B de la forme b dans une base est symétrique alors b est symétrique
$$b(y, x) = {}^{\mathrm{t}}YBX = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}BY = {}^{\mathrm{t}}XBY = b(x, y).$$
La matrice B d'une forme bilinéaire définie b est inversible
$$\det B \neq 0$$
car si $X \in \ker B$ alors
$$b(x, x) = {}^{\mathrm{t}}XBX = 0$$
implique $x = 0$ donc $X = 0$.

Théorème 1.6 (Inégalité de Cauchy-Schwarz)**.** *Soit b une forme bilinéaire symétrique positive alors pour tous vecteurs $x, y \in E$ on a l'inégalité*
$$|b(x, y)| \leq \sqrt{q(x)}\sqrt{q(y)}.$$
Si b est de plus définie positive, il n'y a égalité que lorsque x et y sont liés.

Démonstration. De la positivité de
$$0 \leq q(x - y) = q(x) + q(y) - 2b(x, y)$$
on tire
$$b(x, y) \leq \frac{1}{2}(q(x) + q(y))$$
Et avec le même argument en remplaçant y par $-y$, on tire finalement
$$|b(x, y)| \leq \frac{1}{2}(q(x) + q(y))$$
On applique alors cette inégalité au couple $(\sqrt{t}x, y/\sqrt{t})$ avec $t > 0$ pour obtenir
$$|b(x, y)| = |b(\sqrt{t}x, y/\sqrt{t})| \leq \frac{1}{2}\left(tq(x) + \frac{q(y)}{t}\right).$$
La fonction
$$\begin{aligned}]0, \infty[&\to \mathbf{R}_+ \\ t &\mapsto \frac{1}{2}\left(tq(x) + \frac{q(y)}{t}\right) \end{aligned}$$
admet un minimum égal à $\sqrt{q(x)q(y)}$, ce qui donne l'inégalité de Cauchy-Schwarz.

Le cas d'égalité se produit lorsque $x = 0$ ou $y = 0$ ou lorsque $\sqrt{t}x = y/\sqrt{t}$, c'est à dire lorsque x et y sont liés. $\square$

Remarque 1.7. Voici un autre argument classique pour démontrer l'inégalité de Cauchy-Schwarz. Le polynôme du second degré

$$q(tx+y) = t^2 q(x) + 2tb(x,y) + q(y)$$

est toujours positif, il ne peut donc avoir de racines réelles simples, et son discriminant est donc négatif ou nul

$$\Delta = 4b(x,y)^2 - 4q(x)q(y) \leq 0$$

ce qui donne l'inégalité de Cauchy-Schwarz.

Remarque 1.8. L'inégalité de Cauchy-Schwarz permet de définir dans un espace euclidien $(E, \langle \cdot , \cdot \rangle 1)$ l'écart angulaire entre deux vecteurs non nuls $x, y \in E \setminus \{0\}$

$$\theta = \arccos \underbrace{\frac{\langle x, y\rangle}{\|x\|\|y\|}}_{\in[-1,1]} .$$

Corollaire 1.9 (Inégalité de Minkowski)**.** *Soit q une forme quadratique associée à une form bilinéaire symétrique positive b alors on a l'inégalité de Minkowski*

$$\sqrt{q(x+y)} \leq \sqrt{q(x)} + \sqrt{q(y)}$$

pour tous vecteurs $x, y \in E$.

Démonstration. L'inégalité de Cauchy-Schwarz donne

$$q(x+y) = q(x)+q(y)+2b(x,y) \leq q(x)+q(y)+2\sqrt{q(x)q(y)} = \left(\sqrt{q(x)}+\sqrt{q(y)}\right)^2$$

et on extrait la racine carrée. $\square$

Avec ce dernier résultat, on voit que pour un produit scalaire $b = \langle \cdot , \cdot \rangle$ l'application

$$\begin{aligned} E &\to \mathbf{R} \\ x &\mapsto \sqrt{\langle x, x\rangle} \end{aligned}$$

est une norme. Un espace eudlidien est donc automatiquement un espace vectoirel normé. On la notera par la suite

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

La fonction $d : E \times E \to \mathbf{R}_+$ donnée par

$$d(x,y) = \|x-y\|$$

est une distance entre les vecteurs de E. Le développement quadratique d'une norme s'écrit dans ce contexte

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle.$$

Corollaire 1.10. *Soit q une forme quadratique associée à une forme bilinéaire symétrique positive b alors on a*

$$\sqrt{q(x)} = \max_{y \in E/q(y)=1} |b(x,y)|.$$

En particulier, on peut calculer la norme associée à un produit scalaire de la manière suivante

$$\|x\| = \max_{\|y\|=1} |\langle x, y\rangle| = \max_{y\in E\setminus\{0\}} \frac{|\langle x, y\rangle|}{\|y\|}.$$

Démonstration. L'inégalité de Cauchy-Schwarz donne

$$|b(x,y)| \leq \sqrt{q(x)}$$

pour tout $y \in E$ tel que $q(y) = 1$. Si $q(x) = 0$ alors $\sqrt{q(x)} = 0 = b(x, 0)$ et si $q(x) > 0$ alors

$$\sqrt{q(x)} = b\left(x, \frac{x}{\sqrt{q(x)}}\right)$$

et ainsi $\sqrt{q(x)}$ est le maximum de $\{|b(x,y)| : y \in E, q(y) = 1\}$. □

Définition 1.11. *Une famille de vecteurs $(v_1, \ldots, v_m)$ est dite orthogonale si les vecteurs de la famille sont deux-à-deux orthogonaux*

$$\langle v_j, v_k\rangle = 0 \quad \text{si} \quad j \neq k.$$

Un vecteur est dit unitaire s'il est de norme égale à 1. Une famille est dite orthonormée si elle est orthogonale et si tous les vecteurs ont unitaires.

Un calcul et une récurrence immédiate donne alors le théorème de Pythagore.

Théorème 1.12 (Théorème de Pythagore)**.** *Soit $x, y \in E$ deux vecteurs orthogonaux alors*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

Plus généralement, si $(x_1, \ldots, x_m0$ est une famille orthogonale de vecteurs, on a

$$\left\| \sum_{j=1}^m x_j \right\|^2 = \sum_{j=1}^m \|x_j\|^2.$$

Proposition 1.13. *Toute famille orthogonale de vecteurs non nuls est libre.*

Démonstration. Soit $(v_1, \ldots, v_m)$ une famille de vecteurs orthogonaux, on a

$$\sum_{j=1}^m \lambda_j v_j = 0$$

et en utilisant le théorème de Pythagore

$$\lambda_j \|v_j\|^2 = 0$$

pour tout $1 \leq j \leq m$, soit $\lambda_j = 0$. □

1.2. **Bases orthonormées.** On introduit les symboles de Kronecker

$$\delta_{j,k} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases}.$$

Définition 1.14. *Une base orthonormée $(e_1, \ldots, e_n)$ est une base de vecteurs orthogonaux et normés, i.e.*

$$\langle e_j, e_k\rangle = \delta_{j,k}.$$

Proposition 1.15. *Soit* $(e_1, \dots, e_n)$ *une base orthonormée d'un espace euclidien, tout vecteur se décompose (de manière unique) dans la base*

$$x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j$$

et sa norme est donnée par

$$\|x\|^2 = \sum_{j=1}^{n} \langle x, e_j \rangle^2.$$

Le produit scalaire s'écrit

$$\langle x, y \rangle = \sum_{j=1}^{n} \langle x, e_j \rangle \langle y, e_j \rangle = {}^{\mathrm{t}}YX$$

avec

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Démonstration. Un vecteur $x \in E$ se décompose (de manière unique) dans la base sous la forme

$$x = \sum_{j=1}^{n} x_j e_j$$

et en faisant le produit scalaire avec e_k, on trouve

$$\langle x, e_k \rangle = \sum_{j=1}^{n} x_j \langle e_j, e_k \rangle = x_k.$$

En outre, le théorème de Pythagore donne

$$\|x\|^2 = \sum_{j=1}^{n} x_j^2 \|e_j\|^2 = \sum_{j=1}^{n} \langle x, e_j \rangle^2.$$

Enfin, on a

$$\langle x, y \rangle = \sum_{j,k=1}^{n} \langle x, e_j \rangle \langle y, e_k \rangle \langle e_j, e_k \rangle = \sum_{j=1}^{n} \langle x, e_j \rangle \langle y, e_j \rangle$$

l'expression du produit scalaire sous forme canonique. □

Proposition 1.16 (Procédé d'orthonormalisation de Gram-Schmidt)**.** *Soit une famille libre de vecteurs* $(v_1, \dots, v_m)$ *, alors il existe une unique famille orthonormée* $(e_1, \dots, e_m)$ *vérifiant*

$$\mathrm{vect}(e_1, \dots, e_k) = \mathrm{vect}(v_1, \dots, v_k)$$

pour tout $1 \leq k \leq m$*. En particulier, tout espace euclidien admet une base orthonormée.*

Démonstration. On procède par récurrence sur m. Pour l'initialisation,il suffit de normer le vecteur
$$e_1 = \frac{v_1}{\|v_1\|}$$
puisque v_1 est non nul (sinon le famille $(v_1, \ldots, v_m)$ serait liée). Supposons que l'on ait construit une famille $(e_1, \ldots, e_\ell)$ orthonormée vérifiant

$$\mathrm{vect}(e_1, \ldots, e_k) = \mathrm{vect}(v_1, \ldots, v_k)$$

pour tout $1 \leq k \leq \ell$. Pour garder cette dernière propriété au rang suivant, il faut chercher $e_{\ell+1}$ sous la forme

$$e_{\ell+1} = \lambda_{\ell+1} v_{\ell+1} + \sum_{k=1}^{\ell} \lambda_k e_k.$$

Pour que la famille soit orthogonale au rang suivant, on doit imposer

$$0 = \langle e_{\ell+1}, e_j \rangle = \lambda_{\ell+1} \langle v_{\ell+1}, e_j \rangle + \lambda_j$$

soit

$$\lambda_j = -\lambda_{\ell+1} \langle v_{\ell+1}, e_j \rangle$$

et donc

$$e_{\ell+1} = \lambda_{\ell+1} \left(\sum_{k=1}^{\ell} \langle v_{\ell+1}, e_k \rangle e_k \right).$$

Pour fixer $\lambda_{\ell+1}$, il suffit de normer (en utilisant le théorème de Pyhtagore)

$$\lambda_{\ell+1} = \left(\sum_{k=1}^{\ell} \langle v_{\ell+1}, e_k \rangle^2 \right)^{-\frac{1}{2}}.$$

Pour terminer la récurrence, il suffit de vérifier qu'avec

$$e_{\ell+1} = \left(\sum_{k=1}^{\ell} \langle v_{\ell+1}, e_k \rangle^2 \right)^{-\frac{1}{2}} \left(\sum_{k=1}^{\ell} \langle v_{\ell+1}, e_k \rangle e_k \right)$$

la famille $(e_1, \ldots, e_{\ell+1})$ est orthonormée. On a par construction

$$\mathrm{vect}(e_1, \ldots, e_{\ell+1}) = \mathrm{vect}(e_1, \ldots, e_\ell, v_{\ell+1}) = \mathrm{vect}(v_1, \ldots, v_\ell, v_{\ell+1})$$

ce qui prouve la proposition. Ceci donne également un procédé de construction d'une famille orthonormée à partir d'une famille libre. $\square$

Définition 1.17. *On dit que deux sous-esapces vectoriels F, G de E sont orthogonaux, et on note $F \perp G$ si*

$$\forall f \in F, \quad \forall g \in G, \quad \langle f, g \rangle = 0.$$

1.3. **Orthogonalité.** Développons un peu la notion d'orthogonalité.

Définition 1.18. *Soit $A \subset E$ un sous-ensemble de E, l'orthogonal de A est l'ensemble*

$$A^\perp = \big\{x \in E : \langle x, a\rangle = 0 \quad \forall a \in A\big\}.$$

Remarque 1.19. Attention ! Si deux sous-espaces vectoriels sont orthogonaux $F \perp G$ alors on n'a pas forcément $G = F^\perp$ ou $F = G^\perp$. Par exemple dans $E = R^4$ muni du produit scalaire canonique, $F = \mathrm{vect}(e_1)$ et $G = \mathrm{vect}(e_2)$ sont orhogonaux mais

$$F^\perp = \mathrm{vect}(e_2, e_3, e_4), \quad G^\perp = \mathrm{vect}(e_1, e_3, e_4).$$

Proposition 1.20. *L'orthogonal vérifie les propriétés suivantes*

(*i*) $\{0\}^\perp = E$, $E^\perp = \{0\}$,

(*ii*) $A^\perp$ *est un sous-espace vectoriel de* E,

(*iii*) *Si* $A \subset B$ *alors* $B^\perp \subset A^\perp$,

(*iv*) $A^\perp = \mathrm{vect}(A)^\perp$,

(*v*) $(A^\perp)^\perp = \mathrm{vect}(A)$.

(*vi*) *Si* F *est un sous-espace vectoriel de* E *alors* $\dim F^\perp = n - \dim F$.

Une conséquence du premier point est que si l'on a

$$\langle x, z\rangle = \langle y, z\rangle$$

pour tout $z \in E$ alors $x = y$.

Démonstration.

(*i*) La première égalité $\{0\}^\perp = E$ est claire. Soit $x \in E^\perp$, alors on a

$$\|x\|^2 = \langle x, x\rangle = 0$$

implique $x = 0$ grâce au caractère défini du produit scalaire et ainsi $E^\perp = \{0\}$.

(*ii*) Montrons que $A^\perp$ est un sous-espace vectoriel de E, soit $x, y \in A^\perp$ alors pour tous $\lambda, \mu \in \mathbf{R}$ et tout $a \in A$, on a par linéarité du produit scalaire

$$\langle \lambda x + \mu y, a\rangle = \lambda \underbrace{\langle x, a\rangle}_{=0} + \mu \underbrace{\langle y, a\rangle}_{=0} = 0$$

et ainsi $\lambda x + \mu y \in A^\perp$.

(*iii*) Si $A \subset B$ alors pour tout $x \in B^\perp$ et pour tout $a \in A \subset B$ on a $\langle x, a\rangle = 0$ et ainsi $B^\perp \subset A^\perp$.

(*iv*) Comme $A \subset \mathrm{vect}(A)$, d'après le point précédent $\mathrm{vect}(A)^\perp \subset A^\perp$. Réciproquement si $x \in A^\perp$ alors pour tout

$$v = \sum_{j=1}^{m} \lambda_j a_j \in \mathrm{vect}(A), \quad a_j \in A$$

on a

$$\langle x, v\rangle = \sum_{j=1}^{n} \lambda_j \langle x, a_j\rangle = 0$$

ce qui donne $x \in \mathrm{vect}(A)^\perp$, et finalement $A^\perp = \mathrm{vect}(A)^\perp$.

(v) D'après le point précédent

$$(A^\perp)^\perp = (\mathrm{vect}(A)^\perp)^\perp$$

et il suffit de montrer que pour le sous-espace vectoriel $F = \mathrm{vect}(A)$, on a $(F^\perp)^\perp = F$. Soit $x \in F$ alors pour tout $y \in F^\perp$, on a

$$\langle x, y\rangle = 0$$

soit $x \in (F^\perp)^\perp$ et donc $F \subset (F^\perp)^\perp$. Remarquons que jusqu'à présent, on n'a pas utilisé le fait que E est de dimension finie. On en aura besoin pour montrer l'inclusion inverse. Supposons que le dernier point ait été démontré, alors on a

$$\dim(F^\perp)^\perp = n - \dim F^\perp = n - (n - \dim F) = \dim F$$

et donc F est un sous-espace vectoriel de $(F^\perp)^\perp$ de même dimension, ce qui entraîne

$$F = (F^\perp)^\perp.$$

(vi) Supposons que $m = \dim F$, on choisit $(f_1, \ldots, f_m)$ une base de F que l'on complète en une base $(f_1, \ldots, f_n)$ de E. Par le procédé d'orthonormalisation de Gram-Schmidt, on construit une base orthonormée de E telle que

$$F = \mathrm{vect}(f_1, \ldots, f_m) = \mathrm{vect}(e_1, \ldots, e_m).$$

Montrons que

$$F^\perp = \mathrm{vect}(e_{m+1}, \ldots, e_n)$$

ce qui prouvera

$$\dim F^\perp = n - m = n - \dim F.$$

Soit $x \in F^\perp$

$$x = \sum_{j=1}^{n} \langle x, e_j\rangle e_j = \sum_{j=m+1}^{n} \langle x, e_j\rangle e_j \in \mathrm{vect}(e_{m+1}, \ldots, e_n)$$

et on a donc l'inclusion $F^\perp \subset \mathrm{vect}(e_{m+1}, \ldots, e_n)$. Inversement, par construction $(e_{m+1}, \ldots, e_n) \in F^\perp$ et ainsi puisque $F^\perp$ est un sous-espace vectoriel

$$\mathrm{vect}(e_{m+1}, \ldots, e_n) \subset F^\perp.$$

La proposition est démontrée. □

Proposition 1.21. *Un espace euclidien E se décompose en une somme directe entre un sous-espace vectoriel F et son orthogonal*

$$E = F \oplus F^\perp.$$

Démonstration. Montrons que F et $F^\perp$ sont supplémentaires

(1) soit $x \in F \cap F^\perp$ alors $\|x\|^2 = \langle x, x\rangle = 0$ et donc $x = 0$,

(2) $\dim F^\perp = n - \dim F$. Un autre point de vue consiste à considérer comme dans le point (vi) de la preuve de la proposition 1.20 une base orthonormée $(e_1, \cdots, e_n)$ telle que

$$F = \mathrm{vect}(e_1, \ldots, e_m)$$

et on peut alors décomposer tout vecteur $x \in E$ sous la forme

$$x = \sum_{j=1}^{n} \langle x, e_j\rangle e_j = \underbrace{\sum_{j=1}^{m} \langle x, e_j\rangle e_j}_{=f\in F} + \underbrace{\sum_{j=m+1}^{n} \langle x, e_j\rangle e_j}_{=g\in F^\perp}$$

d'une somme d'un vecteur $f \in F$ et $g \in F^\perp$.

□

Définition 1.22. *Soit F un sous-espace vecotriel de E, la projection orthogonale sur F est l'application*

$$\pi_F : E \to E$$

telle que

$$x - \pi_F(x) \in F^\perp.$$

Cette projection est définie sans ambiguité car $E = F \oplus F^\perp$ et ainsi tout vecteur x se décompose de manière unique sous la forme

$$x = f + g, \quad f \in F, \quad g \in F^\perp$$

et donc $\pi_F(x) = f \in F$ avec $g = x - f = x - \pi_F(x) \in F^\perp$. Notons que l'on a

$$\pi_{F^\perp} = \mathrm{Id}_E - \pi_F$$

ceci est cohérent avec $(F^\perp)^\perp = F$

$$\pi_F = \mathrm{Id}_E - \pi_{F^\perp} = \mathrm{Id}_E - (\mathrm{Id}_E - \pi_F).$$

Remarque 1.23. Si l'on dispose d'une base orthonormée $\mathcal{B} = (e_1, \cdots, e_n)$ telle que

$$F = \mathrm{vect}(e_1, \ldots, e_m)$$

alors on peut facilement projeter orthogonalement sur F

$$\pi_F(x) = \sum_{j=1}^{m} \langle x, e_j\rangle e_j.$$

La matrice d'une projection orthogonale dans une base orthonormée dans cette base est donnée par

$$P_F = \mathrm{Mat}_{\mathcal{B}}(\pi_F) = \begin{pmatrix} \mathrm{I}_m & 0 \\ 0 & 0 \end{pmatrix}.$$

Notons également que l'on peut revisiter le procédé d'orthonormalisation de Gram-Schmidt en constatant que pour orthonormaliser la base $\mathcal{B} = (v_1, \ldots, v_n)$, à l'étape $j+1$, on projète v_{j+1} sur l'orthogonal du sous-espace

$$F_j = \mathrm{vect}(e_1, \ldots, e_j)$$

engendré par les j premiers vecteurs déjà orthonormalisés

$$e_{j+1} = \frac{v_{j+1} - \pi_{F_j}(v_{j+1})}{\|v_{j+1} - \pi_{F_j}(v_{j+1}\|} = \frac{\pi_{F_j^\perp}(v_{j+1})}{\|\pi_{F_j^\perp}(v_{j+1})\|}.$$

Lemme 1.24. *La projection orthogonale est une projection*

$$\pi_F \circ \pi_F = \pi_F,$$

c'est la projection sur F parallèlement à $F^\perp$, i.e.

$$\ker \pi_F = F^\perp, \quad \operatorname{Im} \pi_F = F.$$

Démonstration. Soit $x \in E$, on a

$$x = f + g, \quad f \in F, \quad g \in F^\perp$$

et $f = \pi_F(x)$, or $f = f + 0$ avec $f \in F$ et $0 \in G$ et cette décomposition est unique donc $\pi_F^2(x) = \pi_F(f) = f$. En outre, on a $\operatorname{Im} \pi_F = F$ car $\operatorname{Im} \pi_F \subset F$ et pour tout $f \in F$ on a $f = \pi_F(f) \in \operatorname{Im} \pi_F$. Enfin $F^\perp \subset \ker F$ toujours par unicité de la décomposition dans $F \oplus F^\perp$ et si $x \in \ker F$ alors

$$x = x - \pi_F(x) \in F^\perp$$

ce qui donne $\ker F \subset F^\perp$ et ainsi $\ker F = F^\perp$. □

On peut maintenant calculer la distance à un sous-espace vectoriel $F \subset E$

$$d(x, F) = \inf_{f \in F} \|x - f\|$$

puisque par le théorème de Pythagore

$$\begin{aligned} \|x - f\|^2 = \| \underbrace{x - \pi_F(x)}_{=\pi_{F^\perp}(x) \in F^\perp} + \underbrace{\pi_F(x) - f}_{\in F} \|^2 &= \|x - \pi_F(x)\|^2 + \|\pi_F(x) - f\|^2 \\ &\geq \|x - \pi_F(x)\|^2 \end{aligned}$$

est minimale lorsque $f = \pi_F(x)$. En outre, on a

$$\begin{aligned} \|x - \pi_F(x)\|^2 &= \|x\|^2 + \|\pi_F(x)\|^2 - 2\langle x, \pi_F(x)\rangle \\ &= \|x\|^2 + \|\pi_F(x)\|^2 - 2\langle \pi_F(x), \pi_F(x)\rangle - 2 \underbrace{\langle \pi_{F^\perp}(x), \pi_F(x)\rangle}_{=0} \\ &= \|x\|^2 - \|\pi_F(x)\|^2. \end{aligned}$$

Ainsi a-t-on démontré le résultat suivant.

Proposition 1.25. *Soit F un sous espace vectoriel de E, alors la distance d'un vecteur $x \in E$ à F est atteinte en $\pi_F(x) \in F$ et égale à*

$$d(x, F) = \|x - \pi_F(x)\| = \sqrt{\|x\|^2 - \|\pi_F(x)\|^2}.$$

Exemple 1.26. La distance à une droite

$$F = \operatorname{Vect}(\omega)$$

dirigée par un vecteur unitaire ω est donnée par

$$d(x, F) = \sqrt{\|x\|^2 - \langle x, \omega\rangle^2}.$$

La distance à un hyperplan
$$F = \omega^{\perp}$$
de normale le vecteur unitaire ω est donnée par
$$d(x, F) = |\langle x, \omega\rangle|$$

Remarque 1.27. Il est facile d'obtenir une version affine du résultat précédent : la distance d'un vecteur $x \in E$ au sous-espace affine
$$a + F = \{a + f : f \in F\}$$
est égale à
$$d(x, a + F) = d(x - a, F) = \|x - a - \pi_F(x - a)\|.$$

Définition 1.28. *Soit F un sous-espace vectoriel de E, la symétrie orthogonale par rapport à F est l'application $\sigma_F : E \to E$ donnée par*
$$\sigma_F = \pi_F - \pi_{F^\perp} = \mathrm{Id}_E - 2\pi_F.$$
C'est bien une involution (ou une symétrie vectorielle sur F parallèlement à $F^\perp$)
$$\sigma_F^2 = \sigma_F \circ \sigma_F = \mathrm{Id}_E.$$

En effet, on a
$$\sigma_F^2 = \mathrm{Id}_E + 4\pi_F - 4\pi_F = \mathrm{Id}_E,$$
et alternativement
$$\ker(\sigma_F - \mathrm{Id}_E) = \ker \pi_{F^\perp} = F, \quad \ker(\sigma_F + \mathrm{Id}_E) = \ker \pi_F = F^\perp.$$

1.4. **Adjoints.**

Définition 1.29. *Soit $u \in L(E)$ un endomorphisme, il existe un unique endomorphisme $u^* \in L(E)$ tel que pour tous vecteurs $x, y \in L(E)$*
$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle.$$
Cet endomorphisme est l'appelé l'adjoint de u.

Montrons l'unicité et l'existence d'un tel endomorphisme par analyse et synthèse : soit $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée, on décompose
$$u^*(x) = \sum_{j=1}^{n} \langle u^*(x), e_j\rangle e_j$$
et comme la définition de l'adjoint implique $\langle u^*(x), e_j\rangle = \langle x, u(e_j)\rangle$, on en déduit l'expression de u^*
$$u^*(x) = \sum_{j=1}^{n} \langle x, u(e_j)\rangle e_j.$$
Réciproquement, l'application définie par cette formule vérifie
$$\langle x, u^*(y)\rangle = \sum_{j=1}^{n} \langle y, u(e_j)\rangle \langle x, e_j\rangle = \left\langle y, u\left(\sum_{j=1}^{n} \langle x, e_j\rangle e_j\right)\right\rangle = \langle u(x), y\rangle.$$

Ce qui montre l'existence de l'adjoint et donne l'expression de la matrice de u^* dans la base $\mathcal{B}$

$$\mathrm{Mat}_{\mathcal{B}}(u^*) = (\langle e_j, u(e_k)\rangle)_{1\leq j,j\leq n} = {}^{\mathrm{t}}(\langle u(e_j), e_k\langle)_{1\leq j,j\leq n} = {}^{\mathrm{t}}\,\mathrm{Mat}_{\mathcal{B}}(u).$$

Proposition 1.30. *L'adjonction vérifie les propriétés suivantes*

(*i*) $(\lambda u + \mu v)^* = \lambda u^* + \mu v^*$

(*ii*) $(u \circ v)^* = v^* \circ u^*$

(*iii*) $(u^*)^* = u$

(*iv*) $\|u^*\| = \|u\|$

(*v*) $\det(u^*) = \det u$

(*vi*) $\ker u^* = (\mathrm{Im}\, u)^\perp, \quad \mathrm{Im}\, u^* = (\ker u)^\perp.$

Ainsi l'adjonction $* : L(E) \to L(E)$ *est une involution linéaire qui conserve la norme et le déterminant et inverse l'ordre de composition.*

Remarque 1.31. Rappelons que le déterminant de u peut être défini en choisissant une base $\mathcal{B} = (e_1, \ldots, e_n)$ par

$$\det u = \det_{\mathcal{B}}(u(e_1), \ldots, u(e_n)) = \det A, \quad A = \mathrm{Mat}_{\mathcal{B}}(u)$$

où $\det_{\mathcal{B}}$ est l'unique déterminant tel que $\det_{\mathcal{B}}(e_1, \ldots, e_n) = 1$. Ce déterminant ne dépend pas de la base choisie car si P désigne une matrice de passage d'une base à l'autre

$$\det(PAP^{-1}) = \det A.$$

Démonstration.

(*i*) On a

$$\begin{aligned}\langle x, (\lambda u + \mu v)^*(y)\rangle &= \langle \lambda u(x) + \mu v(x), y\rangle = \lambda\langle u(x), y\rangle + \mu\langle v(x), y\rangle \\ &= \lambda\langle x, u^*(y)\rangle + \mu\langle x, v^*(y)x, \rangle = \langle \lambda u^*(y) + \mu v^*(y)\rangle\end{aligned}$$

pour tous vecteurs x et y et par conséquent $(\lambda u + \mu v)^*(y) = \lambda u^*(y) + \mu u^*(y)$ pour tout vecteur $y \in E$.

(*ii*) On a

$$\begin{aligned}\langle x, (u \circ v)^*(y)\rangle &= \langle u(v(x)), y\rangle = \langle v(x), u^*(y)\rangle = \langle x, v^*(u^*(y))\rangle \\ &= \langle x, v^* \circ u^*(y)\rangle\end{aligned}$$

pour tous vecteurs x et y, et par conséquent $(u \circ v)^*(y) = v^* \circ u^*(y)$ pour tout $y \in E$.

(*iii*) On a

$$\langle x, (u^*)^*(y)\rangle = \langle u^*(x), y\rangle = \langle y, u^*(x)\rangle = \langle u(y), x\rangle = \langle x, u(y)\rangle$$

pour tous vecteurs x et y, et par conséquent $(u^*)^*(y) = u(y)$ pour tout vecteur $y \in E$.

(*iv*) D'après le corollaire 1.10, on a

$$\begin{aligned}\|u\| &= \max_{x\in E/\|x\|=1} \|u(x)\| = \max_{x,y\in E/\|x\|=\|y\|=1} |\langle u(x), y\rangle| \\ &= \max_{x,y\in E/\|x\|=\|y\|=1} |\langle x, u^*(y)\rangle| = \max_{y\in E/\|y\|=1} \|u^*(y)\| = \|u^*\|.\end{aligned}$$

(*v*) Soit A la matrice de u dans une base orthonormée $(e_1, \dots, e_n)$ de E alors
$$\det u = \det A = \det {}^{\mathrm{t}}Q = \det(u^*).$$

(*vi*) Si $\ker u^* = (\operatorname{Im} u)^\perp$ est vrai pour tout endomorphisme u alors en appliquant cette relation à l'endomorphisme adjoint, on obtient
$$\ker u = \ker(u^*)^* = (\operatorname{Im} u^*)^\perp$$
et en passant à l'orthogonal
$$\operatorname{Im} u^* = ((\operatorname{Im} u^*)^\perp)^\perp = (\ker u)^\perp.$$
Il suffit donc de vérifier $\ker u^* = (\operatorname{Im} u)^\perp$: soit $x \in \ker u^*$ alors pour tout $y \in E$
$$\langle x, u(y)\rangle = \langle u^*(x), y\rangle = 0$$
donc $x \in (\operatorname{Im} u)^\perp$. Réciproquement, si $x \in (\operatorname{Im} u)^\perp$ alors on a
$$\langle u^*(x), y\rangle = \langle x, u(y)\rangle = 0$$
pour tout $y \in E$ et ainsi $u^*(x) \in E^\perp$ donc $x \in \ker u^*$.

$\square$

On rappelle que le polynôme caractéristique de u est défini comme
$$\chi_u(t) = \det\big(A - t\mathrm{I}_n\big)$$
où A est la une matrice de u dans une base. Cette définition est bien sûr inépendante du choix de la base car si l'on note P la matrice de passage d'une base à l'autre, on a
$$\det\big(PAP^{-1} - t\mathrm{I}_n\big) = \det P \,\det\big(A - t\mathrm{I}_n\big)\, \det P^{-1} = \det\big(A - t\mathrm{I}_n\big).$$
Le polynôme caractéristique d'un endomorphisme et celui de son adjoint sont égaux
$$\chi_{u^*} = \chi_u$$
car
$$\chi_{u^*}(t) = \det\big({}^{\mathrm{t}}A - t\mathrm{I}_n\big) = \det\big(A - t\mathrm{I}_n\big) = \chi_u(t).$$

Définition 1.32. *On dit qu'un endomorphisme est autoadjoint (ou symétrique) si $u^* = u$ et antiadjoint (ou antisymétrique) si $u^* = -u$. On dit qu'une matrice est symétrique (resp. antisymétrique) si l'endomorphisme associé est autoadjoint (resp. anti-adjoint), i.e. si*
$${}^{\mathrm{t}}A = A \quad (\textit{ resp. } {}^{\mathrm{t}}A = A).$$

Exemple 1.33. La projection orthogonale est une application auto-adjointe
$$\pi_F^* = \pi_F.$$
En effet, on a
$$\langle \pi_F(x), y\rangle = \langle \pi_F(x), \pi_F(y)\rangle = \langle x, \pi_F(y)\rangle$$
pour tous vecteurs $x, y \in E$. Alternativement, la matrice d'une projection orthogonale est symétrique. Par conséquent, une symétrie orthogobale est également autoadjointe
$$\sigma_F^* = \pi_F^* - \pi_{F^\perp}^* = \pi_F - \pi_{F^\perp} = \sigma_F.$$

On s'intéresse maintenant à la question de la réduction des endormorphismes ou matrices symétriques. Commençons par le cas de la dimension 2.

Exemple 1.34. On considère la matrice suivante
$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$
avec $a, b, c \in \mathbf{R}$, et on s'intéresse au cas $b \neq 0$ où la matrice n'est pas diagonale. Le polynôme caractéristique est donné par
$$\chi_M(t) = t^2 - (a+c)t + ac - b^2.$$
Son discriminant est
$$\Delta = (a+c)^2 - 4(ac - b^2) = (a-c)^2 + 4b^2 > 0$$
il y a donc deux racines réelles distinctes et la matrice M est donc diagonalisable
$$\lambda_\pm = \frac{1}{2}\big(a + c \pm \sqrt{(a-c)^2 + 4b^2}\big).$$
Soit $(x, y) \in \mathbf{R}^2 \setminus \{(0,0)\}$ un vecteur propre de M associé à la valeur propre λ_+, ainsi (x, y) est solution du système [1]
$$\begin{cases} (a - \lambda_+)x + by = 0 \\ bx + (c - \lambda_+)y = 0 \end{cases}$$
et comme
$$\lambda_+ + \lambda_- = a + c$$
le vecteur (x, y) est solution de
$$\begin{cases} (\lambda_- - c)x + by = 0 \\ bx + (\lambda_- - a)y = 0 \end{cases}$$
ce qui s'écrit encore
$$\begin{cases} (a - \lambda_-)y - bx = 0 \\ by - (\lambda_- - c)x = 0 \end{cases}$$
ce qui signifie que $(-y, x)$ est un vecteur propre associé à la valeur propre λ_-. Ainsi, si l'on note
$$P = \frac{1}{\sqrt{x^2 + y^2}} \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$
on constate que
$$P^{-1} = {}^{\mathrm{t}}P$$
et par conséquent
$$M = P \begin{pmatrix} \frac{1}{2}\big(a + c + \sqrt{(a-c)^2 + 4b^2}\big) & 0 \\ 0 & \frac{1}{2}\big(a + c - \sqrt{(a-c)^2 + 4b^2}\big) \end{pmatrix} {}^{\mathrm{t}}P.$$
Notons que si l'on choisit
$$\frac{x}{\sqrt{x^2 + y^2}} = \cos\theta, \qquad \frac{y}{\sqrt{x^2 + y^2}} = \sin\theta$$

1. Une solution explicite est par exemple $x = 1, y = (a - \lambda_+)/b$.

alors

$$P = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

est une matrice de rotation. Sur le cas de la dimension 2, on constate donc que les valeurs propres sont réelles, et que les vecteurs propres sont orthogonaux.

On dit qu'une matrice est orthogonale si ses ses vecteurs colonnes forment une base orthonormée. On étudiera plus en détail ces matrices par la suite.

Théorème 1.35. *Soit u un endomorphisme symétrique d'un espace euclidien, alors u est diagonalisable. Il existe une base orthornormée de vecteurs propres. Par conséquent, toute matrice symétrique à coefficients réels est diagonalisable., avec une matrice de passage orthogonale.*

Lemme 1.36. *Les valeurs propres de u sont réelles.*

Démonstration. Commençons par montrer qu'un endomorphisme auto-adjoint ne peut avoir de valeur propre purement imaginaire. Soit $\mu \in \mathbf{R}^*$, supposons que l'on ait

$$u^2(x) + \mu^2 x = 0$$

alors en prenant le produit scalaire avec x

$$\|u(x)\|^2 + \mu^2\|x\|^2 = 0$$

ce qui entraîne $x = 0$. Ainsi l'endomorphisme $u^2 + \mu^2\mathrm{Id}_E$ est inversible et on en tire

$$\begin{aligned}|\chi_u(i\mu)|^2 = \chi_u(i\mu)\overline{\chi_u(i\mu)} &= \det(A + i\mu\mathrm{I}_n)\det(A - i\mu\mathrm{I}_n) \\ &= \det(A^2 + \mu^2\mathrm{I}_n) = \det\left(u^2 + \mu^2\mathrm{Id}_E\right) \neq 0.\end{aligned}$$

Ce qui signifie que $i\mu$ ne peut être valeur propre de u.

Soit λ une valeur propre d'un endomorphisme auto-adjoint u. L'endomorphisme $v = u - \mathrm{Re}(\lambda)\mathrm{Id}_E$ est auto-adjoint et admet $i\,\mathrm{Im}(\lambda)$ comme valeur propre, d'après la première partie de la preuve, on a forcément $\mathrm{Im}\,\lambda = 0$. ☐

Remarque 1.37. Une autre preuve (plus directe) consiste à considérer un vecteur propre $Z \in \mathbf{C}^n$ associée à la valeur propre $\lambda \in \mathbf{C}$ de la matrice symétrique A de u dans une base orthonormée

$$AZ = \lambda Z, \quad Z = (z_1, \dots, z_n) \neq 0$$

et de constater que

$$\lambda = \left(\sum_{j=1}^n |z_j|^2\right)^{-1} \sum_{j,k=1}^n a_{j,k} z_j \bar{z}_k$$

de sorte que

$$\bar{\lambda} = \left(\sum_{j=1}^n |z_j|^2\right)^{-1} \sum_{j,k=1}^n a_{j,k} \bar{z}_j z_k = \left(\sum_{j=1}^n |z_j|^2\right)^{-1} \sum_{j,k=1}^n a_{k,j} z_k \bar{z}_j = \lambda$$

implique que la valeur propre λ est réelle.

Le spectre $\sigma(u)$ de u est l'ensemble des valeurs propres de u, i.e. des racines du polynôme caractéristique χ_u de u. On introduit maintenant l'espace propre associée à une valeur propre $\lambda \in \sigma(u)$

$$E_\lambda = \ker\big(u - \lambda \mathrm{Id}_E\big).$$

Lemme 1.38. *Deux espaces propres E_λ et E_μ associés à deux valeurs propres distinctes $\lambda \neq \mu$ sont orthogonaux.*

Démonstration. Soient x, y deux vecteurs propres associés respectivement à la valeur propre λ et à la valeur propre $\mu \neq \lambda$ alors on a

$$\lambda\langle x, y\rangle = \langle u(x), y\rangle = \langle x, u(y)\rangle = \mu\langle x, y\rangle$$

et par conséquent

$$(\lambda - \mu)\langle x, y\rangle = 0$$

ce qui implique que x et y sont orthogonaux. □

Démonstration du théorème 1.35. On considère l'ensemble des valeurs propres (distinctes) de u

$$\sigma(u) = \big\{\lambda_1, \dots, \lambda_m\big\}$$

et le sous-espace vectoriel

$$F = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}.$$

On a la décomposition en somme direct

$$E = F \oplus F^\perp.$$

Il est clair que F est stable par u

$$u(F) \subset F$$

puisque si $(x_1, \dots, x_m) \in F_1 \times \cdots \times F_m$

$$u\bigg(\sum_{j=1}^m x_j\bigg) = \sum_{j=1}^m \lambda_j x_j \in F.$$

Ceci entraîne que $F^\perp$ est stable par u

$$u(F^\perp) \subset F^\perp$$

car si $x \in F^\perp$ et si $f \in F$

$$\langle u(x), f\rangle = \langle x, \underbrace{u(f)}_{\in F}\rangle = 0.$$

Ceci permet de dire que $v = u|_{F^\perp}$ est un endomorphisme de $F^\perp$, évidemment auto-adjoint. Si $F^\perp \neq \{0\}$ alors $v = u|_{F^\perp}$ est un endomorphisme auto-adjoint de $F^\perp$ qui est un sous-espace vectoriel de dimension ≥ 1. Le polynôme caractéristique de v admet au moins une racine, et toutes ses racines sont réelles puisque v est auto-adjoint. Il admet donc une valeur propre réelle, il existe ainsi $x \in F^\perp$ tel que

$$u(x) = \lambda x$$

ce qui veut dire $\lambda \in \sigma(u)$ et donc que $x \in F$, ce qui est absurde.

Si M est une matrice symétrique, l'endomorphisme associé

$$\begin{aligned} u : \mathbf{R}^n &\to \mathbf{R}^n \\ X &\mapsto MX \end{aligned}$$

est auto-adjoint pour le produit scalaire canonique $\langle X, Y\rangle = {}^{\mathrm{t}}XY$

$$\langle u(X), Y\rangle = {}^{\mathrm{t}}(MX)Y = {}^{\mathrm{t}}X\,{}^{\mathrm{t}}MY = {}^{\mathrm{t}}XMY = \langle X, u(Y)\rangle$$

il existe donc une base $(v_1, \ldots, v_n)$ de vecteurs propres orthonormée, et en notant P la matrice de passage de la base canonique $(e_1, \ldots, e_n)$ à cette base de vecteurs propres

$$Pe_j = v_j$$

on a

$$M = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1}$$

avec $\lambda_1, \ldots, \lambda_n$ les valeurs propres réelles de l'endomorphisme u et donc de la matrice M. On verra par la suite que le caractère orthogonale de la matrice de passage P implique

$$P^{-1} = {}^{\mathrm{t}}P.$$

Ce qui montre que la matrice M est diagonalisable (avec une matrice diagonale réelle et une matrice de passage orthogonale P)

$$M = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} {}^{\mathrm{t}}P.$$

Ceci termine la d'emonstration du théorème de diagonalisation des endomorphismes symétriques. $\square$

2. Groupe orthogonal

2.1. **Changement de bases orthonormées.** Dans un espace euclidien E, on considère deux bases orthonormées $\mathcal{B} = (e_1, \ldots, e_n)$ et $\mathcal{C} = (f_1, \ldots, f_n)$. On peut décomposer les vecteurs de la base $\mathcal{C}$ dans la base $\mathcal{B}$

$$f_j = \sum_{k=1}^{n} \langle f_j, e_k\rangle e_k$$

et on obitent ainsi la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{C}$ (dont les vecteurs colonnes sont les vecteurs $f_1, \ldots, f_n$ exprimés dans la base $\mathcal{B}$ de sorte que $f_j = Pe_j$)

$$P = \mathrm{Mat}_{\mathcal{B}}(\mathcal{C}) = (\langle f_j, e_k\rangle)_{1\leq j,k\leq n}.$$

Le fait que la base $\mathcal{C}$ soit orthonormée s'écrit

$$\delta_{j,k} = \langle f_j, f_k\rangle = \sum_{\ell=1}^{n} \langle f_j, e_\ell\rangle\langle f_k, e_\ell\rangle$$

c'est-à-dire
$$\mathrm{I}_n = P\,{}^{\mathrm{t}}P.$$
Ceci implique que $P^{-1} = {}^{\mathrm{t}}P$. On le voit également en intervertissant les rôles de $\mathcal{B}$ et $\mathcal{C}$
$$P^{-1} = \mathrm{Mat}_{\mathcal{C}}(\mathcal{B}) = (\langle e_j, f_k\rangle)_{1\leq j,k\leq n} = {}^{\mathrm{t}}P.$$
On a alors également
$${}^{\mathrm{t}}PP = \mathrm{I}_n.$$
Inversement, soit P une matrice vérifiant cette relation, alors la famille des vecteurs colonnes $(f_1, \ldots, f_n)$ de la matrice
$$f_j = Pe_j$$
est orthonormée pour le produit scalaire canonique
$$\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j = {}^{\mathrm{t}}XY \quad x = \sum_{j=1}^{n} x_j e_j, \quad y = \sum_{j=1}^{n} y_j e_j$$
(pour lequel la base canonique est orthonormée)
$$\langle f_j, f_k\rangle = {}^{\mathrm{t}}(Pe_j)(Pe_k) = {}^{\mathrm{t}}e_j\,{}^{\mathrm{t}}PPe_k = {}^{\mathrm{t}}e_j e_k = \delta_{j,k}.$$
Cela signifie que les colonnes d'une matrice orthogonale forment une base orthonormée. Comme la transposée d'une matrice orthogonale est elle-même une matrice orthogonale, les lignes d'une matrice orthogonale forment également une base orthonormée.

Définition 2.1. *On dit qu'une matrice $U \in \mathrm{M}(n, \mathbf{R})$ est orthogonale si*
$${}^{\mathrm{t}}UU = U\,{}^{\mathrm{t}}U = \mathrm{I}_n$$
autrement dit si U est une matrice de passage de la base canonique à une base orthonormée (ou encore une matrice dont les colonnes ou les lignes forment une base orthonormée). On note
$$\mathrm{O}(n, \mathbf{R}) = \left\{U \in \mathrm{M}(n, \mathbf{R}) : {}^{\mathrm{t}}UU = \mathrm{I}_n\right\}$$
l'ensemble des matrices orthonomées.

Lemme 2.2. *Les matrices orthogonales forment un sous-groupe du groupe linéaire $\mathrm{GL}(n, \mathbf{R})$ (groupe des matrices inversibles) stable par adjonction.*

Démonstration. L'élément neutre est la matrice identité qui est une matrice orthogonale et le produit de deux matrices orthogonales U et V est une matrice orthogonale
$${}^{\mathrm{t}}(UV)(UV) = {}^{\mathrm{t}}V\,{}^{\mathrm{t}}UUV = {}^{\mathrm{t}}V\,{}^{\mathrm{t}}V = \mathrm{I}_n.$$
et l'inverse $U^{-1} = {}^{\mathrm{t}}U$ d'une matrice orthogonale est orthogonale. $\square$

2.2. **Généralités.**

Définition 2.3. *Dans un espace vectoriel normé, une isométrie (vectorielle) est un endomorphisme $u \in L(E)$ qui préserve la norme*

$$\|u(x)\| = \|x\|.$$

Lemme 2.4. *Un endomorphisme $u \in L(E)$ est une isométrie si et seulement s'il conserve le produit scalaire*

$$\langle u(x), u(y) \rangle = \langle x, y \rangle, \quad x, y \in E$$

ou de manière équivalente si et seulement si

$$u^* \circ u = u \circ u^* = \mathrm{Id}_E$$

ou encore de manière équivalente si et seulement si l'image d'une base orthonormée de E est une base orthonormée de E.

Démonstration. Soit $u \in L(E)$ une isométrie alors

$$\|u(x+y)\|^2 = \|u(x)\|^2 + \|u(y)\|^2 + 2\langle u(x), u(y) \rangle = \|x\|^2 + \|y\|^2 + 2\langle u(x), u(y) \rangle$$

et comme u est une isométrie,

$$\|u(x+y)\|^2 = \|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle$$

on en tire en comparant les deux relations

$$\langle u(x), u(y) \rangle = \langle x, y \rangle, \quad x, y \in E$$

ce qui est équivalent à

$$\langle x, u^* \circ u(y) \rangle = \langle x, y \rangle, \quad x, y \in E$$

soit $u^* \circ u = \mathrm{Id}_E$, i.e. $u^* = u^{-1}$. La réciproque est claire. Enfin, remarquons que si $(e_1, \dots, e_n)$ est une base orthonormée de E alors c'est le cas de $(u(e_1), \dots, u(e_n))$ car

$$\langle u(e_j), u(e_k) \rangle = \langle e_j, e_k \rangle = \delta_{j,k}.$$

Réciproquement si u est un endomorphisme qui envoie une base orthonormée $(e_1, \dots, e_n)$ de E sur une base orthonormée $(u(e_1), \dots, u(e_n))$ de E, alors on a

$$\|u^*(x)\|^2 = \sum_{j=1}^{n} |\langle u^*(x), e_j \rangle|^2 = \sum_{j=1}^{n} |\langle x, u(e_j) \rangle|^2 = \|x\|^2$$

ce qui permet de conclure que u^*, et donc u, est un endomorphisme orthogonal. $\square$

On appelle également une isométrie vectorielle un encomorphisme orthogonal (ou un automorphisme orthogonal) ou une transformation unitaire. On note

$$\mathrm{O}(E) = \left\{ u \in L(E) : u^* \circ u = \mathrm{Id}_E \right\}$$

l'ensemble des endomorphismes orthogonaux de E. Le choix d'une base orthonormée permet de construire un isomorphisme entre le groupe des isométries et le groupe des matrices orthogonales

$$\mathrm{O}(E) \simeq \mathrm{O}(n, \mathbf{R}).$$

En effet, si l'on choisit une base orthonormée $\mathcal{B}$ de E alors un endomorphisme orthogonal u envoie $\mathcal{B}$ sur une base orthonormée de E, et les colonnes de la matrice de u dans $\mathcal{B}$ forment donc une base orthonormée, ce qui signifie que c'est une matrice orthogonale. Par conséquent,

$$\begin{aligned} \varphi : \mathrm{O}(E) &\to \mathrm{O}(n, \mathbf{R}) \\ u &\mapsto \mathrm{Mat}_{\mathcal{B}}(u) \end{aligned}$$

est un isomorphisme de groupe d'inverse donnée par

$$\varphi^{-1}(U)(x) = \sum_{j=1}^{n} (UX)_j e_j, \quad X = \begin{pmatrix} \langle x, e_1 \rangle \\ \vdots \\ \langle x, e_n \rangle \end{pmatrix}.$$

Exemple 2.5. Les symétries orthogonales σ_F

$$\sigma_F^* = \sigma_F = \sigma_F^{-1}$$

sont des endomorphismes orthogonaux (autoadjoints). Dans $E = \mathbf{R}^2$, une matrice de rotation

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \theta \in [0, 2\pi]$$

est une matrice orthogonale.

Résumons la situation : les caractérisations suivantes sont équivalentes

(i) $u \in \mathrm{O}(E)$

(ii) $u^* \circ u = u \circ u^* = \mathrm{Id}_E$

(iii) $\|u(x)\| = \|x\|$ pour tout $x \in E$,

(iv) $\langle u(x), u(y) \rangle = \langle x, y \rangle$ pour tous $x, y \in E$.

(v) u envoie une base orthonormée de E sur une base orthonormée de E.

et en ce qui concerne les matrices

(i) $U \in \mathrm{O}(n, \mathbf{R})$

(ii) ${}^{\mathrm{t}}UU = U\,{}^{\mathrm{t}}U = \mathrm{I}_n$,

(iii) ${}^{\mathrm{t}}(UX)(UX) = {}^{\mathrm{t}}XX$ pour tous $X \in \mathbf{R}^n$,

(iv) ${}^{\mathrm{t}}(UX)(UY) = {}^{\mathrm{t}}XY$ pour tous $X, Y \in \mathbf{R}^n$,

(v) les colonnes de U forment une base orthonormée de $\mathbf{R}^n$.

Avec ces caractérisations, on voit facilement

$$(\det u)^2 = (\det U)^2 = \det({}^{\mathrm{t}}UU) = 1$$

et on peut alors distinguer le sous-ensemble suivant

$$\mathrm{SO}(E) = \big\{u \in L(E) : \det u = 1\big\}$$

et du côté matriciel

$$\mathrm{SO}(n, \mathbf{R}) = \big\{U \in \mathrm{M}(n, \mathbf{R}) : {}^{\mathrm{t}}UU = \mathrm{I}_n\big\}.$$

À nouveau, c'est un sous-groupe ; pour s'en rendre compte, il suffit de remarquer que c'est l'intersection du sous-groupe $\mathrm{O}(n, \mathbf{R})$ de $\mathrm{GL}(n, \mathbf{R})$ et du noyau du morphisme de groupe $\det : (\mathrm{GL}(n, \mathbf{R}), \times) \to (\mathbf{R}^*, \times)$.

Lemme 2.6. *Les valeurs propres*[2] *d'une matrice orthogonale sont contenus dans le disque* $\mathbf{U} = \{\lambda \in \mathbf{C} : |\lambda| = 1\}$ *de* $\mathbf{C}$

$$\sigma(U) \subset \mathbf{U}.$$

Démonstration. On commence par observer

$$(U - \lambda \mathrm{I}_n)({}^{\mathrm{t}}U + \bar{\lambda}) = (1 - |\lambda|^2)\mathrm{I}_n + \bar{\lambda} U - \lambda\, {}^{\mathrm{t}}U$$

de sorte que si $|\lambda| \neq 1$, en passant au déterminant

$$\chi_u(\lambda)\overline{\chi_u(-\lambda)} = (1 - |\lambda|^2)^n \det(\mathrm{I}_n + A)$$

avec

$$A = \bar{\mu} U - \mu\, {}^{\mathrm{t}}U, \quad \mu = \frac{\lambda}{1 - |\lambda|^2} \in \mathbf{C}.$$

Ainsi, il suffit de vérifier que $\mathrm{I}_n + A$ est inversible pour en déduire que λ n'est pas racine du polynôme caractéristique de u lorsque $|\lambda| \neq 1$. Remarquons que

$$A^* = {}^{\mathrm{t}}\bar{A} = -A$$

et ainsi

$$|\det(\mathrm{I}_n + A)|^2 = \det((\mathrm{I}_n + A^*)(\mathrm{I}_n + A)) = \det(I_n + A^* A) > 0.$$

En effet si $Z \in \mathbf{C}^n$ est tel que

$$Z = -A^* A Z$$

alors

$$\sum_{j=1}^n |z_j|^2 = {}^{\mathrm{t}}\bar{Z} Z = {}^{\mathrm{t}}\overline{AZ} A Z = -\sum_{j=1}^n \left| \sum_{k=1}^n a_{j,k} z_k \right|^2 \leq 0$$

implique $Z = 0$. □

Autre démonstration. Soit $\lambda \in \mathbf{C}$ une valeur propre de U, il existe un vecteur $z \in \mathbf{C}^n$ tel que

$$\lambda z_j = \sum_{k=1}^n u_{j,k} z_k$$

et par conséquent

$$|\lambda|^2 |z_j|^2 = \sum_{k,\ell=1}^n u_{j,k} u_{j,\ell} z_k \bar{z}_\ell$$

ce qui donne

$$|\lambda|^2 = \left(\sum_{j=1}^n |z_j|^2 \right)^{-1} \sum_{k,\ell=1}^n \underbrace{\left(\sum_{j=1}^n u_{j,k} u_{j,\ell} \right)}_{=\delta_{k,\ell}} z_k \bar{z}_\ell = 1$$

et la valeur propre λ est donc de module 1. □

Lemme 2.7. *Soit* $u \in \mathrm{O}(E)$ *et soit* F *un sous-espace stable par* F *alors on a*

$$u(F) = F, \quad u(F^\perp) = F^\perp.$$

2. Les valeurs propres complexes ou les racines du polynôme caractéristique.

Démonstration. Montrons d'abord que $F = u(F)$. Pour cela, il suffit de constater que $\dim F = \dim u(F)$ puisque une base orthonormée de F est envoyée par u sur une famille orthonormée de F. Ceci implique

$$F = u^*(u(F)) = u^*(F).$$

Soit $x \in F^\perp$ alors pour tout $f \in F$

$$\langle u(x), f\rangle = \langle u(x), u(u^*(f))\rangle = \langle x, u^*(f)\rangle = 0$$

puisque $u^*(f) \in F$. On en déduit que $F^\perp$ est stable par u et ainsi par le raisonnement du début de cette preuve que $u(F^\perp) = F^\perp$. □

2.3. **En dimension 2.** En dimension $n = 2$, il est relativement facile de déterminer toutes les matrices du groupe spécial orthogonal

$$U = \begin{pmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{pmatrix}.$$

Les relations

$$u_{j,1}^2 + u_{j,2}^2 = 1$$

entraînent

$$U = \begin{pmatrix} \cos\theta_1 & \cos\theta_2 \\ \sin\theta_1 & \sin\theta_2 \end{pmatrix}.$$

Le fait que les deux colonnes sont orthogonales

$$\cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2 = \cos(\theta_1 - \theta_2) = 0$$

et le fait que le déterminant vaut 1

$$\det U = \cos\theta_1\sin\theta_2 - \sin\theta_1\cos\theta_2 = \sin(\theta_1 - \theta_2) = 1$$

impliquent

$$\theta_1 - \theta_2 = \frac{\pi}{2} \mod 2\pi$$

et ainsi toute matrice du groupe spécial orthogonal est une matrice de rotation

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = R_\theta.$$

On a donc

$$\mathrm{SO}(2, \mathbf{R}) = \left\{ R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbf{R} \right\}$$

avec les formules

$$R_\theta R_\omega = R_{\theta+\omega}, \quad R_\theta^{-1} = {}^{\mathrm{t}}R_\theta = R_{-\theta}.$$

Ce qui montre que l'application

$$\begin{aligned} (\mathbf{R}/(2\pi\mathbf{Z}), +) &\to \mathrm{SO}(2, \mathbf{R}) \\ \theta &\mapsto R_\theta \end{aligned}$$

est un isomorphisme de groupe. Notons que le polynôme caractéristique d'une rotation R_θ s'écrit

$$\chi_{R_\theta}(t) = t^2 - 2t\cos\theta + 1$$

et que les valeurs propres sont donc

$$\sigma(R_\theta) = \left\{ \mathrm{e}^{i\theta}, \mathrm{e}^{-i\theta} \right\} \subset \mathbf{U}.$$

Remarque 2.8. Une particularité de la dimension $n = 2$ est qu'il est facile de déterminer un vecteur orthogonal à un vecteur $x \neq 0$ de même norme

$$x_\perp = R_{\pi/2}(x) = (-y, x)$$

le seul autre vecteur orthogonal de même norme est $-x_\perp$.

Pour déterminer l'ensemble des matrices orthogonales, il suffit de considérer la matrice orthogonale

$$S = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \det S = -1$$

(qui est la symétrie orthogonale par rapport à l'axe des ordonnées), en effet si $U \in \mathrm{O}(2, \mathbf{R})$ a pour déterminant -1, alors

$$SU \in \mathrm{SO}(2, \mathbf{R})$$

et donc SU est une rotation R_θ donc

$$U = S(SU) = SR_\theta = \begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Finalement, on a déterminé le groupe des matrices orthogonales 2×2

$$\mathrm{O}(2, \mathbf{R}) = \left\{ R_\theta = \begin{pmatrix} \varepsilon\cos\theta & -\varepsilon\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbf{R},\, \varepsilon \in \{\pm 1\} \right\}.$$

La traduction est la suivante pour les endomorphismes : un endomorphisme orthogonal est soit une rotation, soit la composée d'une symétrie et d'une rotation.

Remarquons pour terminer que si

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

alors comme $J^2 = -\mathrm{I}_2$

$$\exp(\theta J) = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p)!} \theta^{2p} \mathrm{I}_2 + \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!} \theta^{2p+1} J = \cos\theta\, \mathrm{I}_2 + \sin\theta\, J = R_\theta.$$

L'algèbre des matrices symétriques

$$\mathfrak{so}(2, \mathbf{R}) = \left\{ M \in \mathrm{M}(2, \mathbf{R}) : {}^{\mathrm{t}}M = -M \right\}$$

est la droite vectorielle engendrée par J

$$\mathfrak{so}(2, \mathbf{R}) = \mathrm{vect}\,(J).$$

Théorème 2.9. *L'exponentielle*

$$\exp : \mathfrak{so}(2, \mathbf{R}) \to \mathrm{SO}(2, \mathbf{R})$$

est une surjection.

2.4. **En dimension 3.** Le cas de la dimension $n = 3$ est à peine plus compliqué. Soit U une matrice orthogonale de déterminant 1. Les valeurs propres de U sont de module 1, leur produit vaut 1 et comme le polynôme caractéristique est à coefficients réels, il y a deux racines qui sont conjuguées

$$\lambda\mu\bar{\mu} = \lambda|\mu|^2 = \lambda = 1$$

donc 1 est valeur propre et le spectre est de la forme

$$\sigma(U) = \left\{1, \mathrm{e}^{i\theta}, \mathrm{e}^{-i\theta}\right\}$$

avec $\theta \in [0, 2\pi]$. Comme 1 est valeur propre, il existe un vecteur propre $\omega \in \mathbf{R}^3$ (par exemple unitaire)

$$U\omega = \omega$$

l'orthogonal $F = \mathrm{vect}(\omega)^\perp$ de la droite engendrée par e est stable par l'endomorphisme $u : \mathbf{R}^3 \to \mathbf{R}^3$ associé à U

$$u(F) \subset F$$

car pour tout $x \in F$

$$\langle u(x), \omega\rangle = \langle u(x), u(\omega)\rangle = \langle x, \omega\rangle = 0.$$

On peut alors considérer la restriction

$$v = u|_F$$

qui est un endomorphisme orthogonal sur F. Comme F est de dimension 2, la matrice V de v dans n'importe quelle base orthonormée de F est une matrice orthogonal de taille 2×2

$$V \in \mathrm{SO}(2, \mathbf{R})$$

et c'est donc d'après le paragraphe précédent une rotation

$$V = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

On en déduit donc qu'il existe une base orthonormée dans laquele, la matrice de u s'écrit

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

On en tire alors

$$\mathrm{SO}(3, \mathbf{R}) = \left\{ P \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} {}^{\mathrm{t}}P : \theta \in \mathbf{R},\ P \in \mathrm{O}(3, \mathbf{R}) \right\}.$$

Ceci n'est pas vraiment satisfaisant dans la mesure où l'on n'a déterminé une expression d'une matrice de $\mathrm{SO}(3, \mathbf{R})$ que modulo la conjugaison par une matrice orthogonale.

L'exponentielle d'une matrice antisymétrique est unitaire car

$${}^{\mathrm{t}}(\exp A) \exp A = \exp({}^{\mathrm{t}}A) \exp A = \exp(-A) \exp A = \mathrm{I}_n.$$

L'algèbre des matrices antisymétriques

$$\mathfrak{so}(3,\mathbf{R}) = \left\{M \in \mathrm{M}(3,\mathbf{R}) : {}^{\mathrm{t}}M = M\right\}$$

est de dimension 3 et une base est donnée par

$$J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad K = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad L = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Cette base est orthonormée pour le produit scalaire sur $\mathrm{M}(n,\mathbf{R})$ donné par

$$\langle A|B\rangle = \frac{1}{2}\mathrm{tr}({}^{\mathrm{t}}AB).$$

Soit $\omega \in \mathbf{R}^3$ un vecteur unitaire

$$\omega_1^2 + \omega_2^2 + \omega_3^2 = 1$$

on considère la matrice antisymétrique (de norme 1)

$$\begin{aligned} M(\omega) &= \omega_1 J + \omega_2 K + \omega_3 L \\ &= \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}. \end{aligned}$$

Son carré vaut

$$M(\omega)^2 = \begin{pmatrix} -\omega_2^2 - \omega_3^2 & \omega_1\omega_2 & \omega_1\omega_3 \\ \omega_1\omega_2 & -\omega_1^2 - \omega_3^2 & \omega_2\omega_3 \\ \omega_1\omega_3 & \omega_2\omega_3 & -\omega_1^2 - \omega_2^2 \end{pmatrix} = -\mathrm{I}_3 + P(\omega)$$

où $P(\omega)$ est la matrice

$$P(\omega) = \begin{pmatrix} \omega_1^2 & \omega_1\omega_2 & \omega_1\omega_3 \\ \omega_1\omega_2 & \omega_2^2 & \omega_2\omega_3 \\ \omega_1\omega_3 & \omega_2\omega_3 & \omega_3^2 \end{pmatrix}$$

de la projection orthogonale sur la droite engendrée par ω

$$\begin{aligned} \pi_\omega : \mathbf{R}^3 &\to \mathbf{R}^3 \\ x &\mapsto \langle x, \omega\rangle\omega. \end{aligned}$$

Ainsi le carré de $M(\omega)$ est l'opposé de la (matrice de) projection sur $\omega^\perp$. Notons que chacune des colonnes de $M(\omega)$ est orthogonale à ω et ainsi

$$M(\omega)P(\omega) = 0$$

d'où l'on déduit

$$M(\omega)^3 = (-\mathrm{I}_3 + P(\omega))M(\omega) = -M(\omega).$$

On peut alors calculer l'exponentielle

$$\begin{aligned} \exp(\theta M(\omega)) &= \mathrm{I}_3 + \left(\sum_{p=1}^{\infty} \frac{(-1)^p}{(2p)!}\theta^{2p}\right)(\mathrm{I}_3 - P(\omega)) + \left(\sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!}\theta^{2p+1}\right)M(\omega) \\ &= \mathrm{I}_3 + (\cos\theta - 1)(\mathrm{I}_3 - P(\omega)) + \sin\theta\, M(\omega) \\ &= \mathrm{I}_3 + \sin\theta\, M(\omega) + (1 - \cos\theta)M(\omega)^2. \end{aligned}$$

Remarque 2.10. Soit $A \in \mathfrak{so}(3,\mathbf{R})$ une matrice non nulle, alors la matrice de norme 1

$$M = \frac{1}{\sqrt{\langle A|A\rangle}} A$$

se décompose dans la base orthonormée (J, K, L)

$$M = M(\omega), \quad \omega_1 = \frac{\langle A|J\rangle}{\sqrt{\langle A|A\rangle}}, \quad \omega_2 = \frac{\langle A|K\rangle}{\sqrt{\langle A|A\rangle}}, \quad \omega_3 = \frac{\langle A|L\rangle}{\sqrt{\langle A|A\rangle}},$$

et on a donc

$$\exp A = \exp \theta M(\omega), \quad \theta = \sqrt{\langle A|A\rangle}.$$

Le vecteur ω est vecteur propre de $\exp(\theta M(\omega))$ associé à la valeur propre 1

$$\exp(\theta M(\omega))\omega = \omega$$

puisque

$$M(\omega)\omega = 0, \quad M^2(\omega)\omega = 0.$$

Soit alors $x \in \omega^\perp$, on a

$$\langle x, M(\omega)x\rangle = 0$$

car $M(\omega)$ est antisymétrique

$$\langle x, M(\omega)x\rangle = -\langle M(\omega)x, x\rangle = -\langle x, M(\omega)x\rangle$$

ainsi que

$$\|M(\omega)x\|^2 = -\langle M(\omega)^2 x, x\rangle = \|x\|^2$$

et en outre

$$\begin{aligned} \exp(\theta M(\omega))x &= \cos\theta\, x + \sin\theta M(\omega)\, x \\ \exp(\theta M(\omega))M(\omega)x &= -\sin\theta x + \cos\theta M(\omega)x. \end{aligned}$$

Ainsi si l'on choisit $x \in \omega^\perp$ unitaire, dans la base orthonormée $(\omega, x, M(\omega)x)$ la matrice de $\exp(\theta M(\omega))$ prend la forme

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}.$$

Enfin, on a

$$\cos\theta = \frac{1}{2} \operatorname{tr} \exp(\theta M(\omega)).$$

Théorème 2.11. *L'application*

$$\begin{aligned} S^2 \times [0, 2\pi] &\to \mathrm{SO}(3,\mathbf{R}) \\ (\omega, \theta) &\mapsto \exp(\theta M(\omega)) \end{aligned}$$

est surjective et par conséquent également

$$\begin{aligned} \exp : \mathfrak{so}(3,\mathbf{R}) &\to \mathrm{SO}(3,\mathbf{R}) \\ A &\mapsto \exp A. \end{aligned}$$

En outre, toute rotation R de $\mathrm{SO}(3,\mathbf{R})$ vérifie la formule de Rodrigues

$$R = \exp(\theta M(\omega)) = \mathrm{I}_3 + \sin\theta\, M(\omega) + (1 - \cos\theta) M(\omega)^2$$

où $\theta = \arccos\left(\frac{1}{2}\operatorname{tr} R - 1\right)$ *est l'angle de la rotation et* $\omega \in S^2$ *est l'axe de la rotation.*

Démonstration. Soit U une matrice de $\mathrm{SO}(3, \mathbf{R})$, soit ω un vecteur propre unitaire associé à la valeur propre 1 et

$$\theta = \arccos \operatorname{tr} U \in [0, \pi]$$

et soit $x \in \omega^\perp$ un vecteur unitaire orthogonal à ω, alors si 'on considère la matrice bloc U correspondant à la restriction au plan $(x, M(\omega)x)$, c'est une matrice de rotation, de trace $\operatorname{Tr} U - 1 = 2\cos\theta$, donc de la forme R_θ ou $R_{-\theta}$. Si c'est la matrice de rotation $R_{-\theta}$, il suffit de changer ω en $-\omega$ puisque $\exp(\theta M(\omega) = \exp(-\theta M(-\omega))$. Dans la base orthonomrée $(\omega, x, M(\omega)x)$, U et $\exp(\theta M(\omega))$ ont donc même matrice ainsi

$$U = \exp \theta M(\omega) = \exp(-\theta M(-\omega))$$

ce qui montre la surjectivité de la première application, et comme

$$\theta M(\omega) \in \mathfrak{so}(3, \mathbf{R})$$

également la surjectivité de la seconde. $\square$

Définition 2.12. *Le produit vectoriel de deux vecteurs* $x \in \mathbf{R}^3$ *et* $y \in \mathbf{R}^3$ *est l'unique vecteur* $x \wedge y \in \mathbf{R}^3$ *tel que*

$$\langle x \wedge y, z\rangle = \det(x, y, z)$$

pour tous $z \in \mathbf{R}^3$.

Il nous faut prouver qu'un tel vecteur existe et est unique. Il vérifie

$$(x \wedge y)_j = \langle x \wedge y, e_j\rangle = \det(x, y, e_j)$$

ce qui implique

$$x \wedge y = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

Inversement, ce vecteur vérifie bien la définition du produit vectoriel

$$\langle x \wedge y, z\rangle = \sum_{j=1}^{3} z_j \langle x \wedge y, e_j\rangle = \sum_{j=1}^{3} z_j \det(x, y, e_j) = \det(x, y, z).$$

La définition du produit vectoriel implique

$$\langle x \wedge y, x\rangle = \det(x, y, x) = 0, \quad \langle x \wedge y, y\rangle = \det(x, y, y) = 0$$

que le vecteur $x \wedge y$ est orthogonal à x et y

$$x \wedge y \perp x, \quad x \wedge y \perp y.$$

De même, on voit que si x et y sont liés alors

$$\langle x \wedge y, z\rangle = \det(x, y, z) = 0$$

pour tout $z \in \mathbf{R}^3$ et donc le produit vectoriel est nul $x \wedge y = 0$.

Lemme 2.13. *Le produite vectoriel vérifie les propriétés suivantes*

(*i*) *L'application* $(x, y) \mapsto x \wedge y$ *est bilinéaire,* (*ii*) $y \wedge x = -x \wedge y$ *(antisymétrie) pour tout* $x, y \in \mathbf{R}^3$*,*

(*iii*) $x \wedge (y \wedge z) = \langle x, z\rangle y - \langle x, y\rangle z$ *pour tous* $x, y, z \in \mathbf{R}^3$*,*

(*iv*) $x \wedge (y \wedge z) + y \wedge (z \wedge x) + z \wedge (x \wedge y) = 0$ *(identité de Jacobi) pour tous* $x, y, z \in \mathbf{R}^3$*.*

(*v*) $\|x \wedge y\|^2 = \|x\|^2\|y\|^2 \sin^2\theta$ *lorsque* $\langle x, y\rangle = \|x\|\|y\|\cos\theta$*.*

Démonstration.

(*i*) Ceci découle de la bilinéarité de $(x, y) \mapsto (x \wedge y)_j = \det(x, y, e_j)$ pour $1 \leq j \leq 3$.

(*ii*) La définition du produit vectoriel implique pour tout vecteur $z \in \mathbf{R}^3$

$$\langle x \wedge y + y \wedge x, z\rangle = \det(x, y, z) + \det(y, x, z) = 0$$

et ainsi $x \wedge y + y \wedge x = 0$.

(*iii*) Remarquons que l'on peut supposer que (y, z) est un système libre car dans le cas contraire, les deux membres de l'égalité sont nuls. Dans ce cas, l'espace

$$F = \mathrm{vect}(y, z)$$

est un plan et son orthogonal $F^\perp$ une droite. Comme (y, z) est libre, on peut utiliser le procédé d'orthogonalisation de Gram-Schmidt

$$w = z - \frac{\langle z, y\rangle}{\|y\|^2} y$$

pour obtenir une base orthogonale (y, w) de F. Comme on a

$$y \wedge (z + \lambda y) = y \wedge z$$

et pareillement

$$\langle x, z + \lambda y\rangle y - \langle x, y\rangle(z + \lambda y) = \langle x, z\rangle y - \langle x, y\rangle z$$

l'identité à démontrer prend alors la forme

$$x \wedge (y \wedge w) = \langle x, w\rangle y - \langle x, y\rangle w.$$

Comme on l'a remarqué après la construction du produit vectoriel

$$y \wedge w \in F^\perp$$

et ainsi $(y, w, y \wedge w)$ est une base orthogonale avec

$$F = \mathrm{vect}(y, w), \quad F^\perp = \mathrm{vect}(y \wedge w).$$

En outre, $(y/\|y\|, w/\|w\|, y \wedge w/\|y \wedge w\|)$ est une base orthonormée donc

$$\det\left(\frac{y}{\|y\|}, \frac{w}{\|w\|}, \frac{y \wedge w}{\|y \wedge w\|}\right) = \pm 1$$

car c'est le déterminant d'une matrice orthogonale (la matrice de passage de la base canonique à une autre base orthonormée), on en tire alors

$$\det(y, w, y \wedge w) = \|y \wedge w\|^2 = \|y\|\|w\|\|y \wedge w\|$$

et ainsi

$$\|y \wedge w\| = \|y\|\|w\|$$

ce qui implique

$$\langle y \wedge (y \wedge w), w \rangle = \det(y, y \wedge w, w) = -\det(y, w, y \wedge y) = -\|y\|^2 \|w\|^2.$$

Comme $y \wedge (y \wedge w)$ est orthogonal à y et $y \wedge w$, il est colinéaire à w et avec le calcul précédent

$$y \wedge (y \wedge w) = -\|y\|^2 w.$$

De même a-t-on, par permutation

$$w \wedge (y \wedge w) = \|w\|^2 y.$$

On peut alors décomposer tout vecteur $x \in \mathbf{R}^3$ dans la base $(y, w, y \wedge w)$

$$x = \lambda y + \mu w + \varrho y \wedge w$$

et on obtient

$$x \wedge (y \wedge w) = -\lambda \|y\|^2 w + \mu \|w\|^2 y.$$

Or les coordonnées de x sont données par

$$\lambda = \frac{\langle x, y \rangle}{\|y\|^2}, \quad \mu = \frac{\langle x, w \rangle}{\|w\|^2},$$

aussi obtient-on

$$x \wedge (y \wedge w) = \langle x, w \rangle y - \langle x, y \rangle w$$

l'identité à démontrer.

(iv) On utilise la relation précédente

$$\begin{aligned} x \wedge (y \wedge z) &= \langle x, z \rangle y - \langle x, y \rangle z \\ y \wedge (z \wedge x) &= \langle x, y \rangle z - \langle y, z \rangle x \\ z \wedge (x \wedge y) &= \langle y, z \rangle x - \langle x, z \rangle y \end{aligned}$$

et en sommant, les termes se simplifient tous.

(v) Un cas particulier de l'identité (iii)

$$x \wedge (x \wedge y) = \langle x, y \rangle x - \|x\|^2 y$$

donne

$$\begin{aligned} \|x \wedge y\|^2 &= \det(x, y, x \wedge y) = -\det(x, x \wedge y, y) = -\langle x \wedge (x \wedge y), y \rangle \\ &= \|x\|^2 \|y\|^2 - \langle x, y \rangle^2 = \|x\|^2 \|y\|^2 (1 - \cos^2 \theta) = \|x\|^2 \|y\|^2 \sin^2 \theta. \end{aligned}$$

Ce qui termine la preuve des propriétés du produit vectoriel. □

La matrice de l'endomorphisme

$$\begin{aligned} \mathbf{R}^3 &\to \mathbf{R}^3 \\ x &\mapsto x \wedge \omega \end{aligned}$$

est la matrice antisymétrique

$$\omega_1 J + \omega_2 K + \omega_3 L = M(\omega).$$

2.5. **En dimension quelconque.** Soit U une matrice orthogonale à coefficients réels de taille n
$$U \in \mathrm{O}(n, \mathbf{R}).$$

Théorème 2.14. *Pour tout endomorphisme $u \in \mathrm{SO}(E)$ il existe une base orthonormée $\mathcal{B}$ de $\mathbf{R}^n$ dans laquelle sa matrice prend la forme*
$$\begin{pmatrix} \mathrm{I}_p & 0 & 0 & \cdots & 0 \\ 0 & -\mathrm{I}_{2q} & 0 & \cdots & 0 \\ 0 & 0 & R_{\theta_1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & R_{\theta_m} \end{pmatrix}$$
avec
$$\theta_j \in]0, \pi[\cup]\pi, 2\pi[, \quad 1 \leq j \leq m.$$
et
$$n = p + 2q + 2m.$$
Notons que lorsque n est impair, on a forcément $p \geq 1$ et p est impair.

Corollaire 2.15. *Pour tout endomorphisme $u \in \mathrm{O}(E)$ il existe une base orthonormée $\mathcal{B}$ de $\mathbf{R}^n$ dans laquelle sa matrice prend la forme*
$$\begin{pmatrix} \mathrm{I}_p & 0 & 0 & \cdots & 0 \\ 0 & -\mathrm{I}_{q} & 0 & \cdots & 0 \\ 0 & 0 & R_{\theta_1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & R_{\theta_m} \end{pmatrix}$$
avec
$$\theta_j \in]0, \pi[\cup]\pi, 2\pi[, \quad 1 \leq j \leq m.$$
et
$$n = p + q + 2m.$$

Remarque 2.16. Le polynôme caractérstique de $u \in \mathrm{SO}(E)$ s'écrit
$$\begin{aligned} \chi_u(t) &= (t-1)^p \prod_{k=1}^{m} (t - \mathrm{e}^{i\theta_k})(t - \mathrm{e}^{-i\theta_k}) \\ &= (t-1)^p \prod_{k=1}^{m} (t^2 - 2\cos\theta_k + 1). \end{aligned}$$

Démonstration. L'endomorphisme
$$v = \frac{u + u^*}{2}$$
est auto-adjoint donc son spectre est constitué de valeurs propres réelles
$$\sigma(v) = \{\lambda_1, \dots, \lambda_N\}$$
et E se décompose comme la somme directe
$$E = F_1 \oplus \cdots \oplus F_N$$

des espaces propres
$$F_j = \ker(v - \lambda_j \mathrm{Id})$$
deux à deux orthogonaux
$$F_j \perp F_k \quad j \neq k.$$
Ces espaces propres sont stables par u, en effet on a
$$v \circ u = \frac{1}{2}(u^2 + \mathrm{Id}_E) = u \circ v$$
et donc si $x \in F_j$
$$v(u(x)) = u(v(x)) = u(\lambda_j x) = \lambda_j u(x).$$
Il suffit donc d'étudier les restrictions $u_j = u|_{F_j} : F_j \to F_j$.

Commençons par les cas où 1 ou -1 sont valeurs propres de u. Si 1 est valeur propre de u alors on a
$$\ker(u - \mathrm{Id}_E) = \ker(v - \mathrm{Id}_E)$$
car lorsque $u(x) = x$, on a
$$u^*(x) = u^*(u(x)) = x$$
et donc $v(x) = x$ et réciproquement si $v(x) = x$ alors
$$\|x\|^2 = \langle x, v(x)\rangle = \langle x, u(x)\rangle$$
par conséquent
$$\|u(x) - x\|^2 = 2\|x\|^2 - 2\langle u(x), x\rangle = 0$$
et ainsi $u(x) = x$. Il en est de même lorsque -1 est valeur propre de u
$$\ker(u + \mathrm{Id}_E) = \ker(v + \mathrm{Id}_E).$$
Dans le cas où $u \in \mathrm{SO}(E)$, la dimension de cet espace propre est paire car le déterminant de u qui est égal au produit des valeurs propres vaut 1.

On peut donc maintenant s'intéresser aux espaces F_j avec $\lambda_j \neq \pm 1$. Soit un vecteur propre $e_j \in F_j$ unitaire associé à la valeur propre λ_j alors on a
$$\lambda_j = \langle v(e_j), e_j\rangle = \langle u(e_j), e_j\rangle$$
et par Cauchy-Schwarz
$$|\lambda_j| \leq \|u(e_j)\| = 1$$
ce qui implique que la valeur propre λ_j est de la forme
$$\lambda_j = \cos\theta_j, \quad \theta_j \in]0, \pi[\cup]\pi, 2\pi[.$$
En outre, en appliquant u à la relation
$$2v(e_j) = u(e_j) + u^*(e_j) = 2\cos\theta_j e_j$$
on trouve
$$u^2(e_j) - 2\cos\theta_j u(e_j) + e_j = 0.$$
Les vecteurs e_j et $u(e_j)$ sont dans F_j et en outre
$$e_j + u(e_j) \quad \text{et} \quad e_j - u(e_j)$$

sont orthogonaux et non nuls car on est dans le cas où $\lambda_j \neq \pm 1$. Comme on a

$$\|e_j + u(e_j)\|^2 = 2(1 + \langle u(e_j), e_j\rangle) = 2(1+\cos\theta_j) = 4\cos^2\frac{\theta_j}{2}$$
$$\|e_j - u(e_j)\|^2 = 2(1 - \langle u(e_j), e_j\rangle) = 2(1-\cos\theta_j) = 4\sin^2\frac{\theta_j}{2}$$

il en découle que les vecteurs

$$e_j^+ = \frac{e_j + u(e_j)}{2\cos\frac{\theta_j}{2}}, \quad e_j^- = \frac{e_j - u(e_j)}{2\sin\frac{\theta_j}{2}},$$

sont orthonormés et

$$u(e_j^+) = \frac{(1+2\cos\theta_j)u(e_j) - e_j}{2\cos\frac{\theta_j}{2}} = \cos\theta_j e_j^+ - \sin\theta_j e_j^-$$
$$u(e_j^-) = \frac{(1-2\cos\theta_j)u(e_j) + e_j}{2\cos\frac{\theta_j}{2}} = \sin\theta_j e_j^+ + \cos\theta_j e_j^-$$

et la matrice de u restreinte au plan $P_j \mathrm{vect}(e_j^+, e_j^-) = \mathrm{vect}(e_j^+, e_j^-)$ est de la forme

$$\mathrm{Mat}_{(e_j^+, e_j^-)}(u|_{P_j}) = \begin{pmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{pmatrix} = R_{\theta_j}.$$

On décompose alors

$$F_j = P_j \oplus P_j^\perp$$

si $P_j^\perp = \{0\}$, la matrice de $u_j = u|_{F_j}$ est R_{θ_j}, et on a terminé. Et si $P_j^\perp \neq \{0\}$, on choisit un vecteur unitaire dans cet orthogonal et on réitère la construction précédente. Finalement, l'espace F_j est de diemnsion paire et il existe une base orthonorrmée dans laquelle la matrice de $u_j = u|_{F_j}$ est diagonale par blocs avec des blocs tous égaux à R_{θ_j}. $\square$

Corollaire 2.17. *L'exponentielle*

$$\exp : \mathfrak{so}(n, \mathbf{R}) \to \mathrm{SO}(n, \mathbf{R})$$

est une surjection.

Démonstration. Soit $U \in \mathrm{SO}(n, \mathbf{R})$, d'après le théorème de réduction précédent, il existe $P \in \mathrm{O}(n, \mathbf{R})$ telle que

$$U = P \begin{pmatrix} \mathrm{I}_p & 0 & \cdots & 0 \\ 0 & R_{\theta_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R_{\theta_m} \end{pmatrix} {}^{\mathrm{t}}P = P \left(\exp \begin{pmatrix} 0_p & 0 & \cdots & 0 \\ 0 & \theta_1 J & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \theta_m J \end{pmatrix} \right) {}^{\mathrm{t}}P$$
$$= \exp \left(P \begin{pmatrix} 0_p & 0 & \cdots & 0 \\ 0 & \theta_1 J & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \theta_m J \end{pmatrix} {}^{\mathrm{t}}P \right)$$

et ainsi la matrice

$$P\begin{pmatrix} 0_p & 0 & \cdots & 0 \\ 0 & \theta_1 J & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \theta_m J \end{pmatrix}{}^{\mathrm{t}}P$$

est un antécédent de U pour l'exponentielle. $\square$

2.6. **Orientation.**

Définition 2.18. *Dans un espace vectoriel euclidien, on dit que deux bases orthonormées ont la même orientation si la matrice de passage de l'une à l'autre a pour déterminant* 1.

Lemme 2.19. *Avoir la même orientation est une relation d'équivalence sur l'ensemble des bases orthonormées, ne comportant que deux classes d'équivalence.*

Démonstration. Si $\mathcal{B}$ et $\mathcal{C}$ sont deux bases orthonormées, notons

$$\mathcal{B} \sim \mathcal{C} \Leftrightarrow \det P = 1, \quad P = \mathrm{Mat}_{\mathcal{B}}\mathcal{C}.$$

Vérifions alors

(i) la réflexivité : $\mathcal{B} \sim \mathcal{B}$ car $\det \mathrm{I}_n = 1$,

(ii) la symétrie : si $\mathcal{B} \sim \mathcal{C}$ alors $\det P = 1$ et donc $\det P^{-1} = 1$ ce qui implique $\mathcal{C} \sim \mathcal{B}$,

(iii) la transitivité : si $\mathcal{B} \sim \mathcal{C}$ et si $\mathcal{C} \sim \mathcal{D}$ alors comme

$$\det(PQ) = \det P \det Q = 1$$

si P est la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$ et si Q est la matrice de passage de $\mathcal{C}$ à $\mathcal{D}$.

Soit une base orthonormée $\mathcal{B}_+ = (e_1, \dots, e_n)$, on note $\mathcal{B}_- = (-e_1, e_2, \dots, e_n)$, on considère les classes d'équivalence $[\mathcal{B}_+]$ de $\mathcal{B}_+$ et $[\mathcal{B}_-]$ la classe de $\mathcal{B}_-$. L'ensemble des bases orthonormées est partitioonnée sous la forme $[\mathcal{B}_+] \cup [\mathcal{B}_+]$. $\square$

Le choix d'une orientation sur un espace euclidien est le choix d'une classe d'équivalence, c'est à dire le choix d'une base orthonormée. Les bases directes sont celles qui sont dans la classe déquivalence de la base choisie, les bases indirectes sont celles qui sont dans l'autre calsse d'équivalence.

Les endomorphismes orthogonaux de $\mathrm{SO}(E)$ (ou matrices spéciales orthogonales) préservent l'orientation.

3. Formes quadratiques

3.1. **Premières définitions.** On s'intéresse à présent à des formes bilinéaires qui ne sont pas des produits scalaires et des formes quadratiques qui ne sont pas les carrés d'une norme.

Définition 3.1. *Une forme quadratique est une application* $q : E \to \mathbf{R}$ *tel qu'il existe une forme bilinéaire* $b : E \times E \to \mathbf{R}$ *telle que*

$$q(x) = b(x, x), \quad x \in E.$$

Toute forme bilinéaire $b : E \times E \to \mathbf{R}$ se décompose sous la forme

$$b(x,y) = \underbrace{\frac{1}{2}(b(x,y) + b(y,x))}_{=b_s(x,y)} + \underbrace{\frac{1}{2}(b(x,y) - b(y,x))}_{=b_a(x,y)}$$

où b_s est une forme bilinéaire symétrique

$$b_s(x,y) = b_s(y,x)$$

et b_a est une forme bilinéaire antisymétrique

$$b_a(y,x) = -b_a(x,y).$$

Du point de vue algébrique, cela revient à dire que l'espace (vectoriel) des formes bilinéaires est la somme directe de l'espace des formes bilinéaires symétriques et de l'espace des formes bilinéaires antisymétriques. Du point de vue matriciel, cela s'écrit

$$\mathrm{M}(n,\mathbf{R}) = \mathfrak{so}(n,\mathbf{R}) \oplus \mathrm{S}(n,\mathbf{R})$$

où $\mathrm{S}(n,\mathbf{R})$ est l'espace vectoriel des matrices symétriques. On a

$$q(x+y) = q(x) + q(y) + b(x,y) + b(y,x) = q(x) + q(y) + 2b_s(x,y)$$

et la forme bilinéaire (symétrique) est déterminée par la formule

$$b_s(x,y) = \frac{1}{2}(q(x+y) - q(x) - q(y)).$$

Définition 3.2. *Pour toute forme quadratique q, il existe une unique forme bilinéaire symétrique $b : E \to E \to \mathbf{R}$ telle que*

$$q(x) = b(x,x), \quad x \in E$$

déterminée par la formule de polarisation

$$b(x,y) = \frac{1}{2}(q(x+y) - q(x) - q(y), \quad x,y \in E.$$

On appelle b la forme polaire associée à q.

Notons que l'on peut également écrire la formule de polarisation sous la forme (plus symétrique ?)

$$b(x,y) = \frac{1}{4}\big(q(x+y) - q(x-y)\big), \quad x,y \in E.$$

Remarque 3.3. Ce résultat peutêtre reformulé de la façon suivante : l'application entre l'espace (vectoriel) des formes bilinéaires $\mathcal{B}$ et l'espace (vectoriel) des formes quadratiques $\mathcal{Q}$

$$\begin{aligned} \Phi : \mathcal{B} &\to \mathcal{Q} \\ b &\mapsto q \end{aligned}$$

est une application linéaire surjective dont le noyau $\ker \Phi$ est l'espace des formes bilinéaires antisymétriques. La restriction de Φ à l'espace des formes bilinéaires symétriques est un isomorphisme dont l'inverse est donnée par la formule de polarisation.

Définition 3.4. *Le noyau d'une forme bilinéaire est le sous-espace vectoriel*

$$\ker b = \{y \in E : b(x,y) = 0, \quad \forall x \in E\}.$$

On dit que b est non dégénérée si $\ker b = \{0\}$. Le rang de la forme bilinéaire est l'entier

$$\mathrm{rg}(b) = n - \dim \ker b.$$

Proposition 3.5. *Soit $(e_1, \ldots, e_n)$ une base de E et B la matrice de b dans cette base, alors on a*

$$\ker b = \left\{ x \in E : x = \sum_{j=1}^{n} x_j e_j, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \ker B \right\}$$

et $\mathrm{rg}(b) = \mathrm{rg}(B)$. La forme bilinéaire b est non dégénérée si et seulement si $\det B \neq 0$.

Démonstration. Pour tous $x = \sum_{j=1}^{n} x_j e_j$ et $y = \sum_{j=1}^{n} y_j e_j$

$$b(x,y) = {}^{\mathrm{t}}XBY$$

avec

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Ainsi si $y \in \ker b$ alors

$${}^{\mathrm{t}}XBY = 0$$

pour tout $X \in \mathbf{R}^n$ ce qui implique $BY = 0$ et donc $Y \in \ker B$. □

Définition 3.6. *Le cône d'isotropie d'une forme quadratique est l'ensemble contenant l'origine*

$$C(q) = \{x \in E : q(x) = 0\}.$$

On dit que la forme quadratique q est définie si $C(q) = \{0\}$. Le cône d'isotropie contient le noyau de sa forme polaire b

$$\ker b \subset C(q).$$

En particulier, une forme bilinéaire définie est non-dégénérée.

Le cône d'isotropie est un cône, i.e. pour tout scalaire $\lambda \in \mathbf{R}$

$$x \in C(q) \Rightarrow \lambda x \in C(q).$$

Si b est un produit scalaire, i.e. définie positive alors $C(q) = \{0\}$ et donc la forme quadratique $q(x) = \|x\|^2$ est définie. Ce qui explique ce choix de dénomination.

Lemme 3.7. *Une forme bilinéaire b définie est soit définie positive, soit définie négative (i.e. $-b$ est définie positive). Ainsi les seules formes dont le cône d'isotropie est trivial sont les produits scalaires ou leurs opposés.*

Démonstration. Supposons que q ne soit ni définie positive ni définie négative alors il existe $x \in E$ tel que $q(x) < 0$ et $y \in E$ tel que $q(y) > 0$, on considère alors le segment

$$(1-t)x + ty, \quad t \in [0,1]$$

et la restriction de la forme quadratique à ce segment

$$f(t) = q((1-t)x + ty),$$

Comme $f(0) = q(x) < 0$ et $f(1) = q(y) > 0$, le théorème des valeurs intermédiaires permet d'affirmer qu'il existe $t_0 \in]0,1[$ tel que

$$f(t_0) = q((1-t_0)x + t_0 y) = 0$$

ce qui signifie que $(1-t_0)x + t_0 y \in C(q)$, et donc comme q est définie

$$(1-t_0)x = -t_0 y$$

ce qui est absurde car cela impliquerait

$$\underbrace{(1-t_0)^2 q(x)}_{<0} = q((1-t_0)x) = q(t_0 y) = \underbrace{t_0^2 q(y)}_{>0}.$$

La forme quadratique q ne peut donc être définie. □

Lemme 3.8. *Soit b une forme bilinéaire symétrique positive alors le cône d'isotropie de sa forme quadratique q*

$$C(q) = \ker b$$

est le noyau de b.

Démonstration. Il suffit de vérifier $C(q) \subset \ker b$: soit $x \in C(q)$, alors pour tout $y \in E$ on peut appliquer l'inégalité de Cauchy-Schwarz

$$|b(x,y)| \leq \sqrt{q(x)q(y)} = 0$$

ce qui implique $b(x,y) = 0$ et donc $x \in \ker b$. □

3.2. **Deux exemples.**

3.2.1. *Espace de Minkowski.* On munit $\mathbf{R}^{n+1}$ de la forme quadratique

$$q(x) = x_0^2 - \sum_{j=1}^{n} x_j^2, \quad x = (x_0, x_1, \ldots, x_n) \in \mathbf{R}^{n+1}$$

dont la forme polaire a pour matrice

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -\mathrm{I}_n \end{pmatrix}$$

et est par conséquent non dégénérée. Le cône d'isotropie est le cône de lumière donné par

$$C = C(q) = \left\{ x \in \mathbf{R}^{n+1} : x_0^2 = \sum_{j=1}^{n} x_j^2 \right\}$$

et la forme q n'est donc pas définie. On peut décomposer l'espace en

$$\mathbf{R}^{n+1} = T \cup C \cup E$$

où

$$T = \left\{ x \in \mathbf{R}^{n+1} : x_0^2 > \sum_{j=1}^{n} x_j^2 \right\}$$

est l'ensemble des vecteurs de type temps

$$E = \left\{ x \in \mathbf{R}^{n+1} : x_0^2 < \sum_{j=1}^{n} x_j^2 \right\}$$

est l'ensemble des vecteurs de type espace, et C est le cône de lumière (dont les vecteurs sont de type lumière). Les vecteurs de type temps se décomposent en le passé et le futur

$$T = T_- \cup T_+, \quad T_\pm = \big\{ x \in T : \pm x_0 > 0 \big\}.$$

3.2.2. *Espace symplectique.* Une forme symplectique est une forme bilinéaire antisymétrique non dégénérée. L'antisymétrie équivaut au fait que la forme quadratique associée est nulle

$$\omega(x, x) = 0, \quad x \in E.$$

En effet, si la forme quadratique associée est nulle alors

$$\omega(x, y) + \omega(y, x) = \omega(x + y, x + y) - \omega(x, x) - \omega(y, y) = 0.$$

L'existence d'une forme symplectique implique que l'espace est de dimension paire

$$\dim E = 2n.$$

En effet, la matrice Ω d'une forme symplectique dans une base de E est antisymétrique, on a donc

$$\det \Omega = \det {}^{\mathrm{t}}\Omega = \det(-\Omega) = (-1)^n \det \Omega$$

et comme ce déterminant est non nul car une forme symplectique est non dégénérée, on en déduit

$$(-1)^n = 1$$

et donc n est pair. On dit qu'une base $(e_1, \ldots, e_n, \varepsilon_1, \ldots, \varepsilon_n)$ de E est symplectique si l'on a

$$\omega(e_j, e_k) = 0, \quad \omega(\varepsilon_j, \varepsilon_k) = 0, \quad \omega(e_j, \varepsilon_k) = \delta_{j,k}.$$

La matrice Ω d'une forme symplectique ω dans une base symplectique est de la forme

$$\Omega = \begin{pmatrix} 0 & \mathrm{I}_n \\ -\mathrm{I}_n & 0 \end{pmatrix}$$

et vérifie

$$\Omega^2 = -\mathrm{I}_{2n}.$$

Du point de vue de la forme quadratique, l'analyse s'arrête immédiatement puisque la forme quadratique est nulle. Par contre, l'algèbre symplectique revêt de nombreux aspects intéressants.

3.3. **Orthogonalité.** On peut définir comme dans le cas des produits scalaires, l'orthogonal d'un sous-ensemble $A \subset E$ de l'espace vectoriel E muni d'une forme bilinéaire symétrique b

$$A^{\perp} = \big\{x \in E : b(x,a) = 0, \quad \forall a \in A\big\}.$$

Proposition 3.9. *L'orthogonal vérifie les propriétés suivantes*

(*i*) $\{0\}^{\perp} = E$, $E^{\perp} = \ker b$,

(*ii*) $A^{\perp}$ *est un sous-espace vectoriel de* E *contenant* $\ker b$,

(*iii*) *Si* $A \subset B$ *alors* $B^{\perp} \subset A^{\perp}$,

(*iv*) $A^{\perp} = \mathrm{vect}(A)^{\perp}$,

(*v*) $A \cap A^{\perp} \subset C(q)$,

(*vi*) *Si* b *est non dégénérée alors* $(A^{\perp})^{\perp} = \mathrm{vect}(A)$.

(*vii*) *Si* F *est un sous-espace vectoriel de* E *et si* b *est non dégénérée alors*

$$\dim F^{\perp} = n - \dim F.$$

Démonstration. Il s'agit pour les quatre premiers points d'une répétition *mutatis mutandi* des arguments de la proposition 1.20 dans le cas des produits scalaires.

(*i*) La première égalité $\{0\}^{\perp} = E$ est claire. La deuxième est la définition de $\ker b$. Remarquons que l'on voit immediatement que dans le cas d'une forme bilinéaire b dégénérée

$$(\{0\}^{\perp})^{\perp} = \ker b \supsetneq \{0\}.$$

(*ii*) Montrons que $A^{\perp}$ est un sous-espace vectoriel de E, soit $x, y \in A^{\perp}$ alors pour tous $\lambda, \mu \in \mathbf{R}$ et tout $a \in A$, on a par linéarité de b à gauche

$$b(\lambda x + \mu y, a) = \lambda \underbrace{b(x,a)}_{=0} + \mu \underbrace{b(y,a)}_{=0} = 0$$

et ainsi $\lambda x + \mu y \in A^{\perp}$.

(*iii*) L'implication est claire.

(*iv*) Comme $A \subset \mathrm{vect}(A)$, d'après le point précédent $\mathrm{vect}(A)^{\perp} \subset A^{\perp}$. Réciproquement si $x \in A^{\perp}$ alors pour tout

$$v = \sum_{j=1}^{m} \lambda_j a_j \in \mathrm{vect}(A), \quad a_j \in A$$

on a par linéarité à droite

$$b(x,v) = \sum_{j=1}^{n} \lambda_j b(x, a_j) = 0$$

ce qui donne $x \in \mathrm{vect}(A)^{\perp}$, et finalement $A^{\perp} = \mathrm{vect}(A)^{\perp}$.

(*v*) Soit $x \in A \cap A^{\perp}$ alors $q(x) = b(x,x) = 0$.

(vi) D'après le point précédent

$$(A^{\perp})^{\perp} = (\mathrm{vect}(A)^{\perp})^{\perp}$$

et il suffit de montrer que pour le sous-espace vectoriel $F = \mathrm{vect}(A)$, on a $F^{\perp} \cap F = \ker b \cap F$. On sait déjà que $\ker b \cap F \subset F^{\perp} \cap F$, il suffit de démontrer l'autre inclusion. Soit $x \in F \cap F^{\perp}$, on a

$$b($$

(vii) On a

$$(A^{\perp})^{\perp} = (\mathrm{vect}(A)^{\perp})^{\perp}$$

et il suffit de montrer que pour le sous-espace vectoriel $F = \mathrm{vect}(A)$, on a $(F^{\perp})^{\perp} = F$. Ceci découle de l'inclusion évidente $F \subset (F^{\perp})^{\perp}$ (valable y compris lorsuqe b est dégénérée) et du point suivant sur les dimensions.

$(viii)$ On choisit une base $(e_1, \dots, e_n)$ de E et on considère le sous-espace de $\mathbf{R}^n$

$$\tilde{F} = \left\{ X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{R}^n : x = \sum_{j=1}^{n} x_j e_j \in F \right\}$$

isomorphe à F de sorte que

$$\dim F = \dim \tilde{F}.$$

On considère également l'orthogonal de $\tilde{F}$ pour le produit scalaire canonique de $\mathbf{R}^n$

$$\tilde{F}^{\perp} = \left\{ Y \in \mathbf{R}^n : {}^{\mathrm{t}}XY = 0 \quad \forall X \in \mathbf{R}^n \right\}$$

de sorte que

$$\tilde{F}^{\perp} = n - \dim \tilde{F} = n - \dim F.$$

Il suffit de montrer que $F^{\perp}$ est isomorphe à $\tilde{F}^{\perp}$ pour conclure. Soit $y = \sum_{j=1} y_j e_j$ alors si l'on note
on a l'équivalence

$$y \in F^{\perp} \Leftrightarrow b(x, y) = 0 \quad \forall x \in F \Leftrightarrow {}^{\mathrm{t}}XBY = 0 \quad \forall X \in \tilde{F} \Leftrightarrow BY \in \tilde{F}^{\perp}$$

et ainsi $F^{\perp}$ est isomorphe à $B^{-1}(\tilde{F}^{\perp})$ lui même isomorphe à $\tilde{F}^{\perp}$. On a alors

$$\dim \tilde{G} = n - \dim \tilde{F}$$

La proposition est démontrée. □

Définition 3.10. *Soit b une forme bilinéaire symétrique. On dit qu'une base $(e_1, \dots, e_n)$ est b-orthogonale si*

$$b(e_j, e_k) = 0 \quad j \neq k$$

autrement dit si la matrice de b dans cette base est diagonale.

Notons que dans une base b-othogonale, la forme quadratique q prend une forme simple

$$q(x) = \sum_{j=1}^{n} b_{j,j} x_j^2, \quad x = \sum_{j=1}^{n} x_j e_j$$

comme une somme de carrés.

3.4. **Réduction de Gauss.** D'après le théorème de réduction des matrices symétriques, on sait qu'il existe $P \in \mathrm{O}(n, \mathbf{R})$ telle que

$$B = {}^{\mathrm{t}}P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P$$

avec des valeurs propres λ_j réellles, de telle sorte que pour tout $x = \sum_{j=1}^{n} x_j e_j \in E$

$$q(x) = {}^{\mathrm{t}}XBX = \sum_{j=1}^{n} \lambda_j y_j^2, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = PX, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

la forme quadratique s'exprime comme une somme de carrés. Cela nécessite de procéder à une diagonalisation. Y-a-t-il d'autres procédures permettant d'aboutitr à une forme simple ?

Commençons par examiner le cas de la dimension $n = 2$. On considère la forme quadratique

$$q(x) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

avec $a, b, c \in \mathbf{R}$ et $(a, b, c) \neq (0, 0, 0)$. La matrice symétrique de la forme polaire est

$$B = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

dont le déterminant est

$$\det B = ac - b^2 = -\Delta', \quad \Delta' = b^2 - ac.$$

Supposons a non nul, on peut utiliser la méthode de complétion des carrés

$$q(x) = a\left(x_1 + \frac{b}{a}x_2\right)^2 + \left(c - \frac{b^2}{a}\right)x_2^2$$

pour obtenir une forme plus simple

$$q(x) = a\left(y_1^2 - \frac{\Delta}{a^2}y_2^2\right)$$

avec le changement de variables linéaire

$$y = x_1 + \frac{b}{a}, \quad y_2 = x_2$$

qui peut s'écrire sous forme matricielle $Y = PX$ avec une matric inversible (triangulaire supérieure)

$$P = \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix}.$$

Si $a = 0$ et $c = 0$ alors on peut échanger les rôles de x_1 et de x_2

$$q(x) = c\left(x_2 + \frac{b}{c}x_1\right)^2 - \frac{b^2}{c}x_1^2.$$

Le changement de variables est cette fois donnée par la matrice inversible (triangulaire inférieure)

$$P = \begin{pmatrix} 1 & 0 \\ \frac{b}{c} & 1 \end{pmatrix}.$$

Si a et c sont nuls alors

$$q(x) = 2bx_1x_2 = \frac{b}{2}(x_1 + x_2)^2 - \frac{b}{2}(x_1 - x_2)^2$$

et on peut faire le changement de variables

$$y_1 = x_1 + x_2, \quad y_2 = x_1 - x_2$$

donnée par la matrice

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

Le cas général est très similaire. Soit q une forme quadratique dont la forme polaire a pour matrice B dans une base $(e_1, \dots, e_n)$ de E alors

$$q(x) = \sum_{j,k=1}^{n} b_{j,k}x_jx_k = \sum_{j=1}^{n} b_{j,j}x_j^2 + 2\sum_{1\le j<k\le n} b_{j,k}x_jx_k$$

pour tout

$$x = \sum_{j=1}^{n} x_je_j \in E.$$

On choisit alors un terme diagonal non nul s'il y en a un ; pour simplifier les notations supposons que $b_{n,n} \neq 0$ (sinon il suffit de permuter les variables), on a

$$q(x) = \sum_{j,k=1}^{n} b_{j,k}x_jx_k = b_{n,n}x_n^2 + 2\left(\sum_{j=1}^{n} b_{n,j}x_j\right)x_n + \sum_{j,k=1}^{n-1} b_{j,k}x_jx_k$$

et on peut utiliser à nouveau la complétion des carrés

$$q(x) = b_{n,n}\left(x_n + \sum_{j=1}^{n-1}\frac{b_{n,j}}{b_{n,n}}x_j\right)^2 + \underbrace{\sum_{j,k=1}^{n-1} b_{j,k}x_jx_k - \frac{1}{b_{n,n}}\left(\sum_{j=1}^{n-1} b_{n,j}x_j\right)^2}_{=p(x')}$$

où la forme quadratique

$$p(x') = \sum_{j,k=1}^{n-1} \Big(b_{j,k} - \frac{b_{n,j} b_{n,k}}{b_{n,n}} \Big) x_j x_k$$

ne dépend que des $n-1$ premières variables $x' = (x_1, \ldots, x_{n-1})$. On peut alors faire le changement de variables linéaire

$$y_n = x_n + \sum_{j=1}^{n-1} \frac{b_{n,j}}{b_{n,n}} x_j, \quad y' = x'$$

dont la matrice est

$$P = \begin{pmatrix} 1 & & & 0 \\ & \ddots & & \vdots \\ & & 1 & 0 \\ \frac{b_{n,1}}{b_{n,n}} & \cdots & \frac{b_{n,n-1}}{b_{n,n}} & 1 \end{pmatrix}.$$

On itère le même processus sur p' jusqu'à n'avoir que des termes diagonaux nuls et considérer une forme quadratique de la forme

$$q(x) = 2 \sum_{1 \le j < k \le n} b_{j,k} x_j x_k.$$

L'un des coefficients est non nul ; supposons que cela soit $b_{n-1,n}$, alors comme dans le cas de la dimension deux

$$\begin{aligned} q(x) &= 2b_{n-1,n} x_{n-1} x_n + 2 \Big(\sum_{j=1}^{n-2} b_{j,n} x_j \Big) x_n + 2 \Big(\sum_{j=1}^{n-2} b_{j,n} x_j \Big) x_{n-1} \\ &\quad + 2 \sum_{1 \le j < k \le n-2}^{n} b_{j,k} x_j x_k \\ &= 2b_{n-1,n} \underbrace{\Big(x_n + \sum_{j=1}^{n-2} \frac{b_{j,n}}{b_{n-1,n}} x_j \Big) \Big(x_{n-1} + \sum_{j=1}^{n-2} \frac{b_{j,n-1}}{b_{n-1,n}} x_j \Big)}_{=\alpha\beta = \frac{1}{4}(\alpha+\beta)^2 - \frac{1}{4}(\alpha-\beta)^2} \\ &\quad \underbrace{+ 2 \sum_{1 \le j < k \le n-2}^{n} b_{j,k} x_j x_k - 2b_{n-1,n} \Big(\sum_{j=1}^{n-2} \frac{b_{j,n}}{b_{n-1,n}} x_j \Big) \Big(\sum_{j=1}^{n-2} \frac{b_{j,n-1}}{b_{n-1,n}} x_j \Big)}_{=p(x'')} \\ &= \frac{b_{n-1,n}}{2} \Big(x_{n-1} + x_n + \sum_{j=1}^{n-2} \frac{b_{j,n-1} + b_{j,n}}{b_{n-1,n}} x_j \Big)^2 \\ &\quad - \frac{b_{n-1,n}}{2} \Big(x_{n-1} - x_n + \sum_{j=1}^{n-2} \frac{b_{j,n-1} - b_{j,n}}{b_{n-1,n}} x_j \Big)^2 + p(x'') \end{aligned}$$

et $p(x'')$ est une forme quadratique (sans termes diagonaux) dépendant des $n-2$ premières variables $x'' = (x_1, \ldots, x_{n-2})$. On peut donc faire le changement de variable linéaire

$$
\begin{aligned}
y_n &= x_{n-1} + x_n + \sum_{j=1}^{n-2} \frac{b_{j,n-1} + b_{j,n}}{b_{n-1,n}} x_j \\
y_{n-1} &= x_{n-1} - x_n + \sum_{j=1}^{n-2} \frac{b_{j,n-1} - b_{j,n}}{b_{n-1,n}} x_j \\
y'' &= x''
\end{aligned}
$$

dont la matrice (triangulaire inférieure par blocs) est

$$
P = \begin{pmatrix}
1 & & & 0 & 0 \\
& \ddots & & \vdots & \vdots \\
& & 1 & 0 & 0 \\
\frac{b_{n-1,1} - b_{n,1}}{b_{n-1,n}} & \dots & \frac{b_{n-2,n-1} - b_{n-2,n}}{b_{n-1,n}} & 1 & -1 \\
\frac{b_{n-1,1} + b_{n,1}}{b_{n-1,n}} & \dots & \frac{b_{n-2,n-1} + b_{n-2,n}}{b_{n-1,n}} & 1 & 1
\end{pmatrix}.
$$

Lorsque la récurrence se termine, on obtient une forme quadratique de la forme

$$
\sum_{j=1}^{n} \mu_j y_j^2.
$$

Comme les transformations P apparaissant aux différentes étapes de l'algorithme correspondent à des changements de bases, cet algorithme fournit donc le résultat suivant.

Théorème 3.11. *Pour toute forme quadratique q, il existe une base $(g_1, \dots, g_n)$ de E dans laquelle la matrice de la forme quadratique q est diagonale*

$$
\begin{pmatrix}
\mu_1 & & \\
& \ddots & \\
& & \mu_n
\end{pmatrix}.
$$

De manière équivalente, lorsque B est la matrice d'une forme quadratique q dans une base $(e_1, \dots, e_n)$ de E il exitse une matrice inversible $P \in \mathrm{GL}(n, \mathbf{R})$ telle que

$$
B = {}^{\mathrm{t}}P \begin{pmatrix}
\mu_1 & & \\
& \ddots & \\
& & \mu_n
\end{pmatrix} P
$$

de sorte que pour tout $x = \sum_{j=1}^{n} x_j e_j$

$$
q(x) = \sum_{j=1}^{n} \mu_j y_j^2, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = PX, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.
$$

Plus que le résultat, c'est l'algorithme de réduction de Gauss qui est important.

Remarque 3.12. La base $(g_1, \dots, g_n)$ n'est pas forcément orthonormée et la matrice P pas forcément orthogonale ! On n'a donc pas en général ${}^{\mathrm{t}}P = P^{-1}$. L'algorithme de Gauss n'est pas une diagonalisation de la matrice B et les valeurs μ_j ne sont pas en général les valeurs propres de B !

Les valeurs μ_j sont réelles ; μ_j est soit nulle, soit positive, soit négative. Quitte à faire d'autres changement de variables de la forme

$$y_j = \frac{x_j}{\sqrt{|\mu_j|}} \qquad \text{lorsque } \mu_j \neq 0$$

on peut se ramener aux valeurs $0, 1, -1$.

Corollaire 3.13. *Pour toute forme quadratique q, il existe une base $(g_1, \dots, g_n)$ de E dans laquelle la matrice de la forme quadratique q est de la forme*

$$\begin{pmatrix} \mathrm{I}_p & & \\ & -\mathrm{I}_q & \\ & & 0 \end{pmatrix}.$$

De manière équivalente, lorsque B est la matrice d'une forme quadratique q dans une base $(e_1, \dots, e_n)$ de E il exitse une matrice inversible $P \in \mathrm{GL}(n, \mathbf{R})$ telle que

$$B = {}^{\mathrm{t}}P \begin{pmatrix} \mathrm{I}_p & & \\ & -\mathrm{I}_q & \\ & & 0 \end{pmatrix} P$$

de sorte que pour tout $x = \sum_{j=1}^{n} x_j e_j$

$$q(x) = \sum_{j=1}^{p} y_j^2 - \sum_{j=p+1}^{p+q}, \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = PX, \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Remarque 3.14. Cet algorithme de réduction de Gauss s'apparante à l'algorithme du pivot de Gauss pour calculer le rang d'une matrice ou résoudre un système linéaire d'équations.

La question de la dépendence des tailles p et q par rapport à l'algorithme se pose très naturellement. Nous allons voir que ces entiers ne dépendent pas des changements de variables effectués et constituent la signature de la forme quadratique q.

3.5. **Loi d'inertie de Sylvester.** Soit q une forme quadratique, on note

$$\mathcal{E}_+ = \big\{ F \text{ sous-espace vectoriel de } E \text{ tel que } q|_F \text{ est définie négative} \big\}$$
$$\mathcal{E}_- = \big\{ F \text{ sous-espace vectoriel de } E \text{ tel que } q|_F \text{ est définie positive.} \big\}$$

et on définit les indices suivants

$$n_- = \max \big\{ \dim F : F \in \mathcal{E}_+ \big\}$$
$$n_+ = \max \big\{ \dim F : F \in \mathcal{E}_- \big\}.$$

Définition 3.15. *La signature de la forme quadratique q est le couple (n_+, n_-).*

Lemme 3.16. *Soit b une forme bilinéaire symétrique, et soit $(e_1, \dots, e_n)$ une base b-orthogonale, notons*

$$J_0 = \{j \in \{1, \dots, n\} : q(e_j) = 0\}$$

alors on a

$$\ker b = \mathrm{vect}(e_j : j \in J_0)$$

et par conséquent

$$\mathrm{card}\, J_0 = n - \mathrm{rg}\, b.$$

Démonstration. Soit $x = \sum_{j \in J_0} x_j e_j$ alors pour tout $y = \sum_{j=1}^n y_j e_j \in E$ on a

$$b(x, y) = \sum_{j \in J_0} x_j y_j q(e_j) = 0$$

et réciproquement si $x = \sum_{j=1}^n x_j e_j \in \ker b$ alors

$$b(x, e_j) = 0 = x_j q(e_j)$$

et donc $x_j = 0$ pour tout $j \notin J_0$. □

Définition 3.17. *On dit que deux formes quadratiques q et $\tilde{q}$ sont équivalentes s'il existe un endomorphisme inversible $u \in \mathrm{GL}(E)$ tel que*

$$\tilde{q}(x) = q(u(x)), \quad x \in E.$$

On note alors $\tilde{q} \sim q$.

Lemme 3.18. *$\sim$ est une relation d'équivalence.*

Démonstration. Cette relation est

(*i*) réflexive, car il suffit de choisir $u = \mathrm{Id}_E$,

(*ii*) symétrique car si $\tilde{q} \sim q$ alors il existe $u \in \mathrm{GL}(E)$ tel que

$$\tilde{q}(x) = q(u(x)), \quad x \in E$$

et par conséquent

$$q(x) = \tilde{q}(u^{-1}(x)), \quad x \in E.$$

(*iii*) transitive car si $\tilde{q} \sim \hat{q}$ et $\hat{q} \sim q$ alors il existe $u, v \in \mathrm{GL}(E)$ tels que

$$\tilde{q}(x) = \hat{q}(u(x)), \quad \hat{q}(x) = q(v(x)), \quad x \in E$$

et par conséquent

$$\tilde{q}x) = q(v \circ u(x)), \quad x \in E.$$

C'est donc bien une relation d'équivalence. □

Par polarisation, si deux formes quadratiques $\tilde{q}$ et q sont équivalentes alors leurs formes polaires respectives b et $\tilde{b}$ sont rliées par

$$\tilde{b}(x, y) = b(u(x), u(y))$$

où u est un endomorphisme inversible. En particulier, si $(e_1, \dots, e_n)$ est une base b-orthogonale alors $(u^{-1}(e_1), \dots, u^{-1}(e_n))$ est $\tilde{b}$-orthogonale. En outre, en raisonnant sur les matrices, on voit que q et $\tilde{q}$ sont équivalentes si et seulement s'il existe $U \in \mathrm{GL}(n, \mathbf{R})$ telle que

$$\tilde{B} = {}^{\mathrm{t}}UBU.$$

Théorème 3.19. *Soit q une forme quadratique sur E, et b la forme polaire associée alors pour toute base $(e_1,\dots,e_n)$ b-orthogonale, on a*

$$n_- = \operatorname{card}(j \in \{1,\dots,n\} : q(e_j) < 0\big\},$$
$$n_+ = \operatorname{card}(j \in \{1,\dots,n\} : q(e_j) > 0\big\}$$

et $n_- + n_+ = \operatorname{rg}(b)$. Deux formes quadratiques sont équivalentes si et seulement si elles ont même signature.

Démonstration. Soient $(e_1,\dots,e_n)$ et $(f_1,\dots,f_n)$ deux bases b-orthogonales. On note $r = \operatorname{rg}(b)$ et

$$I_\pm = \big\{j \in \{1,\dots,n\} : \pm q(e_j) > 0\big\}$$
$$J_\pm = \big\{j \in \{1,\dots,n\} : \pm q(f_j) > 0\big\}.$$

D'après le lemme 3.16, on a

$$\operatorname{card} I_+ + \operatorname{card} I_- = r, \quad \operatorname{card} J_+ + \operatorname{card} J_- = r.$$

On considère alors la famille $\mathcal{B} = ((e_j)_{j\in I_+}, (f_j)_{j\in J_-})$ et F le sous-espace vectoriel engendré par cette famille. La famille $\mathcal{B}$ est libre car si l'on a

$$\sum_{j\in I_+} \lambda_j e_j + \sum_{j\in J_-} \mu_j f_j = 0$$

alors on en déduit

$$\sum_{j\in I_+} \lambda_j e_j = -\sum_{j\in J_-} \mu_j f_j$$

et en appliquant la forme quadratique aux deux membres de l'égalité

$$\underbrace{\sum_{j\in I_+} \lambda_j^2 q(e_j)}_{\geq 0} = \underbrace{\sum_{j\in J_-} \mu_j^2 q(f_j)}_{\leq 0}$$

on tire

$$\sum_{j\in I_+} \lambda_j^2 q(e_j) = \sum_{j\in J_-} \mu_j^2 q(f_j) = 0$$

et ainsi $\lambda_j = 0$ pour tout $j \in I_+$ et $\mu_j = 0$ pour tout $j \in I_-$. En outre, on a

$$F \cap \ker b = \{0\}.$$

En effet supposons que

$$x = \underbrace{\sum_{j\in I_+} \lambda_j e_j}_{=x_+} + \underbrace{\sum_{j\in J_-} \mu_j f_j}_{=x_-}$$

soit dans le noyau $\ker b$ de b alors on a

$$0 = b(x, x_+) = q(x_+) + b(x_-, x_+)$$
$$0 = b(x, x_-) = q(x_-) + b(x_+, x_-)$$

ce qui implique

$$q(x_-) = q(x_+)$$

soit comme auparavant

$$\underbrace{\sum_{j\in I_+}\lambda_j^2 q(e_j)}_{\geq 0} = \underbrace{\sum_{j\in J_-}\mu_j^2 q(f_j)}_{\leq 0}$$

ce qui n'est possible que si $x_+ = 0$ et $x_- = 0$ et ainsi $x = 0$. On en déduit

$$\dim F + \dim\ker b = \dim(F \oplus \ker b) \leq n$$

soit

$$\operatorname{card} I_+ + \operatorname{card} J_- + n - r = \operatorname{card} I_+ + n - \operatorname{card} J_+ \leq n$$

qui implique

$$\operatorname{card} I_+ \leq \operatorname{card} J_+.$$

En intervertissant les rôles des bases $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_n)$ on obtient l'inégalité dans l'autre sens et donc finalement l'égalité

$$\operatorname{card} I_+ = \operatorname{card} J_+$$

qui donne également

$$\operatorname{card} I_- = \operatorname{card} J_-.$$

Montrons maintenant que pour une base b-orthogonale bien choisie — et donc pour toute base b-orthogonale, d'après la première partie de cette preuve —, on a $\operatorname{card} I_+ = n_+$ et $\operatorname{card} I_- = n_-$. Considérons un sous-espace vectoriel $F \in \mathcal{E}_+$ de dimension n_+, choisissons[3] une base b-orthogonale $(e_1, \ldots, e_{n_+})$ de F, comme tout vecteur se décompose sous la forme

$$x = \underbrace{\sum_{j=1}^{n_+}\frac{b(x,e_j)}{q(e_j)}e_j}_{\in F} + \underbrace{x - \sum_{j=1}^{n_+}\frac{b(x,e_j)}{q(e_j)}e_j}_{\in F^\perp}$$

et comme

$$F \cap F^\perp \subset C(q) \cap F = \{0\}$$

on a

$$E = F \oplus F^\perp.$$

On peut alors compléter[4] $(e_1, \ldots, e_{n_+})$ en une base b-orthogonale $(e_1, \ldots, e_n)$. Ainsi, on a

$$F = \operatorname{vect}(e_1, \ldots, e_{n_+}), \quad F^\perp = \operatorname{vect}(e_{n_++1}, \ldots, e_n)$$

et de plus

$$q(e_j) \leq 0, \quad j \geq n_+ + 1$$

car si l'on avait $q(e_{j_0}) > 0$ pour un $j_0 \geq n_+ + 1$ alors q serait définie positive sur $F \oplus \mathbf{R}e_{j_0}$ ce qui contredirait le caractère maximal de n_+. La première partie de la preuve implique ainsi

$$n_+ = \dim F = \operatorname{card} I_+.$$

Les mêmes arguments, en remplaçant q par $-q$ donnent

$$n_- = \operatorname{card} I_-.$$

3. Par exemple, par une réduction de Gauss de $q|_F$.
4. Par exemple, par une réduction de Gauss de $q|_{F^\perp}$.

Supposons maintenant que deux formes quadratiques q et $\tilde{q}$ sont équivalentes, alors il existe un endomorphisme inversible $u \in \mathrm{GL}(E)$ tel que

$$\tilde{q}(x) = q(u(x)), \quad x \in E.$$

Soient b et $\tilde{b}$ les formes polaires respectives de q et $\tilde{q}$. Si $(e_1, \ldots, e_n)$ est une base b-orthogonale alors $(\tilde{e}_1 = u^{-1}(e_1), \ldots, \tilde{e}_n = u^{-1}(e_n))$ est une base $\tilde{b}$-orthogonale. Le fait que

$$\tilde{q}(\tilde{e}_j) = q(u(\tilde{e}_j)) = q(e_j)$$

implique donc que q et $\tilde{q}$ aient même signature. Réciproquement supposons que deux formes quadratiques q et $\tilde{q}$ aient même signature (n_+, n_-). Considérons les formes polaires respectives b et $\tilde{b}$ de q et $\tilde{q}$. Alors par réduction de Gauss (corollaire 3.13), il existe une base b-orthogonale $(g_1, \ldots, g_n)$ et une base $\tilde{b}$-orthogonale $(\tilde{g}_1, \ldots, \tilde{g}_n)$ de E dans lesquelles les matrices des formes polaires b et $\tilde{b}$ sont égales à

$$\begin{pmatrix} \mathrm{I}_{n_+} & & \\ & -\mathrm{I}_{n_-} & \\ & & 0 \end{pmatrix}.$$

Soit l'endomorphisme $u : E \to E$ qui envoie la base $(\tilde{g}_1, \ldots, \tilde{g}_n)$ sur la base $(g_1, \ldots, g_n)$ alors

$$\tilde{b}(x, y) = b(u(x), u(y))$$

et on en déduit l'équivalence de q et $\tilde{q}$. □

Corollaire 3.20. *Soit q une forme quadratique de signature (n_+, n_-), b sa forme polaire et B la matrice de b dans une base de E. Le nombre de valeurs propres de B (comptées avec multiplicité) positives est n_+, le nombre de valeurs propres (comptées avec multiplicité) positives est n_-.*

Démonstration. Notons $(e_1, \ldots, e_n)$ la base de E dans laquelle la matrice de la forme polaire b est B. Il suffit de diagonaliser B, il existe une matrice orthogonale $P = (p_{j,k})_{1 \leq j,k \leq n} \in \mathrm{O}(n, \mathbf{R})$ telle que

$$B = {}^{\mathrm{t}}P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P.$$

La base $(g_1, \ldots, g_n)$ donnée par la matrice P

$$g_j = \sum_{k=1}^{n} p_{j,k} e_k$$

est b-orthogonale, et dans cette nouvelle base, la matrice de b est

$$\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Le résultat découle donc de la loi d'inertie de Sylvester. □

4. Dualité

Dans cette section, on choisit la notation $(e^1, \ldots, e^n)$ (indice en exposant) pour une base de vecteurs.

4.1. **Formes linéaires.**

Définition 4.1. *Une forme linéaire sur E (ou un covecteur) est une application linéaire à valeurs dans le corps des scalaires. Le dual (algébrique) de E est l'espace vectoriel des formes linéaires sur E. On note E^* cet espace vectoriel. Le bidual de E est le dual E^{**} du dual.*

Le dual est de même dimension que E

$$\dim E^* = \dim L(E, \mathbf{R}) = n \times 1 = \dim E.$$

Exemple 4.2.

1. La trace est une forme linéaire sur l'espace vectoriel des matrices carrées.
2. L'évaluation $P \to P(a)$ en un nombre $a \in \mathbf{R}$ est une forme linéaire sur l'espace $\mathbf{R}_n[X]$ des polynômes de degré inférieur ou égal à n.
3. Si $(E, \langle \cdot\,, \cdot \rangle)$ est un espace euclidien alors $x \mapsto \langle x, a \rangle$ est une forme linéaire.
4. Si E est un espace vectoriel normé (de dimension finie) et $f : E \to \mathbf{R}$ est une fonction différentiable en x_0 alors la différentielle $D_{x_0} f$ est une forme linéaire.
5. Dans une vectoriel E, l'application $x \mapsto x_j$ qui à un vecteur $x = \sum_{j=1}^{n} x_j e^j$ fait correspondre sa j-ième coordonnée dans une base $(e^1, \ldots, e^n)$ de E est une forme linéaire.

Lemme 4.3. *Un vecteur $x \in E$ qui annule*

$$\xi(x) = 0$$

toute forme linéaire $\xi \in E^*$ est nul.

Démonstration. On choisit une base $(e_1, \ldots, e_n)$ de E, l'application qui à un vecteur fait correspondre sa j-ième coordonnée dans la base est une forme linéaire. Les coordonnés de x dans cette base sont donc toutes nulles. □

Théorème 4.4. *Soit $x \in E$, on note $\delta_x : E^* \to \mathbf{R}$ la forme d'évaluation*

$$\delta_x(\xi) = \xi(x)$$

d'une forme $\xi \in E^*$ en x. L'application

$$\begin{aligned} \Delta : E &\to E^{**} \\ x &\mapsto \delta_x \end{aligned}$$

est un isomorphisme. Le bidual est donc (canoniquement) isomorphe à l'espace E.

Démonstration. L'application Δ est linéaire car

$$\delta_{\alpha x + \beta y}(\xi) = \xi(\alpha x + \beta y) = \alpha \xi(x) + \beta \xi(y) = (\alpha \delta_x + \beta \delta_y)(\xi)$$

et son noyau est trivial

$$\ker \Delta = \{0\}.$$

En effet, si $x \in \ker \Delta$ alors
$$\xi(x) = 0$$
pour toute forme linéaire $\xi \in E^*$ ce qui entraîne $x = 0$ d'après le lemme précédent. Comme E et E^{**} ont même dimension, c'est bien un isomorphisme. □

De même, l'espace et son dual sont isomorphes mais l'isomorphisme n'est pas canonique et dépend d'une base choisie sur E.

4.2. **Bases duales et préduales.** Soit $(e^1, \dots, e^n)$ une base de E alors tout vecteur $x \in E$ se décompose de manière unique dans la base
$$x = \sum_{j=1}^{n} x_j e^j$$
l'application
$$\begin{aligned} \varepsilon_j : E &\to \mathbf{R} \\ x &\mapsto x_j \end{aligned}$$
qui à un vecteur associe sa j-ième coordonnée est une forme linéaire. Elle vérifie
$$\varepsilon_j(e^k) = \delta_{j,k}.$$
où rappelons que
$$\delta_{j,k} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{si } j \neq k \end{cases}.$$
La famille $(\varepsilon_1, \dots, \varepsilon_n)$ est une base de E^*. En effet, supposons que
$$\sum_{j=1}^{n} \lambda_j \varepsilon_j = 0$$
alors en évaluant en $x = e_k$, on obtient
$$\lambda_k = 0$$
et en outre comme on a pour tout $\xi \in E$ et tout $x \in E$
$$\xi(x) = \sum_{j=1}^{n} x_j \xi(e^j) = \sum_{j=1}^{n} \xi(e^j) \varepsilon_j(x)$$
on en déduit la décomposition de toute forme linéaire $\xi \in E^*$ dans la base
$$\xi = \sum_{j=1}^{n} \xi(e^j) \varepsilon_j.$$
De même, la décomposition d'un vecteur $x \in E$ s'écrit
$$x = \sum_{j=1}^{n} \varepsilon_j(x) e^j.$$

Définition 4.5. *La base $(\varepsilon_1, \dots, \varepsilon_n)$ est appelée la base duale de $(e^1, \dots, e^n)$.*

Exemple 4.6. Soit E un espace euclidien et $(e_1, \dots, e_n)$ une base orthonormée alors la base duale est donnée par

$$\varepsilon_j = \langle \cdot \,, e_j \rangle.$$

On définit à présent les isomorphismes musicaux construit en identifiant les bases et leurs duales. L'application dièse $\sharp : E^* \to E$ est celle qui envoie la base duale sur la base, elle hausse les indices

$$(\varepsilon_j)^\sharp = e^j$$

elle est donc donnée par

$$\xi^\sharp = \sum_{j=1}^n \xi(e^j) e^j$$

et l'application bémol $\flat : E \to E^*$ est celle qui envoie la base de E sur sa duale, elle abaisse les indices

$$(e^j)^\flat = \varepsilon_j$$

et est donnée par

$$x^\flat = \sum_{j=1}^n \varepsilon_j(x) \varepsilon_j.$$

Par construction, les ismorphismes musicaux sont réciproques l'un de l'autre.

Exemple 4.7. Soit E un espace euclidien et $(e_1, \dots, e_n)$ une base orthonormée alors

$$x^\flat(y) = \sum_{j=1}^n \langle x, e_j \rangle \langle y, e_j \rangle = \langle x, y \rangle.$$

On voit que l'on peut donc également représenter toute forme linéaire ξ (de manière unique) par un vecteur x (comme un produit scalaire avec ce vecteur). C'est une autre manière de construire un isomorphisme (qui dépend cette fois du produit scalaire) entre un espace et son dual.

Théorème 4.8 (Riesz). *L'application linéaire*

$$\begin{aligned} E &\mapsto E^* \\ x &\mapsto \langle x, \cdot \rangle \end{aligned}$$

est un isomorphisme.

Démonstration. Pour des raions de dimension, il suffit de vérifier l'injectivité ; si un vecteur x est dans le noyau alors

$$\langle x, y \rangle = 0$$

pour tout vecteur $y \in E$ et donc $x = 0$. □

Proposition 4.9. *Soit* $(e^1, \dots, e^n)$ *et* $(h^1, \dots, h^n)$ *deux bases de* E *et* P *la matrice de passage de la base* $(e^1, \dots, e^n)$ *à la base* $(h^1, \dots, h^n)$*, donnée par*

$$h^j = \sum_{k=1} p_{k,j} e^k$$

alors la matrice de passage de la base $(\varepsilon_1, \dots, \varepsilon_n)$ *duale de* $(e^1, \dots, e^n)$ *à la base* $(\eta_1, \dots, \eta_n)$ *duale de* $(h^1, \dots, h^n)$ *est*

$$Q = {}^{\mathrm{t}}P^{-1}.$$

Démonstration. On décompose la forme η_j dans la base duale de $(e^1, \dots, e^n)$

$$\eta_j = \sum_{k=1}^{n} \eta_j(e^k)\varepsilon_k$$

or par définition $\eta_j(e^k)$ est le j-ième coefficient de la décomposition de e^k dans la base $(h^1, \dots, h^n)$,

$$e^k = \sum_{j=1}^{n} q_{j,k} h^j$$

c'est à dire $q_{j,k}$ où $Q = P^{-1} = (q_{j,k})_{1 \le j,k \le n}$ est la matrice de passage de la base $(h^1, \dots h^n)$ à la base $(e^1, \dots, e^n)$. La matrice de passage recherchée est donc ${}^{\mathrm{t}}Q = {}^{\mathrm{t}}P^{-1}$. □

Soit $(\varepsilon_1, \dots, \varepsilon_n)$ une base du dual E^*, on peut considérer la base duale $(\varepsilon_1^*, \dots, \varepsilon_n^*)$ dans E^{**} puis identifier cette base dans E via l'isomorphisme canonique Δ

$$e^j = \Delta^{-1}(\varepsilon_j^*).$$

Lemme 4.10. *La base duale de la base* $(e^1, \dots, e^n)$ *construite de cette façon est* $(\varepsilon_1, \dots, \varepsilon_n)$.

Démonstration. En effet, on a

$$\Delta(e^j) = \varepsilon_j^*$$

donc

$$\delta_{e_j} = \varepsilon_j^*$$

et ainsi

$$\varepsilon_k(e^j) = \delta_{e^j}(\varepsilon^k) = \varepsilon_j^*(\varepsilon_k) = \delta_{j,k}$$

ce qui signifie que base duale de $(e^1, \dots, e^n)$ est $(\varepsilon_1, \dots, \varepsilon_n)$. □

On a donc montré que tout base du dual est la base duale d'une base de E. Il n'y en a pas d'autre car si $(e^1, \dots, e^n)$ et $(\tilde{e}^1, \dots, \tilde{e}^n)$ ont toutes les deux $(\varepsilon_1, \dots, \varepsilon_n)$ comme base duale alors

$$\varepsilon_j(\mathrm{e}^k - \tilde{e}^k) = 0$$

pour tout $j, k \in \{1, \dots, n\}$, ce qui implique que

$$\xi(\mathrm{e}^k - \tilde{e}^k) = 0$$

pour toute forme linéaire ξ et donc $\Delta(\mathrm{e}^k - \tilde{e}^k) = 0$ qui entraîne $\mathrm{e}^k = \tilde{e}^k$.

Définition 4.11. *L'unique base* $(\mathrm{e}_1, \dots, \mathrm{e}_n)$ *dont la base duale est* $(\varepsilon_1, \dots, \varepsilon_n)$ *ainsi construite est appelée la base préduale de* $(\mathrm{e}_1, \dots, \mathrm{e}_n)$.

4.3. **Hyperplans.**

Définition 4.12. *Un hyperplan d'un espace vectoriel E est le noyau d'une forme linéaire $\nu \in E^*$ non nulle.*

Si $H = \ker \nu$ est un hyperplan alors par le théorème du rang

$$\dim H = \dim \ker \nu = n - 1$$

car l'image $\nu(E)$ d'une forme linéaire ν non nulle est $\mathbf{R}$. Si E est un espace euclidien alors par le théorème de Riesz, il existe $N \in E \setminus \{0\}$ tel que

$$\nu(x) = \langle N, x \rangle$$

et donc

$$H = \ker \nu = N^{\perp}.$$

Le vecteur N est la normale à l'hyperplan H est la forme (ou covecteur) ν est la conormale à H.

Lemme 4.13. *Soit $\nu \in E^* \setminus \{0\}$ alors pour tout $x \in E \setminus \ker \nu$ on a*

$$E = \ker \nu \oplus \mathbf{R}x.$$

Démonstration. Montrons que $\ker \nu \cap \mathbf{R}x = \{0\}$. Soit $\lambda x \in \ker \nu$ alors

$$\nu(\lambda x) = \lambda \underbrace{\nu(x)}_{\neq 0} = 0$$

donc $\lambda = 0$. Montrons que $E = \ker \nu + \mathbf{R}x$. Cela peut se faire par un argumenet dimensionnel mais nous allons faire une preuve directe ; tout vecteur $y \in E$ se décompose sous la forme

$$y = \underbrace{\frac{\nu(y)}{\nu(x)} x}_{\in \mathbf{R}x} + \underbrace{y - \frac{\nu(y)}{\nu(x)} x}_{\in \ker \nu}$$

ce qui conclut la preuve. □

Corollaire 4.14. *Deux formes linéaires sont proportionnelles si et seulement si elles ont même noyau.*

Démonstration. Soit ξ et ζ deux formes linéaires qui ont même noyau $H = \ker \xi = \ker \eta$. Si les formes sont nulles alors elles sont proportionnelles. Si elles ne sont pas nulles alors il existe $x \in E \setminus \{0\}$ tel que $\xi(x) \neq 0$ et $\eta(x) \neq 0$. Montrons que

$$\xi = \frac{\xi(x)}{\eta(x)} \eta.$$

D'après le lemme précédent 4.13, on a

$$E = H \oplus \mathbf{R}x$$

et donc pour tout $y = h + \lambda x \in E$ avec $h \in H$ et $\lambda \in \mathbf{R}$, on vérifie

$$\xi(h + \lambda x) = \lambda \xi(x) = \frac{\xi(x)}{\eta(x)} \lambda \eta(x) = \frac{\xi(x)}{\eta(x)} \eta(h + \lambda x)$$

ce qui démontre le résultat. □

Proposition 4.15. *Tout sous-espace vectoriel de E de dimension $n-1$ est un hyperplan.*

Démonstration. Soit H un espace vectoriel de dimension $n-1$ alors il existe une droite $\mathbf{R}x$ (avec $x \neq 0$) supplémentaire de H

$$E = H \oplus \mathbf{R}x$$

et on définit $u : E \to \mathbf{R}$ de la manière suivante

$$\nu(h + \lambda x) = \lambda$$

pour tout vecteur $y = h + \lambda x \in E$ avec $h \in H$ et $\lambda \in \mathbf{R}$. On a

$$\ker \nu = H$$

et H est donc un hyperplan. □

4.4. **Quelques commentaires sur les formes quadratiques.** Soit b une forme bilinéaire alors on peut considérer la forme linéaire $\beta_y \in E^*$

$$\beta_y(x) = b(x, y)$$

et l'application

$$\begin{aligned} \beta : E &\to E^* \\ y &\mapsto \beta_y. \end{aligned}$$

Notons que

$$\ker b = \big\{ y \in E : \beta_y = 0 \big\} = \ker \beta$$

donc b est non dégénérée si et seulement si β est un isomorphisme. Ainsi toute forme bilinéaire non dégénérée induit un isomorphisme entre l'espace et son dual.

Si $(e^1, \ldots, e^n)$ est une base de E et $(\varepsilon_1, \ldots, \varepsilon_n)$ sa base duale alors pour tous vecteurs $x = \sum_{j=1}^n x_j e^j$ et $y = \sum_{j=1}^n y_j e^j$

$$b(x, y) = \sum_{j,k=1}^n b_{j,k} x_j y_k$$

et ainsi

$$b = \sum_{j,k=1}^n b_{j,k}\, \varepsilon_j \otimes \varepsilon_k$$

où le produit tensoriel $\varepsilon_j \otimes \varepsilon_k : E \times E \to \mathbf{R}$ est la forme bilinéaire définie par

$$\varepsilon_j \otimes \varepsilon_k(x, y) = \varepsilon_j(x)\varepsilon_k(y) = x_j y_k.$$

La forme quadratique associée est donc donnée par

$$q = \sum_{j,k=1}^n b_{j,k}\, \varepsilon_j\, \varepsilon_k$$

Si l'on revient sur l'endomorphisme $\beta : E \to E^*$ associé à la forme bilinéaire b, la décomposition de $\beta_y \in E^*$ dans la base duale s'écrit

$$\beta_y = \sum_{k=1}^n \beta_y(e^k)\varepsilon_k = \sum_{k=1}^n b(e^k, y)\varepsilon_k$$

et ainsi

$$\beta_{e_j} = \sum_{k=1}^{n} b(e^k, e^j)\varepsilon_k$$

et donc la matrice de l'endomorphisme β dans la base $(e^1, \ldots, e^n)$ de E et sa base duale $(\varepsilon_1, \ldots, \varepsilon_n)$ de E^* est la matrice

$$B = (b(e^j, e^k))_{1 \leq j,k \leq n}$$

de la forme bilinéaire b.

On pourrait décrire la réduction de ce Gauss dans ce contexte, mais on se contente de ces dernières remarques impressionnistes.