ALGÈBRE BILINÉAIRE

STÉPHANE LAMY

Table des matières

Quelques sources :

- [Per96], chapitre V, VI et VIII.
- [CG13], chapitres V, XI et XII.
- [RW07], modules II.3 et II.4.
- [Szp09], chapitre 4.

1. Formes quadratiques

1.1. **Dualité.** Soit E un $\mathbf{k}$-espace vectoriel de dimension finie n. Une application linéaire $E \to \mathbf{k}$ est appelé une *forme linéaire*. L'espace vectoriel de toutes les formes linéaires sur E est appelé l'espace dual de E, noté E^*.

Étant donné une base (e_i) sur E, il existe une base préférée de E^*, appelée *base duale* de (e_i), et notée (e_i^*).

Date: 13 décembre 2019.

Définition abstraite : e_i^* est l'unique forme linéaire telle que $e_i^*(e_i) = 1$ et $e_i^*(e_j) = 0$ pour tout $j \neq i$.

Définition en terme de coordonnées. Si $x \in E$ est un vecteur de coordonnées $(x_1, \ldots, x_n)$ dans la base (e_i), e_i^* est la forme linéaire qui à x associe la $i$ème coordonnées x_i ("projection sur le $i$ème axe parallèlement aux autres axes").

Définition en terme de produit matriciel : si on représente les éléments de E comme des vecteurs colonnes, alors les élements de E^* sont des vecteurs lignes, et l'évaluation $\ell(x)$ pour $x \in E$ et $\ell \in E^*$ correspond au produit d'un vecteur ligne par un vecteur colonne, ce qui donne une matrice 1×1 que l'on identifie à un scalaire. La base (e_i) correspond aux vecteurs colonnes

$$\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \cdots, \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}$$

et la base duale (e_i^*) correspond aux vecteurs lignes

$$(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1).$$

Remarque. Une matrice rectangulaire A peut-être interprétée comme un ensemble de colonnes (images des vecteurs de base) ou comme une union de vecteurs lignes (forme linéaire correspondant à la $i$ème coordonnée de l'application linéaire associée à A).

Remarque. La notation e_i^* pourrait faire penser que e_i^* ne dépend que de e_i : ce n'est pas le cas, e_i^* dépend de la base entière $(e_1, \ldots, e_n)$.

Remarque. Si $\ell \in E^*$ et $x \in E$, on note indifféremment $\ell(x)$ ou $\langle \ell, x \rangle$ la valeur de la forme ℓ évaluée en le vecteur x. Pour la deuxième notation on parle de "crochet de dualité". La notation est justifiée par la ressemblance avec le produit scalaire usuel, quand on utilise une base de E et sa base duale.

1.2. **Formes bilinéaires.** Soit E un $\mathbf{k}$-espace vectoriel. Une application $b\colon E \times E \to \mathbf{k}$ est appelée une *forme bilinéaire* si pour tout $y \in E$, chacune des applications $x \mapsto b(x, y)$ et $x \mapsto b(y, x)$ est une forme linéaire.

On dit que b est symétrique (resp. antisymétrique) si pour tout $x, y \in E$ on a $b(x, y) = b(y, x)$ (resp. $b(x, y) = -b(y, x)$).

À une forme bilinéaire b on associe une application linéaire $\varphi : E \to E^*$ via l'égalité :

$$\forall x, y \in E, \langle \varphi(x), y \rangle = b(x, y).$$

On appelle noyau de b le noyau de φ, on dit que b est non dégénérée si son noyau $\ker b$ est réduit à $\{0\}$.

NB : pour une forme quelconque on aurait en fait des notions de noyaux à droite où à gauche, on se limite en pratique au cas des formes symétriques ou antisymétriques où les deux notions coïncident.

Soit b une forme bilinéaire symétrique ou antisymétrique. On dit que $x, y \in E$ sont *orthogonaux* (sous-entendu : pour la forme b), si $b(x, y) = 0$. On note alors $x \perp y$. Si $A \subset E$, on note $A^\perp$ le sous-espace des vecteurs orthogonaux à tout éléments de A. Le noyau de b se réinterprète comme l'ensemble des $x \in E$ qui sont orthogonaux à tout autre vecteur $y \in E$:

$$\ker b = E^\perp = \{x \in E \mid x \perp y \text{ pour tout } y \in E\}$$

Matrice d'une forme bilinéaire : Si b est une forme bilinéaire, et (e_i) une base de E, on appelle matrice de b dans la base (e_i) la matrice $M = (b(e_i, e_j))$.

Lemme 1. *La matrice $M = (b(e_i, e_j))$ est aussi la transposée de la matrice de $\varphi\colon E \to E^*$ dans les base (e_i) et (e_i^*).*

Preuve. Par définition, la $j$ème colonne de la matrice de $\varphi\colon E \to E^*$ dans les base (e_i) et (e_i^*) s'obtient en exprimant $\varphi(e_j)$ en termes des e_i^* : si

$$\varphi(e_j) = a_1 e_1^* + \cdots + a_n e_n^*$$

alors la $j$ème colonne est

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Et par ailleurs, pour $1 \leqslant i \leqslant n$:

$$a_i = \langle a_1 e_1^* + \cdots + a_n e_n^*, e_i \rangle = \langle \varphi(e_j), e_i \rangle = b(e_j, e_i) \qquad \square$$

NB : pour éviter d'avoir à dire "la transposée" il aurait fallu définir φ comme $\langle y, \varphi(x) \rangle = b(x, y)$, pas terrible. Du coup il faut vraiment la transposée, [CG13, A.1.6] l'oublient tranquilement, en laissant la preuve au lecteur... C'est corrigé dans les "nouvelles histoires".

1.3. **Relation de congruence, problème de la classification.** Si X, Y sont des vecteurs colonne correspondant à des vecteurs $x, y \in E$, on a

$$b(x, y) = X^t M Y.$$

(vérification en exo, il suffit de le faire pour des couples de vecteurs de base).

Si (e_i') est une autre base, et que P est la matrice de passage, de telle sorte que $X = PX', Y = PY'$ où X', Y' sont les vecteurs colonnes correspondant à x, y exprimés dans cette nouvelle base, alors

$$b(x, y) = (PX')^t M (PY') = X'^t (P^t M P) Y'.$$

et donc $P^t M P$ est la matrice de b dans la nouvelle base. On dit que les matrices M et $P^t M P$ sont *congruentes*.

Action par congruence :

$$\begin{array}{ccccc} \mathrm{GL}_n(\mathbf{k}) & \times & \mathrm{Mat}_n(\mathbf{k}) & \to & \mathrm{Mat}_n(\mathbf{k}) \\ (P & , & M) & \mapsto & PMP^t \end{array}$$

est une action (à gauche). Noter que le t a changé de côté.

Comme le rang d'une matrice est invariant par multiplication à droite et à gauche par des matrices inversibles, on voit que le rang est aussi un invariant pour la relation de congruence.

Comme $\det PMP^t = \det M \cdot (\det P)^2$, on obtient un autre invariant (disons pour les non dégénérés) : le déterminant de M modulo un carré dans $\mathbf{k}^*$. La classe de $\det M$ dans le groupe $\mathbf{k}^*/(\mathbf{k}^*)^2$ s'appelle le *discriminant* de b. (Pour une forme dégénérée on commence par quotienter par le noyau pour avoir une définition intéressante).

Exemple.
- $\mathbb{C}^*/(\mathbb{C}^*)^2$ est le groupe trivial (tout complexe est un carré).
- $\mathbb{R}^*/(\mathbb{R}^*)^2$ est un groupe à deux éléments (un réel est un carré si et seulement si il est positif).

- $\mathbb{F}_q^*/(\mathbb{F}_q^*)^2$ est un groupe à deux éléments (Le morphisme $F_q^* \to (\mathbb{F}_q^*)^2$ est surjectif et admet ± 1 pour noyau).
- $\mathbb{Q}^*/(\mathbb{Q}^*)^2$ est un groupe infini dont tous les éléments sont d'ordre 2.

On peut considérer le problème de la classification à congruence près pour diverses classes de formes bilinéaires :

- Aucune condition sur la forme bilinéaire : c'est en fait un problème très dur, c'est pour ça qu'on en parle jamais !
- Formes bilinéaires antisymétriques : c'est intéressant, voir [CG13, page 160]. On en redira un mot dans la §2.3.
- Formes bilinéaires symétriques : le sujet d'étude standard, on se concentre dessus maintenant.

On dit que $q\colon E \to \mathbf{k}$ est une *forme quadratique* s'il existe b une forme bilinéaire tel que $q(x) = b(x,x)$ pour tout $x \in E$. Noter qu'en remplaçant b par $b + b'$ avec b' antisymétrique on ne change pas q. Comme (en caractéristique $\neq 2$) toute matrice est somme d'une matrice symétrique et d'une matrice antisymétrique, on peut toujours se ramener à b symétrique. Une telle forme b symétrique définissant q s'appelle la *forme polaire* de q, et est uniquement déterminée par q via l'une des deux formules (qui nécessitent caractéristique du corps $\neq 2$) :

$$b(x,y) = \frac{1}{4}(q(x+y) - q(x-y)) = \frac{1}{2}(q(x+y) - q(x) - q(y)).$$

On applique tout le vocabulaire précédent indifféremment à q ou à b : on parle de noyau de q, de discriminant de q, etc...

Spoiler alert sur la classification des formes quadratiques (ou formes bilinéaires symétriques) :

(1) Sur $\mathbb{C}$, le rang détermine complètement la classe de congruence (le discriminant ne sert à rien).
(2) Sur $\mathbb{R}$, le rang et le discriminant ne suffisent pas, il faut introduire la signature.
(3) Sur un corps fini $\mathbb{F}_q$, le rang et le discriminant (qui peut prendre deux valeurs) déterminent complètement la classe de congruence.

fin cours no 1

Un peu de vocabulaire supplémentaire :

Définition. Soit b une forme bilinéaire symétrique, et q la forme quadratique associée. Un vecteur $x \in E$ est *isotrope* si $q(x) = 0$. L'ensemble des vecteurs isotropes s'appelle le *cône isotrope*. Un sous-espace vectoriel $F \subseteq E$ est *régulier* (resp. *singulier*) si $q|_F$ est non-dégénérée (resp. dégénérée), ce qui revient à dire $F \cap F^\perp = \{0\}$ (resp. $F \cap F^\perp \supsetneq \{0\}$). F est dit *totalement singulier* (ou *totalement isotrope*) si tous les vecteurs de F sont isotropes, ce qui revient à dire que $q|_F$ est la forme quadratique nulle, ou encore $F \subseteq F^\perp$.

Remarque. Certains auteurs disent qu'un sous espace $F \subseteq E$ est non-isotrope/isotrope au lieu de régulier/singulier, par exemple [Per96, p. 123].

Remarque. Un vecteur est isotrope s'il est orthogonal à lui-même, alors qu'un vecteur est dans le noyau s'il est orthogonal à tout le monde. Donc clairement $\ker b$ est inclu dans le cône isotrope. L'inclusion est stricte en général, et d'ailleurs $\ker b$ est un sev, alors que le cône isotrope n'est pas forcément un sev (c'est juste un cône).

Exo : trouver un exemple de forme quadratique où on a deux inclusions strictes

$$\{0\} \subsetneq \ker q \subsetneq \text{ cône isotrope.}$$

1.4. **Bases orthogonales, diagonalisation.** Notion de base orthogonale pour une forme quadratique q. Matriciellement, cela revient à demander que la matrice soit diagonale. Attention : diagonaliser les formes quadratiques ou les endomorphismes n'a a priori rien à voir (même si plus tard on "diagonalisera en base orthonormée")...

Réf pour les trois lemmes suivants : [CG13, p. 174]. NB : ils marchent aussi pour une forme antisymétrique.

Lemme 2. *Soit b une forme bilinéaire symétrique non dégénérée sur E, et $F \subseteq E$ un sous-espace. Alors* $\dim E = \dim F + \dim F^\perp$.

Preuve. On utilise les deux applications linéaires

$$\varphi_b\colon \begin{array}{ccc} E & \to & E^* \\ x & \mapsto & b(x,\cdot) \end{array} \quad \text{et} \quad r\colon \begin{array}{ccc} E^* & \to & F^* \\ \ell & \mapsto & \ell|_F \end{array}$$

Le théorème du rang appliqué à l'application surjective r donne $\dim E^* = \dim\ker(r) + \dim F^*$. Si $\ell = b(x,\cdot)$, on a $\ell|_F = 0$ ssi $x \in F^\perp$. Ainsi l'isomorphisme φ_b identifie $F^\perp \subset E$ avec $\ker(r) \subset E^*$. Comme de plus $\dim E = \dim E^*$ et $\dim F = \dim F^*$, on obtient l'égalité attendue. □

Lemme 3. *Soit b une forme bilinéaire symétrique quelconque sur E, et $F \subseteq E$ un sous-espace. Alors* $\dim E + \dim(F \cap \ker b) = \dim F + \dim F^\perp$.

Preuve. On considère l'application $\pi\colon E \to E/\ker b$, et on applique le lemme 2 sur le quotient (pour $\bar{x}, \bar{y}$ deux classes dans $E/\ker b$, on définit $\bar{b}(\bar{x},\ bary) := b(x,y)$ et on vérifie que cla ne dépend pas du choix de représentants) :

$$\dim(E/\ker b) = \dim \pi(F) + \dim \pi(F)^\perp.$$

Comme $\pi(F)^\perp = \pi(F^\perp)$ et $\ker b \subseteq F^\perp$ (le noyau est dans l'orthogonal de tout sous-espace), on a

$$\begin{aligned} \dim \pi(F)^\perp &= \dim F^\perp - \dim\ker b, \\ \dim \pi(F) &= \dim F - \dim(F \cap \ker b), \\ \dim(E/\ker b) &= \dim E - \dim\ker b, \end{aligned}$$

d'où le résultat. □

Lemme 4. *Soit b une forme bilinéaire symétrique quelconque sur E, et $F \subseteq E$ un sous-espace régulier. Alors $E = F \oplus F^\perp$.*

Preuve. Par définition F est régulier si $F \cap F^\perp = \{0\}$, et donc $F \cap \ker b = \{0\}$ également. Le lemme 3 donne $\dim E = \dim F + \dim F^\perp$, d'où la somme directe. □

Le théorème qui suit est vraiment spécial aux formes symétriques : on utilise b non nulle implique existence de vecteurs non isotropes. (Par contraste, pour une forme antisymétrique on a $b(x,x) = 0$ pour tout x, même si b est non dégénérée...)

Théorème 5. *Tout forme bilinéaire symétrique b (ou forme quadratique q) admet une base orthogonale.*

Preuve. [CG13, p. 179]. Le cas de la dimension 1 étant clair, montrons le cas de la dimension $n \geqslant 2$, en supposant le résultat acquis en dimension $n-1$. Si q est nulle, n'importe quelle base convient. Sinon, il existe au moins un vecteur e_1 avec $q(e_1) \neq 0$. Comme la droite $\mathrm{Vect}(e_1)$ est régulière, par le lemme 4 on a $E = \mathrm{Vect}(e_1) \oplus \mathrm{Vect}(e_1)^\perp$. Par hypothèse de récurrence q restreint à $\mathrm{Vect}(e_1)^\perp$ admet une base orthogonale, et en y adjoignant e_1 on obtient la base attendue. □

Cette preuve était super courte, mais non effective.

Théorème 6 (Carrés de Gauss)**.** *Il existe un algorithme qui produit une base orthogonale pour une forme quadratique*

$$q(x) = \sum_{1 \leqslant i \leqslant n} a_i x_i^2 + \sum_{1 \leqslant i < j \leqslant n} 2a_{ij} x_i x_j$$

Preuve. Pour une forme nulle, ou en dimension 1, toute base convient. Supposons maintenant $n \geqslant 2$ et q non nulle.

Premier cas : l'un des a_i est non nul. On peut supposer $a_1 = 1$ (permuter les indices, diviser par a_i). On écrit

$$\begin{aligned} q(x) &= x_1^2 + 2x_1\ell + q' \\ &= (x_1 + \ell)^2 + q' - \ell^2 \end{aligned}$$

avec ℓ, q' respectivement linéaire et quadratique en les variables $x_2, \ldots, x_n$.

Deuxième cas : tous les a_i sont nuls, alors l'un des a_{ij} est non nul. On peut supposer $a_{12} = 1$, et on écrit

$$\begin{aligned} q(x) &= x_1x_2 + x_1\ell_1 + x_2\ell_2 + q' \\ &= (x_1 + \ell_2)(x_2 + \ell_1) + q' - \ell_1\ell_2 \\ &= \frac{1}{4}(x_1 + \ell_2 + x_2 + \ell_1)^2 - \frac{1}{4}(x_1 + \ell_2 - x_2 - \ell_1)^2 + q' - \ell_1\ell_2 \end{aligned}$$

avec ℓ_1, ℓ_2, q' respectivement linéaires et quadratique en les variables $x_3, \ldots, x_n$.

Par récurrence, dans les deux cas on a écrit

$$q(x) = c_1 x_1'^2 + \cdots + c_n x_n'^2$$

avec les c_i des scalaires, et les x_i' des formes linéaires indépendantes, et formant donc une base de E^*. La base cherchée est la base antéduale. □

Exemple.

$$q(x_1, x_2) = x_1^2 + 4x_1x_2 = (x_1 + 2x_2)^2 - 4x_2^2 = x_1'^2 - 4x_2'^2,$$

où $x_1' = x_1 + 2x_2$, $x_2' = x_2$. La base antéduale est e_1', e_2' avec $e_1' = (1, 0)$, $e_2' = (-2, 1)$, qui sont bien les solutions des systèmes

$$\begin{cases} x_1 + 2x_2 &= 1 \\ x_2 &= 0 \end{cases} \qquad \begin{cases} x_1 + 2x_2 &= 0 \\ x_2 &= 1 \end{cases}$$

Une autre façon d'obtenir e_1' et e_2' est d'exprimer les x_i en termes des x_i' ("inverser la matrice de passage"), puis lire les vecteurs sur les colonnes :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1' - 2x_2' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}.$$

1.5. **Classification des formes quadratiques.**

1.5.1. *Formes quadratiques sur* $\mathbb{C}$. [CG13, p. 151]

Proposition 7. *Sur* $\mathbb{C}$*, deux formes quadratiques sont congruentes si et seulement si elles sont de même rang.*

Preuve. Soit E un $\mathbb{C}$-espaces vectoriel de dimension n, et q une forme quadratique de rang r sur E. Il suffit de montrer qu'il existe une base (e'_i) de E où la matrice de q est

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

Par le théorème général de diagonalisation on sait qu'il existe une base (e_i) tel que

$$q(x) = a_1 x_1^2 + \cdots + a_r x_r^2,$$

pour certains $a_i \in \mathbb{C}^*$. Il suffit alors de choisir des b_i tel que $b_i^2 = a_i$, et de poser $e'_i = b_i^{-1} e_i$. Dans les coordonnées $x'_i = b_i x_i$ correspondantes on a bien

$$q(x) = x_1'^2 + \cdots + x_r'^2. \qquad \square$$

1.5.2. *Formes quadratiques sur* $\mathbb{R}$. Sur le corps des réels, on a une notion de nombre positif. Du coup on peut faire les définitions suivantes :

Définition. Soit q une forme quadratique sur un espace vectoriel réel E. On dit que q est définie positive si $q(x) > 0$ pour tout vecteur $x \neq 0$. De même on dit que q est définie négative si $q(x) < 0$ pour tout vecteur $x \neq 0$.

Les formes définies positives ou négatives sont des formes quadratiques non dégénérées particulières. On peut aussi parler de formes quadratiques positives ou négatives en mettant des inégalités larges, mais ces formes peuvent alors être dégénérée.

Exemple. Sur $\mathbb{R}^2$:

- $q(x_1, x_2) = x_1^2 + x_2^2$ est définie positive.
- $q(x_1, x_2) = -x_1^2 - x_2^2$ est définie négative.
- $q(x_1, x_2) = x_1 x_2$ est non dégénérée mais ni positive ni négative.
- $q(x_1, x_2) = x_1^2$ est positive mais non définie (c'est-à-dire dégénérée).

NB : les deux premières sont de même rang, de même discriminant, mais ne sont pas congruentes.

Définition. Soit q une forme quadratique sur un espace vectoriel réel E. Notons s la dimension maximale d'un sous-espace $F \subseteq E$ tel que $q|_F$ soit définie positive, et de façon similaire notons t la dimension maximale où la restriction de q est définie négative. Le couple (s, t) s'appelle la *signature* de q.

Théorème 8 (d'inertie de Sylvester)**.** *Soit q une forme quadratique de signature (s, t) sur un espace vectoriel réel E, et (e_i) une base orthogonale. Alors la matrice diagonale de q dans la base (e_i) admet s coefficients > 0, et t coefficients < 0.*

Preuve. Soit s' et t' les nombres de coefficients respectivement > 0 et < 0. On a donc $\dim E = s' + t' + \dim \ker q$. En prenant les sous-espaces engendrés par les vecteurs de base correspondants, on obtient $s' \leqslant s$ et $t' \leqslant t$. Mais si F_+, F_- sont des sous-espaces sur lesquels la restriction de q est respectivement définie positive et négative, on a F_+, F_- et $\ker q$ en somme directe, donc en appliquant ceci à des sous-espaces maximaux on obtient $s + t + \dim \ker q \leqslant \dim E$. Donc finalement on obtient la chaîne d'inégalités

$$\dim E = s' + t' + \dim \ker q \leqslant s + t + \dim \ker q \leqslant \dim E,$$

donc les inégalités sont des égalités, et $s = s'$, $t = t'$. $\square$

Corollaire 9. *Sur $\mathbb{R}$, deux formes quadratiques sont congruentes si et seulement si elles sont de même signature.*

Preuve. Soit E un $\mathbb{R}$-espaces vectoriel de dimension n, et q une forme quadratique de signature (s,t) sur E. Il suffit de montrer qu'il existe une base (e'_i) de E où la matrice de q est

$$\begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

C'est immédiat à partir du théorème de Sylvester, en procédant comme pour la proposition 7. □

1.5.3. *Formes quadratiques sur un corps fini.*

Proposition 10. *Sur un corps fini, deux formes quadratiques non dégénérées sont congruentes si et seulement si elles ont même discriminant.*

Remarque. Pour inclure le cas dégénéré il faudrait introduire la notion de discriminant réduit (discriminant après avoir quotienté par le noyau), et on pourrait énoncer "deux formes quadratiques sont congruentes si et seulement si elles ont même rang et même discriminant réduit".

NB : voir [CG13, p. 153] pour la preuve en dimension 2, le cas général s'en déduit.

1.6. **Diagonalisation en base orthonormée.**

Lemme 11. *Soit $A \in \mathrm{Mat}_n(\mathbb{R})$ une matrice symétrique. Alors les valeurs propres de A sont réelles.*

Preuve. On considère A comme un endomorphisme de $\mathbb{C}^n$, et on muni $\mathbb{C}^n$ du produit hermitien standard $\langle \cdot, \cdot \rangle$. Si λ est une valeur propre de A, de vecteur propre $v \in \mathbb{C}^n$, on a

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle$$

d'où $\lambda = \bar{\lambda}$, autrement dit $\lambda \in \mathbb{R}$. □

Remarque. Cette preuve super courte a aussi l'avantage de marcher sans changer une seule virgule dans le cas hermitien.

fin cours no 2

Lemme 12. *Soit $A \in \mathrm{Mat}_n(\mathbb{R})$ une matrice symétrique, et munissons $\mathbb{R}^n$ du produit scalaire standard $\langle \cdot, \cdot \rangle$. Si $F \subseteq \mathbb{R}^n$ est un sous-espace invariant par A, alors $F^\perp$ est aussi invariant par A.*

Preuve. Soit $w \in F^\perp$, on veut montrer $Aw \in F^\perp$. Pour tout $v \in F$, on a

$$\langle Aw, v\rangle = \langle w, Av\rangle = 0$$

car $Av \in F$. □

Remarque. Ce lemme est juste un cas particulier de : si F est invariant par A, alors $F^\perp$ est invariant par A^T. (même preuve).

Théorème 13. *Toute matrice symétrique réelle A est diagonalisable dans une base orthonormée.*

Preuve. Soit $v \in \mathbb{R}^n$ un vecteur propre de A (il y en a par le lemme 11), que l'on peut supposer de norme 1. Par le lemme 12 l'hyperplan $v^\perp$ est aussi stable. Si $w_1, w_2 \in v^\perp$, alors $\langle Aw_1, w_2\rangle = \langle w_1, Aw_2\rangle$ ce qui montre que la restriction de A à $v^\perp$ est toujours auto-adjointe (et donc la matrice de la restriction est symétrique), et on conclut par récurrence, le cas d'une matrice de taille 1 étant clair. □

Corollaire 14. *Soit A une matrice symétrique réelle, et notons s (resp. t) le nombre de valeurs propres > 0 (resp. < 0) de A. Alors A est de signature (s, t).*

Preuve. Immédiat en écrivant $A = P^tDP = P^{-1}DP$ avec D diagonale. □

Corollaire 15. *Soit E un espace vectoriel réel, et q_1, q_2 deux formes quadratiques, avec q_1 définie positive. Alors il existe une base qui diagonalise simultanément q_1 et q_2.*

Preuve. Comme q_1 est définie positive, il existe une base où la matrice de q_1 est la matrice identité, c'est-à-dire q_1 devient le produit scalaire standard dans cette base. On applique alors le théorème à q_2 : il existe un changement de base orthonormé (donc préservant q_1), qui diagonalise q_2. □

2. Groupe orthogonal

Référence : [Per96, p.123-126]

Dans toute cette section on se donne $\mathbf{k}$ un corps de caractéristique différente de 2, E un $\mathbf{k}$-espace vectoriel de dimension n, q une forme quadratique non dégénérée sur E, et b sa forme polaire.

2.1. Groupe préservant une forme quadratique.

Lemme 16. *Soit $u \in \mathcal{L}(E)$. Les deux propriétés suivantes sont équivalentes :*

(1) *u préserve la forme quadratique q :*
$$\forall x \in E, q(u(x)) = q(x).$$

(2) *u préserve la forme bilinéaire b :*
$$\forall x, y \in E, b(u(x), u(y)) = b(x, y).$$

De plus tout endomorphisme u vérifiant ces propriétés est inversible.

Preuve. Le sens (2) $\implies$ (1) est clair, et la réciproque suit de l'égalité suivante, qui utilise l'hypothèse $\operatorname{car}\mathbf{k} \neq 2$:
$$b(x, y) = \frac{1}{4}(q(x + y) - q(x - y)).$$
Si $x \in \ker u$, pour tout $y \in E$ on a $b(x, y) = b(u(x), u(y)) = b(0, u(y)) = 0$, et comme q et non dégénérée on obtient $x = 0$. □

Définition. On appelle *groupe orthogonal* associé à q, noté $\mathrm{O}(q)$, le sous-groupe des $u \in \mathrm{GL}(E)$ préservant la forme quadratique q. Matriciellement, si dans une base donnée q a pour matrice symétrique A et u a pour matrice M, on a
$$u \in \mathrm{O}(q) \iff M^tAM = A.$$
(car on doit avoir $x^tM^tAMy = x^tAy$ pour tous $x, y \in E$, en particulier pour tout couple de vecteurs de base). Le groupe *spécial orthogonal* $\mathrm{SO}(q)$ est le sous-groupe de $\mathrm{O}(q)$ des isométries de déterminant 1.

Remarque. (1) Les seules homothéties dans $\mathrm{O}(q)$ sont $\pm\mathrm{id}$: en effet si x est un vecteur non isotrope

$$q(\lambda x) = q(x) \iff \lambda^2 = 1.$$

(2) Tout élément $u \in \mathrm{O}(q)$ est de déterminant ± 1 :

$$\det(M^t A M) = \det A \implies (\det M)^2 = 1.$$

(3) Dans le cas $A = \mathrm{id}$ la condition devient $M^t M = \mathrm{id}$ ou encore $M^t = M^{-1}$, ce n'est qu'un cas particulier, que l'on approfondira plus tard (sur le corps $\mathbb{R}$).

Lemme 17. *Si $u \in \mathrm{O}(q)$ et $F \subseteq E$ est un sous-espace vectoriel, alors $u(F^\perp) = u(F)^\perp$. En particulier si F est stable par u, alors $F^\perp$ est aussi stable par u.*

Preuve. Pour tous $x \in F$, $y \in F^\perp$ on a

$$0 = b(x, y) = b(u(x), u(y)).$$

Cela implique $u(F^\perp) \subseteq u(F)^\perp$ (et aussi $u(F^\perp)^\perp \supseteq u(F)$, mais celle-ci ne va pas nous servir). De plus on a égalité des dimensions :

$$\dim u(F^\perp) = \dim F^\perp = \dim E - \dim F = \dim E - \dim u(F) = \dim u(F)^\perp,$$

d'où l'égalité attendue $u(F^\perp) = u(F)^\perp$. Si $u(F) = F$, on obtient donc $u(F^\perp) = u(F)^\perp = F^\perp$. □

Remarque. Une notion voisine est celle d'isométrie entre deux espaces quadratiques (E, q) et (E', q') : on dit qu'une application linéaire bijective $u\colon E \to E'$ est une *isométrie* si pour tout $x \in E$ on a $q'(u(x)) = q(x)$.

2.2. **Symétries, réflexions, renversements.**

Rappel. Si $u \in \mathrm{GL}(E)$ vérifie $u^2 = \mathrm{id}$, ou autrement dit si $X^2 - 1$ est un polynôme annulateur pour u, alors $E = E^+ \oplus E^-$ est la somme directe des espaces propres associés aux valeurs propres ± 1 (lemme des noyaux), et E admet une base pour laquelle la matrice de u est de la forme

$$\begin{pmatrix} +1 & & & & & \\ & \ddots & & & & \\ & & +1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{pmatrix}$$

Si $\dim E^- > 0$ (c'est-à-dire si $u \neq \mathrm{id}$) on dit que u est une *involution* (ou une *symétrie*), si $\dim E^- = 1$ on dit que u est une *réflexion*, et si $\dim E^- = 2$ on dit que u est un *renversement*. Si de plus $u \in \mathrm{O}(q)$, on parle alors de symétrie orthogonale, réflexion orthogonale ou renversement orthogonal.

Exemple. $(x, y) \mapsto (-x + y, y)$ est une réflexion de $E = \mathbb{R}^2$, avec $E^- = \mathrm{Vect}(1, 0)$ et $E^+ = \mathrm{Vect}(1, 2)$, mais n'est pas une réflexion orthogonale pour le produit euclidien standard. Par contre c'est une réflexion orthogonale par rapport à la forme quadratique définie par la matrice $\left(\begin{smallmatrix} -2 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Proposition 18. *Soit $u \in \mathrm{GL}(E)$ vérifiant $u^2 = \mathrm{id}$. Alors $u \in \mathrm{O}(q)$ ssi E^+ et E^- sont orthogonaux (et en particulier E^+ et E^- sont réguliers). De plus, si $F \subsetneq E$ est un sous-espace régulier, il existe une unique symétrie $u \in \mathrm{O}(q)$ tel que $E^+ = F$.*

Preuve. Soit $u \in \mathrm{O}(q)$ une symétrie, et $x \in E^+$, $y \in E^-$. Alors

$$b(x,y) = b(u(x),u(y)) = b(x,-y) = -b(x,y),$$

ce qui donne $b(x,y) = 0$.

Réciproquement, si $E^+ \perp E^-$ et $x, y \in E$, on écrit $x = x^+ + x^-$, $y = y^+ + y^-$ les décompositions dans la somme directe $E = E^+ \oplus E^-$, et on obtient

$$u(x) = x^+ - x^-, \qquad u(y) = y^+ - y^-.$$

Ainsi (les termes croisés étant nuls)

$$b(x,y) = b(x^+,y^+) + b(x^-,y^-) = b(u(x),u(y)).$$

Enfin, si F est régulier, ce qui implique $E = F \oplus F^\perp$, u est uniquement définie par les conditions $u|_F = \mathrm{id}$ et $u|_{F^\perp} = -\mathrm{id}$. □

Notation. Une réflexion orthogonale u est entièrement déterminée par son hyperplan H de point fixe, ou par la droite régulière $D = H^\perp$, ou encore par un vecteur non nul $x \in D$. On notera $u = u_H = u_D = u_x$ suivant le contexte. Explicitement, on a

$$u_x(y) = y - 2\frac{b(x,y)}{b(x,x)}x.$$

2.3. **Plan hyperbolique.** La notion de plan hyperbolique [Per96, p.186] intervient souvent dans des raisonnements par récurrence sur la dimension, et également dans la preuve du (difficile) théorème de Witt.

Exercice. Par contre à l'aide de la notion de signature le théorème de Witt sur $\mathbb{R}$ est très abordable (voir [CG13, D.12 p. 191]). Voici une version :

Soit E un espace vectoriel de dimension finie sur $\mathbb{R}$, muni d'une forme quadratique non dégénérée q. On suppose que $F, F' \subseteq E$ sont deux sous-espaces isométriques. Il s'agit de montrer que $F^\perp$ et $F'^\perp$ sont aussi isométriques (on se convaincra que cela équivaut à dire : "l'isométrie de F vers F' est induite par un élément $u \in \mathrm{O}(q)$"). On pourra commencer par le cas où F, F' sont réguliers.

Définition. Soit P un espace vectoriel de dimension 2 sur $\mathbf{k}$ et q une forme quadratique sur P de forme polaire b. On dit que (P,q) est un *plan hyperbolique* s'il existe une base (e_1,e_2) de P tel que $q(e_1) = q(e_2) = 0$, $b(e_1,e_2) = 1$. La matrice de q dans cette base est donc $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.

Exemple. Dans $E = \mathbb{R}^2$ avec $q\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = 2x_1x_2$, les niveaux $q\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = r$ sont des hyperboles pour tout $r \neq 0$, d'où le nom.

Remarque. Un plan hyperbolique contient exactement deux droites singulières, car $q\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) = 2x_1x_2 = 0$ équivaut à $x_1 = 0$ ou $x_2 = 0$.

Lemme 19. *Soit (P,q) un plan muni d'une forme quadratique. Alors (P,q) est un plan hyperbolique ssi q est non dégénérée et P admet un vecteur isotrope non nul.*

Preuve. Le sens direct est clair, supposons donc q non dégénérée et P admettant un vecteur isotrope non nul e_1. Comme q est non dégénérée, il existe $v \in P$ tel que $b(e_1,v) = a \neq 0$. Quitte à remplacer v par $v - \frac{q(v)}{2a}e_1$, on peut supposer v isotrope sans changer la valeur $a = b(e_1,v)$. En posant $e_2 = v/a$, on obtient la base (e_1,e_2) attendue. □

Remarque. Il vaut la peine de méditer un moment sur le lemme précédent. Prenons un corps de base un peu compliqué, par exemple $\mathbf{k} = \mathbb{Q}$.

(1) Les matrices $\begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ pour $p \in \mathbb{N}$ premier correspondent à une famille infinie de formes quadratiques non dégénérées non congruentes sur $\mathbb{Q}^2$ (les $q(x_1, x_2) = x_1^2 + px_2^2$). Elles sont toutes sans vecteurs isotropes non nuls.
(2) Par contraste, le lemme dit qu'à congruence près il n'existe qu'*une seule* forme quadratique non dégénérée sur $\mathbb{Q}^2$ admettant des vecteurs isotropes.

Remarque. On peut faire le lien en passant avec la classification à congruence près des formes antisymétriques. Le seul invariant dans ce cas est le rang, et les briques élémentaires (à partir desquelles on reconstruit tout par somme directe) sont des plans avec une base dans laquelle la forme admet pour matrice $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Cela a un air de famille avec un plan hyperbolique, et par analogie cela donne envie de tenter de dévisser toute forme bilinéaire symétrique admettant au moins un vecteur isotrope comme somme directe d'un plan hyperbolique et de son orthogonal. C'est l'idée de départ de la théorie de Witt.

2.4. Centre du groupe orthogonal.

Lemme 20. *Soit $u \in \mathrm{GL}(E)$.*

(1) *Si pour toute droite $D \subset E$ on a $u(D) = D$, alors u est une homothétie.*
(2) *Si $\dim E = 2$ et il existe 3 vecteurs propres deux à deux non colinéaires pour u, alors u est une homothétie.*

Preuve. Exercice (de révision !) □

fin cours no 3

Proposition 21 ([Per96, p.186-187])**.** *Le centre de $\mathrm{O}(q)$ est $\{\mathrm{id}, -\mathrm{id}\}$, sauf dans le cas d'un plan hyperbolique sur le corps $\mathbb{F}_3$.*

En fait dans la preuve on va commencer par montrer que le centralisateur de $\mathrm{O}(q)$ dans $\mathrm{GL}(E)$ est le groupe $\mathbf{k}^*$ des homothéties :

Preuve. Soit D une droite régulière, et τ_D la réflexion orthogonale associée. Si $u \in \mathrm{GL}(E)$ centralise $\mathrm{O}(q)$, on a $\tau_D = u\tau_D u^{-1} = \tau_{u(D)}$, et donc u préserve la droite D.

Si $\dim E = 2$, alors ou bien toutes les droites de E sont régulières et on conclut par le lemme 20(1), ou bien E est un plan hyperbolique avec donc exactement 2 droites singulières, et donc au moins 3 droites régulières qui permettent encore de conclure par le lemme 20(2), sauf dans le cas du corps de base $\mathbb{F}_3$ (il y a $q+1$ droites vectorielles dans le plan sur le corps fini $\mathbb{F}_q$).

Supposons maintenant $\dim E \geqslant 3$. Si $P \subset E$ est un plan régulier, alors il est déterminé par deux de ses droites régulières (par exemple une base orthogonale), et donc un tel plan est aussi laissé invariant par u. Remarquons alors que toute droite singulière D est intersection de deux plans réguliers P_1, P_2 :

En effet si $D = \mathbf{k}v$, on peut trouver $v_1 \in E$ tel que $b(v, v_1) \neq 0$, et donc $P_1 = \mathbf{k}v + \mathbf{k}v_1$ est un premier plan régulier contenant D. Soit v_2 un vecteur non nul dans $P_1^\perp$. Alors $P_2 = \mathbf{k}v + \mathbf{k}(v_1 + v_2)$ est un autre plan régulier contenant D.

On conclut de nouveau par le lemme 20(1) que u est une homothétie, et si de plus $u \in \mathrm{O}(q)$, alors $u = \pm\mathrm{id}$. □

Remarque (Cas du plan hyperbolique sur $\mathbb{F}_3$)**.** Si $q = xy$ (ou $2xy$) sur le corps à 3 éléments $\mathbb{F}_3$, alors

$$\mathrm{O}(q) = \left\{ \begin{pmatrix} +1 & 0 \\ 0 & +1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & +1 \\ +1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \right\},$$

qui est isomorphe au groupe abélien $\mathbb{Z}/2 \times \mathbb{Z}/2$.

2.5. **Générateurs.**

Théorème 22 ([Per96, p.187])**.** *Le groupe* $\mathrm{O}(q)$ *est engendré par les réflexions orthogonales.*

Preuve. Pour $n = 1$, on a $q(x) = ax^2$ pour un certain $a \neq 0$, et $\mathrm{O}(q) = \{+1, -1\}$ d'ordre 2 engendré par la réflexion orthogonale $x \mapsto -x$ (parce qu'on est en dimension 1 !). Supposons maintenant $n > 1$, et le résultat prouvé jusqu'en dimension $n - 1$.

Traitons d'abord le cas particulier où il existe un vecteur x non isotrope et fixe par u. Posons $L = x^\perp$ qui est stable par u par le lemme 17, alors $q|_L$ est non dégénérée, et on peut appliquer l'hypothèse de récurrence pour écrire $u|_L$ comme un produit de réflexions sur L :

$$u|_L = \tau_r \circ \cdots \circ \tau_1.$$

En prolongeant chaque τ_i en la réflexion σ_i telle que $\sigma_i(x) = x$ et $\sigma_i|_L = \tau_i$ on obtient le résultat attendu.

Finalement, montrons qu'on peut toujours en composant u par au plus deux réflexions se ramener au cas où il existe un vecteur fixe non isotrope. Tout d'abord comme q est non dégénérée il existe un vecteur x non isotrope. Si $y = u(x)$ est distinct de x, alors au moins l'un des vecteurs $x - y$ ou $x + y$ est non isotrope, car ces deux vecteurs sont orthogonaux et (le cas facile $y = -x$ mis à part) forment une base du plan $\mathbf{k}x + \mathbf{k}y$ qui n'est pas totalement isotrope.

Si $x-y$ est non isotrope, alors $H = (x-y)^\perp$ contient $x+y$, et la réflexion orthogonale τ_H satisfait

$$\tau_H(x + y) = x + y, \qquad \tau_H(x - y) = y - x,$$

d'où $x = \tau_H(y) = (\tau_H \circ u)(x)$.

Si $x + y$ est non isotrope, alors cette fois en posant $H = (x+y)^\perp$ on obtient $\tau_H(y) = -x$. En notant $L = x^\perp$, on a $\tau_L(x) = -x$ et donc $(\tau_L \circ \tau_H \circ u)(x) = x$. □

Corollaire 23. *Pour* $n \geqslant 3$*, le groupe* $\mathrm{SO}(q)$ *est engendré par les renversements.*

Preuve. Par le théorème, tout élément de $\mathrm{SO}(q)$ est la composée d'un nombre pair de réflexions. Il reste à voir que si τ_1, τ_2 sont des réflexions orthogonales (distinctes), alors il existe des renversements orthogonaux σ_1, σ_2 tels que $\sigma_1\sigma_2 = \tau_1\tau_2$. Si $n = 3$, il suffit de poser $\sigma_i = -\tau_i$ pour $i = 1, 2$. Dans le cas $n \geqslant 4$, notons $H_i = x_i^\perp$ l'hyperplan de la réflexion τ_i, on a donc $(H_1 \cap H_2)^\perp = \mathbf{k}x_1 + \mathbf{k}x_2$ de dimension 2. On a $\operatorname{Ker} q|_{H_1 \cap H_2} = (H_1 \cap H_2) \cap (\mathbf{k}x_1 + \mathbf{k}x_2)$ qui est de dimension au plus 1 (par exemple x_1 n'est pas dans cette intersection, car $x_1 \notin H_1$), et donc il existe $V \subset H_1 \cap H_2$ régulier et de dimension $n-3$. On applique alors le raisonnement précédent aux restrictions de τ_1, τ_2 à $V^\perp$, et on étend trivialement à V, pour obtenir les renversements attendus sur $E = V^\perp \oplus V$. □

Remarque. Méditer sur l'analogie des résultats (et des preuves) précédents avec les résultats concernant le groupe symétrique : les transpositions engendrent S_n, et les 3-cycles engendrent A_n.

2.6. **Le cas de la dimension 2.** On termine avec des propriétés familières pour les isométries du plan euclidien $\mathbb{R}^2$, mais qui sont en fait valables pour un plan sur un corps arbitraire muni d'une forme quadratique non dégénérée :

Proposition 24. *Soit* P *un plan sur* $\mathbf{k}$*, et* q *une forme quadratique non dégénérée.*

(1) *Si* $u \in \mathrm{O}(q)$ *avec* $\det u = -1$*, alors* u *est une réflexion orthogonale.*

(2) *Si $u \in \mathrm{SO}(q)$ fixe un vecteur x non nul, alors $u = \mathrm{id}$.* [NB : [Per96, p. 188] suppose x non isotrope, c'est inutile]

(3) *Tout $u \in \mathrm{SO}(q)$ s'écrit comme un produit de deux réflexions orthogonales $u = \tau_1\tau_2$, et on peut même imposer un choix arbitraire pour l'une des deux.*

(4) *Si $u \in \mathrm{SO}(q)$ et τ est une réflexion, alors $\tau u \tau = u^{-1}$.*

(5) *Le groupe $\mathrm{SO}(q)$ est abélien.*

Preuve. (1) Choisissons x non nul et non isotrope. On a vu dans la preuve du théorème qu'il existe une réflexion orthogonale τ telle que $\tau \circ u(x) = \pm x$. Comme $\det(\tau \circ u) = 1$ et $\varepsilon = \pm 1$ est valeur propre de $\tau \circ u$, on déduit que ε est la seule valeur propre de $\tau \circ u$. Enfin comme la droite $x^\perp$ est stable par $\tau \circ u$ et distincte de $\mathbf{k}x$, on conclut que $\tau \circ u = \pm\mathrm{id}$, et donc $u = \pm\tau$. Reste à remarquer qu'en dimension 2 l'isométrie $-\tau$ est encore une réflexion.

(2) Par le même argument on obtient que 1 est la seule valeur propre de u. Si x est non isotrope, on obtient une seconde droite propre avec $x^\perp$. Si x est isotrope, alors P est un plan hyperbolique, donc contient exactement 2 droites singulières, qui sont donc chacune des droites propres.

(3) Si disons τ_1 est donnée, alors $\det \tau_1 u = -1$ qui est donc une réflexion τ_2.

(4) On écrit $u = \tau\tau'$ pour une certaine réflexion τ', et

$$\tau u \tau = \tau^2 \tau' \tau = \tau' \tau = u^{-1}.$$

(5) Soit $u, v \in \mathrm{SO}(q)$, on peut écrire $u = \tau_1\tau_2$ comme un produit de deux réflexions, et

$$uvu^{-1} = \tau_1\tau_2 v \tau_2\tau_1 = \tau_1 v^{-1}\tau_1 = v. \qquad \square$$

Exemple. Un développement qui a du succès est l'étude du groupe $\mathrm{O}(p, q)$. Si on le présente il est bon d'avoir décortiqué le cas de $\mathrm{O}(1, 1)$, c'est-à-dire des isométries du plan hyperbolique sur $\mathbb{R}$. Prenons la forme quadratique $q(x, y) = xy$. On a

$$\mathrm{SO}(1, 1) = \{(x, y) \mapsto (\lambda x, \lambda^{-1} y) \mid \lambda \in \mathbb{R}^*\}.$$

et $\mathrm{O}(1, 1)$ est le produit semi-direct (et pas direct !) du groupe distingué $\mathrm{SO}(1, 1)$ et du sous-groupe d'ordre 2 engendré par $(x, y) \mapsto (y, x)$.

3. Isométries d'un espace euclidien

Un espace vectoriel réel E de dimension finie n muni d'une forme quadratique q définie positive est appelé un *espace euclidien*. Autrement dit, E est isométrique à $\mathbb{R}^n$ muni du produit scalaire standard, on pourra donc s'autoriser (via un choix de base) à identifier E au modèle $\mathbb{R}^n$ standard. On notera $\langle \cdot, \cdot \rangle$ le produit scalaire (forme polaire de q), et $\mathrm{O}(E)$ ou $\mathrm{O}_n(\mathbb{R})$ le groupe orthogonal.

3.1. **Dimension 2.** Une matrice $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ est dans $\mathrm{O}_2(\mathbb{R})$ si et seulement si $A^t A = \mathrm{id}$, ce qui équivaut à

$$\begin{cases} a^2 + c^2 = 1 \\ ab + cd = 0 \\ b^2 + d^2 = 1 \end{cases}$$

La première condition donne l'existence d'un θ tel que $a = \cos\theta$, $c = \sin\theta$; la deuxième condition donne l'existence d'un $\lambda \in \mathbb{R}^*$ tel que $b = -\lambda c$, $d = \lambda a$; et enfin la troisième

condition donne $\lambda = \pm 1$. Remarquons que $\lambda = \pm 1$ est le déterminant de la matrice. Finalement on obtient que

$$\mathrm{SO}_2(\mathbb{R}) = \left\{ R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mid \theta \in \mathbb{R} \right\}$$

est le groupe des rotations, et toute matrice dans $\mathrm{O}_2(\mathbb{R})$ de déterminant -1 est de la forme

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

c'est à dire composition de la symétrie S d'axe les abscisses avec la rotation R_θ. Comme on l'a déjà noté de façon générale dans la proposition 24, une telle matrice S_θ est une réflexion orthogonale par rapport à une certaine droite. On peut le retrouver et préciser la conjugante :

$$S_\theta = R_\theta \circ S = R_{\theta/2} \circ S \circ (S \circ R_{\theta/2} \circ S) = R_{\theta/2} \circ S \circ R_{\theta/2}^{-1}.$$

fin cours no 4

Remarque. On a des isomorphismes

$$\begin{array}{ccc} & \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \in \mathrm{SO}_2(\mathbb{R}) & \\ \swarrow^{\simeq} & & \searrow^{\simeq} \\ \bar{\theta} \in \mathbb{R}/2\pi\mathbb{Z} & \xrightarrow{\quad\simeq\quad} & e^{i\theta} \in S^1 \end{array}$$

Lemme 25. *Tout sous-groupe fini $G \subset \mathrm{SO}_2(\mathbb{R})$ est cyclique, de la forme $G = \langle R_\theta \rangle$ avec $\theta = \frac{2\pi}{n}$ où $n = |G|$.*

Preuve. Grâce à l'isomorphisme de la remarque, il suffit de montrer l'énoncé analogue pour un sous-groupe $G \subset S^1 \subset \mathbb{C}^*$ d'ordre n. Pour tout $z \in G$, ordre z divise n et donc $z^n = 1$. Autrement dit z est une racine nième de l'unité, et par égalité des cardinaux l'inclusion $G \subseteq U_n$ est en fait une égalité :

$$G = U_n = \{ e^{2i\pi \frac{k}{n}} \mid k = 0, \ldots, n-1 \} = \langle e^{\frac{2i\pi}{n}} \rangle. \qquad \square$$

Remarque. Remarquer la simplification apportée par l'usage des nombres complexes : on s'est servi du fait qu'un polynôme de degré n sur $\mathbb{C}$ a exactement n racines (ici le polynôme $X^n - 1$).

Rappelons qu'on appelle groupe *diédral* D_n le groupe des isométries du plan préservant le polygone régulier dont les sommets sont les racines $n$ièmes de l'unité. Explicitement, en utilisant les nombres complexes et en notant $\alpha = e^{\frac{2i\pi}{n}}$:

$$D_n = \{ z \mapsto \alpha^k z \} \cup \{ z \mapsto \alpha^k \bar{z} \}$$

où $k = 0, \ldots, n-1$. C'est un groupe non abélien sauf quand $n = 1$ ($D_1 \simeq \mathbb{Z}/2$) ou $n = 2$ ($D_2 \simeq \mathbb{Z}/2 \times \mathbb{Z}/2$).

Lemme 26. *Tout sous-groupe fini $G \subset \mathrm{O}_2(\mathbb{R})$ non inclus dans $\mathrm{SO}_2(\mathbb{R})$ est isomorphe (et même conjugué) au groupe diédral D_n pour un certain $n \geqslant 1$.*

Preuve. Par hypothèse il existe $\tau \in G$ de déterminant -1. C'est une réflexion orthogonale, et il existe une rotation $r \in \mathrm{SO}(2, \mathbb{R})$ qui conjugue cette réflexion à la réflexion $z \mapsto \bar{z}$ (réflexion par rapport à l'axe des abscisses).

De plus $G \cap \mathrm{SO}_2(\mathbb{R})$ est d'indice 2 dans G, et cyclique d'un certain ordre n, engendré par $z \mapsto \alpha^k z$ avec $\alpha = e^{2i\pi/n}$.

Finalement rGr^{-1} contient le groupe diédral D_n, donc lui est égal par égalité des cardinaux. □

3.2. **Dimension n.** Forme normale des éléments de $\mathrm{O}_n(\mathbb{R})$:

Théorème 27 ([Com98, p. 155], [Per96, p. 147])**.** *Soit $v \in \mathrm{O}_n(\mathbb{R})$. Alors $\mathbb{R}^n = \oplus F_k$ est somme directe orthogonale de sous-espaces F_k non nuls invariants par v, et minimaux pour cette propriété. Ces sous-espaces sont de dimension 1 ou 2. Ceux de dimension 1 sont engendrés par des vecteurs propres associés aux valeurs propres -1 ou $+1$. Pour ceux de dimension 2, et pour tout choix d'une base orthonormée d'un tel F_k, la matrice de la restriction de v à F_k est de la forme (matrice de rotation)*

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

avec θ non multiple de π.

Preuve. On procède par récurrence sur la dimension n. Le cas $n = 1$ est clair, et le cas $n = 2$ a été traité plus haut. Supposons $n \geqslant 3$, et montrons qu'il existe un sous-espace F non-trivial stable par v, ce qui permettra de conclure en appliquant l'hypothèse de récurrence à chacun des facteurs de la décomposition orthogonale $F \oplus F^\perp$. Si v admet une valeur propre réelle, nécessairement égale à ± 1, c'est terminé. Sinon, soit $\lambda, \bar{\lambda}$ un couple de valeurs propres complexes conjuguées, et $x \in \mathbb{C}^n$ un vecteur propre complexe pour λ. Alors $\bar{x}$ est un vecteur propre pour $\bar{\lambda}$, et le plan $P = \mathbb{C}x + \mathbb{C}\bar{x}$ est invariant par u. Comme u est réel, u préserve $F = P \cap \mathbb{R}^n$, et $P \cap \mathbb{R}^n$ est le plan réel engendré par les vecteurs réels $x + \bar{x}$ et $i(x - \bar{x})$. □

On reprend et précise le théorème 22 dans le cas d'un espace euclidien E, que l'on identifie à $\mathbb{R}^n$ muni du produit scalaire standard.

Théorème 28 ([Per96, p. 143])**.** *Le groupe $\mathrm{O}_n(\mathbb{R})$ est engendré par les réflexions orthogonales. Précisément, si $u \in \mathrm{O}_n(\mathbb{R})$ avec espace des points fixes $\operatorname{Fix} u$ de codimension r, alors u s'écrit comme un produit de r réflexions orthogonales.*

Preuve. Clairement si $u = \tau_1 \circ \ldots \tau_k$ est un produit de k réflexions orthogonales, alors $k \geqslant \operatorname{codim} \operatorname{Fix} u$, c'est-à-dire il faut au moins $\operatorname{codim} \operatorname{Fix} u$ réflexions pour écrire u. Il suffit donc de montrer que u s'écrit comme un produit d'au plus $\operatorname{codim} \operatorname{Fix} u$ réflexions orthogonales.

On procède par récurrence sur la codimension r, le cas $r = 0$ correspondant à $u = \mathrm{id}$.

Si $r > 0$, prenons $x \in (\operatorname{Fix} u)^\perp$ non nul, et notons $y = u(x)$, qui est aussi dans $(\operatorname{Fix} u)^\perp$ par le lemme 17. Soit τ la réflexion de vecteur $x - y$. On remarque que

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

Donc $\tau(x - y) = y - x$, $\tau(x + y) = x + y$, et donc $\tau(x) = y$. Comme $x - y \in (\operatorname{Fix} u)^\perp$, $\tau|_{\operatorname{Fix} u} = \mathrm{id}$. On a donc une inclusion $\operatorname{Fix} u \subset \operatorname{Fix}(\tau u)$, et cette inclusion est stricte car x est fixé par τu mais pas par u. Par hypothèse de récurrence, on peut écrire τu comme un produit d'au plus $\operatorname{codim} \operatorname{Fix}(\tau u) < r$ réflexions orthogonales. Donc u s'écrit comme un produit d'au plus $\operatorname{codim} \operatorname{Fix}(\tau u) + 1 \leqslant r$ réflexions orthogonales. □

3.3. **Dimension 3.** ref : [Com98, chapitre 8]

On prend maintenant $n = 3$, $\mathbf{k} = \mathbb{R}$. De l'égalité $M^{-1} = M^t$ on déduit $\det M = \pm 1$; de plus les valeurs propres sont de module 1 : écrire $||v||^2 = v^T v = v^T M^T M v = ||\lambda v||^2$. Une matrice dans $\mathrm{O}_3(\mathbb{R})$ a une ou trois valeurs propres réelles. On obtient la liste (c'est aussi un cas particulier du théorème 27)

(1) $M = \mathrm{id}$;

(2) M est conjuguée à $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$: réflexion orthogonale par rapport à un plan ;

(3) M est conjuguée à $\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$: rotation d'angle θ (noter le cas particulier $\theta = \pi$, renversement) ;

(4) M est conjuguée à $\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}$: composée des deux cas précédents.

On étudie maintenant plus en détail le groupe spécial orthogonal $\mathrm{SO}_3(\mathbb{R})$. Le groupe $\mathrm{SO}_3(\mathbb{R})$ est donc le groupe des rotations de $\mathbb{R}^3$ dont l'axe passe par l'origine. Les renversements sont des éléments de $\mathrm{SO}_3(\mathbb{R})$, et correspondent dans ce contexte aux rotations d'angle π.

Lemme 29. *Quelques propriétés du groupe* $\mathrm{SO}_3(\mathbb{R})$ *:*

(1) *Les renversements engendrent* $\mathrm{SO}_3(\mathbb{R})$.

(2) *Les renversements sont deux à deux conjugués dans* $\mathrm{SO}_3(\mathbb{R})$.

(3) *Le groupe* $\mathrm{SO}_3(\mathbb{R})$ *est connexe.*

(4) *Le centre du groupe* $\mathrm{SO}_3(\mathbb{R})$ *est trivial.*

Preuve. (1) *(faire un dessin !)* Soit D une droite vectorielle de $\mathbb{R}^3$, et θ un angle. Considérons H le plan orthogonal à D, et $D_1, D_2 \subset H$ deux droites formant un angle $\theta/2$. Alors la composée des renversements d'axe D_1 et D_2 est une rotation d'angle θ et d'axe D : en effet tout point de D est fixe, et en restriction à H on a la composée de deux symétries axiales, c'est-à-dire une rotation.

(2) Soit $D_1 \neq D_2$ les axes de deux renversements σ_1, σ_2. Dans le plan P engendré par D_1, D_2, soit D l'une des bissectrices de ces deux droites. Alors le renversements σ d'axe D échange D_1 et D_2, et donc $\sigma\sigma_1\sigma = \sigma_2$.

(NB : on a montré au passage que l'action de $\mathrm{SO}_3(\mathbb{R})$ sur les droites vectorielles de $\mathbb{R}^3$ est transitive).

(3) Il est facile de voir que $\mathrm{SO}_3(\mathbb{R})$ est connexe par arc, en reliant toute matrice $\left(\begin{smallmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$ à l'identité en faisant varier l'angle de θ à 0.

Un argument plus sophistiqué (mais qui s'applique à des cas plus généraux) est donné dans [CG13, p. 37] : le groupe $\mathrm{SO}_3(\mathbb{R})$ agit transitivement sur la sphère S^2, avec stabilisateur d'un point $\mathrm{SO}_2(\mathbb{R}) \simeq S^1$. Alors la connexité découle du fait général que si G/H et H sont connexe, alors G est connexe.

(4) Soit g dans le centre de $\mathrm{SO}_3(\mathbb{R})$, et soit v un vecteur non nul de $\mathbb{R}^3$. Notons r le renversement d'axe donné par v. Comme $grg^{-1} = r$, l'axe de r est préservé par g, et donc v est un vecteur propre pour g. Ceci étant vrai pour tout v, g est une homothétie,

or l'identité est la seule homothétie dans $\mathrm{SO}_3(\mathbb{R})$ (argument tiré de [CG13, Lemme A.1.23 p. 17]).

□

Théorème 30 ([CG13, p. 239])**.** *Le groupe* $\mathrm{SO}_3(\mathbb{R})$ *est simple.*

Preuve. Soit $N \triangleleft \mathrm{SO}_3(\mathbb{R})$ un sous-groupe distingué non trivial, on veut montrer que $N = \mathrm{SO}_3(\mathbb{R})$. D'après le lemme précédent, il suffit de montrer que N contient un renversement. Soit $h \in N$ une rotation non triviale. On considère l'application continue

$$\begin{array}{rccc} \varphi\colon & \mathrm{SO}_3(\mathbb{R}) & \to & \mathbb{R} \\ & g & \mapsto & \mathrm{Tr}(ghg^{-1}h^{-1}). \end{array}$$

La trace d'une rotation d'angle θ de $\mathbb{R}^3$ est $1+2\cos\theta$, ainsi l'application φ est à valeur dans $[-1,3]$. Plus précisément, l'image de φ est un intervalle (car $\mathrm{SO}_3(\mathbb{R})$ est connexe), de la forme $[a,3]$ (car $\varphi(h)=3$, ou $\varphi(\mathrm{id})=3$, au choix), et avec $a<3$ (car le centre de $\mathrm{SO}_3(\mathbb{R})$ est trivial). Pour tout n entier assez grand, on a $a \leqslant 1+2\cos(\pi/n) < 3$, et il existe alors g tel que $\mathrm{Tr}(ghg^{-1}h^{-1}) = 1+2\cos(\pi/n)$, autrement dit $u = ghg^{-1}h^{-1}$ est une rotation d'angle $\pm\pi/n$. De plus $u \in N$, car ghg^{-1} et h^{-1} sont dans N. Finalement $u^n \in N$ est le renversement cherché. □

4. Compléments

4.1. **Résolution au sens des moindres carrés.** Si un système linéaire n'a pas de solution, on peut chercher à le résoudre au sens des moindres carrés, on tombe sur un système avec une matrice symétrique définie positive (cadre de Cholesky) : voir [RW07, p. 180].

Soit $M \in \mathrm{Mat}_{n,p}(\mathbb{R})$ une matrice rectangulaire de rang p (application linéaire injective $\mathbb{R}^p \to \mathbb{R}^n$), correspondant à un système linéaire de n équations à p inconnues (avec $n > p$) :

$$MX = Y$$

Supposons $Y \notin \mathrm{Im}\, M$, de telle sorte que le système n'a pas de solution, et cherchons X tel que $\|MX - Y\|$ soit minimal.

Lemme 31. $\|MX-Y\|$ *est minimal ssi* $MX-Y$ *est orthogonal à* $\mathrm{Im}\, M$*.*

Preuve. $MX-Y$ est orthogonal à $\mathrm{Im}\, M$ revient à dire que MX est le projeté orthogonal de Y sur $\mathrm{Im}\, X$. Le fait que cela équivaut à la minimalité de $\|MX-Y\|$ vient du théorème de Pythagore (faire une figure). □

Lemme 32. *L'orthogonal de* $\mathrm{Im}\, M$ *est* $\ker M^t$*.*

Preuve.

$$\forall X, \langle MX, Y\rangle = 0 \iff \forall X, \langle X, M^tY\rangle = 0 \iff M^tY = 0$$ □

Lemme 33. *Pour une matrice réelle symétrique positive, le noyau et le cône isotrope coïncident.*

Preuve. En général on a le noyau inclu dans le le cône isotrope, il s'agit de vérifier l'inclusion inverse sous l'hypothèse de positivité. Pour une base orthogonale (e_i) la forme quadratique (de rang r) associée s'écrit

$$q(x) = a_1x_1^2 + \cdots + a_rx_r^2$$

pour certains $a_i > 0$. Soit $x \in E$, on écrit $x = \sum_{i=1}^{r} x_i e_i + x_0$ avec $x_0 \in \ker A$. Si x est dans le cône isotrope, on a $q(x) = 0 \implies x_i = 0$ pour chaque $i = 1, \dots, r$, et donc $x = x_0 \in \ker A$ comme attendu. □

Lemme 34. *La matrice $A = M^t M$ est symétrique définie positive.*

Preuve. La positivité vient de $X^t M^t M X = \|MX\|^2 \geqslant 0$. Reste à vérifier que $M^t M$ est non dégénérée, c'est-à-dire de rang p. Par le lemme précédent on sait que son cône isotrope est égal à son noyau, ce qui donne la première équivalence dans la ligne qui suit :

$$M^t M X = 0 \iff X^t M^t M X = 0 \iff \|MX\| = 0 \iff MX = 0.$$

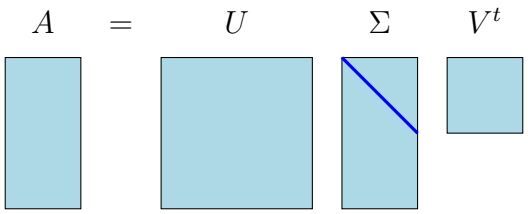
Donc $\ker M^t M = \ker M$. Par le théorème du rang on conclut $\operatorname{rang} M^t M = \operatorname{rang} M = p$. □

On cherche donc X tel que $M^t M X - M^t Y = 0$. Posons $A = M^t M$ et $B = M^t Y$, on est ramené au système $AX = B$ avec A symétrique définie positive, et on passe le bébé à Cholesky.

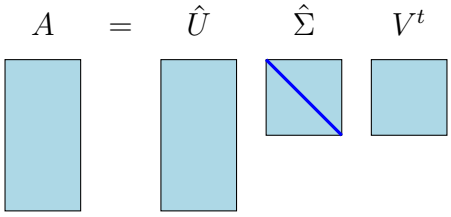
fin cours no 5

4.2. **Décomposition en valeurs singulières.** Le théorème de décomposition en valeurs singulières et le B-A-BA de l'algèbre linéaire pour les maths appliquées, et semble complètement ignoré par les livres d'algèbre plus fondamentaux... à tort ! Cette section est fortement inspirée du livre [TB97]. Si on veut aller directement à l'essentiel il suffit de regarder cette video youtube. Voici l'énoncé suivi de quelques détails supplémentaires, si la chanson n'est pas assez claire :

Théorème 35. *Soit $M \in \mathrm{Mat}_{m,n}$ la matrice d'une application linéaire $\varphi\colon \mathbb{R}^n \to \mathbb{R}^m$, avec $m \geqslant n$. Alors il existe $U \in \mathrm{O}_m(\mathbb{R})$ et $V \in \mathrm{O}_n(\mathbb{R})$ des matrices orthogonales, et $\Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_n) \in \mathrm{Mat}_{m,n}$ une matrice rectangulaire diagonale réelle avec $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n$ uniquement déterminées, tel que*



Remarque. Si on oublie les $m - n$ dernières lignes (nulles) de Σ pour la rendre carrée, et les $m - n$ dernières colonnes de U pour que le produit matriciel reste bien défini, on obtient la décomposition en valeur singulière *réduite* :



Définition. Les σ_i s'appellent les *valeurs singulières* de la matrice A. Quand la matrice A est carrée on peut aussi parler de ses valeurs propres, qui sont distinctes des valeurs singulières en général.

On appelle *ellipsoïde* un sous-ensemble de $\mathbb{R}^n$ défini dans une base orthonormée par une équation de la forme

$$a_1x_1^2 + \cdots + a_nx_n^2 = 1$$

avec les $a_i > 0$. Autrement dit un ellipsoïde est le niveau $q(x) = 1$ d'une forme quadratique définie positive sur $\mathbb{R}^n$.

Lemme 36. *Soit $A = B^{-1} \in \mathrm{GL}_n(\mathbb{R})$, et $S \subset \mathbb{R}^n$ la sphère unité de $\mathbb{R}^n$. Alors*

(1) *$A(S)$ est un ellipsoïde, associé à la forme quadratique q de matrice B^tB.*

(2) *Deux vecteurs $y_1 = A(x_1), y_2 = A(x_2)$ sont orthogonaux pour q ssi x_1, x_2 sont orthogonaux pour le produit scalaire euclidien.*

(3) *En particulier les préimages par A des demi-axes principaux de l'ellipsoïde $A(S)$ forment une base orthogonale (euclidienne).*

Preuve. (1) Si $y \in \mathbb{R}^n$, on a

$$y \in A(S) \iff \exists x \in S, y = Ax \iff \exists x \in S, x = By \iff y^tB^tBy = 1$$

Or la forme quadratique q associée à la matrice B^tB est définie positive, donc cette dernière condition revient à dire que y est dans l'ellipsoïde $q(y) = 1$.

(2) C'est juste l'équivalence

$$y_1^tB^tBy_2 = 0 \iff x_1^tx_2 = 0.$$

(3) Les demi-axes principaux sont orthogonaux pour q (et aussi pour le produit euclidien, mais cela ne sert pas ici). C'est une façon géométrique de paraphraser le corollaire 15 (diagonalisation simultanée). □

Preuve géométrique du théorème 35, dans le cas d'une matrice carrée inversible.

Supposons A inversible de taille n, et notons (e_i) la base canonique de $\mathbb{R}^n$. Par le lemme, $A(S)$ est un ellipsoïde. On note $u_1, \ldots, u_n$ une base orthonormée de $\mathbb{R}^n$ dans laquelle l'équation de l'ellipsoïde $A(S)$ est

$$a_1x_1^2 + \cdots + a_nx_n^2 = 1$$

avec $0 < a_1 \leqslant a_2 \leqslant \cdots \leqslant a_n$. On pose $\sigma_i = \frac{1}{\sqrt{a_i}}$, ainsi les vecteurs σ_iu_i sont sur

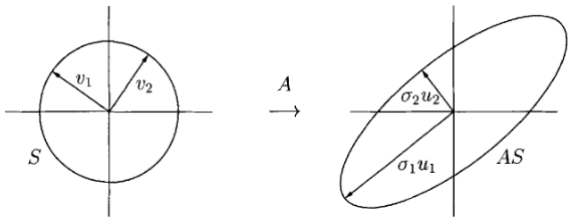
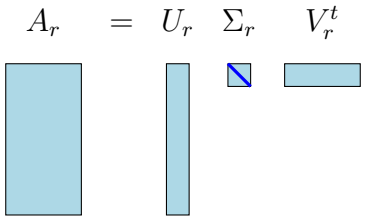


FIGURE 1. (extraite de [TB97])

l'ellipsoïde et sont ses demi-axes principaux (ordonnés par longueur décroissante). On a $\sigma_iu_i = Av_i$ pour des $v_i \in S$, et par le lemme les v_i forment une base orthogonale (donc orthonormée). Notons $\Sigma = \mathrm{diag}(\sigma_i)$, U la matrice dont les colonnes sont les $u_i = Ue_i$, et V celle dont les colonnes sont les $v_i = Ve_i$. Alors on vérifie que les matrices A et $U\Sigma V^t$ sont égales car coïncident sur la base des v_i :

$$Av_i = \sigma_iu_i = U\sigma_ie_i = U\Sigma e_i = U\Sigma V^tv_i.$$ □

Remarque. La décomposition en valeur singulière répond à la question naturelle : étant donnée une matrice rectangulaire A, et pour $r < \operatorname{rang} A$, quelle est la matrice A_r de rang r et de même taille la plus proche de A (pour la norme euclidienne matricielle) ? On peut montrer qu'il suffit de prendre le produit $A_r = U_r \Sigma_r V_r^t$, où U_r est formé des r premières colonnes de U, V_r^t des r premières lignes de V^t, et $\Sigma_r = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$:



Références

[CG13] P. Caldero & J. Germoni. *Histoires hédonistes de groupes et de géométries, Tome premier*. Calvage & Mounet, 2013.

[Com98] F. Combes. *Algèbre et géométrie*. Bréal, 1998.

[Per96] D. Perrin. *Cours d'algèbre*. Ellipses, 1996.

[RW07] J.-P. Ramis & A. Warusfel. *Mathématiques Tout-en-un pour la Licence, Niveau L2*. Dunod, 2007.

[Szp09] A. Szpirglas, editor. *Mathématiques L3 Algèbre*. Pearson Education, 2009.

[TB97] L. Trefethen & D. Bau. *Numerical Linear Algebra*. SIAM, 1997.