Décomposition polaire

Mohamed NASSIRI

Références:

Histoires hédonistes de groupes et de géométries - Tome 1, Philippe Caldero et Jérôme Germoni - p.202 → 204

Recasage :

●●●●● 106 : Groupe linéaire d'un espace vectoriel de dimension finie E, sous-groupes de GL(E). Applications.
●●●●● 155 : Endomorphismes diagonalisables en dimension finie.
●●●●● 158 : Matrices symétriques réelles, matrices hermitiennes.
●●●●○ 150 : Exemples d'actions de groupes sur les espaces de matrices.
●●●●○ 160 : Endomorphismes remarquables d'un espace vectoriel euclidien (de dimension finie).

Résumé:

C'est un développement incontournable de l'agrégation ! La décomposition polaire est un théorème remarquable et pas si "perché" que ça. En effet, pour $n = 1$, on a que $\mathcal{H}_n^{++} = \mathbb{R}^{+*}$ (réels de la forme $\rho > 0$, U_n sont les complexes de module 1 (de la forme $e^{i\theta}$) et $\mathrm{GL}_n(\mathbb{C})$ sont les complexes non nuls ($z \in \mathbb{C}\backslash\{0\}$). La décomposition polaire nous dit que $z = \rho e^{i\theta}$, qui est la "forme exponentielle" (ou "forme polaire") d'un nombre complexe. La décomposition polaire généralise donc en dimension supérieure cette écriture.

Prérequis:

Groupe linéaire - Groupe orthogonal - Matrices symétriques réelles définie positives

---

Théorème : On a les homéomorphismes suivants :

$$\mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R}) \xrightarrow{\sim} \mathrm{GL}_n(\mathbb{R}),\ (O,S) \mapsto OS$$

$$\mathrm{U}_n \times \mathcal{H}_n^{++} \xrightarrow{\sim} \mathrm{GL}_n(\mathbb{C}),\ (U,H) \mapsto UH$$

Démonstration.
Pour la démonstration, on se limite au cas réel. Pour montrer l'homéomorphisme, on procède classiquement : on montre que l'application est bien définie, continue, bijective et que sa réciproque est continue.

Posons $\mu : \mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R}) \xrightarrow{\sim} \mathrm{GL}_n(\mathbb{R}),\ (O,S) \mapsto OS$

Etape 1 - Définition et continuité:
Les éléments O et S détermine de façon unique $\mu(O,S) = OS$.
De plus, $OS \in \mathrm{GL}_n(\mathbb{R})$ car $\det(OS) = \underbrace{\det O}_{=1}\underbrace{\det S}_{>0} \neq 0$.

L'application μ est manifestement continue (il s'agit de la multiplication matricielle).

Etape 2 - Surjectivité :
Soit $M \in \mathrm{GL}_n(\mathbb{R})$. La matrice ${}^tMM \in \mathcal{S}_n^{++}(\mathbb{R})$ car $\forall X \in \mathbb{R}^n \backslash \{0\}$, ${}^tX^tMMX =^t (MX)MX = ||MX||^2 > 0$. On peut donc diagonaliser ${}^tMM$ dans une base orthonormée. Ainsi, il existe $P \in \mathrm{O}_n(\mathbb{R})$ telle que :

$$ {}^tMM = P \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} P^{-1} = P\mathrm{diag}(\lambda_1, ..., \lambda_n)P^{-1}, \quad \text{avec } \lambda_i > 0 \text{ pour tout } i. $$

On pose alors :

$$ S = P\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})P^{-1} $$

On a $S \in \mathcal{S}_n^{++}(\mathbb{R})$. En effet, comme $P \in \mathrm{O}_n(\mathbb{R})$, $P^{-1} =^t P$, on a

$$ {}^tS =^t (P\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})^tP) =^t ({}^tP)^t(\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}))^tP = P\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})t^P = S $$

et de plus, les valeurs propres de S sont strictement positives (ce sont les $\sqrt{\lambda_1}, ..., \sqrt{\lambda_n}$, avec $\lambda_i > 0$ pour tout i).
On pose $O := MS^{-1}$, et en remarquant que $S^2 =^t MM$ il vient :

$$ {}^tOO =^t (MS^{-1})MS^{-1} =^t S^{-1} \underbrace{{}^tMM}_{=S^2} S^{-1} = S^{-1}S^2S^{-1} = I_n $$

Par conséquent, $M = OS$ avec $O \in \mathrm{O}_n(\mathbb{R})$ et $S \in \mathcal{S}_n^{++}(\mathbb{R})$, donc μ est surjective.

Etape 3 - Injectivité :
Supposons que l'on ait $M = OS = O'S'$, avec $O, O' \in \mathrm{O}_n(\mathbb{R})$ et $S, S' \in \mathcal{S}_n^{++}(\mathbb{R})$. Alors, on a :

$$ S^2 =^t MM =^t (O'S')O'S' =^t S' \underbrace{{}^tO'O'}_{=I_n} S' = S'^2 $$

Soit Q le polynôme d'interpolation de Lagrange tel que pour tout i, $Q(\lambda_i) = \sqrt{\lambda_i}$. Alors :

$$ \begin{aligned} S = P\mathrm{diag}(\sqrt{\lambda_1}, ..., \sqrt{\lambda_n})P^{-1} &= P\mathrm{diag}(Q(\lambda_1), ..., Q(\lambda_n))P^{-1} \\ &= PQ(\mathrm{diag}(\lambda_1, ..., \lambda_n))P^{-1} \\ &= Q(P\mathrm{diag}(\lambda_1, ..., \lambda_n)P^{-1}) \\ &= Q(S^2) = Q(S'^2). \end{aligned} $$

Comme S' commute avec S'^2, et donc avec $Q(S'^2) = S$, il vient que S' et S commutent, et sont donc diagonalisables dans une même base. Il existe ainsi une matrice de passage P_0 telle que

$$ S' = P_0\mathrm{diag}(\mu_1', ..., \mu_n')P_0^{-1} \text{ et } S = P_0\mathrm{diag}(\mu_1, ..., \mu_n)P_0^{-1} $$

Par suite :

$$ \begin{aligned} S'^2 = S^2 &\Rightarrow P_0\mathrm{diag}(\mu_1'^2, ..., \mu_n'^2)P_0^{-1} = P_0\mathrm{diag}(\mu_1^2, ..., \mu_n^2)P_0^{-1} \\ &\Rightarrow \forall 1 \leq i \leq n,\ \mu_i'^2 = \mu_i^2 \\ &\Rightarrow \forall 1 \leq i \leq n,\ \mu_i' = \mu_i \quad (\mu_i', \mu_i \in \mathbb{R}^{+*} \text{ car } S, S' \in \mathcal{S}_n^{++}(\mathbb{R})) \\ &\Rightarrow S = S' \end{aligned} $$

Ainsi, on a $O = O'$, d'où l'injectivité de μ.

Etape 4 - Continuité de μ^{-1} :

Soit $(M_p)_{p\in\mathbb{N}}$ une suite de $\mathrm{GL}_n(\mathbb{R})$ qui converge vers M. On note, pour tout entier p, $\mu^{-1}(M_p) = (O_p, S_p) \in \mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$ (donc $M_p = O_p S_p$) et $\mu^{-1}(M) = (O, S) \in \mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$.
Montrons que les suites (O_p) et (S_p) convergent respectivement vers O et S.
Comme $\mathrm{O}_n(\mathbb{R})$ est compact, par le théorème de Bolzano-Weierstrass, il existe $\overline{O}$ valeur d'adhérence de $(O_p)_{p\in\mathbb{N}}$ (*i.e.*) il existe une sous-suite $(O_{p_k})_{k\in\mathbb{N}}$ de (O_p) converge vers $\overline{O}$. Alors la sous-suite S_{p_k} converge vers $\overline{O}^{-1}M \in \mathcal{S}_n^{++}(\mathbb{R})$ car

$$\overline{S} := \overline{O}^{-1}M \in \mathrm{GL}_n(\mathbb{R}) \cap \overline{\mathcal{S}_n^{++}(\mathbb{R})} = \mathrm{GL}_n(\mathbb{R}) \cap \mathcal{S}_n^{+}(\mathbb{R}) = \mathcal{S}_n^{++}(\mathbb{R})$$

On a donc $M = \overline{O}\ \overline{S}$, mais par unicité de la décomposition (injectivité de μ), on en déduit que $\overline{O} = O$ et $\overline{S} = S$.
Une suite ne possédant qu'une valeur d'adhérence dans un compact converge vers cette valeur d'adhérence. Donc (O_p) converge vers O et par suite (S_p) converge S.

□

Remarques :

- Dans la démonstration, on introduit Q le polynôme d'interpolation de Lagrange tel que pour tout i, $Q(\lambda_i) = \sqrt{\lambda_i}$. Il faut faire attention ici au fait que les λ_i ne sont pas tous distincts. Quitte à réindexer les λ_i, on peut supposer que $\lambda_1, ..., \lambda_r$ sont distincts et que $\lambda_{r+1}, ..., \lambda_n$ apparaissent parmi les $\lambda_1, ..., \lambda_r$. On prend alors le polynôme
$$Q(X) = \sum_{i=1}^{r} \sqrt{\lambda_i} \prod_{\substack{i=1 \\ i\neq j}}^{r} \frac{X - \lambda_j}{\lambda_i - \lambda_j}$$
- On a utilisé la très indispensable **diagonalisation simultanée** :

> Soient E un espace vectoriel de dimension finie n sur un corps $\mathbb{K}$ et deux endomorphismes f et g de E diagonalisables. Alors
> $$f \circ g = g \circ f \Leftrightarrow f \text{ et } g \text{ sont simultanément diagonalisables.}$$

Démontrons rapidement cette proposition :

Démonstration :
Raisonnons par récurrence sur n. Pour $n = 1$, la démonstration est triviale.
Supposons qu'il soit vrai sur tout espace de dimension inférieure ou égale à $n-1$. Soit E un espace vectoriel de dimension n et soit $\lambda_1, ..., \lambda_p$ les valeurs propres de f distinctes deux à deux.

- Si f est une homothétie, le résultat est vrai car toute base qui diagonalise g diagonalise f.
- Si f n'est pas une homothétie, E est somme directe de ses sous-espaces propres qui sont tous de dimensions strictement inférieures à n (bah oui car f n'est pas une homothétie...).
Soit E_λ un sous-espace propre de f. Montrons qu'il est stable par g.
Soit $x \in E_\lambda$. On a :

$$(f - \lambda Id) \circ g(x) \underbrace{=}_{f\circ g = g\circ f} g \circ \underbrace{(f - \lambda Id)(x)}_{=0} = 0$$

D'où $g(x) \in E_\lambda$.
En appliquant l'hypothèse de récurrence à chacun des sous-espaces propres de f, on obtient le résultat.

□

- Dans la démonstration de la continuité de μ^{-1}, on a utilisé le théorème de Bolzano-Weierstrass car $\mathrm{O}_n(\mathbb{R})$ est compact. Montrons cette dernière affirmation.

  Démonstration :
  On rappelle que
  $$O_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid {}^tMM = I\}$$
  Soit l'application
  $$\varphi : \mathcal{M}_n(\mathbb{R}) \to \mathcal{M}_n(\mathbb{R}),\ M \mapsto^t MM$$
  On a donc que $O_n(\mathbb{R}) = \varphi^{-1}(\{I\})$. L'application φ étant continue, $O_n(\mathbb{R})$ est fermé comme image réciproque d'un fermé (le singleton $\{I\}$).
  De plus,
  $$O_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid {}^tMM = I\} = \{M = (m_{ij}) \in \mathcal{M}_n(\mathbb{R}) \mid \sum_{k=1}^{n} m_{ki}m_{kj} = \delta_{ij} \ \ \forall i, j\}$$
  Ainsi, $\sum_{k=1}^n m_{ki}^2 = 1$, implique $|m_{ki}| \leq 1$, $\forall i, j$ et donc tout élément de $O_n(\mathbb{R})$ donc $O_n(\mathbb{R})$ est borné. D'après le théorème de Borel-Lebesgue, puisque $O_n(\mathbb{R})$ est un fermé borné de l'espace de dimension finie $\mathcal{M}_n(\mathbb{R})$, $O_n(\mathbb{R})$ est un compact de $\mathcal{M}_n(\mathbb{R})$.

  $\square$

- On a également utilisé le résultat suivant :

Une suite (x_n) ne possédant qu'une valeur d'adhérence a dans un compact K converge vers cette valeur d'adhérence.

  Démonstration :
  En effet, supposons que ce ne soit pas le cas. Alors $\exists \epsilon > 0,\ \forall n_0 \in \mathbb{N}, \exists n \geq n_0, d(x_n, a) > \epsilon$.
  Fixons un tel ϵ. Il existe alors une extraction $\varphi : \mathbb{N} \to \mathbb{N}$ telle que $\forall n, d(x_{\varphi(n)}, a) > \epsilon$.
  Par compacité de K, $(x_{\varphi(n)})$ possède une valeur d'adhérence $b \in K$. Comme c'est aussi une valeur d'adhérence de (x_n), on a par l'hypothèse d'unicité $a = b$, contre le fait que pour tout $n \in \mathbb{N}$, $d(x_{\varphi(n)}, a) > \epsilon$. Contradiction !

  $\square$

  Au passage, on peut donner un contre-exemple à ce phénomène quand K n'est pas compact. On va donc construire une suite non convergente qui possède une unique valeur d'adhérence. On considère donc la suite (u_n) définit sur $\mathbb{R}$ par
  $$u_n = \begin{cases} n & \text{si } n \text{ est pair} \\ 0 & \text{si } n \text{ est impair} \end{cases}$$
  Il est clair que la suite (u_n) possède une unique valeur d'adhérence (qui est 0) mais ne converge pas (car la suite extraite (u_{2n+1}) tend vers $+\infty$).

- Montrons également un dernier résultat utilisé : $\overline{\mathcal{S}_n^{++}(\mathbb{R})} = \mathcal{S}_n^{+}(\mathbb{R})$. Montrons donc que tout élément de $\mathcal{S}_n^{+}(\mathbb{R})$ est limite d'une suite d'élements de $\mathcal{S}_n^{++}(\mathbb{R})$.

  Démonstration :
  Soit $S \in \mathcal{S}_n^{+}(\mathbb{R})$. Alors il existe une matrice $P \in O_n(\mathbb{R})$ telle que :
  $$S = P\mathrm{diag}(\lambda_1, ..., \lambda_n)P^{-1}, \ \text{ avec } \lambda_i \geq 0 \text{ pour tout } i.$$

Posons

$$S_k = P\text{diag}(\lambda_1 + \frac{1}{k}, ..., \lambda_n + \frac{1}{k})P^{-1}, \quad \text{pour tout } k \in \mathbb{N}.$$

De ce fait, on a :

$$\lim_{k \to +\infty} S_k = S$$

De plus, comme le polynôme $(X - \lambda_1 - \frac{1}{k})...(X - \lambda_n - \frac{1}{k})$ admet un nombre fini de racines, il existe donc un rang k_0 à partir duquel $\lambda_i + \frac{1}{k} > 0$ pour tout i, et pour tout $k \geq k_0$. Par suite, pour tout $k \geq k_0$, $S_k \in \mathcal{S}_n^{++}(\mathbb{R})$.

$\square$

Quelques conséquences de la décomposition polaire :

- **Corollaire :** Pour tout $M \in \mathrm{GL}_n(\mathbb{R})$, $||M||_2 = \sqrt{\rho({}^tMM)}$, où ρ est le rayon spectral.

  Démonstration :
  Soit $M = OS$. Comme $||OSx||_2 = ||Sx||_2$ pour tout $x \in \mathbb{R}^n$, on a donc $||M||_2 = ||S||_2$. Comme S est symétrique réelle, elle est diagonalisable dans une base orthonormée $(v_1, ..., v_n)$ associés respectivements aux valeurs propres $\lambda_1, ..., \lambda_n$ telles que $\rho(S) = |\lambda_1| \geq ... \geq |\lambda_n|$.
  Soit $x = \sum_{i=1}^n x_i v_i \in \mathbb{R}^n$ et de norme 1, alors

  $$||Sx||_2 = ||\sum_{i=1}^n \underbrace{\lambda_i}_{\leq \rho(S)} x_i v_i||_2 \leq \rho(S) \underbrace{||\sum_{i=1}^n x_i v_i||_2}_{=1} = \rho(S)$$

  La borne étant atteinte pour $x = v_1$, on a bien $||S||_2 = \rho(S)$, et donc

  $$||M||_2 = ||S||_2 = \rho(S) = |\lambda_1| = \sqrt{|\lambda_1|^2} = \sqrt{\rho(S^2)} = \sqrt{\rho({}^tMM)}$$

  $\square$

- **Corollaire (Maximalité de $O_n(\mathbb{R})$) :** Tout sous-groupe compact de $\mathrm{GL}_n(\mathbb{R})$ qui contient $O_n(\mathbb{R})$ est $O_n(\mathbb{R})$ lui-même.

  Démonstration :
  Soit G un groupe compact contenant $O_n(\mathbb{R})$ et soit $g \in G$. Par la décomposition polaire, on peut écrire $g = os$ avec $(o, s) \in \mathrm{O}_n(\mathbb{R}) \times \mathcal{S}_n^{++}(\mathbb{R})$.
  Par hypothèse, $s = o^{-1}g \in G$, et donc $s^k \in G$ pour tout $k \in \mathbb{Z}$. Or s est diagonalisable et son spectre est contenu dans $\mathbb{R}^{+*}$.
  Comme G est compact, les valeurs propres de s ne peuvent être strictement supérieures à 1 sinon la suite $(||s^k||_2)_{k \in \mathbb{N}}$ ne serait pas bornée, ni strictement inférieures à 1 sinon la suite $(||s^{-k}||_2)_{k \in \mathbb{N}}$ ne serait pas bornée. Donc $s = Id$ et $g = o \in \mathrm{O}_n(\mathbb{R})$.

  $\square$

- Un dernier résultat en lien avec le précédent mais qui n'utilise pas la décomposition est le suivant :
  Corollaire : Tout sous-groupe fini G de $\mathrm{GL}_n(\mathbb{R})$ est conjugué à un sous-groupe de $O_n(\mathbb{R})$.

  Démonstration :
  Considérons la forme euclidienne canonique $\langle ., . \rangle$ de $\mathbb{R}^n$ et posons

  $$q(x) = \sum_{g \in G} \langle g(x), g(x) \rangle$$

q est une forme quadratique définie positive. De plus, si $h \in G$, on a

$$q(h(x)) = \sum_{g\in G} \langle g(h(x)), g(h(x))\rangle = \sum_{k\in G} \langle k(x), k(x)\rangle$$

La dernière égalité s'explique par le fait que quand g parcourt G, gh parcourt aussi G. Donc on peut poser le changement de variable $k = gh$.
Ainsi $q(h(x)) = q(x)$, ce qui implique $h \in O(q)$ et donc $G \in O(q)$.

Maintenant montrons que $O(q)$ est conjugué à $O_n(\mathbb{R})$ et on aura terminé.
Soit $S \in \mathcal{S}_n^{++}(\mathbb{R})$ la matrice de q. Alors, il existe $T \in \mathcal{S}_n^{++}(\mathbb{R})$ telle que $S = T^2$.
Soit $f \in O(q)$ associé à la matrice M. , on a :

$$\begin{aligned} & M =^t MSM \\ \Leftrightarrow\ & M =^t MT^2M =^t MTTM \\ \Leftrightarrow\ & I_n = (T^{-1t}MT)(TMT^{-1}) \\ \Leftrightarrow\ & I_n =^t (TMT^{-1})(TMT^{-1}) \end{aligned}$$

Ainsi $TMT^{-1} \in O_n(\mathbb{R})$ $(i.e.)$ $TO(q)T^{-1} \subset O_n(\mathbb{R})$

$\square$

- Mises en garde sur le développement : Ce développement est très proche, dans sa démonstration, de celui sur l'homéomorphisme entre $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{S}_n^{++}(\mathbb{R})$. Il est donc déconseillé de présenter les deux comme **seuls** développements.