The Polar Decomposition of Matrices

based on the "Les maths en tête - Algèbre" by Xavier Gourdon

December 5, 2025

Statement of the Polar Decomposition Theorem

Recall that

- a matrix $U \in \mathcal{M}_n(\mathbb{C})$ is **unitary** if $U^*U = I_n$ (finite dimension). Equivalently,

$$\forall\,(X, Y) \in (\mathbb{C}^n)^2, \quad \langle UX, UY \rangle = \langle X, U^*UY \rangle = \langle X, Y \rangle .$$

- a matrix $H \in \mathcal{M}_n(\mathbb{C})$ is **Hermitian/self-adjoint** if $H^* = H$. Equivalently,

$$\forall\,(X, Y) \in (\mathbb{C}^n)^2, \quad \langle HX, Y \rangle = \langle X, HY \rangle .$$

- a matrix $P \in \mathcal{M}_n(\mathbb{C})$ is **positive** if $\exists\, R \in \mathcal{M}_n(\mathbb{C})$ s.t. $P = R^*R$. Equivalently,

$$\forall\, X \in \mathbb{C}^n, \quad \langle X, PX \rangle \geq 0.$$

Exercise 6 (p.249 in the 2nd Edition, 2009), or (p.261 in the 3rd Edition, 2021)

Let $A \in \mathcal{M}_n(\mathbb{C})$. Show that there exists a pair (U, H) of matrices, where U is **unitary** *and H is* **positive** *(***Hermitian***), such that $\boxed{A = UH}$.*
If A is invertible, show that such a pair (U, H) is then **unique**.

Outline

1. Find H such that $A^*A = H^2$.
2. If A is invertible, then so is H, and one can define $U := AH^{-1}$.
3. Otherwise, method 1: define U "by parts" on $\mathrm{Ker}(A) \oplus \mathrm{Ker}(A)^\perp$.
   method 2: one can write $A = \lim_{k\to\infty} U_k H_k = U_\infty H_\infty$.

$\boxed{\text{Proof:}}$ **Step 1, the square-root matrix.** if a polar decomposition $A = UH$ exists with U unitary and H self-adjoint, then

$$A^*A = \left(UH\right)^* UH = H^*U^*UH = H^2 \tag{1}$$

Remark : In the abstract framework of C^*-algebras, one can make sense of a continuous function applied to a normal element, here $N := A^*A$. More precisely, the Gelfand map $\Phi : \begin{cases} \mathcal{C}(\mathrm{Sp}(N)) \longrightarrow \langle N, N^* \rangle \subseteq \mathcal{M}_n(\mathbb{C}) \\ \qquad f \longmapsto f(N) \end{cases}$, where $\mathrm{Sp}(N) := \left\{\lambda \in \mathbb{C},\ N - \lambda I_n \notin GL_n(\mathbb{C})\right\}$, is $*$-isomorphism. We'll see that $\mathrm{Sp}(N) \subseteq \mathbb{R}_+$, so one can make sense of $\sqrt{A^*A}$.

Step 1 (continued)

More concretely, A^*A is self-adjoint $\left((A^*A)^* = A^*A^{**} = A^*A \right)$, so it is diagonalizable in an orthonormal basis, i.e. $\exists\ P \in GL_n(\mathbb{C})$ and $D \in \mathcal{M}_n(\mathbb{C})$ diagonal such that

$$A^*A = P \cdot D \cdot P^{-1} = P \cdot \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix} \cdot P^{-1} \tag{2}$$

$N := A^*A$ diagonalizable means that there exists a basis consisting of eigenvectors. Let $\mathcal{B} = \left(C_1, C_2, \cdots, C_n\right) \in \mathbb{C}^{n \times n}$ be the corresponding column vectors and $\left(\lambda_1, \lambda_2, \cdots, \lambda_n\right) \in \mathbb{C}^n$ such that $\boxed{\forall\ i \in [\![1, n]\!],\ N C_i = \lambda_i C_i}$.
We interpret them as columns of a matrix P, which is thus such that if $\Xi \in \mathbb{C}^n$ are the coordinates of a vector in the new basis, those in the original one are given by $X = P\Xi$. $\mathcal{B}$ being a free and generating family means that $\Xi \longmapsto P\Xi$ is a bijection, so one can write $P^{-1}X = \Xi$. The boxed equalities can thus be rewritten $\boxed{NP = PD}$ or $\boxed{N = PDP^{-1}}$ or $\boxed{P^{-1}NP = D}$.

Step 1 (end)

"N self-adjoint" implies that it is **semi-simple**, i.e. if $E \subseteq \mathbb{C}^n$ is a stable subspace then it admits a supplementary subspace F. Here it is even better, one can take $F := E^\perp$. Indeed, we know that $NE \subseteq E$, let us show that $NF \subseteq F$:

$$\forall\, X \in E,\, Y \in F, \quad \left\langle X, NY \right\rangle = \left\langle NX, Y \right\rangle = 0 \quad \text{so} \;\; NY \in F.$$

If $\lambda_i = \lambda_j$, one can simply choose an orthonormal basis of the corresponding eigenspace and if $\lambda_i \neq \lambda_j$, one can show that the eigenspaces are orthogonal to each other. So ultimately, P can in fact be chosen unitary.

So $N = P \cdot D \cdot P^{-1}$, cf. (2). Moreover N is positive: for all $X \in \mathbb{C}^n$,

$$\left\langle X, A^* AX \right\rangle = \left\langle AX, AX \right\rangle = \|AX\|^2 \geq 0.$$

In particular this is true for $X := C_i$:

$$\forall\, i \in [\![1, n]\!], \quad \left\langle C_i, A^* A C_i \right\rangle = \left\langle C_i, \lambda_i C_i \right\rangle = \lambda_i \,\|C_i\|^2 \geq 0.$$

Hence all the eigenvalues are positive.

Step 1 (end); Step 2: if A is invertible

Define thus

$$H := P \cdot \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sqrt{\lambda_n} \end{pmatrix} \cdot P^{-1} \qquad (3)$$

It is indeed such that $A^*A = H^2$. There is a unique positive such matrix.

Step 2: If A is invertible then so is $A^*A = H^2$, and therefore so is H.

A first argument for the last step is the equivalence

$$\text{“}A^*A - 0I_n \quad \text{invertible iff} \quad 0 \notin \mathrm{Sp}(A^*A)\text{”}$$

Hence for all $i \in [\![1, n]\!]$, $\lambda_i \neq 0$. A second argument is the fact that if $f \circ g$ is bijective, then f is surjective and g is injective. One can also reason with the determinant.

Step 2 (end); Step 3

Since we are looking (U, H) such that $\boxed{A = UH}$, we naturally define

$$U := AH^{-1} \tag{4}$$

Let us check that it is unitary:

$$U^*U = \left(H^{-1}\right)^* A^*AH^{-1} = H^{-1}H^2H^{-1} = I_n.$$

Unicity: H is the unique ≥ 0 matrix s.t. $A^*A = H^2$. U also as it is necessarily by (4).

Step 3, if A is non-invertible. We still have H positive such that $A^*A = H^2$. Furthermore, for all $X \in \mathbb{C}^n$,

$$\|AX\|^2 = \langle AX, AX \rangle = \langle X, A^*AX \rangle = \langle X, H^2X \rangle = \langle HX, HX \rangle = \|HX\|^2 \tag{5}$$

so $\mathrm{Ker}(A) = \mathrm{Ker}(H)$. Since H is self-adjoint (and even positive), $\mathbb{C}^n$ can be written as the direct sum of eigenspaces or just $\boxed{\mathbb{C}^n = \mathrm{Ker}(H) \oplus \mathrm{Im}(H) = \mathrm{Ker}(H) \oplus \mathrm{Ker}(H)^\perp}$.

Step 3, constructive method

Since we are looking for U s.t. $A = UH$, we must have $\forall X \in \mathbb{C}^n$, $U\big(HX\big) = AX$.

$$U|_{\mathrm{Im}(H)} : \begin{cases} \mathrm{Im}(H) \longrightarrow \mathrm{Im}(A) \\ \quad Y \longmapsto AX \text{ where } X \in \mathbb{C}^n \text{ is s.t. } Y = HX \end{cases} \quad \text{is well-defined.}$$

If indeed $X' \in \mathbb{C}^n$ is another vector such that $Y = HX = HX'$, then $H(X - X') = 0$. But since $\mathrm{Ker}(H) = \mathrm{Ker}(A)$, we also have $A(X - X') = 0$, i.e. $AX = AX'$.
A systematic choice of X can be made as follows: if we restrict H to $\mathrm{Ker}(A)^\perp$ (or to any other supplementary space to $\mathrm{Ker}(A)$), then $H|_{\mathrm{Ker}(A)^\perp} : \mathrm{Ker}(A)^\perp \longrightarrow \mathrm{Im}(H)$ is a bijection, and $U|_{\mathrm{Im}(H)} = A\, H|^{-1}_{\mathrm{Ker}(A)^\perp}$.
As for the definition of the restriction of U to $\mathrm{Ker}(H)$, there is no unicity. Since we look for a unitary matrix U, if it exists, it must satisfy

$$\forall\ X \in \mathrm{Ker}(H),\ \forall\ Y \in \mathrm{Ker}(H)^\perp = \mathrm{Im}(H), \quad 0 = \big\langle X, Y \big\rangle = \big\langle UX, UY \big\rangle$$

Here, one can choose any isometry $J : \mathrm{Ker}(H) \longrightarrow \mathrm{Im}(A)^\perp$.These two spaces have the same dimension since $\mathrm{Ker}(H) = \mathrm{Ker}(A)$. Together with the rank theorem, one obtains that $\dim \mathrm{Ker}(H) = \dim \mathrm{Ker}(A) = n - \dim \mathrm{Im}(A) = \dim \big(\mathrm{Im}(A)^\perp\big)$.

Step 3, Constructive method, U is unitary

Let us show that

$$U : \begin{cases} \mathbb{C}^n = \mathrm{Ker}(H) \oplus \mathrm{Im}(H) \longrightarrow \mathbb{C}^n = \mathrm{Im}(A)^\perp \oplus \mathrm{Im}(A) \\ \qquad\qquad W + Y \quad \longmapsto \quad JW + A\, H|^{-1}_{\mathrm{Ker}(A)^\perp} \end{cases} \tag{6}$$

satisfies $A = UH$ (obvious) and that it is unitary: let $Z \in \mathbb{C}^n$. $\exists\ W \in \mathrm{Ker}(H)$, $X \in \mathbb{C}^n$ such that $Z = W + HX$

$$\|UZ\|^2 = \|JW + UHX\|^2 = \|JW + AX\|^2 = \|JW\|^2 + \|AX\|^2$$

$$\overset{(5)}{=} \|W\|^2 + \|HX\|^2 = \|Z\|^2$$

By the polarization identity (the version where the inner product is linear on the right)

$$\langle UX, UY \rangle = \frac{\|U(X+Y)\|^2 - \|U(X-Y)\|^2 - i\,\|U(X+iY)\|^2 + i\,\|U(X-iY)\|^2}{4}$$

$$= \frac{\|X+Y\|^2 - \|X-Y\|^2 - i\,\|X+iY\|^2 + i\,\|X-iY\|^2}{4} = \langle X, Y \rangle$$

Step 3, Topological method

Method 2: We'll show that an arbitrary matrix $A \in \mathcal{M}_n(\mathbb{C})$ can be approximated by a sequence $(A_p)_{p\in\mathbb{N}}$ of invertible ones,for which we have already established the polar decomposition, $A_p = U_p H_p$. We then use compactness of the group $\mathcal{U}_n(\mathbb{C})$ of unitary matrices to extract a convergence subsequence of $(U_{\varphi(p)})$ and finally show that the decomposition still holds in the limit.

$GL_n(\mathbb{C})$ is dense in $\mathcal{M}_n(\mathbb{C})$: (Proposition 2, p.183 (2009), or p. 193 (2021)). Let $A \in \mathcal{M}_n(\mathbb{C})$. Its characteristic polynomial $\chi_A(X) := \det(A - X I_n)$ has at maximum n roots, this means that $\mathrm{Sp}(A)$ is finite. Hence the minimal distance ρ between any two different roots/eigenvalue is necessarily strictly positive. No matter that 0 is or isn't a root, one always has,

$$\forall\ \lambda \in \mathbb{C},\ 0 < |\lambda| < \rho, \quad \chi_A(\lambda) \neq 0$$

or equivalently, $A - \lambda I_n$ is invertible. But $\lim_{\lambda \to 0} (A - \lambda I_n) = A$.

Take for example $A_p := A - \dfrac{\rho}{p+2} I_n,\ p \in \mathbb{N}$. By step 2, $\exists\ (U_p, H_p)$ s.t. $A_p = U_p H_p$.

Step 3, Topological method (continued)

$\mathcal{U}_n(\mathbb{C})$ is compact: the "polynomial" map $\Phi : \begin{cases} \mathcal{M}_n(\mathbb{C}) \longrightarrow \mathcal{M}_n(\mathbb{C}) \\ M \longmapsto M^* M \end{cases}$ is continuous and the singleton $\{I_n\}$ is a closed subset of $\mathcal{M}_n(\mathbb{C})$ so $\mathcal{U}_n(\mathbb{C}) = \Phi^{-1}\left(\{I_n\}\right)$ is closed.

Moreover, $\mathcal{U}_n(\mathbb{C})$ lies in the unit sphere of $\mathcal{M}_n(\mathbb{C})$ endowed with the **operator norm**:

$$|||M||| := \sup_{\|X\|=1} \|MX\| \quad \text{where} \quad \|X\| := \sqrt{\sum_{i=1}^{n} \overline{x_i} x_i}$$

Indeed, if U is unitary, then it is an isometry: for all $X \in \mathbb{C}^n$, $\|UX\| = \|X\|$ so the operator norm $|||U||| = 1$. As a closed and bounded subset of a finite dimensional vector space, $\mathcal{U}_n(\mathbb{C})$ is compact.

By the Bolzano-Weierstraß theorem, the sequence $(U_p)_{p\geq 0} \in \mathcal{U}_n(\mathbb{C})^{\mathbb{N}}$ admits a convergent subsequence. One can express this as follows: there exists $\varphi : \mathbb{N} \to \mathbb{N}$, strictly increasing such that $\lim_{k\to\infty} U_{\varphi(k)}$ exists and is in $\mathcal{U}_n(\mathbb{C})$. Let us denote it U_∞.

Topological method (end)

- $(A_p)_{p\geq 0}$ is convergent, hence, so any of its subsequence.
- we have just chosen $\left(U_{\varphi(k)}\right)_{k\in\mathbb{N}}$ convergent, so finally

$$\exists\ H_\infty \in \mathcal{M}_n(\mathbb{C}), \quad H_{\varphi(k)} = U_{\varphi(k)}^{-1} A_{\varphi(k)} \underset{k\to\infty}{\longrightarrow} H_\infty .$$

Let us now show that H_∞ is still a positive hermitian matrix:

- $$\forall\ X \in \mathbb{C}^2,\ \forall\ k \in \mathbb{N}, \quad \left\langle X, H_{\varphi(k)} X \right\rangle \geq 0,$$
  $$\text{so} \quad \lim_{k\to\infty} \left\langle X, H_{\varphi(k)} X \right\rangle = \left\langle X, H_\infty X \right\rangle \geq 0 .$$
- $$\Psi : \begin{cases} \mathcal{M}_n(\mathbb{C}) \longrightarrow \mathcal{M}_n(\mathbb{C}) \\ \qquad M \longmapsto M^* - M \end{cases} \quad \text{is continuous}$$
  $$\text{so} \quad \lim_{k\to\infty} \Psi\left(H_{\varphi(k)}\right) = \lim_{k\to\infty} 0 = 0 = \Psi\left(H_\infty\right) .$$

Finally, by continuity of the product $A = \lim\limits_{k\to\infty} U_{\varphi(k)} H_{\varphi(k)} = U_\infty \cdot H_\infty$ with U_∞ unitary and H_∞ positive.