

13.10 Polar Decomposition and Takagi's Factorization

Every nonzero complex number is a rotation times a stretch, $z = e^{i\theta}r$ with $r = |z| > 0$. Section 9 has put the matrix version within reach: if $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, then inserting $\mathbf{V}^*\mathbf{V} = \mathbf{I}$ in the middle splits the factorization into a unitary part and a positive semidefinite part. This section carries that out, proves that the unitary part is the closest unitary matrix to $\mathbf{A}$, and then treats a factorization that looks like the singular value decomposition but is not: Takagi's, for complex **symmetric** matrices.

The field is $\mathbb{C}$ unless stated otherwise, and $|\mathbf{A}| = (\mathbf{A}^*\mathbf{A})^{1/2}$ is the matrix absolute value of Definition 13.2.1.

13.10.1 Rotation times stretch

Theorem 13.10.1. Polar Decomposition

Let $\mathbf{A} \in M_n(\mathbb{C})$. Then there is a unitary $\mathbf{W} \in M_n(\mathbb{C})$ with

$$\mathbf{A} = \mathbf{W}|\mathbf{A}|.$$

The positive semidefinite factor $|\mathbf{A}|$ is determined by $\mathbf{A}$. The unitary factor $\mathbf{W}$ is determined by $\mathbf{A}$ if and only if $\mathbf{A}$ is invertible, and then $\mathbf{W} = \mathbf{A}|\mathbf{A}|^{-1}$. Moreover the same $\mathbf{W}$ satisfies the right-handed identity

$$\mathbf{A} = |\mathbf{A}^*| \mathbf{W}, \quad |\mathbf{A}^*| = \mathbf{W}|\mathbf{A}|\mathbf{W}^*.$$

Idea. The singular value decomposition already separates the two kinds of behavior: $\mathbf{U}$ and $\mathbf{V}^*$ rotate, and $\mathbf{\Sigma}$ stretches along the axes. Writing $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^* = (\mathbf{U}\mathbf{V}^*)(\mathbf{V}\mathbf{\Sigma}\mathbf{V}^*)$ puts all the rotating in the first bracket and all the stretching in the second, and the second bracket is a positive semidefinite matrix whose square is $\mathbf{A}^*\mathbf{A}$. Uniqueness of the positive semidefinite square root does the rest. For the uniqueness of $\mathbf{W}$, notice that $\mathbf{W}$ is pinned down only where $|\mathbf{A}|$ is invertible; when $\mathbf{A} = \mathbf{O}$, every unitary works.

Proof. Existence. By Theorem 13.9.1 write $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ with $\mathbf{U}, \mathbf{V} \in M_n(\mathbb{C})$ unitary and $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$ with $\sigma_i \geq 0$. Put

$$\mathbf{W} := \mathbf{U}\mathbf{V}^*, \quad \mathbf{P} := \mathbf{V}\mathbf{\Sigma}\mathbf{V}^*.$$

Then $\mathbf{W}$ is unitary, being a product of unitary matrices, and

$$\mathbf{W}\mathbf{P} = \mathbf{U}\mathbf{V}^*\mathbf{V}\mathbf{\Sigma}\mathbf{V}^* = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^* = \mathbf{A}.$$

The matrix $\mathbf{P}$ is self-adjoint, since $\mathbf{P}^* = \mathbf{V}\mathbf{\Sigma}^*\mathbf{V}^* = \mathbf{P}$, and positive semidefinite: its eigenvalues are the $\sigma_i \geq 0$ because $\mathbf{V}^*\mathbf{P}\mathbf{V} = \mathbf{\Sigma}$, so Theorem 13.1.4 applies. Finally

$$\mathbf{P}^2 = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^*\mathbf{V}\mathbf{\Sigma}\mathbf{V}^* = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^* = \mathbf{A}^*\mathbf{A},$$

the last equality because $\mathbf{A}^* \mathbf{A} = \mathbf{V} \Sigma \mathbf{U}^* \mathbf{U} \Sigma \mathbf{V}^* = \mathbf{V} \Sigma^2 \mathbf{V}^*$. A positive semidefinite matrix has exactly one positive semidefinite square root (Theorem 13.2.1), so $\mathbf{P} = (\mathbf{A}^* \mathbf{A})^{1/2} = |\mathbf{A}|$, and $\mathbf{A} = \mathbf{W}|\mathbf{A}|$.

The factor $|\mathbf{A}|$ is determined, since $\mathbf{A}^* \mathbf{A}$ is and the square root is unique.

Uniqueness of $\mathbf{W}$. ($\Rightarrow$) Suppose $\mathbf{A}$ is invertible. Then $|\mathbf{A}|^2 = \mathbf{A}^* \mathbf{A}$ is invertible, so $|\mathbf{A}|$ is, and any $\mathbf{W}$ with $\mathbf{A} = \mathbf{W}|\mathbf{A}|$ satisfies $\mathbf{W} = \mathbf{A}|\mathbf{A}|^{-1}$: there is only one. ($\Leftarrow$) Suppose $\mathbf{A}$ is not invertible, so there is a nonzero $\mathbf{z} \in \ker \mathbf{A}$. Let $\mathbf{H} = \mathbf{H}_{\mathbf{z}}$ be the Householder reflection of Definition 11.8.1, which is unitary with $\mathbf{H}\mathbf{z} = -\mathbf{z}$ and $\mathbf{H}\mathbf{v} = \mathbf{v}$ for every $\mathbf{v} \perp \mathbf{z}$ (Proposition 11.8.2 (c), (d)). By Proposition 13.2.2, $\text{nul}|\mathbf{A}| = \text{nul} \mathbf{A}$, and $|\mathbf{A}|$ is self-adjoint, so

$$\text{col}|\mathbf{A}| = \text{col}(|\mathbf{A}|^*) = (\text{nul}|\mathbf{A}|)^\perp = (\text{nul} \mathbf{A})^\perp \subseteq \mathbf{z}^\perp$$

by Theorem 11.6.5 (b). Hence $\mathbf{H}$ fixes every vector of $\text{col}|\mathbf{A}|$, that is, $\mathbf{H}|\mathbf{A}| = |\mathbf{A}|$, and therefore $(\mathbf{W}\mathbf{H})|\mathbf{A}| = \mathbf{A}$ as well, with $\mathbf{W}\mathbf{H}$ unitary and $\mathbf{W}\mathbf{H} \neq \mathbf{W}$ because $\mathbf{H} \neq \mathbf{I}$ and $\mathbf{W}$ is invertible. So $\mathbf{W}$ is not determined.

The right-handed form. Set $\mathbf{Q} := \mathbf{W}|\mathbf{A}|\mathbf{W}^*$. It is self-adjoint and positive semidefinite, being a unitary congruence of a positive semidefinite matrix: $\mathbf{x}^* \mathbf{Q} \mathbf{x} = (\mathbf{W}^* \mathbf{x})^* |\mathbf{A}| (\mathbf{W}^* \mathbf{x}) \geq 0$. And

$$\mathbf{Q}^2 = \mathbf{W}|\mathbf{A}|^2 \mathbf{W}^* = \mathbf{W} \mathbf{A}^* \mathbf{A} \mathbf{W}^* = \mathbf{A} \mathbf{A}^*,$$

using $\mathbf{A} = \mathbf{W}|\mathbf{A}|$, so $\mathbf{A}^* = |\mathbf{A}|\mathbf{W}^*$ and $\mathbf{A} \mathbf{A}^* = \mathbf{W}|\mathbf{A}|^2 \mathbf{W}^*$. By uniqueness of the square root again, $\mathbf{Q} = (\mathbf{A} \mathbf{A}^*)^{1/2} = |\mathbf{A}^*|$. Therefore $|\mathbf{A}^*| \mathbf{W} = \mathbf{W}|\mathbf{A}|\mathbf{W}^* \mathbf{W} = \mathbf{W}|\mathbf{A}| = \mathbf{A}$. This proves the theorem. ■

Remark. There is a second route to existence, the one Proposition 13.2.2 was pointing at: the identity $\| |\mathbf{A}| \mathbf{x} \| = \| \mathbf{A} \mathbf{x} \|$ says that $|\mathbf{A}| \mathbf{x} \mapsto \mathbf{A} \mathbf{x}$ is a well-defined isometry from $\text{col}|\mathbf{A}|$ onto $\text{col} \mathbf{A}$, and extending it by any isometry between the orthogonal complements produces $\mathbf{W}$ directly. That proof shows more clearly *where* the freedom in $\mathbf{W}$ lives, but it needs the extension step written out; the route above borrows all of that from Theorem 13.9.1.

For $n = 1$ the theorem is exactly $z = e^{i\theta}|z|$: a 1×1 unitary matrix is a complex number of modulus 1, and $|\mathbf{A}|$ is the modulus. The nonuniqueness for singular $\mathbf{A}$ is the matrix version of the fact that θ is undefined when $z = 0$.

Example. A polar decomposition

Find the left and right polar decompositions of $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$.

Solution. Exercise (B1: A 2 by 2 decomposition) gives the singular value decomposition

$$\mathbf{U} = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}, \quad \Sigma = \sqrt{5} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{V} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Hence

$$\mathbf{W} = \mathbf{U}\mathbf{V}^\top = \frac{1}{\sqrt{20}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix},$$

$$|\mathbf{A}| = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^\top = \sqrt{5} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Check: $|\mathbf{A}|^2 = 5 \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix} = \mathbf{A}^\top \mathbf{A}$, the columns of $\mathbf{W}$ are orthonormal, and $\mathbf{W}|\mathbf{A}| = \mathbf{A}$. Here $\mathbf{W}$ is a rotation, since $\det \mathbf{W} = (4 + 1)/5 = 1$.

For the right-handed form, $\mathbf{A}\mathbf{A}^\top = \begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix}$ and

$$|\mathbf{A}^*| = \mathbf{W}|\mathbf{A}|\mathbf{W}^\top = \frac{\sqrt{5}}{5} \begin{pmatrix} 6 & 3 \\ 3 & 14 \end{pmatrix},$$

whose square is indeed $\frac{1}{5} \begin{pmatrix} 45 & 60 \\ 60 & 205 \end{pmatrix} = \mathbf{A}\mathbf{A}^\top$. The two absolute values differ, and that is not an accident.

Remark. $|\mathbf{A}| = |\mathbf{A}^*|$ exactly when $\mathbf{A}$ is normal. Indeed $\mathbf{A}^*\mathbf{A} = \mathbf{A}\mathbf{A}^*$ if and only if the two positive semidefinite square roots agree, by uniqueness. So for a normal matrix — and only then — the stretching looks the same before and after the rotation, and $\mathbf{W}$ commutes with $|\mathbf{A}|$ by the identity $|\mathbf{A}^*| = \mathbf{W}|\mathbf{A}|\mathbf{W}^*$.

Quick check. What is the polar decomposition of a unitary matrix? Of a positive semidefinite matrix?

Solution. If $\mathbf{A}$ is unitary then $\mathbf{A}^*\mathbf{A} = \mathbf{I}$, so $|\mathbf{A}| = \mathbf{I}$ and $\mathbf{W} = \mathbf{A}$: all rotation, no stretch. If $\mathbf{A} \succeq \mathbf{O}$ then $\mathbf{A}^*\mathbf{A} = \mathbf{A}^2$ and $\mathbf{A}$ is itself a positive semidefinite square root of $\mathbf{A}^2$, so $|\mathbf{A}| = \mathbf{A}$ by uniqueness (Theorem 13.2.1) and $\mathbf{W} = \mathbf{I}$ whenever $\mathbf{A}$ is invertible: all stretch, no rotation.

13.10.2 The nearest unitary matrix

The polar factor is not merely a unitary matrix attached to $\mathbf{A}$. It is the one closest to $\mathbf{A}$, measured in the Frobenius norm of Chapter 12 §11. This is the reason the polar decomposition is used in practice: given a matrix that ought to be unitary but has drifted, $\mathbf{W}$ is how to put it back.

Corollary 13.10.2. The Polar Factor Is the Nearest Unitary Matrix

Let $\mathbf{A} \in M_n(\mathbb{C})$ with polar decomposition $\mathbf{A} = \mathbf{W}|\mathbf{A}|$. Then

$$\|\mathbf{A} - \mathbf{W}\|_F \leq \|\mathbf{A} - \mathbf{Q}\|_F \quad \text{for every unitary } \mathbf{Q} \in M_n(\mathbb{C}),$$

and the minimum value is $(\|\mathbf{A}\|_F^2 + n - 2 \sum_i \sigma_i)^{1/2}$. If $\mathbf{A}$ is invertible, $\mathbf{W}$ is the **only** unitary matrix attaining the minimum.

Idea. Two moves. First, the Frobenius norm does not see unitary factors (Lemma 11.7.5), so sandwiching by $\mathbf{U}^*$ and $\mathbf{V}$ turns the problem into: among unitary $\mathbf{Z}$, which is closest to the diagonal matrix $\mathbf{\Sigma}$? Second, expanding $\|\mathbf{\Sigma} - \mathbf{Z}\|_F^2$ leaves only the diagonal entries of $\mathbf{Z}$ to be chosen, and each of them has modulus at most 1 because it sits in a unit column. So the best possible $\mathbf{Z}$ is $\mathbf{I}$, and translating back gives $\mathbf{Q} = \mathbf{U}\mathbf{V}^* = \mathbf{W}$.

Proof. Fix a singular value decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ as in Theorem 13.9.1, so that $\mathbf{W} = \mathbf{U}\mathbf{V}^*$ as in the proof of Theorem 13.10.1. Let $\mathbf{Q}$ be unitary and put $\mathbf{Z} := \mathbf{U}^*\mathbf{Q}\mathbf{V}$, which is unitary; as $\mathbf{Q}$ runs over all unitary matrices so does $\mathbf{Z}$, since $\mathbf{Q} = \mathbf{U}\mathbf{Z}\mathbf{V}^*$. By Lemma 11.7.5, applied with the unitary factors $\mathbf{U}^*$ on the left and $\mathbf{V}$ on the right,

$$\|\mathbf{A} - \mathbf{Q}\|_F = \|\mathbf{U}^*(\mathbf{A} - \mathbf{Q})\mathbf{V}\|_F = \|\mathbf{\Sigma} - \mathbf{Z}\|_F.$$

Expanding the square in the Frobenius inner product $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{tr}(\mathbf{Y}^*\mathbf{X})$ of Chapter 11 §01,

$$\begin{aligned} \|\mathbf{\Sigma} - \mathbf{Z}\|_F^2 &= \|\mathbf{\Sigma}\|_F^2 + \|\mathbf{Z}\|_F^2 - 2 \text{Re tr}(\mathbf{Z}^*\mathbf{\Sigma}) \\ &= \sum_{i=1}^n \sigma_i^2 + n - 2 \sum_{i=1}^n \sigma_i \text{Re } z_{ii}, \end{aligned} \tag{13.10.1}$$

where $\|\mathbf{Z}\|_F^2 = \text{tr}(\mathbf{Z}^*\mathbf{Z}) = \text{tr } \mathbf{I}_n = n$, and $\text{tr}(\mathbf{Z}^*\mathbf{\Sigma}) = \sum_i (\mathbf{Z}^*)_{ii} \sigma_i = \sum_i \overline{z_{ii}} \sigma_i$, whose real part is $\sum_i \sigma_i \text{Re } z_{ii}$.

Column i of $\mathbf{Z}$ is a unit vector, so $|z_{ii}| \leq 1$ and hence $\text{Re } z_{ii} \leq 1$. Since every $\sigma_i \geq 0$,

$$\sum_{i=1}^n \sigma_i \text{Re } z_{ii} \leq \sum_{i=1}^n \sigma_i,$$

so Eq. 13.10.1 gives $\|\mathbf{A} - \mathbf{Q}\|_F^2 \geq \sum_i \sigma_i^2 + n - 2 \sum_i \sigma_i$ for every unitary $\mathbf{Q}$. Taking $\mathbf{Q} = \mathbf{W} = \mathbf{U}\mathbf{V}^*$ makes $\mathbf{Z} = \mathbf{U}^*\mathbf{U}\mathbf{V}^*\mathbf{V} = \mathbf{I}_n$, for which $\text{Re } z_{ii} = 1$ for every i and the bound is attained. Finally $\|\mathbf{A}\|_F = \|\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*\|_F = \|\mathbf{\Sigma}\|_F$ by Lemma 11.7.5, applied once on each side, and the only nonzero entries of $\mathbf{\Sigma}$ are $\sigma_1, \dots, \sigma_n$ on its diagonal, so $\|\mathbf{A}\|_F^2 = \sum_i \sigma_i^2$ (Exercise (C1: The Frobenius norm and the largest singular value)). The minimum value is therefore as stated.

Now suppose $\mathbf{A}$ is invertible, so $\sigma_i > 0$ for every i , and suppose $\mathbf{Q}$ attains the minimum. Then $\sum_i \sigma_i(1 - \text{Re } z_{ii}) = 0$ is a sum of nonnegative terms, so $\text{Re } z_{ii} = 1$ for every i ; with $|z_{ii}| \leq 1$ this forces $z_{ii} = 1$. Column i of $\mathbf{Z}$ then has a coordinate equal to 1 and total squared length 1, so all its other coordinates vanish and the column is $\mathbf{e}_i$. Hence $\mathbf{Z} = \mathbf{I}_n$ and $\mathbf{Q} = \mathbf{U}\mathbf{Z}\mathbf{V}^* = \mathbf{W}$. This proves the corollary. ■

For $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$ of Example (A polar decomposition), with $\sigma_1 = 3\sqrt{5}$, $\sigma_2 = \sqrt{5}$, and $\|\mathbf{A}\|_F^2 = 50$,

the nearest unitary matrix is $\mathbf{W} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ and the distance to the unitary group is $\sqrt{52 - 8\sqrt{5}}$, about 5.84.

Warning. The corollary is a statement about the **Frobenius** norm. A different way of measuring distance can have a different minimizer, and “nearest” always needs its yardstick named. It is true, though not proved here, that $\mathbf{W}$ also minimizes the distance in every unitarily invariant norm; Chapter 21 takes up that family of norms.

13.10.3 Takagi's factorization

A complex matrix can be symmetric, $\mathbf{A}^\top = \mathbf{A}$, without being Hermitian, $\mathbf{A}^* = \mathbf{A}$. Over $\mathbb{R}$ the two coincide and Chapter 12 handled them; over $\mathbb{C}$ they are genuinely different conditions, and symmetry is much the weaker one. It does not imply normality, it does not imply diagonalizability, and it says nothing about the eigenvalues. Here is the standing example:

$$\mathbf{N} = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, \quad \mathbf{N}^\top = \mathbf{N}, \quad \mathbf{N}^2 = \mathbf{O}.$$

A nonzero matrix with $\mathbf{N}^2 = \mathbf{O}$ has minimal polynomial x^2 and is therefore not diagonalizable, so no version of the spectral theorem can apply to $\mathbf{N}$. Nevertheless something survives, and it is a statement about singular values rather than eigenvalues.

Theorem 13.10.3. Takagi's Factorization

Let $\mathbf{A} \in M_n(\mathbb{C})$ be **symmetric**, $\mathbf{A}^\top = \mathbf{A}$, with singular values $\sigma_1 \geq \cdots \geq \sigma_n \geq 0$. Then there is a unitary $\mathbf{U} \in M_n(\mathbb{C})$ with

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top, \quad \mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n).$$

Read the statement against the two theorems it resembles.

- **Against the singular value decomposition.** Theorem 13.9.1 gives $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ with two unrelated unitary matrices. Since $\mathbf{U}^\top = (\bar{\mathbf{U}})^*$, Takagi's identity says $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}(\bar{\mathbf{U}})^*$: it is a singular value decomposition in which the right factor is forced to be $\mathbf{V} = \bar{\mathbf{U}}$. So the content is not a new diagonal matrix but a relation between $\mathbf{U}$ and $\mathbf{V}$.
- **Against the spectral theorem.** Corollary 12.4.3 writes a normal $\mathbf{A}$ as $\mathbf{U}\mathbf{D}\mathbf{U}^*$, a *similarity*, with the eigenvalues of $\mathbf{A}$ on the diagonal. Takagi writes a symmetric $\mathbf{A}$ as $\mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top$, which is a *congruence* $\mathbf{X} \mapsto \mathbf{U}\mathbf{X}\mathbf{U}^\top$, with the singular values on the diagonal. The hypotheses are different (symmetric, not normal), the diagonal holds different numbers, and the transpose is not the conjugate transpose.

The difference is visible even for a real matrix. Take $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$, which is real symmetric with eigenvalues 5 and -5 . Corollary 12.5.2 gives $\mathbf{A} = \mathbf{Q} \text{diag}(5, -5) \mathbf{Q}^\top$ with $\mathbf{Q}$ real orthogonal. Takagi insists on the **nonnegative** diagonal $\text{diag}(5, 5)$, and no real $\mathbf{U}$ can achieve that, since $\mathbf{U} \text{diag}(5, 5) \mathbf{U}^\top = 5\mathbf{U}\mathbf{U}^\top = 5\mathbf{I}$ for real orthogonal $\mathbf{U}$. A genuinely complex $\mathbf{U}$ is needed, and [Example \(Takagi for a real symmetric matrix\)](#) finds one.

Idea. Four steps, and the first two are the polar decomposition doing the work. ① Symmetry makes $\mathbf{A}$ intertwine $\mathbf{A}^*\mathbf{A}$ with its conjugate: $\mathbf{A}(\mathbf{A}^*\mathbf{A}) = \overline{(\mathbf{A}^*\mathbf{A})}\mathbf{A}$. Because the positive square root is a polynomial in the matrix, the same relation passes to $|\mathbf{A}|$. ② That relation, combined with $\mathbf{A}^\top = \mathbf{A}$, forces the polar factor $\mathbf{W}$ to be a **symmetric** unitary matrix. ③ A symmetric unitary matrix factors as $\mathbf{V}\mathbf{V}^\top$: split it into real and imaginary parts, which turn out to be commuting real symmetric matrices, and use Chapter 12's simultaneous diagonalization (Corollary 12.7.3). ④ Conjugating $\mathbf{A}$ by that $\mathbf{V}$ in the right way produces a matrix that is both symmetric and Hermitian, hence **real** symmetric, and the real spectral theorem in matrix form (Corollary 12.5.2) finishes the job. We prove all of this for invertible $\mathbf{A}$, where the polar factor is unique and $|\mathbf{A}|^{-1}$ exists.

Proof. We prove the theorem for **invertible** $\mathbf{A}$. Throughout, $\mathbf{A}^\top = \mathbf{A}$, so $\mathbf{A}^* = \overline{\mathbf{A}}^\top = \overline{\mathbf{A}}$. Write $\mathbf{M} := \mathbf{A}^*\mathbf{A} = \overline{\mathbf{A}}\mathbf{A}$ and $\mathbf{P} := |\mathbf{A}| = \mathbf{M}^{1/2}$. Since $\mathbf{A}$ is invertible, so are $\mathbf{M}$ and $\mathbf{P}$, and $\mathbf{M} \succ \mathbf{O}$.

Step 1: $\mathbf{A}\mathbf{M} = \overline{\mathbf{M}}\mathbf{A}$ and $\mathbf{A}\mathbf{P} = \overline{\mathbf{P}}\mathbf{A}$. For the first, $\overline{\mathbf{M}} = \overline{\overline{\mathbf{A}}\mathbf{A}} = \mathbf{A}\overline{\mathbf{A}}$, so

$$\mathbf{A}\mathbf{M} = \mathbf{A}\overline{\mathbf{A}}\mathbf{A} = \overline{\mathbf{M}}\mathbf{A}.$$

For the second we need a polynomial.

Claim. There is a polynomial p with **real** coefficients such that $p(\mathbf{M}) = \mathbf{P}$ and $p(\overline{\mathbf{M}}) = \overline{\mathbf{P}}$.

Proof. By Theorem 13.2.1 (a) there is a polynomial p with real coefficients such that $p(\mathbf{M}) = \mathbf{M}^{1/2} = \mathbf{P}$. Conjugating every entry of the identity $\mathbf{P} = p(\mathbf{M})$, and using that the coefficients of p are real so that conjugation passes through them, gives $\overline{\mathbf{P}} = p(\overline{\mathbf{M}})$. ■

Iterating Step 1's first identity gives $\mathbf{A}\mathbf{M}^j = \overline{\mathbf{M}}^j\mathbf{A}$ for every $j \geq 0$, hence $\mathbf{A}p(\mathbf{M}) = p(\overline{\mathbf{M}})\mathbf{A}$ for every polynomial p , and with the claim's p this reads

$$\mathbf{A}\mathbf{P} = \overline{\mathbf{P}}\mathbf{A}. \quad (13.10.2)$$

Step 2: the polar factor is symmetric. By Theorem 13.10.1, $\mathbf{A} = \mathbf{W}\mathbf{P}$ with $\mathbf{W} = \mathbf{A}\mathbf{P}^{-1}$ unitary. Substituting $\mathbf{A} = \mathbf{W}\mathbf{P}$ into Eq. 13.10.2 gives $\mathbf{W}\mathbf{P}^2 = \overline{\mathbf{P}}\mathbf{W}\mathbf{P}$, and canceling the invertible $\mathbf{P}$ on the right,

$$\mathbf{W}\mathbf{P} = \overline{\mathbf{P}}\mathbf{W}.$$

On the other hand $\mathbf{P}$ is Hermitian, so $\mathbf{P}^\top = \overline{\mathbf{P}}$, and transposing $\mathbf{A} = \mathbf{W}\mathbf{P}$ while using $\mathbf{A}^\top = \mathbf{A}$ gives

$$\mathbf{W}\mathbf{P} = \mathbf{A} = \mathbf{A}^\top = \mathbf{P}^\top\mathbf{W}^\top = \overline{\mathbf{P}}\mathbf{W}^\top.$$

Comparing the two displays, $\overline{\mathbf{P}}\mathbf{W} = \overline{\mathbf{P}}\mathbf{W}^\top$, and $\overline{\mathbf{P}}$ is invertible, so $\mathbf{W}^\top = \mathbf{W}$.

Step 3: a symmetric unitary matrix is $\mathbf{V}\mathbf{V}^\top$.

Claim. Let $\mathbf{W} \in M_n(\mathbb{C})$ be unitary with $\mathbf{W}^\top = \mathbf{W}$. Then $\mathbf{W} = \mathbf{V}\mathbf{V}^\top$ for some unitary $\mathbf{V}$.

Proof. From $\mathbf{W}^{-1} = \mathbf{W}^* = \overline{\mathbf{W}}^\top = \overline{\mathbf{W}}$ we get $\overline{\mathbf{W}}\mathbf{W} = \mathbf{I}$. Write $\mathbf{W} = \mathbf{X} + i\mathbf{Y}$ with $\mathbf{X} = \frac{1}{2}(\mathbf{W} + \overline{\mathbf{W}})$ and $\mathbf{Y} = \frac{1}{2i}(\mathbf{W} - \overline{\mathbf{W}})$, both **real** matrices, and both symmetric because $\mathbf{W}$

and $\overline{\mathbf{W}}$ are. Then

$$\mathbf{I} = \overline{\mathbf{W}}\mathbf{W} = (\mathbf{X} - i\mathbf{Y})(\mathbf{X} + i\mathbf{Y}) = \mathbf{X}^2 + \mathbf{Y}^2 + i(\mathbf{XY} - \mathbf{YX}),$$

and since $\mathbf{X}^2 + \mathbf{Y}^2$ and $\mathbf{XY} - \mathbf{YX}$ are real matrices, comparing real and imaginary parts gives

$$\mathbf{X}^2 + \mathbf{Y}^2 = \mathbf{I}, \quad \mathbf{XY} = \mathbf{YX}.$$

So $\mathbf{X}$ and $\mathbf{Y}$ are commuting real symmetric matrices, and Corollary 12.7.3 supplies $\mathbf{Q} \in \mathrm{O}(n)$ with $\mathbf{Q}^\top \mathbf{X} \mathbf{Q} = \mathrm{diag}(x_1, \dots, x_n)$ and $\mathbf{Q}^\top \mathbf{Y} \mathbf{Q} = \mathrm{diag}(y_1, \dots, y_n)$. Hence $\mathbf{Q}^\top \mathbf{W} \mathbf{Q} = \mathrm{diag}(x_1 + iy_1, \dots, x_n + iy_n)$, and $x_j^2 + y_j^2 = 1$ for every j by the first relation. Each $x_j + iy_j$ therefore has modulus 1. The polynomial $x^2 - (x_j + iy_j)$ has a root $c_j \in \mathbb{C}$ by Theorem 6.5.1, and $|c_j|^2 = |c_j^2| = 1$, so $|c_j| = 1$. Put $\mathbf{C} = \mathrm{diag}(c_1, \dots, c_n)$ and $\mathbf{V} := \mathbf{Q}\mathbf{C}$, which is unitary as a product of unitary matrices. Since $\mathbf{Q}$ is real orthogonal, $\mathbf{Q}^\top = \mathbf{Q}^{-1}$, and $\mathbf{C}^\top = \mathbf{C}$, so

$$\mathbf{V}\mathbf{V}^\top = \mathbf{Q}\mathbf{C}\mathbf{C}^\top \mathbf{Q}^\top = \mathbf{Q}\mathbf{C}^2 \mathbf{Q}^\top = \mathbf{Q} \mathrm{diag}(x_j + iy_j) \mathbf{Q}^\top = \mathbf{W}.$$

This proves the claim. ■

Step 4: conjugate to a real matrix. Take $\mathbf{V}$ from Step 3, so $\mathbf{W} = \mathbf{V}\mathbf{V}^\top$, and set

$$\mathbf{B} := \mathbf{V}^* \mathbf{A} \overline{\mathbf{V}}.$$

Three observations. First, $\mathbf{B}^\top = \mathbf{B}$: using $(\overline{\mathbf{V}})^\top = \mathbf{V}^*$, $(\mathbf{V}^*)^\top = \overline{\mathbf{V}}$, and $\mathbf{A}^\top = \mathbf{A}$,

$$\mathbf{B}^\top = (\overline{\mathbf{V}})^\top \mathbf{A}^\top (\mathbf{V}^*)^\top = \mathbf{V}^* \mathbf{A} \overline{\mathbf{V}} = \mathbf{B}.$$

Second, $\mathbf{B} \succ \mathbf{O}$: since $\mathbf{A} = \mathbf{W}\mathbf{P} = \mathbf{V}\mathbf{V}^\top \mathbf{P}$ and $\mathbf{V}^* \mathbf{V} = \mathbf{I}$,

$$\mathbf{B} = \mathbf{V}^* \mathbf{V} \mathbf{V}^\top \mathbf{P} \overline{\mathbf{V}} = \mathbf{V}^\top \mathbf{P} \overline{\mathbf{V}} = (\overline{\mathbf{V}})^* \mathbf{P} (\overline{\mathbf{V}}),$$

a unitary congruence of the positive definite $\mathbf{P}$, so $\mathbf{x}^* \mathbf{B} \mathbf{x} = (\overline{\mathbf{V}} \mathbf{x})^* \mathbf{P} (\overline{\mathbf{V}} \mathbf{x}) > 0$ for $\mathbf{x} \neq \mathbf{0}$; in particular $\mathbf{B}^* = \mathbf{B}$. Third, $\mathbf{B}$ is **real**, because

$$\overline{\mathbf{B}} = \overline{\mathbf{B}^\top} = \mathbf{B}^* = \mathbf{B},$$

the first equality using $\mathbf{B}^\top = \mathbf{B}$, the second the definition of the conjugate transpose, and the third Hermitian symmetry.

So $\mathbf{B}$ is a real symmetric positive definite matrix. By Corollary 12.5.2 there is $\mathbf{Q}_0 \in \mathrm{O}(n)$ with $\mathbf{B} = \mathbf{Q}_0 \mathbf{\Sigma} \mathbf{Q}_0^\top$, where $\mathbf{\Sigma}$ is the diagonal matrix of eigenvalues of $\mathbf{B}$, which we list in decreasing order. Finally, using $\mathbf{V}\mathbf{V}^* = \mathbf{I}$ and its conjugate $\overline{\mathbf{V}}\mathbf{V}^\top = \mathbf{I}$,

$$\mathbf{V}\mathbf{B}\mathbf{V}^\top = \mathbf{V}\mathbf{V}^* \mathbf{A} \overline{\mathbf{V}}\mathbf{V}^\top = \mathbf{A},$$

so with $\mathbf{U} := \mathbf{V}\mathbf{Q}_0$, which is unitary,

$$\mathbf{A} = \mathbf{V}\mathbf{Q}_0 \mathbf{\Sigma} \mathbf{Q}_0^\top \mathbf{V}^\top = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\top.$$

It remains to identify the diagonal. Conjugate-transposing, $\mathbf{A}^* = \overline{\mathbf{U}}\mathbf{\Sigma}\mathbf{U}^*$, so

$$\mathbf{A}^*\mathbf{A} = \overline{\mathbf{U}}\mathbf{\Sigma}\mathbf{U}^*\mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top = \overline{\mathbf{U}}\mathbf{\Sigma}^2(\overline{\mathbf{U}})^*.$$

This exhibits $\mathbf{\Sigma}^2$ as the eigenvalue list of $\mathbf{A}^*\mathbf{A}$ in decreasing order, so $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$ by Definition 13.9.1. This proves the theorem for invertible $\mathbf{A}$. ■

Remark. What is left out. The theorem as stated is true for every complex symmetric $\mathbf{A}$, singular ones included, and the example below exhibits a singular case. Our proof used invertibility twice: to know that $\mathbf{P}$ is invertible, so that $\mathbf{W} = \mathbf{A}\mathbf{P}^{-1}$ is determined and the cancellations of Step 2 are legal. One way to remove the hypothesis is to apply the proved case to $\mathbf{A} + t\mathbf{I}$, which is symmetric and invertible for all but finitely many t , and let $t \rightarrow 0$ along a sequence for which the unitary factors converge — a step that needs the compactness of the unitary group, which is analysis this book has not set up. We state the general theorem and prove the invertible case.

Example. Takagi for a real symmetric matrix

Find a Takagi factorization of $\mathbf{A} = \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}$.

Solution. First the singular values: $\mathbf{A}^*\mathbf{A} = \mathbf{A}^\top\mathbf{A} = \mathbf{A}^2 = 25\mathbf{I}$, so $\sigma_1 = \sigma_2 = 5$ and $\mathbf{\Sigma} = 5\mathbf{I}$. We therefore need a unitary $\mathbf{U}$ with

$$\mathbf{A} = 5\mathbf{U}\mathbf{U}^\top, \quad \text{that is} \quad \mathbf{U}\mathbf{U}^\top = \frac{1}{5}\mathbf{A}.$$

The matrix $\frac{1}{5}\mathbf{A}$ is real symmetric and orthogonal, with eigenvalues 1 and -1 . Its eigenvectors: $(\mathbf{A} - 5\mathbf{I})\mathbf{x} = \mathbf{0}$ reads $-2x_1 + 4x_2 = 0$, giving $\mathbf{q}_1 = \frac{1}{\sqrt{5}}(2, 1)$, and orthogonally $\mathbf{q}_2 = \frac{1}{\sqrt{5}}(1, -2)$.

So with $\mathbf{Q} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix} \in \text{O}(2)$ we have $\frac{1}{5}\mathbf{A} = \mathbf{Q} \text{diag}(1, -1)\mathbf{Q}^\top$. The diagonal entries 1 and -1 have square roots 1 and i , so, following Step 3 of [Theorem 13.10.3](#), set

$$\mathbf{U} = \mathbf{Q} \text{diag}(1, i) = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & i \\ 1 & -2i \end{pmatrix}.$$

Then $\mathbf{U}$ is unitary — its columns $\frac{1}{\sqrt{5}}(2, 1)$ and $\frac{i}{\sqrt{5}}(1, -2)$ are orthonormal, since multiplying a real orthonormal pair by unit scalars keeps it orthonormal — and

$$\mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top = 5\mathbf{U}\mathbf{U}^\top = \mathbf{Q} \text{diag}(5, -5)\mathbf{Q}^\top = \mathbf{A}.$$

Note where the complex entries were needed: $\text{diag}(1, i)$ converted the eigenvalue -5 of $\mathbf{A}$ into the singular value $+5$.

For a singular example, return to $\mathbf{N} = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$. Here $\mathbf{N}^*\mathbf{N} = \begin{pmatrix} 2 & 2i \\ -2i & 2 \end{pmatrix}$, with eigenval-

ues 4 and 0, so $\sigma_1 = 2$ and $\sigma_2 = 0$. With $\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$, which is unitary, one checks $\mathbf{U} \operatorname{diag}(2, 0) \mathbf{U}^\top = 2 \mathbf{u}_1 \mathbf{u}_1^\top = \mathbf{N}$, where $\mathbf{u}_1 = \frac{1}{\sqrt{2}}(1, i)$ is the first column of $\mathbf{U}$. The theorem holds here even though our proof did not reach it, and even though $\mathbf{N}$ has no eigenbasis at all.

Warning. Symmetric is not Hermitian, and $\mathbf{U}^\top$ is not $\mathbf{U}^*$. Takagi's hypothesis is $\mathbf{A}^\top = \mathbf{A}$ with no conjugation, which over $\mathbb{C}$ is a different and weaker condition than $\mathbf{A}^* = \mathbf{A}$; the matrix $\mathbf{N}$ above is symmetric, is not Hermitian, is not normal, and is not diagonalizable. The conclusion likewise carries $\mathbf{U}^\top$, not $\mathbf{U}^*$: writing $\mathbf{U}\Sigma\mathbf{U}^*$ by mistake would assert that $\mathbf{A}$ is positive semidefinite, which $\mathbf{N}$ certainly is not.

13.10.4 Exercises

13.10.4.1 A. Check your understanding

Exercise. A1

- State the polar decomposition, and say which factor is unique and when.
- What is $|\mathbf{A}|$, and what identity relates $\| |\mathbf{A}| \mathbf{x} \|$ to $\| \mathbf{A} \mathbf{x} \|$?
- What does the polar decomposition say when $n = 1$?
- True or false: $|\mathbf{A}| = |\mathbf{A}^*|$ for every square $\mathbf{A}$. Justify your answer.
- State Takagi's factorization, and name two ways in which it differs from the singular value decomposition and from the complex spectral theorem.
- True or false: a complex symmetric matrix is diagonalizable. Justify your answer.

Solution.

- Every $\mathbf{A} \in M_n(\mathbb{C})$ is $\mathbf{A} = \mathbf{W}|\mathbf{A}|$ with $\mathbf{W}$ unitary (Theorem 13.10.1). The positive semidefinite factor $|\mathbf{A}|$ is always unique; $\mathbf{W}$ is unique exactly when $\mathbf{A}$ is invertible, and then $\mathbf{W} = \mathbf{A}|\mathbf{A}|^{-1}$.
- $|\mathbf{A}| = (\mathbf{A}^* \mathbf{A})^{1/2}$, the unique positive semidefinite square root of $\mathbf{A}^* \mathbf{A}$ (Definition 13.2.1), and $\| |\mathbf{A}| \mathbf{x} \| = \| \mathbf{A} \mathbf{x} \|$ for every $\mathbf{x}$.
- It says $z = e^{i\theta}|z|$: a complex number is a number of modulus 1 times its modulus, with θ undetermined exactly when $z = 0$.
- False in general, true exactly for normal $\mathbf{A}$. For $\mathbf{A} = \begin{pmatrix} 3 & 0 \\ 4 & 5 \end{pmatrix}$, $\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 25 & 20 \\ 20 & 25 \end{pmatrix} \neq \begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix} = \mathbf{A} \mathbf{A}^\top$, so the two square roots differ.
- Theorem 13.10.3: a symmetric $\mathbf{A} \in M_n(\mathbb{C})$ is $\mathbf{U}\Sigma\mathbf{U}^\top$ with $\mathbf{U}$ unitary and Σ the diagonal matrix of singular values. Against the singular value decomposition: the right factor is not free but forced to be $\overline{\mathbf{U}}$. Against the spectral theorem: the hypothesis is symmetry rather than normality, the operation is the congruence $\mathbf{X} \mapsto \mathbf{U}\mathbf{X}\mathbf{U}^\top$ rather than a similarity, and the diagonal carries singular values rather than eigenvalues.
- False. $\mathbf{N} = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$ is symmetric and nonzero with $\mathbf{N}^2 = \mathbf{O}$, so its minimal poly-

mial is x^2 , which has a repeated root; it is not diagonalizable.

13.10.4.2 B. Practice

Exercise. B1: A polar decomposition by hand

Find both polar decompositions $\mathbf{A} = \mathbf{W}|\mathbf{A}| = |\mathbf{A}^*|\mathbf{W}$ of

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

and say what $\mathbf{W}$ is geometrically.

Solution. $\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 9 & 0 \\ 0 & 4 \end{pmatrix}$, which is diagonal with nonnegative entries, so $|\mathbf{A}| = \text{diag}(3, 2)$: this is positive semidefinite and squares to $\mathbf{A}^\top \mathbf{A}$, which by Theorem 13.2.1 identifies it. Since $\mathbf{A}$ is invertible,

$$\mathbf{W} = \mathbf{A}|\mathbf{A}|^{-1} = \begin{pmatrix} 0 & -2 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

the quarter turn of $\mathbb{R}^2$. Check: $\mathbf{W}^\top \mathbf{W} = \mathbf{I}$ and $\mathbf{W}|\mathbf{A}| = \mathbf{A}$.

For the right-handed form, $\mathbf{A}\mathbf{A}^\top = \text{diag}(4, 9)$, so $|\mathbf{A}^*| = \text{diag}(2, 3)$, and indeed

$$|\mathbf{A}^*|\mathbf{W} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 \\ 3 & 0 \end{pmatrix} = \mathbf{A}.$$

The two absolute values are different, which is consistent with $\mathbf{A}$ not being normal. In words: $\mathbf{A}$ stretches by 3 horizontally and 2 vertically, then rotates by a quarter turn; equivalently it rotates first and then stretches by 2 horizontally and 3 vertically.

Exercise. B2: The nearest unitary matrix

Let $\mathbf{A} = \text{diag}(2, -3) \in M_2(\mathbb{R})$. Find the unitary matrix nearest to $\mathbf{A}$ in the Frobenius norm, compute the distance, and verify the value against the formula of [Corollary 13.10.2](#).

Solution. $\mathbf{A}^* \mathbf{A} = \text{diag}(4, 9)$, so $|\mathbf{A}| = \text{diag}(2, 3)$ and $\mathbf{W} = \mathbf{A}|\mathbf{A}|^{-1} = \text{diag}(1, -1)$, which is unitary. By [Corollary 13.10.2](#), $\mathbf{W}$ is the nearest unitary matrix, and it is the only one since $\mathbf{A}$ is invertible. Directly,

$$\mathbf{A} - \mathbf{W} = \text{diag}(1, -2), \quad \|\mathbf{A} - \mathbf{W}\|_F^2 = 1 + 4 = 5.$$

The formula gives $\|\mathbf{A}\|_F^2 + n - 2(\sigma_1 + \sigma_2) = 13 + 2 - 2(3 + 2) = 5$, the same. So the distance is $\sqrt{5}$.

Exercise. B3: A Takagi factorization

Find a Takagi factorization of $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and explain why no **real** $\mathbf{U}$ can work.

Solution. $\mathbf{A}$ is symmetric and $\mathbf{A}^* \mathbf{A} = \mathbf{A}^2 = \mathbf{I}$, so $\sigma_1 = \sigma_2 = 1$ and $\Sigma = \mathbf{I}$. We need a unitary $\mathbf{U}$ with $\mathbf{U}\mathbf{U}^\top = \mathbf{A}$.

Diagonalize $\mathbf{A}$ over $\mathbb{R}$: its eigenvalues are 1 and -1 , with unit eigenvectors $\frac{1}{\sqrt{2}}(1, 1)$ and $\frac{1}{\sqrt{2}}(1, -1)$, so $\mathbf{A} = \mathbf{Q} \text{diag}(1, -1) \mathbf{Q}^\top$ with $\mathbf{Q} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. Following Step 3 of [Theorem 13.10.3](#), take square roots 1 and i of the diagonal entries:

$$\mathbf{U} = \mathbf{Q} \text{diag}(1, i) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}, \quad \mathbf{U}\mathbf{U}^\top = \mathbf{Q} \text{diag}(1, -1) \mathbf{Q}^\top = \mathbf{A}.$$

$\mathbf{U}$ is unitary, and $\mathbf{U}\Sigma\mathbf{U}^\top = \mathbf{U}\mathbf{U}^\top = \mathbf{A}$.

No real $\mathbf{U}$ works: a real $\mathbf{U}$ with $\mathbf{U}^* \mathbf{U} = \mathbf{I}$ is orthogonal, and then $\mathbf{U}\Sigma\mathbf{U}^\top = \mathbf{U}\mathbf{U}^\top = \mathbf{I} \neq \mathbf{A}$.

13.10.4.3 C. Going deeper

Exercise. C1: The positive factor is forced

Let $\mathbf{A} \in M_n(\mathbb{C})$.

- Suppose $\mathbf{A} = \mathbf{Q}\mathbf{P}$ with $\mathbf{Q}$ unitary and $\mathbf{P} \succeq \mathbf{O}$. Prove that $\mathbf{P} = |\mathbf{A}|$. (Invertibility is **not** assumed.)
- Deduce that $\mathbf{A}$ is unitary if and only if $|\mathbf{A}| = \mathbf{I}$.

Solution.

- From $\mathbf{A} = \mathbf{Q}\mathbf{P}$ we get $\mathbf{A}^* \mathbf{A} = \mathbf{P}^* \mathbf{Q}^* \mathbf{Q} \mathbf{P} = \mathbf{P}^* \mathbf{P} = \mathbf{P}^2$, using $\mathbf{Q}^* \mathbf{Q} = \mathbf{I}$ and $\mathbf{P}^* = \mathbf{P}$. So $\mathbf{P}$ is a positive semidefinite square root of $\mathbf{A}^* \mathbf{A}$. Such a square root is unique (Theorem 13.2.1), and $|\mathbf{A}|$ is one by definition (Definition 13.2.1). Hence $\mathbf{P} = |\mathbf{A}|$.
- ($\Rightarrow$) If $\mathbf{A}$ is unitary then $\mathbf{A}^* \mathbf{A} = \mathbf{I}$, and $\mathbf{I} \succeq \mathbf{O}$ with $\mathbf{I}^2 = \mathbf{I}$, so $|\mathbf{A}| = \mathbf{I}$ by uniqueness. ($\Leftarrow$) If $|\mathbf{A}| = \mathbf{I}$ then $\mathbf{A}^* \mathbf{A} = |\mathbf{A}|^2 = \mathbf{I}$, which for a square matrix says $\mathbf{A}$ is unitary (Definition 11.7.2).

Exercise. C2: Normality through the polar factors

Let $\mathbf{A} \in M_n(\mathbb{C})$ be invertible with polar decomposition $\mathbf{A} = \mathbf{W}|\mathbf{A}|$. Prove that the following are equivalent.

- $\mathbf{A}$ is normal.
- $|\mathbf{A}| = |\mathbf{A}^*|$.
- $\mathbf{W}|\mathbf{A}| = |\mathbf{A}|\mathbf{W}$.

Solution.

- (a) $\Rightarrow$ (b). If $\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^*$, then $|\mathbf{A}|$ and $|\mathbf{A}^*|$ are positive semidefinite square roots of the same matrix, hence equal by Theorem 13.2.1.
- (b) $\Rightarrow$ (c). By Theorem 13.10.1, $|\mathbf{A}^*| = \mathbf{W}|\mathbf{A}|\mathbf{W}^*$. If this equals $|\mathbf{A}|$, multiply on the right by $\mathbf{W}$ to get $|\mathbf{A}|\mathbf{W} = \mathbf{W}|\mathbf{A}|$.
- (c) $\Rightarrow$ (a). Write $\mathbf{P} = |\mathbf{A}|$, so $\mathbf{A} = \mathbf{W}\mathbf{P} = \mathbf{P}\mathbf{W}$. Then

$$\mathbf{A}\mathbf{A}^* = \mathbf{W}\mathbf{P}\mathbf{P}^*\mathbf{W}^* = \mathbf{W}\mathbf{P}^2\mathbf{W}^* = \mathbf{P}^2\mathbf{W}\mathbf{W}^* = \mathbf{P}^2,$$

using $\mathbf{P}^* = \mathbf{P}$, then $\mathbf{W}\mathbf{P}^2 = \mathbf{P}^2\mathbf{W}$ (which follows from $\mathbf{W}\mathbf{P} = \mathbf{P}\mathbf{W}$), then $\mathbf{W}\mathbf{W}^* = \mathbf{I}$. And $\mathbf{A}^* \mathbf{A} = \mathbf{P}^* \mathbf{W}^* \mathbf{W} \mathbf{P} = \mathbf{P}^2$. So $\mathbf{A}\mathbf{A}^* = \mathbf{A}^* \mathbf{A}$.

Exercise. C3: Complex symmetric matrices are transposed squares

Let $\mathbf{A} \in M_n(\mathbb{C})$ be symmetric and invertible.

- (a) Using Theorem 13.10.3, prove that $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$ for some invertible $\mathbf{B} \in M_n(\mathbb{C})$.
- (b) Show that over $\mathbb{R}$ the analogous statement is false unless $\mathbf{A}$ is positive definite, and give a 1×1 witness.
- (c) Explain in one sentence why (a) and (b) do not contradict each other.

Solution.

- (a) By Theorem 13.10.3, $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top$ with $\mathbf{U}$ unitary and $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$, where every $\sigma_i > 0$ because $\mathbf{A}$ is invertible (Theorem 13.9.1). Let $\mathbf{\Sigma}^{1/2} = \text{diag}(\sqrt{\sigma_1}, \dots, \sqrt{\sigma_n})$ and put $\mathbf{B} := \mathbf{\Sigma}^{1/2}\mathbf{U}^\top$, which is invertible as a product of invertible matrices. Then

$$\mathbf{B}^\top \mathbf{B} = \mathbf{U} \mathbf{\Sigma}^{1/2} \mathbf{\Sigma}^{1/2} \mathbf{U}^\top = \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\top = \mathbf{A},$$

using $(\mathbf{\Sigma}^{1/2})^\top = \mathbf{\Sigma}^{1/2}$ and $(\mathbf{U}^\top)^\top = \mathbf{U}$.

- (b) Over $\mathbb{R}$, if $\mathbf{A} = \mathbf{B}^\top \mathbf{B}$ with $\mathbf{B} \in M_n(\mathbb{R})$ then $\mathbf{x}^\top \mathbf{A} \mathbf{x} = \|\mathbf{B} \mathbf{x}\|^2 \geq 0$ for every $\mathbf{x}$, so $\mathbf{A} \succeq \mathbf{O}$, and $\mathbf{A} \succ \mathbf{O}$ when $\mathbf{B}$ is invertible. So no real $\mathbf{A}$ with a negative eigenvalue is of this form: for $n = 1$, $\mathbf{A} = (-1)$ is symmetric and invertible but $b^2 = -1$ has no real solution.
- (c) The factor $\mathbf{B}$ produced in (a) is complex, and over $\mathbb{C}$ the quantity $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ is not a sum of squared lengths — the transpose carries no conjugation, so $\mathbf{B}^\top \mathbf{B}$ says nothing about positivity. Indeed $b = i$ solves $b^2 = -1$.