The Schur decomposition

Nicholas Hu · Last updated on 2025-05-28

The complex Schur decomposition

Let $A \in \mathbb{C}^{n \times n}$. The **(complex) Schur decomposition** is a factorization of A as UTU^{-1}, where $U \in \mathbb{C}^{n \times n}$ is *unitary* and $T \in \mathbb{C}^{n \times n}$ is *upper triangular*.

Such a factorization always exists and can be constructed recursively: let $\lambda \in \mathbb{C}$ be an eigenvalue of A (which exists by the fundamental theorem of algebra) and $v_1 \in \mathbb{C}^n$ be a corresponding normalized eigenvector. Extending $\{v_1\}$ to an orthonormal basis $\{v_j\}_{j=1}^n$ and defining $V := \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \in \mathbb{C}^{n \times n}$, we obtain

$$A = V \begin{bmatrix} \lambda & b^* \\ & \hat{A} \end{bmatrix} V^{-1}$$

for some $\hat{A} \in \mathbb{C}^{(n-1)\times(n-1)}$ and $b \in \mathbb{C}^{n-1}$. Thus, if $\hat{A}$ has a Schur decomposition $\hat{U}\hat{T}\hat{U}^{-1}$, then

$$A = \underbrace{V \begin{bmatrix} 1 & \\ & \hat{U} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \lambda & b^*\hat{U} \\ & \hat{T} \end{bmatrix}}_{T} \underbrace{\begin{bmatrix} 1 & \\ & \hat{U}^{-1} \end{bmatrix} V^{-1}}_{U^{-1}}$$

is a Schur decomposition of A, and its existence in the base case $n = 1$ is trivial.

Complex normal matrices

If a normal matrix $T \in \mathbb{C}^{n \times n}$ can be partitioned as $\begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix}$ with $T_{11} \in \mathbb{C}^{m \times m}$, then by definition

$$\begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix} \begin{bmatrix} T_{11}^* & \\ T_{12}^* & T_{22}^* \end{bmatrix} = \begin{bmatrix} T_{11}^* & \\ T_{12}^* & T_{22}^* \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} \\ & T_{22} \end{bmatrix},$$

so $T_{11}T_{11}^* + T_{12}T_{12}^* = T_{11}^*T_{11}$. Taking traces, we obtain $\|T_{11}\|_{\mathrm{F}}^2 + \|T_{12}\|_{\mathrm{F}}^2 = \|T_{11}\|_{\mathrm{F}}^2$, which implies that $T_{12} = 0$ and hence that T_{11} and T_{22} are normal.

Now if A is normal, then so is the upper triangular matrix $T = U^*AU$ in its complex Schur factorization. It follows that T must in fact be *diagonal*, which yields the **complex spectral theorem**: a (square) complex matrix is unitarily diagonalizable if and only if it is normal.

The real Schur decomposition

Let $A \in \mathbb{R}^{n \times n}$. The **real Schur decomposition** is a factorization of A as QTQ^{-1}, where $Q \in \mathbb{R}^{n \times n}$ is *orthogonal* and $T \in \mathbb{R}^{n \times n}$ is *upper quasi-triangular* – that is, block upper triangular with diagonal blocks that are 1×1 or 2×2.

As in the complex case, such a factorization always exists; arguing as above, it suffices to show that A has a one- or two-dimensional invariant subspace. Indeed, when regarded as a complex matrix, A has an eigenvalue $\lambda = \alpha + i\beta \in \mathbb{C}$ (where $\alpha, \beta \in \mathbb{R}$) and an eigenvector $v_1 = x_1 + iy_1 \in \mathbb{C}^n$ (where $x_1, y_1 \in \mathbb{R}^n$).

If $\lambda \in \mathbb{R}$, then $Ax_1 + iAy_1 = \lambda x_1 + i\lambda y_1$, so x_1 or y_1 is an eigenvector with eigenvalue λ (at least one of these vectors must be nonzero) and therefore spans a one-dimensional invariant subspace of A. On the other hand, if $\lambda \in \mathbb{C} \setminus \mathbb{R}$, then $Ax_1 + iAy_1 = (\alpha x_1 - \beta y_1) + i(\beta x_1 + \alpha y_1)$. Moreover, $\overline{v_1}$ is an eigenvector of A with eigenvalue $\overline{\lambda}$ and $\lambda \neq \overline{\lambda}$, so $\{v_1, \overline{v_1}\}$ is linearly independent. Hence $x_1 = \frac{1}{2}(v_1 + \overline{v_1})$ and $y_1 = \frac{1}{2i}(v_1 - \overline{v_1})$ span a two-dimensional invariant subspace of A.

Real normal matrices

If A is normal, then $T = Q^* AQ$ will be “*quasi-diagonal*” – block diagonal with diagonal blocks that are 1×1 or 2×2 – and each diagonal block will itself be normal. Thus, these blocks will be of the form

$$[\alpha] \quad (\alpha \in \mathbb{R}) \qquad \text{or} \qquad \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \quad (\alpha \in \mathbb{R},\, \beta \in \mathbb{R} \setminus \{0\}).$$

In particular, if A is symmetric, there will only be 1×1 blocks, which yields the **real spectral theorem**: a (square) real matrix is orthogonally diagonalizable if and only if it is symmetric. If A is orthogonal, we will have $\alpha = \pm 1$ in the 1×1 blocks and $\alpha = \cos(\theta)$ and $\beta = \sin(\theta)$ for some $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$ in the 2×2 blocks.