Schur's Triangularization Theorem

Math 422

The characteristic polynomial $p(t)$ of a square complex matrix A splits as a product of linear factors of the form $(t-\lambda)^m$. Of course, finding these factors is a difficult problem, but having factored $p(t)$ we can triangularize A whether or not A is diagonalizable.

Example 1 *The characteristic polynomial $p(t) = t^2$ of the triangular matrix*

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

has the single root $\lambda = 0$, which is an eigenvalue of algebraic multiplicity 2. The eigenspace of λ is one dimensional and is spanned by the single vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, so the geometric multiplicity of λ is 1. Therefore A is defective and is not diagonalizable (one needs two linearly independent eigenvectors to construct a transition matrix P that diagonalizes A).

Let V^n be an n-dimensional complex inner product space with Euclidean inner product.

Definition 2 *A hyperplane in V^n is a translation of an $(n-1)$-dimensional subspace.*

Note that the orthogonal complement $\mathbf{u}^\perp$ of a non-zero vector $\mathbf{u} \in \mathbb{C}^n$ is a hyperplane through the origin. Consider the matrix

$$P = I - \frac{1}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*;$$

then $Q = P - \frac{1}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^* = I - \frac{2}{\|\mathbf{u}\|^2}\,\mathbf{u}\mathbf{u}^*$ is the Householder matrix associated with $\mathbf{u}$.

Proposition 3 *$N(P) = span\{\mathbf{u}\}$ and multiplication by P is orthogonal projection on $\mathbf{u}^\perp$, i.e., for all $\mathbf{x} \in \mathbb{C}^n$,*

$$P\mathbf{x} = \mathbf{proj}_{\mathbf{u}^\perp}\mathbf{x}.$$

Proof. If $P\mathbf{x} = 0$, then $\mathbf{x} - \frac{\mathbf{u}^*\mathbf{x}}{\|\mathbf{u}\|^2}\mathbf{u} = 0$ or equivalently $\mathbf{x} = \frac{\mathbf{u}^*\mathbf{x}}{\|\mathbf{u}\|^2}\mathbf{u}$. Thus $\mathbf{x} = t\mathbf{u}$ for some $t \in \mathbb{C}$, and $N(P) = span\{\mathbf{u}\}$. Furthermore, for all $\mathbf{x} \in \mathbb{C}^n$, $P\mathbf{x} = \mathbf{x} - \mathbf{proj}_{\mathbf{u}}\mathbf{x} = \mathbf{proj}_{\mathbf{u}^\perp}\mathbf{x}$. ■

Definition 4 *Let $\mathbf{x}, \mathbf{y}, \mathbf{u} \in \mathbb{C}^n$ with $\mathbf{u} \neq 0$. Then $\mathbf{y}$ is the reflection of $\mathbf{x}$ in the hyperplane $\mathbf{u}^\perp$ iff*

$$\mathbf{x} - \mathbf{y} = 2\ \mathbf{proj}_{\mathbf{u}}\mathbf{x}.$$

Proposition 5 *Let $\mathbf{u} \in \mathbb{C}^n$ be a non-zero vector. The Householder transformation associated with $\mathbf{u}$ is reflection in the hyperplane $\mathbf{u}^\perp$.*

Proof. Note that for all $\mathbf{x} \in C^n$, $Q\mathbf{x} = \mathbf{x} - 2\left(\frac{\mathbf{u}^*\mathbf{x}}{\|\mathbf{u}\|^2}\mathbf{u}\right) = \mathbf{x} - 2\ \mathbf{proj}_{\mathbf{u}}\mathbf{x}$. Thus $\mathbf{x} - Q\mathbf{x} = 2\ \mathbf{proj}_{\mathbf{u}}\mathbf{x}$. ■

Exercise 6 *Earlier we proved that a real Householder matrix Q is symmetric and orthogonal, i.e.,$Q^T = Q$ and $Q^{-1} = Q^T$. Generalize this result for complex matrices: Prove that complex Householder matrices Q are Hermitian and unitary.*

Let $\mathbf{v} = (v_1, \ldots, v_n)$ be a non-zero vector in $\mathbb{C}^n$ and set $\mathbf{x} = \frac{\overline{v_1}\mathbf{v}}{\|\overline{v_1}\mathbf{v}\|}$. Then $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{C}^n$ is a unit vector with $x_1 \in \mathbb{R}$. Let $\mathbf{u} = \mathbf{x} - \mathbf{e}_1$ then

$$\begin{aligned}
\|\mathbf{u}\|^2 &= \langle \mathbf{x} - \mathbf{e}_1, \mathbf{x} - \mathbf{e}_1 \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{e}_1 \rangle - \langle \mathbf{e}_1, \mathbf{x} \rangle + \langle \mathbf{e}_1, \mathbf{e}_1 \rangle = 2 - 2x_1 = 2(1 - x_1) \\
\mathbf{u}^*\mathbf{x} &= \langle \mathbf{x}, \mathbf{u} \rangle = \langle \mathbf{x}, \mathbf{x} - \mathbf{e}_1 \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - \langle \mathbf{x}, \mathbf{e}_1 \rangle = 1 - x_1.
\end{aligned}$$

If $\mathbf{x} \neq \mathbf{e}_1$, then $\mathbf{u} \neq 0$ and we may apply the Householder transformation Q associated with $\mathbf{u}$ to $\mathbf{x}$ and $\mathbf{e}_1$:

$$Q\mathbf{x} = \mathbf{x} - \frac{2\mathbf{u}^*\mathbf{x}}{\|\mathbf{u}\|^2}\mathbf{u} = \mathbf{x} - \frac{2(1-x_1)}{2(1-x_1)}\mathbf{u} = \mathbf{x} - \mathbf{u} = \mathbf{x} - (\mathbf{x} - \mathbf{e}_1) = \mathbf{e}_1; \tag{1}$$

applying Q to both sides of (1) and using the fact that $Q^2 = I$ we have

$$\mathbf{x} = Q^2\mathbf{x} = Q\mathbf{e}_1.$$

If $\mathbf{x} = \mathbf{e}_1$, set $Q = I$; then in either case

$$\mathbf{x} = Q\mathbf{e}_1 \text{ and } \mathbf{e}_1 = Q\mathbf{x}$$

are reflection of each other in the hyperplane $(\mathbf{x} - \mathbf{e}_1)^\perp$. For a unit vector $\mathbf{x} \in \mathbb{R}^2$, the line $(\mathbf{x} - \mathbf{e}_1)^\perp$ bisects the angle between $\mathbf{x}$ and $\mathbf{e}_1$. We are ready to prove our main theorem in this lecture:

Theorem 7 *(Schur's Triangularization Theorem) Every $n \times n$ complex matrix A is unitarily similar to an upper-triangular matrix T, i.e., there exists a unitary matrix U such that $U^*AU = T$.*

Proof. Use induction on the size of A. For $n = 1$ there is nothing to prove. So assume $n > 1$ and that the result holds for all matrices of size less than n. Since every complex matrix has an eigenvalue, choose an eigenvalue λ of A and an associated eigenvector $\mathbf{v} = (v_1, \ldots, v_n)$. Let $\mathbf{x} = \frac{\overline{v_1}\mathbf{v}}{\|\overline{v_1}\mathbf{v}\|}$ and set $\mathbf{u} = \mathbf{x} - \mathbf{e}_1$; if $\mathbf{x} \neq \mathbf{e}_1$, let Q be the Householder matrix associated with $\mathbf{u}$; if $\mathbf{x} = \mathbf{e}_1$ let $Q = I$. Then $\mathbf{x} = Q\mathbf{e}_1$ by the discussion above, so $\mathbf{x}$ is the first column of Q. By Exercise 6, Q is Hermitian and unitary, so $\mathbf{x}^*$ is the first row of Q. Since $Q = Q^{-1} = Q^*$ we have $Q = [\mathbf{x} \mid V] = \begin{bmatrix} \mathbf{x}^* \\ V^* \end{bmatrix}$ and

$$QAQ = QA[\mathbf{x} \mid V] = Q[\lambda\mathbf{x} \mid AV] = \left[\lambda\mathbf{e}_1 \mid \begin{bmatrix} \mathbf{x}^* \\ V^* \end{bmatrix} AV\right] = \begin{bmatrix} \lambda & \mathbf{x}^*AV \\ 0 & V^*AV \end{bmatrix}.$$

Now apply the induction hypothesis to V^*AV, which is an $(n-1) \times (n-1)$ matrix, and obtain an $(n-1) \times (n-1)$ unitary matrix R such that $T_{n-1} = R^*(V^*AV)R$ is upper-triangular. Let

$$U = Q\begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix};$$

then

$$U^*U = \begin{bmatrix} 1 & 0 \\ 0 & R^* \end{bmatrix} Q^*Q \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & R^*R \end{bmatrix} = I$$

so U is unitary. Hence

$$\begin{aligned} T &= U^*AU = \begin{bmatrix} 1 & 0 \\ 0 & R^* \end{bmatrix} QAQ \begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & R^* \end{bmatrix}\begin{bmatrix} \lambda & \mathbf{x}^*AV \\ 0 & V^*AV \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & R^* \end{bmatrix}\begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ 0 & V^*AVR \end{bmatrix} \\ &= \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ 0 & R^*V^*AVR \end{bmatrix} = \begin{bmatrix} \lambda & \mathbf{x}^*AVR \\ 0 & T_{n-1} \end{bmatrix}, \end{aligned}$$

which is upper triangular as desired. ■

Remark 8 *Since similar matrices have the same eigenvalues, the eigenvalues of A are the diagonal entries of every Schur triangularization $T = U^*AU$.*

When all eigenvalues of A are real, Schur's Triangularization Theorem tells us that A is orthogonally similar to a triangular matrix. Our next example demonstrates this.

Example 9 *Let's numerically approximate the Schur triangularization of*

$$A = \begin{bmatrix} -1 & -1 & -2 \\ 8 & -11 & -8 \\ -10 & 11 & 7 \end{bmatrix}.$$

The eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = -3$ and $\lambda_3 = -3$. Arbitrarily choose an eigenvalue, say $\lambda_1 = 1$, then

$$A - I = \begin{bmatrix} -2 & -1 & -2 \\ 8 & -12 & -8 \\ -10 & 11 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

and $\mathbf{x} = \begin{bmatrix} -1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ *is an associated unit eigenvector. Let* $\mathbf{u} = \mathbf{x} - \mathbf{e}_1 = \begin{bmatrix} -4/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ *and let Q be the associated Householder matrix, i.e.,*

$$Q = I - \frac{3}{4}\mathbf{u}\mathbf{u}^T = \frac{1}{3}\begin{bmatrix} -1 & -2 & 2 \\ -2 & 2 & 1 \\ 2 & 1 & 2 \end{bmatrix} = [\mathbf{x} \mid V],$$

where

$$V = \frac{1}{3}\begin{bmatrix} -2 & 2 \\ 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Then

$$QAQ = \frac{1}{3}\begin{bmatrix} 3 & 64 & 13 \\ 0 & -13 & -1 \\ 0 & 16 & -5 \end{bmatrix} \text{ and } V^TAV = \frac{1}{3}\begin{bmatrix} -13 & -1 \\ 16 & -5 \end{bmatrix}.$$

Now triangularize the 2×2 matrix V^TAV, which has the single eigenvalue -3. The vector $\mathbf{x} = \frac{1}{\sqrt{17}}\left[\begin{smallmatrix} 1 \\ -4 \end{smallmatrix}\right]$ is a unit vector associated with -3. Let $\mathbf{u} = \mathbf{x} - \mathbf{e}_1 = \frac{1}{\sqrt{17}}\left[\begin{smallmatrix} 1-\sqrt{17} \\ -4 \end{smallmatrix}\right]$ and let R be the Householder matrix associated with $\mathbf{u}$, i.e.,

$$R = \begin{bmatrix} 0.242\,54 & -0.970\,14 \\ -0.970\,14 & -0.242\,54 \end{bmatrix}.$$

Then

$$RV^TAVR = \begin{bmatrix} -3 & 17/3 \\ 0 & -3 \end{bmatrix}$$

is a Schur triangularization of V^TAV. Finally, let

$$U = Q\begin{bmatrix} 1 & 0 \\ 0 & R \end{bmatrix}$$

then

$$U^TAU = \begin{bmatrix} 1 & 0.970\,25 & -21.747 \\ 0 & -3.000 & 5.6667 \\ 0 & 0 & -3.000 \end{bmatrix}$$

is a (numerically approximate) Schur triangularization of A.

In summary, every matrix is triangularizable but only non-defective matrices are diagonalizable.

Exercise 10 *Show that the matrix*

$$A = \begin{bmatrix} -1 & -1 & -2 \\ 8 & -11 & -8 \\ -10 & 11 & 7 \end{bmatrix}.$$

in Example 9 is defective and hence not diagonalizable.

Exercise 11 *Following the proof of Schur's Triangularization Theorem, find an orthogonal matrix P such that $P^T AP$ is upper triangular:*

a. $A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}$

b. $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

c. $A = \begin{bmatrix} 13 & -9 \\ 16 & -11 \end{bmatrix}$

11-15-10