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 $\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, so the geometric multiplicity of λ is 1. Therefore A is defective and is not diagonalizable (one needs three 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 $u^\perp$ of a non-zero vector $u \in \mathbb{C}^n$ is a hyperplane through the origin. Consider the matrix

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

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

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

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

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

$$Px = x - \mathbf{proj}_u x = \mathbf{proj}_{u^\perp} x.$$

■

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

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

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

Proof. Note that for all $x \in C^n$,

$$Qx = x - 2\left(\frac{x \bullet u}{\|u\|^2} u\right) = x - 2\ \mathbf{proj}_u x.$$

Thus

$$x - Qx = 2\ \mathbf{proj}_u x.$$

■

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 $v = (v_1, \ldots, v_n)$ be a non-zero vector in $\mathbb{C}^n$ and set $x = \dfrac{\overline{v_1} v}{\|\overline{v_1} v\|}$. Then $x = (x_1, \ldots, x_n) \in \mathbb{C}^n$ is a unit vector with $x_1 \in \mathbb{R}$. Let $u = x - e_1$ then

$$\begin{aligned} \|u\|^2 &= (x - e_1) \bullet (x - e_1) = x \bullet x - x \bullet e_1 - e_1 \bullet x + e_1 \bullet e_1 = 2 - 2x_1 \\ &= 2\,(1 - x \bullet e_1) \end{aligned}$$

$$\begin{aligned} x \bullet u &= x \bullet (x - e_1) = x \bullet x - x \bullet e_1 \\ &= 1 - x \bullet e_1. \end{aligned}$$

If $x \neq e_1$, then $u \neq 0$ and we may apply the Householder transformation Q associated with u to x and e_1:

$$Qx = x - 2\frac{x \bullet u}{\|u\|^2} u = x - \frac{2\,(1 - x \bullet e_1)}{2\,(1 - x \bullet e_1)} u = x - u = x - (x - e_1) = e_1;$$

applying Q to both sides we have

$$Q^2 x = x = Q e_1.$$

If $x = e_1$, set $Q = I$; then in either case

$$x = Q e_1 \text{ and } e_1 = Qx$$

are reflection of each other in the hyperplane $(x - e_1)^{\perp}$. For a unit vector $x \in \mathbb{R}^2$, the line $(x - e_1)^{\perp}$ bisects the angle between x and e_1. We are ready to prove our main theorem in this lecture:

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

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 $v = (v_1, \ldots, v_n)$. Let $x = \dfrac{\overline{v_1} v}{\|\overline{v_1} v\|}$ and set $u = x - e_1$; if $x \neq e_1$, let Q be the Householder matrix associated with u; if $x = e_1$ let $Q = I$. Then $x = Q e_1$ by the discussion above, so x is the first column of Q. By Exercise 6, Q is Hermitian and unitary, so x^* is the first row of Q. Since $Q = Q^{-1} = Q^*$ we have $Q = [x \mid V] = \left[\begin{smallmatrix} x^* \\ V^* \end{smallmatrix}\right]$ and

$$QAQ = QA\,[x \mid V] = Q\,[\lambda x \mid AV] = \left[\lambda e_1 \mid \begin{bmatrix} x^* \\ V^* \end{bmatrix} AV\right] = \begin{bmatrix} \lambda & x^* A V \\ 0 & V^* A V \end{bmatrix}.$$

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

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

is triangular as claimed. ■

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 find a Schur triangularization of the matrix*

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

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

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

and $x = \begin{bmatrix} -1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ *is an associated unit eigenvector. Let* $u = x - e_1 = \begin{bmatrix} -4/3 \\ -2/3 \\ 2/3 \end{bmatrix}$ *and let* Q *be the associated Householder matrix, i.e.,*

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

where

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

Then

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

Now triangularize the 2×2 *matrix* $V^T AV$, *which has the single eigenvalue* -3. *The vector* $x = \frac{1}{\sqrt{17}}\left[\begin{smallmatrix} 1 \\ -4 \end{smallmatrix}\right]$ *is a unit vector associated with* -3. *Let* $u = x - e_1 = \frac{1}{\sqrt{17}}\left[\begin{smallmatrix} 1-\sqrt{17} \\ -4 \end{smallmatrix}\right]$ *and let* R *be the Householder matrix associated with* u, *i.e.,*

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

Then

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

is a Schur triangularization of $V^T AV$. *Finally, let*

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

then

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

is a (numerically approximate) Schur triangularization of A.

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.

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

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

Exercise 11 $A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}$

Exercise 12 $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

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