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 $\lambda$ is one dimensional and is spanned by the single vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, so the geometric multiplicity of $\lambda$ 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 $u^\perp$ 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** $N(P) = \text{span} \{u\}$ and multiplication by $P$ is orthogonal projection on $u^\perp$, i.e., for all $x \in \mathbb{C}^n$,

$$Px = \text{proj}_u x.$$  

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

**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

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

**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 $x \in \mathbb{C}^n$, $Qx = x - 2 \left( \frac{u^* x}{\|u\|^2} u \right) = x - 2 \text{proj}_u x$. Thus $x - Qx = 2 \text{proj}_u x$. ■

**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 $x = \overline{v_1} v$. 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

$$\|u\|^2 = \langle x - e_1, x - e_1 \rangle = \langle x, x \rangle - \langle x, e_1 \rangle - \langle e_1, x \rangle + \langle e_1, e_1 \rangle = 2 - 2x_1 = 2(1 - x_1)$$

$$u^* x = \langle x, u \rangle = \langle x, x - e_1 \rangle = \langle x, x \rangle - \langle x, e_1 \rangle = 1 - x_1.$$
If $x \neq e_1$, then $u \neq 0$ and we may apply the Householder transformation $Q$ associated with $u$ to $x$ and $e_1$:

$$
Qx = x - \frac{2u^*x}{\|u\|^2}u = x - \frac{2(1-x_1)}{2(1-x_1)}u = x - u = x - (x - e_1) = e_1;
$$

(1)

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

$$
x = Q^2x = Qe_1.
$$

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

$$
x = Qe_1 \text{ and } e_1 = Qx
$$

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 $n \times n$ 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 $\lambda$ of $A$ and an associated eigenvector $v = (v_1, \ldots, v_n)$. Let $x = \frac{v}{\|v\|}$ 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 $x = Qe_1$ by the discussion above, so $x$ is the first column of $Q$. By Exercise 6, $Q$ is Hermitian and unitary, so $x^*\perp$ is the first row of $Q$. Since $Q = Q^{-1} = Q^\dagger$, we have $Q = [x \mid V] = [\begin{smallmatrix} x^* \n ab \end{smallmatrix}]$ and

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

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

$$
U = Q \left[ \begin{array}{cc} 1 & 0 \\ 0 & R \end{array} \right];
$$

then

$$
U^*U = \left[ \begin{array}{cc} 1 & 0 \\ 0 & R^* \end{array} \right] Q^\dagger Q \left[ \begin{array}{cc} 1 & 0 \\ 0 & R \end{array} \right] = \left[ \begin{array}{cc} 1 & 0 \\ 0 & R^*R \end{array} \right] = I
$$

so $U$ is unitary. Hence

$$
T = U^*AU = \left[ \begin{array}{cc} 1 & 0 \\ 0 & R^* \end{array} \right] QAQ \left[ \begin{array}{cc} 1 & 0 \\ 0 & R \end{array} \right] = \left[ \begin{array}{cc} 1 & 0 \\ 0 & R^* \\ \lambda & x^*AV \end{array} \right] \left[ \begin{array}{cc} 1 & 0 \\ 0 & V^*AV \end{array} \right] = \left[ \begin{array}{cc} 1 & 0 \\ 0 & R^*AV \end{array} \right] = \left[ \begin{array}{cc} \lambda & x^*AVR \\ 0 & R^*V^*AVR \end{array} \right] = \left[ \begin{array}{cc} \lambda & x^*AVR \\ 0 & T_{n-1} \end{array} \right],
$$

which is upper triangular as desired. \hfill \blacksquare

**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

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

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.,

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

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 \( 2 \times 2 \) matrix \( V^T AV, \) which has the single eigenvalue \(-3.\) The vector \( \mathbf{x} = \frac{1}{\sqrt{17}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \) is a unit vector associated with \(-3.\) Let \( \mathbf{u} = \mathbf{x} - \mathbf{e}_1 = \frac{1}{\sqrt{17}} \begin{bmatrix} 1 - \sqrt{17} \\ -4 \end{bmatrix} \) and let \( R \) be the Householder matrix associated with \( \mathbf{u}, \) i.e.,

\[ R = \begin{bmatrix} 0.24254 & -0.97014 \\ -0.97014 & -0.24254 \end{bmatrix}. \]

Then

\[ R V^T AV R = \begin{bmatrix} -3 & 17/3 \\ 0 & -3 \end{bmatrix} \]

is a Schur triangularization of \( V^T AV. \) Finally, let

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

then

\[ U^T AU = \begin{bmatrix} 1 & 0.97025 & -21.747 \\ 0 & -3.000 & 5.6667 \\ 0 & 0 & -3.000 \end{bmatrix} \]

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 $P^TAP$ 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