Math 223 Symmetric and Hermitian Matrices. Richard Anstee

An \( n \times n \) matrix \( Q \) is orthogonal if \( Q^T = Q^{-1} \). The columns of \( Q \) would form an orthonormal basis for \( \mathbb{R}^n \). The rows would also form an orthonormal basis for \( \mathbb{R}^n \).

A matrix \( A \) is symmetric if \( A^T = A \).

**Theorem 1** Let \( A \) be a symmetric \( n \times n \) matrix of real entries. Then there is an orthogonal matrix \( Q \) and a diagonal matrix \( D \) so that

\[
AQ = QD, \quad \text{i.e.} \quad Q^T AQ = D.
\]

Note that the entries of \( Q \) and \( D \) are real.

There are various consequences to this result:

- A symmetric matrix \( A \) is diagonalizable.
- A symmetric matrix \( A \) has an orthonormal basis of eigenvectors.
- A symmetric matrix \( A \) has real eigenvalues.

We have proven this in a previous set of notes.

Recall that for a complex number \( z = a + bi \), the conjugate \( \overline{z} = a - bi \). We may extend the conjugate to vectors and matrices. We would like some notation for the conjugate transpose. For a vector, define \( v^H = v^T \) (so that \( z^H = \overline{z} \)). Some use the dagger in place of \( H \). When we consider extending inner products to \( \mathbb{C}^n \) we must define \( \langle x, y \rangle = x^H y \) so that \( \langle x, x \rangle = x^H x \geq 0 \). Note that \( \langle y, x \rangle = \langle x, y \rangle \) and so we don’t have commutivity. Thus we have made a choice for the definition of the complex inner product

\[
\langle x, y \rangle = x^H y
\]

so that \( \langle x, x \rangle \in \mathbb{R} \) and \( \langle x, x \rangle \geq 0 \). Note that \( \langle y, x \rangle = \overline{\langle x, y \rangle} \) and so we don’t have commutivity. Thus we have made a choice for the definition of the complex inner product \( \langle x, y \rangle = x^H y \) which we use in what follows. We define \( A^H = (A^*)^T \).

We define two vectors \( x, y \) to be orthogonal if \( x^H y = 0 \). We need to do Gram Schmidt process and so need the projection. Define:

\[
\text{proj}_x y = \frac{x^H y}{x^H x} x
\]

Then

\[
\text{proj}_x y = \frac{x^H y}{x^H x} x \quad \text{so that} \quad \text{proj}_x y \quad \text{and} \quad y - \text{proj}_x y \quad \text{are orthogonal},
\]

namely

\[
(\text{proj}_x y)^H (y - \text{proj}_x y) = (\frac{x^H y}{x^H x} x^H) (y - \frac{x^H y}{x^H x} x) = \frac{y^H x}{x^H x} x^H y - \frac{y^H x}{x^H x} (\frac{x^H y}{x^H x}) x^H x = 0
\]

Using this inner product one can perform Gram Schmidt on complex vectors (but remain careful with the order since in general \( \langle u, v \rangle \neq \langle v, u \rangle \)). You are determining an orthogonal set of vectors \( v_1, v_2, \ldots, v_k \) from \( u_1, u_2, \ldots, u_k \) and so we need \( v_i^H v_j = 0 \) for all pairs \( i \neq j \). We need not worry about order in this setting after computing \( v_i \)'s since if \( v_i^H v_j = 0 \) then \( v_j^H v_i = 0 \). This may not be immediate but you note that \( (v_i^H v_j)^H = v_j^H v_i \) as well as \( 0^H = 0 \) and so if \( v_i^H v_j = 0 \) then \( v_j^H v_i = 0 \).

Our Gram-Schmidt process carries on as before.
\[ \begin{align*}
v_1 &= u_1, \\
v_2 &= u_2 - \text{proj}_{v_1}u_2 \\
v_3 &= u_3 - \text{proj}_{v_1}u_3 - \text{proj}_{v_2}u_3 \\
&\vdots \\
v_k &= u_k - \text{proj}_{v_1}u_k - \text{proj}_{v_2}u_k - \cdots - \text{proj}_{v_{k-1}}u_k.
\end{align*} \]

A matrix \( A \) is \textit{hermitian} if \( \overline{A}^T = A \). For example any symmetric matrix of real entries is also hermitian. The follow matrix is hermitian:

\[
\begin{bmatrix}
3 & 1 - 2i \\
1 + 2i & 4
\end{bmatrix}
\]

Sensibly, Hermitian matrices are allowed to have complex entries. One has interesting identities such as \( \langle x, Ay \rangle = \langle Ax, y \rangle \) when \( A \) is hermitian. The following Theorem is essentially a generalization of the result for symmetric matrices. Note that a Unitary matrix \( U \) is an orthogonal matrix if the entries of \( U \) are real.

**Theorem** Let \( A \) be a hermitian matrix. Then there is a unitary matrix \( U \) with entries in \( \mathbb{C} \) and a diagonal matrix \( D \) of real entries so that

\[ AU = UD, \quad A = UDU^{-1} \]

**Proof**: We follow the proof of the theorem for symmetric matrices. The proof begins with an appeal to the fundamental theorem of algebra applied to \( \det(A - \lambda I) \) which asserts that the polynomial factors into linear factors and one of which yields an eigenvalue \( \lambda \) which may not be real.

Our second step it to show \( \lambda \) is real. Let \( x \) be an eigenvector for \( \lambda \) so that \( Ax = \lambda x \). Again, if \( \lambda \) is not real we must allow for the possibility that \( x \) is not a real vector.

Now \( x^Hx \geq 0 \) with \( x^Hx = 0 \) if and only if \( x = 0 \). We compute \( x^HAx = x^H(\lambda x) = \lambda x^Hx \). Now taking complex conjugates and transpose \((x^H Ax)^H = x^H A^H x \) using that \((x^H)^H = x \). Then \((x^H Ax)^H = x^H A x = \lambda x^H x \) using \( A^H = A \). It is important to use our hypothesis that \( A \) is Hermitian. But also \((x^H Ax)^H = \lambda x^H x = \lambda x^H x \) (using \( x^H x \in \mathbb{R} \)). Knowing that \( x^H x > 0 \) (since \( x \neq 0 \)) we deduce that \( \lambda = \overline{\lambda} \) and so we deduce that \( \lambda \in \mathbb{R} \).

The rest of the proof uses induction on \( n \). The result is easy for \( n = 1 \) \((U = [1])\). Note that an orthogonal matrix is unitary. Assume we have a real eigenvalue \( \lambda_1 \) and an eigenvector \( x_1 \) (not necessarily real) with \( Ax_1 = \lambda_1 x_1 \) and \( ||x_1|| = 1 \). We can extend \( x_1 \) to an orthonormal basis \( \{x_1, x_2, \ldots, x_n\} \) using Gram Schmidt applied as described above so that \( x_i^H x_j = 0 \) for all pairs \( i \neq j \). Let \( M = [x_1 \ x_2 \ \cdots \ x_n] \) be the unitary matrix formed with columns \( x_1, x_2, \ldots, x_n \). Then

\[ AM = M \begin{bmatrix} \lambda_1 & B \\ 0 & C \end{bmatrix} \quad \text{or} \quad M^{-1}AM = \begin{bmatrix} \lambda_1 & B \\ 0 & C \end{bmatrix}. \]

which is the sort of result from our assignments. But the matrix on the right is hermitian since it is equal to \( M^{-1}AM = M^H AM \) (since the basis was orthonormal) and we note \((M^H AM)^H = M^H AM \) (using \( A^H = A \) since \( A \) is hermitian). Then \( B \) is a \( 1 \times (n-1) \) zero matrix and \( C \) is a hermitian \((n-1) \times (n-1) \) matrix.
By induction there exists a unitary \((n-1) \times (n-1)\) matrix \(N\) (with \(N^H = N^{-1}\)) and a diagonal \((n-1) \times (n-1)\) matrix \(E\) with \(N^{-1}CN = E\). We form a new unitary matrix

\[
P = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & N \\
\end{bmatrix}
\]

which is seen to be unitary since

\[
P^H = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & N^H \\
\end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & N^{-1} \\
\end{bmatrix} = P^{-1}.
\]

We obtain

\[
P^{-1} \begin{bmatrix} \lambda_1 & 0^T \\ 0 & C \end{bmatrix} P = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & E \\
\end{bmatrix}
\]

This becomes

\[
P^{-1}M^{-1}AMP = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & E \\
\end{bmatrix}
\]

which is a \(n \times n\) diagonal matrix \(D\). We note that \((MP)^H = P^HM^H = P^{-1}M^{-1}\) and so \(U = MP\) is an Unitary matrix with \(U^HAU = D\). This proves the result by induction. ■

As an example let

\[
A = \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}
\]

We compute

\[
\det(A - \lambda I) = \begin{bmatrix} 1 - \lambda & i \\ -i & 1 - \lambda \end{bmatrix} = \lambda^2 - 2\lambda
\]

and thus the eigenvalues are 0, 2 (Note that they are real which is a consequence of the theorem). We find that the eigenvectors are

\[
\lambda_1 = 2 \quad v_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}, \quad \lambda_2 = 0 \quad v_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}
\]

Not surprisingly \(<v_1, v_2> = v_1^Hv_2 = 0\), another consequence of the theorem. We would have to make them of unit length to obtain an orthonormal basis:

\[
U = \begin{bmatrix} \frac{1}{\sqrt{2}}i & -\frac{1}{\sqrt{2}}i \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \quad AU = UD
\]

Note that \(U^HU = I\) and so \(U^H = U^{-1}\). Such matrices are called unitary.

The following matrix has orthogonal columns:

\[
\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}
\]

since

\[
\begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} 1 \\ -i \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ -i \end{bmatrix}^T \begin{bmatrix} 1 & 1 \\ -i & -i \end{bmatrix} = 0
\]

thus \(\begin{bmatrix} 1 \\ i \end{bmatrix}^H \begin{bmatrix} 1 \\ -i \end{bmatrix} = 0\). To make this unitary we need to normalize the vectors:

\[
\begin{bmatrix} \frac{1}{\sqrt{2}}i & \frac{1}{\sqrt{2}}i \\ \frac{1}{2}i & -\frac{1}{2}i \end{bmatrix}
\]
Here is an example of Gram Schmidt obtaining a unitary matrix but using more ‘complicated’ vectors.

\[ u_1 = \begin{bmatrix} 2 \\ 1 + i \end{bmatrix}, \quad u_2 = \begin{bmatrix} i \\ 1 + i \end{bmatrix}, \quad < u_1, u_2 > = u_1^H u_2 = [2 \ 1 - i] \begin{bmatrix} i \\ 1 + i \end{bmatrix} = 2 + 2i \neq 0. \]

\[ v_1 = u_1 \]
\[ v_2 = u_2 - \text{proj}_{v_1} u_2 = u_2 - \frac{v_1^H u_2}{v_1^H v_1} v_1 = \begin{bmatrix} i \\ 1 + i \end{bmatrix} - \frac{2 + 2i}{6} \begin{bmatrix} 2 \\ 1 + i \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} + \frac{1}{3}i \\ 1 + \frac{1}{3}i \end{bmatrix} \]

You may check

\[ < u_2, u_1 > = u_1^H u_2 = [2 \ 1 - i] \begin{bmatrix} -\frac{2}{3} + \frac{1}{3}i \\ 1 + \frac{1}{3}i \end{bmatrix} = -\frac{4}{3} + \frac{2}{3}i + \frac{4}{3} - \frac{2}{3}i = 0. \]

Obtaining this was a mess for me keeping track of the terms. I will not test you on such a computation. To form a unitary matrix we must normalize the vectors.

\[ \begin{bmatrix} 2 \\ 1 + i \end{bmatrix} \rightarrow \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}}i \end{bmatrix}, \quad \begin{bmatrix} -\frac{2}{3} + \frac{1}{3}i \\ 1 + \frac{1}{3}i \end{bmatrix} \rightarrow \begin{bmatrix} -2 + i \\ 3 + i \end{bmatrix} \rightarrow \begin{bmatrix} -\frac{2}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \\ \frac{2}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \end{bmatrix} \]

\[ U = \begin{bmatrix} \frac{2}{\sqrt{6}} + \frac{1}{\sqrt{6}}i \\ \frac{2}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \end{bmatrix} \]

where we can check \( U^T U = I \). Best to let a computer do these calculations!