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 $\mathbf{R}^n$. The rows would also form an orthonormal basis for $\mathbf{R}^n$.

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*

$$AQ = QD, \qquad i.e.\ Q^T AQ = D.$$

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 othonormal 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 $\mathbf{v}^H = \overline{\mathbf{v}}^T$ (so that $z^H = \overline{z}$). Some use the dagger in place of H. When we consider extending inner products to $\mathbf{C}^n$ we must define

$$< \mathbf{x}, \mathbf{y} >= \mathbf{x}^H \mathbf{y}$$

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

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

$$\text{proj}_{\mathbf{x}} \mathbf{y} = \frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}} \mathbf{x}$$

Then

$$\text{proj}_{\mathbf{x}} \mathbf{y} = \frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}} \mathbf{x} \text{ so that } \text{proj}_{\mathbf{x}} \mathbf{y} \text{ and } \mathbf{y} - \text{proj}_{\mathbf{x}} \mathbf{y} \text{ are orthogonal,}$$

namely

$$(\text{proj}_{\mathbf{x}} \mathbf{y})^H (\mathbf{y} - \text{proj}_{\mathbf{x}} \mathbf{y}) = (\left(\frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}}\right)^H \mathbf{x}^H)(\mathbf{y} - \frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}} \mathbf{x})$$

$$= (\frac{\mathbf{y}^H \mathbf{x}}{\mathbf{x}^H \mathbf{x}} \mathbf{x}^H)(\mathbf{y} - \frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}} \mathbf{x}) = \frac{\mathbf{y}^H \mathbf{x}}{\mathbf{x}^H \mathbf{x}} \mathbf{x}^H \mathbf{y} - (\frac{\mathbf{y}^H \mathbf{x}}{\mathbf{x}^H \mathbf{x}})(\frac{\mathbf{x}^H \mathbf{y}}{\mathbf{x}^H \mathbf{x}}) \mathbf{x}^H \mathbf{x} = 0$$

Using this inner product one can perform Gram Schmidt on complex vectors (but remain careful with the order since in general $< \mathbf{u}, \mathbf{v} >\neq< \mathbf{v}, \mathbf{u} >$. You are determining an orthogonal set of vectors $\mathbf{v}_1, \mathbf{v}_2 \ldots, \mathbf{v}_k$ from $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$ and so we need $\mathbf{v}_i^H \mathbf{v}_j = 0$ for all pairs $i \neq j$. We need not worry about order in this setting after computing $\mathbf{v}_i$'s since if $\mathbf{v}_i^H \mathbf{v}_j = 0$ then $\mathbf{v}_j^H \mathbf{v}_i = 0$. This may not be immediate but you note that $(\mathbf{v}_i^H \mathbf{v}_j)^H = \mathbf{v}_j^H \mathbf{v}_i$ as well as $0^H = 0$ and so if $\mathbf{v}_i^H \mathbf{v}_j = 0$ then $\mathbf{v}_j^H \mathbf{v}_i = 0$.

Our Gram-Schmidt process carries on as before.

$$\begin{array}{llll} \mathbf{v}_1 = \mathbf{u}_1. & & & \\ \mathbf{v}_2 = \mathbf{u}_2 & -\mathrm{proj}_{\mathbf{v}_1}\mathbf{u}_2 & & \\ \mathbf{v}_3 = \mathbf{u}_3 & -\mathrm{proj}_{\mathbf{v}_1}\mathbf{u}_3 & -\mathrm{proj}_{\mathbf{v}_2}\mathbf{u}_3 & \\ & \vdots & & \\ \mathbf{v}_k = \mathbf{u}_k & -\mathrm{proj}_{\mathbf{v}_1}\mathbf{u}_k & -\mathrm{proj}_{\mathbf{v}_2}\mathbf{u}_k \quad \cdots & -\mathrm{proj}_{\mathbf{v}_{k-1}}\mathbf{u}_k \end{array}$$

A matrix A is *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 $<\mathbf{x}, A\mathbf{y}>=<A\mathbf{x}, \mathbf{y}>$ 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 $\mathbf{C}$ and a diagonal matrix D of real entries so that

$$AU = UD, \qquad A = UDU^{-1}$$

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 λ *which may not be real.*

Our second step it to show λ is real. Let $\mathbf{x}$ be an eigenvector for λ so that $A\mathbf{x} = \lambda\mathbf{x}$. Again, if λ is not real we must allow for the possibility that $\mathbf{x}$ is not a real vector.

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

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 λ_1 and an eigenvector $\mathbf{x}_1$ (not necessarily real) with $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$ and $||\mathbf{x}_1|| = 1$. We can extend $\mathbf{x}_1$ to an orthonormal basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ using Gram Schmidt applied as described above so that $\mathbf{x}_i^H\mathbf{x}_j = 0$ for all pairs $i \neq j$. Let $M = [\mathbf{x}_1\, \mathbf{x}_2\, \cdots \mathbf{x}_n]$ be the unitary matrix formed with columns $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. Then

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

which is the sort of result from our assignments. But the matrix on the right is hemitian since it is equal to $M^{-1}AM = M^HAM$ (since the basis was orthonormal) and we note $(M^HAM)^H = M^HAM$ (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

$$P = \begin{bmatrix} 1 & 0\,0\cdots 0 \\ \mathbf{0} & N \end{bmatrix}$$

which is seen to be unitary since

$$P^H = \begin{bmatrix} 1 & 0\,0\cdots 0 \\ \mathbf{0} & N^H \end{bmatrix} = \begin{bmatrix} 1 & 0\,0\cdots 0 \\ \mathbf{0} & N^{-1} \end{bmatrix} = P^{-1}.$$

We obtain

$$P^{-1}\begin{bmatrix} \lambda_1 & \mathbf{0}^T \\ \mathbf{0} & C \end{bmatrix} P = \begin{bmatrix} \lambda_1 & 0\,0\cdots 0 \\ \mathbf{0} & E \end{bmatrix}$$

This becomes

$$P^{-1}M^{-1}AMP = \begin{bmatrix} \lambda_1 & 0\,0\cdots 0 \\ \mathbf{0} & E \end{bmatrix}$$

which is a $n\times n$ diagonal matrix D. We note that $(MP)^H = P^H M^H = P^{-1}M^{-1}$ and so $U = MP$ is an Unitary matrix with $U^H AU = D$. 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

$$\lambda_1 = 2 \quad \mathbf{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}, \quad \lambda_2 = 0 \quad \mathbf{v}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

Not surprisingly $<\mathbf{v}_1, \mathbf{v}_2> = \mathbf{v}_1^H \mathbf{v}_2 = 0$, another consequence of the theorem. We would have to make them of unit length to obtain an orthonormal basis:

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

Note that $U^H U = I$ 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 $\overline{\begin{bmatrix} 1 \\ i \end{bmatrix}} = \begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -i \end{bmatrix}^T \begin{bmatrix} 1 \\ -i \end{bmatrix} = 0$ 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:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2}i & -\frac{1}{2}i \end{bmatrix}$$

Here is an example of Gram Schmidt obtaining a unitary matrix but using more 'complicated' vectors.

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1+i \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} i \\ 1+i \end{bmatrix}, \quad < \mathbf{u}_1, \mathbf{u}_2 >= \mathbf{u}_1^H \mathbf{u}_2 = [2 \;\; 1-i] \begin{bmatrix} i \\ 1+i \end{bmatrix} = 2 + 2i \neq 0.$$

$$\mathbf{v}_1 = \mathbf{u}_1$$

$$\mathbf{v}_2 = \mathbf{u}_2 - \text{proj}_{v_1} \mathbf{u}_2 = \mathbf{u}_2 - \frac{\mathbf{v}_1^H \mathbf{u}_2}{\mathbf{v}_1^H \mathbf{v}_1} \mathbf{v}_1 = \begin{bmatrix} i \\ 1+i \end{bmatrix} - \frac{2+2i}{6} \begin{bmatrix} 2 \\ 1+i \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} + \frac{1}{3}i \\ 1 + \frac{1}{3}i \end{bmatrix}$$

You may check

$$< \mathbf{u}_2, \mathbf{u}_1 >= \mathbf{u}_1^H \mathbf{u}_2 = [2 \;\; 1-i] \begin{bmatrix} -\frac{2}{3} + \frac{1}{3}i \\ 1 + \frac{1}{3}i \end{bmatrix} = -\frac{4}{3} + \frac{2}{3}i + \frac{4}{3} - \frac{2}{3}i = 0.$$

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.

$$\begin{bmatrix} 2 \\ 1+i \end{bmatrix} \to \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}}i \end{bmatrix}, \quad \begin{bmatrix} -\frac{2}{3} + \frac{1}{3}i \\ 1 + \frac{1}{3}i \end{bmatrix} \to \begin{bmatrix} -2 + i \\ 3 + i \end{bmatrix} \to \begin{bmatrix} -\frac{2}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \\ \frac{3}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \end{bmatrix}$$

$$U = \begin{bmatrix} \frac{2}{\sqrt{6}} & -\frac{2}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \\ \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}}i & \frac{3}{\sqrt{15}} + \frac{1}{\sqrt{15}}i \end{bmatrix}$$

where we can check $\overline{U}^T U = I$. Best to let a computer do these calculations!