Lecture 3x

Hermitian Matrices

(pages 432-435)

We end the course by looking at the complex equivalent of symmetric matrices. As usual, we add complex conjugation to the definition from the reals to get the complex equivalent.

Definition: An $n \times n$ matrix A with complex entries is called **Hermitian** if $A^* = A$ or, equivalently, if $\overline{A} = A^T$.

Example: Let $A = \begin{bmatrix} 1 & 2 & -i \\ 2 & -1 & 1+i \\ i & 1-i & 1 \end{bmatrix}$. Then $A^* = A$, so A is Hermitian.

Let $B = \begin{bmatrix} 3 & 1+i & 2-i \\ 1+i & 2i & 4+2i \\ 2-i & 4+2i & 5 \end{bmatrix}$. Then B is symmetric, but B is not Hermitian, since $B^* = \begin{bmatrix} 3 & 1-i & 2+i \\ 1-i & -2i & 4-2i \\ 2+i & 4-2i & 5 \end{bmatrix} \neq B$.

One thing worth noting about Hermitian matrices is that they must have real numbers on the main diagonal, as these entries must equal their own conjugate. So, we can immediately see that B is not Hermitian by noting the complex entry $2i$ on the main diagonal.

As we did with the symmetric matrices in the reals, we will find not only that every Hermitian matrix is diagonalizable, but it can be diagonalized by a unitary matrix (just as symmetric matrices could be diagonalized by orthogonal matrices). It will simply take us few steps to prove it. First, we want to notice the following:

Theorem 9.6.1: An $n \times n$ matrix A is Hermitian if and only if for all $\vec{z}, \vec{w} \in \mathbb{C}^n$, we have

$$\langle \vec{z}, A\vec{w} \rangle = \langle A\vec{z}, \vec{w} \rangle$$

Proof of Theorem 9.6.1: First, let's assume that A is Hermitian, and show that $\langle \vec{z}, A\vec{w} \rangle = \langle A\vec{z}, \vec{w} \rangle$ for all $\vec{z}, \vec{w} \in \mathbb{C}^n$. This follows easily from Theorem 9.5.2 (1), which says that for any matrix A, $\langle \vec{z}, A^*\vec{w} \rangle = \langle A\vec{z}, \vec{w} \rangle$ for all $\vec{z}, \vec{w} \in \mathbb{C}^n$. And since A is Hermitian, we know that $A = A^*$, and we have our result.

Now let's assume that $\langle \vec{z}, A\vec{w} \rangle = \langle A\vec{z}, \vec{w} \rangle$ for all $\vec{z}, \vec{w} \in \mathbb{C}^n$, and show that A is Hermitian. We do this by plugging in $\vec{z} = \vec{e}_j$ and $\vec{w} = \vec{e}_k$ to this equality.

Why those vectors? Because these standard basis vectors will cause these inner products to single out the jk-th entry of A:

$$\begin{aligned}
\langle A\vec{e}_j, \vec{e}_k\rangle &= (A\vec{e}_j)\cdot\overline{\vec{e}_k} \\
&= (A\vec{e}_j)\cdot\vec{e}_k && \text{(since } \vec{e}_k \text{ has only real components)} \\
&= \vec{a}_j\cdot\vec{e}_k && \text{(where } \vec{a}_j \text{ is the } j\text{-th column of } A) \\
&= (\vec{a}_j)_k && \text{(the } k\text{-th component of } \vec{a}_j) \\
&= a_{jk} && \text{(the } jk\text{-th entry of } A)
\end{aligned}$$

And now, let's look at $\langle\vec{e}_j, A\vec{e}_k\rangle$:

$$\begin{aligned}
\langle\vec{e}_j, A\vec{e}_k\rangle &= \vec{e}_j\cdot\overline{A\vec{e}_k} \\
&= \vec{e}_j\cdot\overline{A}\ \overline{\vec{e}_k} \\
&= \vec{e}_j\cdot\overline{A}\vec{e}_k && \text{(again, since } \vec{e}_k \text{ has only real entries)} \\
&= \vec{e}_j\cdot\overline{\vec{a}_k} && \text{the conjugate of the } k\text{-th column of } A \\
&= (\overline{\vec{a}_k})_j && \text{(the } j\text{-the component of } \overline{\vec{a}_k}) \\
&= \overline{a_{kj}} && \text{(the } kj\text{-th entry of } \overline{A})
\end{aligned}.$$

And since we have that $\langle A\vec{e}_j, \vec{e}_k\rangle = \langle\vec{e}_j, A\vec{e}_k\rangle$, we have that $a_{jk} = \overline{a_{kj}}$. But $\overline{a_{kj}}$ is the jk-th entry of A^*. So we see that the jk-th entry of A is the same as the jk-th entry of A^*, for all $1 \le j, k \le n$, and thus we have shown that $A = A^*$, which means that A is Hermitian.

We use this theorem to prove the following key facts, which you should recognize form our work with symmetric matrices.

Theorem 9.6.2: Suppose that A is an $n \times n$ Hermitian matrix. Then
(1) All eigenvalues of A are real
(2) Eigenvectors corresponding to distinct eigenvalues are orthogonal to each other.

Proof of Theorem 9.6.2: To prove part (1), let λ be an eigenvalue of A with corresponding eigenvector $\vec{z}$, so that $A\vec{z} = \lambda\vec{z}$. Then by Theorem 9.6.1, $\langle A\vec{z}, \vec{z}\rangle = \langle\vec{z}, A\vec{z}\rangle$. But $\langle A\vec{z}, \vec{z}\rangle = \langle\lambda\vec{z}, \vec{z}\rangle = \lambda\langle\vec{z}, \vec{z}\rangle$, while $\langle\vec{z}, A\vec{z}\rangle = \langle\vec{z}, \lambda\vec{z}\rangle = \overline{\lambda}\langle\vec{z}, \vec{z}\rangle$. So we see that $\lambda\langle\vec{z}, \vec{z}\rangle = \overline{\lambda}\langle\vec{z}, \vec{z}\rangle$. Since $\vec{z}$ is an eigenvector, we know that $\vec{z} \ne \vec{0}$, so $\langle\vec{z}, \vec{z}\rangle \ne 0$. So we can divide both sides by $\langle\vec{z}, \vec{z}\rangle$ to get that $\lambda = \overline{\lambda}$, which means that λ must be real.

To prove part (2), let λ_1 and λ_2 be eigenvalues of A such that $\lambda_1 \ne \lambda_2$, and let $\vec{z}_1$ and $\vec{z}_2$ be eigenvectors corresponding to λ_1 and λ_2. Since A is Hermitian, Theorem 9.6.1 tells us that

$$\langle A\vec{z}_1, \vec{z}_2\rangle = \langle\vec{z}_1, A\vec{z}_2\rangle$$

This means that

$$\langle\lambda_1\vec{z}_1, \vec{z}_2\rangle = \langle\vec{z}_1, \lambda_2\vec{z}_2\rangle$$

and from this we get that

$$\lambda_1\langle \vec{z}_1, \vec{z}_2\rangle = \overline{\lambda_2}\langle \vec{z}_1, \vec{z}_2\rangle$$

In part (1) we showed that the eigenvalues of A are real, so $\overline{\lambda_2} = \lambda_2$, which means that

$$\lambda_1\langle \vec{z}_1, \vec{z}_2\rangle = \lambda_2\langle \vec{z}_1, \vec{z}_2\rangle$$

For this equality to hold, we either have that $\langle \vec{z}_1, \vec{z}_2\rangle = 0$ or $\lambda_1 = \lambda_2$. Since we know that $\lambda_1 \neq \lambda_2$, we get that $\langle \vec{z}_1, \vec{z}_2\rangle = 0$, which means that our eigenvectors are orthogonal.

And now we can prove the main result:

Theorem 9.6.3 (Spectral Theorem for Hermitian Matrices): Suppose that A is an $n \times n$ Hermitian matrix. Then there exists a unitary matrix U and a diagonal matrix D such that $U^*AU = D$.

The proof of this works exactly the same as its counterpart in the reals (The Principal Axis Theorem), so I won't bother repeating it here. As with symmetric matrices, the main idea is that we can find an orthonormal basis for the eigenspaces each eigenvalue, and use these eigenvectors (instead of any random eigenvector) to form the matrix U we use to diagonalize A, and this U will in fact be unitary.

Example: Let $A = \begin{bmatrix} 6 & \sqrt{3}-i \\ \sqrt{3}+i & 3 \end{bmatrix}$. Then $A = A^*$, so A is Hermitian. Let's find the eigenvalues of A:

$$\begin{aligned} \det\,(A - \lambda I) &= \det \begin{bmatrix} 6-\lambda & \sqrt{3}-i \\ \sqrt{3}+i & 3-\lambda \end{bmatrix} \\ &= (6-\lambda)(3-\lambda) - (3 - \sqrt{3}i + \sqrt{3}i - i^2) \\ &= 18 - 6\lambda - 3\lambda + \lambda^2 - 4 \\ &= 14 - 9\lambda + \lambda^2 \\ &= (7-\lambda)(2-\lambda) \end{aligned}$$

So, the eigenvalues of A are $\lambda = 2, 7$. Now let's find the eigenspaces for these eigenvectors:

The eigenspace for $\lambda = 2$ is the nullspace of $\begin{bmatrix} 6-\lambda & \sqrt{3}-i \\ \sqrt{3}+i & 3-\lambda \end{bmatrix} = \begin{bmatrix} 4 & \sqrt{3}-i \\ \sqrt{3}+i & 1 \end{bmatrix}$, which is row equivalent to $\begin{bmatrix} 4 & \sqrt{3}-i \\ 0 & 0 \end{bmatrix}$. So, the eigenvectors for $\lambda = 2$ satisfy $4z_1 + (\sqrt{3}-i)z_2 = 0$, so we have that $z_1 = (-1/4)(\sqrt{3}-i)z_2$. If we replace the variable z_2 with the parameter α, we see that the eigenvectors of $\lambda = 2$ are $\left\{ \begin{bmatrix} (-1/4)(\sqrt{3}-i)\alpha \\ \alpha \end{bmatrix} \mid \alpha \in \mathbb{C} \right\}$, which we can write as $\text{Span}\left\{ \begin{bmatrix} (\sqrt{3}-i)/\sqrt{20} \\ -4/\sqrt{20} \end{bmatrix} \right\}$. And so have that $\left\{ \begin{bmatrix} (\sqrt{3}-i)/\sqrt{20} \\ -4/\sqrt{20} \end{bmatrix} \right\}$ is an orthonormal basis for the eigenspace for $\lambda = 2$.

The eigenspace for $\lambda = 7$ is the nullspace of $\begin{bmatrix} 6-\lambda & \sqrt{3}-i \\ \sqrt{3}+i & 3-\lambda \end{bmatrix} = \begin{bmatrix} -1 & \sqrt{3}-i \\ \sqrt{3}+i & -4 \end{bmatrix}$, which is row equivalent to $\begin{bmatrix} -1 & \sqrt{3}-i \\ 0 & 0 \end{bmatrix}$. So, the eigenvectors for $\lambda = 7$ satisfy $-z_1 + (\sqrt{3}-i)z_2 = 0$, so we have that $z_1 = (\sqrt{3}-i)z_2$. If we replace the variable z_2 with the parameter α, we see that the eigenvectors of $\lambda = 7$ are $\left\{ \begin{bmatrix} (\sqrt{3}-i)\alpha \\ \alpha \end{bmatrix} \mid \alpha \in \mathbb{C} \right\}$, which we can write as $\text{Span}\left\{ \begin{bmatrix} (\sqrt{3}-i)/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \right\}$. And so have that $\left\{ \begin{bmatrix} (\sqrt{3}-i)/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \right\}$ is an orthonormal basis for the eigenspace for $\lambda = 7$.

And so we have that the matrix $U = \begin{bmatrix} (\sqrt{3}-i)/\sqrt{20} & (\sqrt{3}-i)/\sqrt{5} \\ -4/\sqrt{20} & 1/\sqrt{5} \end{bmatrix}$ is a unitary matrix such that $U^*AU = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 7 \end{bmatrix}$.