# 28. Lecture 28: 8.1 Diagonalizing Symmetric Matrices

As we have seen it is natural to consider complex vectors. Recall that for complex numbers $z = a + ib$ we have $|z|^2 = a^2 + b^2 = \overline{z}z$, where the complex conjugate is $\overline{z} = a - ib$. The length of a vector with complex components is $\|\mathbf{z}\|^2 = |z_1|^2 + \cdots + |z_n|^2 = \overline{z}_1 z_1 + \cdots + \overline{z}_n z_n = \overline{\mathbf{z}}^T\mathbf{z}$. We define the complex inner product by

$$\langle \mathbf{w}, \mathbf{z} \rangle = \mathbf{z}^H \mathbf{w} = \overline{z}_1 w_1 + \cdots + \overline{z}_n w_n, \qquad \text{where} \quad \mathbf{z}^H = \overline{\mathbf{z}}^T.$$

If $\mathbf{z}$ and $\mathbf{w}$ are real then this reduces to the dot product of vectors in $\mathbf{R}^n$. Just like for the real inner product it is linear in the first argument but the symmetry is replaced by

$$\langle \mathbf{w}, \mathbf{z} \rangle = \overline{\langle \mathbf{z}, \mathbf{w} \rangle}. \tag{28.1}$$

We want to take the transpose and conjugate of matrices so we introduce a notation:

$$A^H = \overline{A}^T.$$

(Also denoted by $A^*$.) As for the transpose $(AB)^H = B^H A^H$ and $(A^H)^H = A$. It follows that

$$\langle A\mathbf{w}, \mathbf{z} \rangle = \mathbf{z}^H A \mathbf{w} = (A^H \mathbf{z})^H \mathbf{w} = \langle \mathbf{w}, A^H \mathbf{z} \rangle. \tag{28.2}$$

There is an analog of symmetric matrices called **Hermitian** matrices

$$A^H = A. \tag{28.3}$$

**Lem** $\langle A\mathbf{z}, \mathbf{z} \rangle$ is real also for complex $\mathbf{z}$ if $A^H = A$.
**Pf** By (28.2), (28.3) and (28.1) $\langle A\mathbf{z}, \mathbf{z} \rangle = \langle A\mathbf{z}, A^H\mathbf{z} \rangle = \langle \mathbf{z}, A\mathbf{z} \rangle = \overline{\langle A\mathbf{z}, \mathbf{z} \rangle}$ so it must be real.
**Th** Eigenvalues of $A$ are real if $A^H = A$, in particular if $A$ is real and symmetric.
**Pf** By the (proof of the) previous lemma $\langle A\mathbf{z}, \mathbf{z} \rangle = \overline{\langle A\mathbf{z}, \mathbf{z} \rangle}$. Applying this to an eigenvector $A\mathbf{z} = \lambda\mathbf{z}$ and using the linearity in the first argument and the fact that $\langle \mathbf{z}, \mathbf{z} \rangle$ is real gives $\lambda\langle \mathbf{z}, \mathbf{z} \rangle = \overline{\lambda}\langle \mathbf{z}, \mathbf{z} \rangle$. It follows that $\lambda = \overline{\lambda}$ so $\lambda$ must be real.

**Rem** The sole purpose of the complex numbers was to prove that the eigenvalues of real symmetric matrices are real. Since the conclusion has nothing to do with complex numbers one may ask if there could be a proof that doesn't use them and this is indeed the case.

**Lem** Eigenvectors of a real symmetric matrix with different eigenvalues are orthogonal.
**Pf** If $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$ and $A\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$ then since $A^T = A$

$$\lambda_1 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle = \langle A\mathbf{x}_1, \mathbf{x}_2 \rangle = \langle \mathbf{x}_1, A\mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1, \mathbf{x}_2 \rangle.$$

It therefore follows that $\langle \mathbf{x}_1, \mathbf{x} \rangle_2 = 0$ if $\lambda_1 \neq \lambda_2$.
**Spectral Theorem** If $A$ is real symmetric with different eigenvalues then it can be factorized $A = QDQ^T$, where $Q$ is orthogonal and $D$ is diagonal.
**Proof** We previously showed that we can write $A = SDS^{-1}$, where the columns of $S$ are the eigenvectors. Since the eigenvectors are orthogonal we can normalize them so they are orthonormal and then $Q = S$ is an orthogonal matrix and $Q^T = S^{-1}$.
**Remark** The theorem holds even if the eigenvalues are repeated as we will show.
**Remark** If $A = QDQ^T$, then $A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$.

**Spectral Theorem** Suppose that $A$ is real and symmetric. Then there is an orthogonal matrix $U$ such that $U^{-1}AU = D$ is diagonal.

**Proof** We will just do the $4 \times 4$ case since it is clear from that how to do it in general. Every symmetric matrix has at least one real eigenvalue $\lambda_1$ and for this eigenvalue we pick an orthonormal eigenvector $\mathbf{u}_1$. Then we use Gram Schmidt to pick three other vectors so $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ form an orthonormal set and we set $U_1 = \left[\mathbf{u}_1\mathbf{u}_2\mathbf{u}_3\mathbf{u}_4\right]$. Then $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ and $A\mathbf{u}_i = \sum_{j=1}^{4} c_{ij}\mathbf{u}_j$ so

$$AU_1 = \left[A\mathbf{u}_1 A\mathbf{u}_2 A\mathbf{u}_3 A\mathbf{u}_4\right] = \left[\mathbf{u}_1\mathbf{u}_2\mathbf{u}_3\mathbf{u}_4\right] \begin{bmatrix} \lambda_1 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{bmatrix}.$$

Since $U_1$ is orthogonal $U_1^{-1} = U_1^T$ and since $A$ is symmetric it follows that $U_1^{-1}AU_1$ is symmetric and hence

$$U_1^{-1}AU_1 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & A_2 & \\ 0 & & & \end{bmatrix},$$

where $A_2$ is symmetric as well. Hence, the $3 \times 3$ matrix $A_2$ has a real eigenvalue $\lambda_2$ and we can form $M_2$ such that

$$M_2^{-1}A_2M_2 = \begin{bmatrix} \lambda_2 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

Hence if $U_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2 & \\ 0 & & & \end{bmatrix}$ then $U_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2^{-1} & \\ 0 & & & \end{bmatrix}$ and

$$U_2^{-1}U_1^{-1}AU_1U_2 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & M_2^{-1}A_2M_2 & \\ 0 & & & \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{bmatrix}.$$

Continuing in this way we get

$$U_4^{-1}U_3^{-1}U_2^{-1}U_1^{-1}AU_1U_2U_3U_4 = \begin{bmatrix} \lambda_1 & & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix} = D.$$

and the lemma follows with $U = U_1U_2U_3U_4$.

**Ex** Diagonalize $A = \begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$ with an orthogonal transformation.

**Sol** $A$ is symmetric so it can be diagonalized by an orthogonal transformation.
The eigenvalues are $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 5$. The eigenspace corresponding to eigenvalue $-1$; $(A+I)\mathbf{x}=\mathbf{0}$ satisfy $x_1 + 2x_2 - x_3 = 0$ so $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the eigenspace corresponding to $\lambda = -1$. We can apply the Gram-Schmidt process to obtain an orthonormal basis. Let

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{p}_1 = \mathbf{v}_2 \cdot \mathbf{u}_1 \, \mathbf{u}_1 = -\sqrt{2}\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{\mathbf{v}_2 - \mathbf{p}_1}{\|\mathbf{v}_2 - \mathbf{p}_1\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

The eigenspace corresponding to $\lambda_3 = 5$ is spanned by $\mathbf{v}_3 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$ and we set

$$\mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

Hence $A = UDU^T$ where

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{3} & -1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix}, \qquad D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

## Summary

We define the complex inner product by

$$\langle \mathbf{w}, \mathbf{z} \rangle = \mathbf{z}^H \mathbf{w} = \overline{z}_1 w_1 + \cdots + \overline{z}_n w_n, \qquad \text{where} \quad \mathbf{z}^H = \overline{\mathbf{z}}^T.$$

With this notation $\mathbf{z}$ and $\mathbf{w}$ are called **orthogonal** if

$$\mathbf{z}^H \mathbf{w} = 0.$$

If $\mathbf{z}$ and $\mathbf{w}$ are real then this reduces to orthogonality in the sense of vectors in $\mathbf{R}^n$.
We want to take the transpose and conjugate of matrices so we introduce a notation:

$$A^H = \overline{A}^T$$

**Th** $\mathbf{z}^H A\mathbf{z}$ is real also for complex $\mathbf{z}$ if $A^H = A$.
**Pf** $(\mathbf{z}^H A\mathbf{z})^H = \mathbf{z}^H A^H (\mathbf{z}^H)^H = \mathbf{z}^H A\mathbf{z}$. Since $\mathbf{z}^H A\mathbf{z}$ is just a complex number $(\mathbf{z}^H A\mathbf{z})^H = \overline{\mathbf{z}^H A\mathbf{z}}$ this shows that $\mathbf{z}^H A\mathbf{z}$ is its own conjugate so it must be real.

**Th** Eigenvalues of $A$ are real if $A^H = A$.
**Pf** Multiply $A\mathbf{z} = \lambda \mathbf{z}$ by $\mathbf{z}^H$; $\mathbf{z}^H A\mathbf{z} = \lambda \mathbf{z}^H \mathbf{z}$. Since the left is real by the previous theorem and $\mathbf{z}^H \mathbf{z}$ is real it follows that $\lambda$ must be real.

**Theorem** Eigenvectors for different eigenvalues are orthogonal if $A^H = A$.
**Proof** If $A\mathbf{x}_1 = \lambda_1 \mathbf{x}_1$ and $A\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$ then sine $A^H = A$ and $\lambda_1$ is real

$$\lambda_1 \mathbf{x}_1^H \mathbf{x}_2 = (A\mathbf{x}_1)^H \mathbf{x}_2 = \mathbf{x}_1^H A\mathbf{x}_2 = \lambda_2 \mathbf{x}_1^H \mathbf{x}_2$$

It therefore follows that $\mathbf{x}_1^H \mathbf{x}_2 = 0$ since we assumed that $\lambda_1 \neq \lambda_2$.

**Spectral Theorem** Suppose that $A$ is real and symmetric. Then there is an orthogonal matrix $U$ such that $U^{-1}AU = D$ is diagonal.

**Ex** Diagonalize $A = \begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$ with an orthogonal transformation $U$.
**Sol** $A$ is symmetric so it can be diagonalized by an orthogonal transformation.
The eigenvalues are $\lambda_1 = \lambda_2 = -1$ and $\lambda_3 = 5$. The eigenspace corresponding to eigenvalue $-1$; $(A + I)\mathbf{x} = \mathbf{0}$ satisfy $x_1 + 2x_2 - x_3 = 0$ so $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ form a basis for the eigenspace corresponding to $\lambda = -1$. We can apply the Gram-Schmidt process to obtain an orthonormal basis. Let

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \qquad \mathbf{u}_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

The eigenspace corresponding to $\lambda_3 = 5$ is spanned by $\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$ and we set $\mathbf{u}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$.
and we set $U = [\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3]$ so $A = UDU^T$, where $D = \operatorname{diag}\{\lambda_1, \lambda_2, \lambda_3\}$.