MATH-UA 140 - Linear Algebra

Lecture 23: Symmetric Matrices

Among all matrices, symmetric matrices (i.e such that $A^T = A$) play a particularly important role in linear algebra, and in applications of linear algebra in science, economics, etc.

Symmetric matrices have special properties regarding their eigenvalues their eigenvectors, and their diagonalization. We cover these properties in this lecture.

I] Eigenvalues and eigenvectors of symmetric matrices

1) Eigenvalues

Theorem: All the eigenvalues of a real symmetric matrix are real.

Proof: Suppose A is symmetric, with real entries, and $A\vec{x} = \lambda\vec{x}$
As we have seen in Lecture 21, λ can in principle be a complex number: $\lambda = a + ib$, where a and b are real numbers.
The eigenvector $\vec{x}$ may also have components which are complex numbers.
We take the complex conjugate of $A\vec{x} = \lambda\vec{x}$: $\overline{A\vec{x}} = \overline{\lambda\vec{x}}$
Now, the complex conjugate of a product is the product of the complex conjugates, so $\bar{A}\,\bar{\vec{x}} = \bar{\lambda}\,\bar{\vec{x}}$
A has real entries, so $A\bar{\vec{x}} = \bar{\lambda}\bar{\vec{x}}$
Taking the transpose of this equality, $\bar{\vec{x}}^T A = \bar{\vec{x}}^T \bar{\lambda}$ (since $A^T = A$)

To sum up to this point, we have $A\vec{x} = \lambda \vec{x}$

$$\overline{\vec{x}}^T A = \overline{\vec{x}}^T \overline{\lambda}$$

Let us multiply the first equation by $\overline{\vec{x}}^T$ on the left, and the second equation by $\vec{x}$ on the right:

$$\overline{\vec{x}}^T A\vec{x} = \lambda \overline{\vec{x}}^T\vec{x} \quad \text{and} \quad \overline{\vec{x}}^T A\vec{x} = \overline{\lambda}\, \overline{\vec{x}}^T\vec{x}$$

We conclude that

$$\lambda \overline{\vec{x}}^T\vec{x} = \overline{\lambda}\, \overline{\vec{x}}^T\vec{x}$$

Now, $\overline{\vec{x}}^T\vec{x}$ is the square of the magnitude of $\vec{x}$: $\|\vec{x}\|^2$

Since $\vec{x}$ is an eigenvector, its magnitude is not zero.

Therefore: $\lambda = \overline{\lambda}$

i.e $a+ib = a-ib$

i.e $b=0$

λ is a real number.

2) In passing: a result for real, nonsymmetric matrices

A step in the proof above gives us a nice result for nonsymmetric matrices with real entries: indeed, we had

$$A\vec{x} = \lambda\vec{x} \implies A\overline{\vec{x}} = \overline{\lambda}\,\overline{\vec{x}}$$

In other words, for real matrices, if $(\lambda, \vec{x})$ is an eigenvalue/eigenvector pair, then so is $(\overline{\lambda}, \overline{\vec{x}})$

For real matrices, complex eigenvalues and eigenvectors come in conjugate pairs.

3) Back to symmetric matrices: eigenvalues and pivots

Pivots and eigenvalues are very different mathematical concepts, calculated in two different ways: elimination for the pivots, $\det(A-\lambda I)=0$ for the eigenvalues. However, we have already seen a link between them, through the determinant of A:

$$\text{Det } A = \text{product of its pivots} = \text{product of its eigenvalues}$$

Here is another theorem linking the concepts even more closely, for symmetric matrices:

For symmetric matrices, the pivots and the eigenvalues have the same signs.

Example: $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ The pivots are 1 and -3

The eigenvalues are given by $\begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = 0$

$$\Leftrightarrow (1-\lambda)^2 - 4 = -(3-\lambda)(1+\lambda)$$

$$\Rightarrow \lambda = 3 \text{ and } \lambda = -1$$

4) Eigenvectors of symmetric matrices

Theorem: Eigenvectors of a real symmetric matrix corresponding to distinct eigenvalues are orthogonal.

<u>Proof:</u> Consider a symmetric matrix A, and an eigenvector $\vec{x}_1$ for the eigenvalue λ_1, and an eigenvector $\vec{x}_2$ for the eigenvalue λ_2, with $\lambda_1 \neq \lambda_2$

$$A\vec{x}_1 = \lambda_1 \vec{x}_1 \qquad A\vec{x}_2 = \lambda_2 \vec{x}_2$$

From the first equality, we can write $(\lambda_1 \vec{x}_1)^T \vec{x}_2 = (A\vec{x}_1)^T \vec{x}_2$

$$= \vec{x}_1^T A \vec{x}_2 \underset{\text{second equality}}{=} \lambda_2 \vec{x}_1^T \vec{x}_2$$

In other words, $\lambda_1 \vec{x}_1^T \vec{x}_2 = \lambda_2 \vec{x}_1^T \vec{x}_2$

Since $\lambda_2 \neq \lambda_1$, $\vec{x}_1^T \vec{x}_2 = 0$, i.e. $\vec{x}_1$ and $\vec{x}_2$ are orthogonal.

<u>Example:</u> Let us return to $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

with eigenvalues $\lambda = -1$ and $\lambda = 3$

An eigenvector for $\lambda = -1$ is in the nullspace of $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$

$\vec{x}_1 = (-1, 1)$ is an eigenvector

An eigenvector for $\lambda = 3$ is in the nullspace of $\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$

$\vec{x}_2 = (1, 1)$ is an eigenvector

$\vec{x}_1^T x_2 = -1 + 1 = 0$

II) Diagonalization of symmetric matrices

1) Simple case: n distinct eigenvalues

We just saw that a symmetric matrix with n distinct eigenvalues has n eigenvectors $\vec{x}_1, \ldots, \vec{x}_n$ which are orthogonal to each other. Dividing each $\vec{x}_i$ by its magnitude, we obtain n eigenvectors $\vec{q}_1, \ldots, \vec{q}_n$ which form an orthonormal basis.

The eigenvector matrix S is then $\begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix}$

and $S^T S = \begin{bmatrix} \vec{q}_1^{\,T} \\ \vdots \\ \vec{q}_n^{\,T} \end{bmatrix} \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix} = I$

We conclude that $S = Q$ is an orthogonal matrix, with $Q^T Q = I$, i.e $Q^T = Q^{-1}$

A symmetric matrix with n distinct eigenvalues can be diagonalized as

$$A = Q \Lambda Q^T$$

Let us look more closely at how this looks:

$$A = \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{q}_1^{\,T} \\ \vdots \\ \vec{q}_n^{\,T} \end{bmatrix}$$

$$= \begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix} \begin{bmatrix} \lambda_1 \vec{q}_1^{\,T} \\ \lambda_2 \vec{q}_2^{\,T} \\ \vdots \\ \lambda_n \vec{q}_n^{\,T} \end{bmatrix} = \lambda_1 \vec{q}_1 \vec{q}_1^{\,T} + \lambda_2 \vec{q}_2 \vec{q}_2^{\,T} + \cdots + \lambda_n \vec{q}_n \vec{q}_n^{\,T}$$

So A is the sum of n rank one matrices $\lambda_i \vec{q_i} \vec{q_i}^T$.
Each of these is a projection matrix onto the subspace spanned by the eigenvector $\vec{q_i}$.

2) General case

The question is: if A is a symmetric matrix with eigenvalues that may not be distinct, can we still write $A = Q \Lambda Q^T$, with $Q^T = Q^{-1}$?

This answer is yes, but we will not prove it here. Instead, to understand why it is true, we use a result that applies to all square matrices, not only symmetric ones, called Schur's theorem:

Every square matrix A factors into $A = QTQ^{-1}$ where T is upper triangular and $\bar{Q}^T = Q^{-1}$.

In this theorem, T may have complex numbers in its entries, and Q too, which is why there is a $-$ in $\bar{Q}^T$.

If A has only real eigenvalues, then Q and T can be chosen so that they only have real entries

$A = QTQ^{-1}$ with $Q^T = Q^{-1}$

We are now ready for our general result for symmetric matrices:

* Symmetric matrices only have real eigenvalues, so Schur's theorem tells us that any symmetric A can be written as

$$A = Q T Q^T$$

* This means that $T = Q^T A Q$

$T^T = Q^T A Q$ since $A^T = A$

So T is upper triangular and symmetric: T is in fact diagonal, $T = \Lambda$ and

$$\underline{A = Q \Lambda Q^T} \quad \blacksquare$$

In other words, <u>all</u> symmetric matrices are diagonalizable.

Example: $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ $\qquad \begin{vmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = 0$

$$\Leftrightarrow -\lambda(\lambda^2 - 1) + (\lambda+1) + (\lambda+1) = 0$$
$$\Leftrightarrow (\lambda+1)[2 - \lambda(\lambda-1)] = 0$$
$$\Leftrightarrow (\lambda+1)(\lambda^2 - \lambda - 2) = 0$$
$$\Leftrightarrow (\lambda+1)^2(\lambda-2) = 0$$

A has two eigenvalues: $\lambda = -1$ and $\lambda = 2$

The eigenvector for $\lambda = 2$ is in the nullspace of

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

The row reduced echelon form is obtained as follows:

$$\begin{bmatrix} -2 & 1 & 1 \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 1 & 1 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} -2 & 1 & 1 \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 0 & \frac{3}{2} & -\frac{3}{2} \end{bmatrix} \longrightarrow \begin{bmatrix} -2 & 1 & 1 \\ 0 & -\frac{3}{2} & \frac{3}{2} \\ 0 & 0 & 0 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

An eigenvector for $\lambda = 2$ is $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Similarly, we look for eigenvectors for $\lambda = -1$ in the nullspace of

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Two linearly independent eigenvectors are $\vec{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

and $\vec{x}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

$\vec{x}_1^T \vec{x}_2 = \vec{x}_1^T \vec{x}_3 = 0$. However, $\vec{x}_2^T \vec{x}_3 = 1 \neq 0$

We construct $\vec{X}_3$ orthogonal to $\vec{x}_2$ by the Gram-Schmidt process:

$$\vec{X}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

We conclude that $\vec{q}_1 = \frac{\vec{x}_1}{\|\vec{x}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ is a unit eigenvector with eigenvalue $\lambda = 2$, and $\vec{q}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{q}_3 = \sqrt{\frac{2}{3}}\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$ are unit eigenvectors with eigenvalue $\lambda = -1$.

By construction, $\vec{q}_1$, $\vec{q}_2$, and $\vec{q}_3$ are orthogonal to one another. We can write:

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix}$$