Lecture 2l

Symmetric Matrices

(pages 363-366)

Definition A matrix A is symmetric if $A^T = A$ or, equivalently, if $a_{ij} = a_{ji}$ for all i and j.

Note that the definition of a symmetric matrix requires that it be a square matrix, since if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix, so $A^T = A$ means that $m = n$.

Examples: Some examples of symmetric matrices are:

$$\begin{bmatrix} 3 & -5 \\ -5 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & -3 \\ 0 & -3 & -5 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

As a practice problem I will have you look at some of the properties of symmetric matrices, but the result we are going to focus on is the fact that every symmetric matrix is diagonalizable. At the end of Math 106, we learned how to find eigenvalues and eigenspaces, and we learned that not every matrix is diagonalizable. In the next lecture, we will prove that not only can every symmetric matrix be diagonalized, they can all be diagonalized by an orthogonal matrix.

Definition: A matrix A is said to be **orthogonally diagonalizable** if there exists an orthogonal matrix P and a diagonal matrix D such that $P^T AP = D$.

Recall that, if P is an orthogonal matrix, then $P^T = P^{-1}$, so this definition is just the same as saying there is an orthogonal matrix P that diagonalizes A.

Before we prove that every symmetric matrix is orthogonally diagonalizable, we will do some examples (and assignment problems) of diagonalizing symmetric matrices, as this hands on work will help us understand the theoretical proof.

Example: Let's diagonalize the symmetric matrix $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$. To do this, we first need to find the eigenvalues, which means we need to find the roots of the characteristic polynomial $\det(A - \lambda I)$

$$\begin{aligned} \det \begin{bmatrix} 1-\lambda & 2 \\ 2 & -2-\lambda \end{bmatrix} &= (1-\lambda)(-2-\lambda) - (2)(2) \\ &= -6 + \lambda + \lambda^2 \\ &= (2-\lambda)(-3-\lambda) \end{aligned}$$

So we see that the eigenvalues for A are $\lambda_1 = 2$ and $\lambda_2 = -3$. The next step is to find a basis for the eigenspaces.

For $\lambda_1 = 2$, we need to find the nullspace of $\begin{bmatrix} 1-\lambda_1 & 2 \\ 2 & -2-\lambda_1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix}$.

Row reducing, we see that $\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$, so our nullspace consists of all solutions to $x_1 - 2x_2 = 0$. If we replace the variable x_2 with the parameter s, then we have $x_1 = 2s$ and $x_2 = s$, so the general solution is $s\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, which means that $\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$ is a basis for the eigenspace for λ_1.

For $\lambda_2 = -3$, we need to find the nullspace of $\begin{bmatrix} 1-\lambda_2 & 2 \\ 2 & -2-\lambda_2 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$. Row reducing, we see that $\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$, so our nullspace consists of all solutions to $2x_1 + x_2 = 0$. If we replace the variable x_1 with the parameter s, then we have $x_1 = s$ and $x_2 = -2s$, so the general solution is $s\begin{bmatrix} 1 \\ -2 \end{bmatrix}$, which means that $\left\{\begin{bmatrix} 1 \\ -2 \end{bmatrix}\right\}$ is a basis for the eigenspace for λ_2.

Our study of diagonalization (specifically Theorem 6.2.2–the Diagonalization Theorem) tells us that that A can be diagonalized by a matrix P whose columns are the basis vectors for the eigenspaces. That is, we know that $P = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$ is an invertible matrix such that $P^{-1}AP = D$, where $D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$.

So that settles the issue of A being diagonalizable, but what about being orthogonally diagonalizable? A quick check shows that the columns of P are orthogonal, but not orthonormal. But what if we normalize these vectors? Well, since $\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$ is a basis for the eigenspace for λ_1, we also have that $\left\{s\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$ is a basis for the eigenspace for λ_1 for any scalar s, which means that $\left\{\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}\right\}$ is a basis for the eigenspace for λ_1. Similarly, $\left\{\begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}\right\}$ is a basis for the eigenspace for λ_2. And going back to our knowledge of diagonalization, this means that $Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}$ is an invertible matrix such that $Q^{-1}AQ = D$, where D is still $\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$. And moreover, we see that Q is orthogonal, so we have shown that A is orthogonally diagonalizable.

Example: Let's diagonalize the symmetric matrix $A = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$. To do this, we first need to find the eigenvalues, which means we need to find the roots of the characteristic polynomial $\det(A - \lambda I)$

$$\det \begin{bmatrix} 1-\lambda & -2 & -2 \\ -2 & 1-\lambda & 2 \\ -2 & 2 & 1-\lambda \end{bmatrix} \begin{aligned} &= (1-\lambda) \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} + 2 \begin{vmatrix} -2 & 2 \\ -2 & 1-\lambda \end{vmatrix} - 2 \begin{vmatrix} -2 & 1-\lambda \\ -2 & 2 \end{vmatrix} \\ &= (1-\lambda)(-3-2\lambda+\lambda^2) + 2(2+\lambda) - 2(-2-2\lambda) \\ &= 5 + 9\lambda + 3\lambda^2 - \lambda^3 \end{aligned}$$

A quick check shows that $\lambda = -1$ is a root of this equation. If we factor out $(-1-\lambda)$, we are left with $-5 - 4\lambda + \lambda^2$, which factors as $(-1-\lambda)(5-\lambda)$. So, the characteristic polynomial for A is $(-1-\lambda)^2(5-\lambda)$, and the eigenvalues for A are $\lambda_1 = -1$ and $\lambda_2 = 5$. The next step is to find a basis for the eigenspaces.

For $\lambda_1 = -1$, we need to find the nullspace of $\begin{bmatrix} 1-\lambda_1 & -2 & -2 \\ -2 & 1-\lambda_1 & 2 \\ -2 & 2 & 1-\lambda_1 \end{bmatrix} =$ $\begin{bmatrix} 2 & -2 & -2 \\ -2 & 2 & 2 \\ -2 & 2 & 2 \end{bmatrix}$. Row reducing, we see that $\begin{bmatrix} 2 & -2 & -2 \\ -2 & 2 & 2 \\ -2 & 2 & 2 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, so our nullspace consists of all solutions to $x_1 - x_2 - x_3 = 0$. If we replace the variable x_2 with the parameter s and the variable x_3 with the parameter t, then we have $x_1 = s+t$, $x_2 = s$, and $x_3 = t$. So the general solution is $s \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, which means that $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigenspace for λ_1.

For $\lambda_2 = 5$, we need to find the nullspace of $\begin{bmatrix} 1-\lambda_2 & -2 & -2 \\ -2 & 1-\lambda_2 & 2 \\ -2 & 2 & 1-\lambda_2 \end{bmatrix} =$ $\begin{bmatrix} -4 & -2 & -2 \\ -2 & -4 & 2 \\ -2 & 2 & -4 \end{bmatrix}$. Row reducing, we see that

$$\begin{bmatrix} 4 & -2 & -2 \\ -2 & -4 & 2 \\ -2 & 2 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & -6 & 6 \\ 0 & -6 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

So our nullspace consists of all solutions to the system

$$\begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 0 \end{aligned}$$

If we replace the variable x_3 with the parameter s, then we have $x_1 = -s$, $x_2 = s$, and $x_3 = s$. So the general solution is $s \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, which means that

$\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is a basis for the eigenspace for λ_2.

Our study of diagonalization tells us that that A can be diagonalized by a matrix P whose columns are the basis vectors for the eigenspaces. That is, we know that $P = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ is an invertible matrix such that $P^{-1}AP = D$, where $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

Again, we have shown that A is diagonalizable, but what about being orthogonally diagonalizable? In this case, we do not even have that columns of P are orthogonal, much less orthonormal. We could apply the Gram-Schmidt Procedure to the columns of P, but our theory of diagonalization requires that the columns of P be basis vectors for the corresponding eigenspaces, not simply any basis vector for $\mathbb{R}^3$. So, as we did in the previous example, we will go back to our eigenspaces and find orthonormal bases for them.

Looking at the eigenspace for λ_2 first, since it is the span of a single vector, we simply need to normalize this vector to get an orthonormal basis. And so, we see that $\left\{ \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \right\}$ is an orthonormal basis for the eigenspace for λ_2.

But the eigenspace for λ_1 has two basis vectors that are not orthogonal, so we will need to use the Gram-Schmidt Procedure to find first an orthogonal, then an orthonormal, basis for this eigenspace. We keep our first vector as $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, and then our second vector becomes

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\|}$$

$$= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{(1)(1)+(0)(1)+(1)(1)}{(1^2+1^2+0^2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

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

$$= \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix}$$

So, $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/2 \\ -1/2 \\ 1 \end{bmatrix} \right\}$ is an orthogonal basis for our eigenspace, and by normalizing the vectors we get that $\left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} \right\}$ is an orthonormal basis for the eigenspace for λ_1.

Again, returning to our knowledge of diagonalization, we know that the matrix $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$ whose columns are these new basis vectors will diagonalize A. But is Q orthonormal? A quick check verifies that Q is in fact orthonormal, and so we have shown that A is orthogonally diagonalizable.

Again, I am assigning the problems on showing that various similar matrices are orthogonally diagonalizable before we prove that this is always the case, as I feel that working hands on with these matrices will help motivate the steps in the proof. At this point, all you need to know is that if you find an orthonormal basis for all of the eigenspaces, then you WILL end up with an orthogonal matrix P to diagonalize A.