22m:033 Notes:
7.1 Diagonalization of Symmetric Matrices

Dennis Roseman
University of Iowa
Iowa City, IA

http://www.math.uiowa.edu/∼roseman

May 3, 2010

1 Symmetric matrices

Definition 1.1 *A symmetric matrix is a matrix A such that* $A = A^T$.

In other words a symmetric matrix is a square matrix A such that $a_{ij} = a_{ji}$.

Example 1.2

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 0 & 5 \\ 3 & 0 & 2 & -7 \\ 4 & 5 & -7 & -2 \end{pmatrix}$$

is symmetric but

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & 0 & -5 \\ 3 & 0 & 2 & -7 \\ 4 & 5 & -7 & -2 \end{pmatrix}$$

is not.

2 Diagonalization of Symmetric Matrices

We will see that any symmetric matrix is diagonalizable. This is surprising enough, but we will also see that in fact a symmetric matrix is similar to a diagonal matrix in a very special way.

Recall that, by our definition, a matrix A is diagonalizable if and only if there is an invertible matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix. We make a stronger definition.

Definition 2.1 *A matrix A is orthogonally diagonalizable if and only if there is an orthogonal matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.*

Remark 2.2 Recall that any orthogonal matrix A *is* invertible and also that $A^{-1} = A^T$. Thus we can say that A matrix A is orthogonally diagonalizable if there is a square matrix P such that $A = PDP^T$ where D is a diagonal matrix.

Remark 2.3 Recall (see page 115) the formula for transpose of a product: $(MN)^T = N^T M^T$. Using this we can see that any orthogonally diagonalizable A must be symmetric. This is because

$$A^T = (PDP^T)^T = ((P^T)^T D^T P^T = PDP^T = A.$$

Although we do not prove Proposition 2.1 the following theorem used in the proof will help us find the matrix P. in its own right:

Proposition 2.4 *If A is symmetric then any two eigenvalues from different eigenspaces are orthogonal*

For example if all of the eigenspaces are one dimensional then the set of eigenvectors is an orthogonal set.

Example 2.5 Suppose

$$A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

We calculate the characteristic polynomial by expanding the determinant along top row (or left column):

$$\begin{aligned}
A &= \left| \begin{pmatrix} 3-\lambda & 2 & 0 \\ 2 & 2-\lambda & 2 \\ 0 & 2 & 1-\lambda \end{pmatrix} \right| = \\
&= (3-\lambda)\left((2-\lambda)(1-\lambda)-4\right) - 2\left(2(1-\lambda)\right) + 0 \cdot 4 \\
&= (3-\lambda)(2-3\lambda+\lambda^2-4) - 2(2-2\lambda) \\
&= (3-\lambda)(\lambda^2-3\lambda-2) - 4 + 4\lambda \\
&= 3\lambda^2 - 9\lambda - 6 - \lambda^3 + 3\lambda^2 + 2\lambda - 4 + 4\lambda \\
&= -\lambda^3 + 6\lambda^2 - 3\lambda - 10
\end{aligned}$$

The eigenvalues are the roots of

$$\lambda^3 - 6\lambda^2 + 3\lambda + 10 = 0$$

We could use the usual test for rational roots or just stare at this and note that $\lambda = -1$ is a root.

We then do long division of polynomials and see that:

$$\frac{\lambda^3 - 6\lambda^2 + 3\lambda + 10}{\lambda + 1} = \lambda^2 - 7\lambda + 10 = (\lambda - 5)(\lambda - 2)$$

So we have three distinct eigenvectors and we know the matrix is diagonalizable.

Next we find eigenvectors for these values. We recall that these vectors are not unique but are all multiples of each other.

To make a long story short here are three such vectors:

$$\left\{ \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \right\}$$

First we should at least verify this. We note that

$$\begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \\ 5 \end{pmatrix}$$

So the first vector is an eigenvector with eigenvalue 5.

Next

$$\begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \\ 4 \end{pmatrix}$$

So the first vector is an eigenvector with eigenvalue 2.

and finally

$$\begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -2 \end{pmatrix}$$

So the first vector is an eigenvector with eigenvalue -1.

At this point we should check our calculations by directly verifying that this is an orthogonal set of vectors.

But now to get our orthogonal matrix we need an *orthonormal* basis. It is clear that each of our vectors has length 3.

So construct our matrix

$$P = \begin{pmatrix} \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

How would we check that this matrix is orthogonal? One way is to calculate P^{-1}—make our augmented matrix and row reduce, etc. However if we only want to verify this and not do all that calculation we only need to check that $PP^T = I$, which is a lot easier. We would expect that

$$PAP^T = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

. Why?

We leave it as an exercise to check this.

Example 2.6 In general, we have found that if the characteristic polynomial has a root λ_0 of order $n > 2$ or more we do not know if we can find n linearly independent eigenvalues for this eigenspace. But for symmetric matrices, we always can. In fact we can sum all this up in what the text calls the Spectral Theorem:

Proposition 2.7 *(The Spectral Theorem) An $n \times n$ symmetric matrix has the following properties:*

1. *A has n real eigenvalues if we count multiplicity*
2. *For each eigenvalue the dimension of the corresponding eigenspace is equal to the algebraic multiplicity of that eigenvalue.*
3. *The eigenspaces are mutually orthogonal*
4. *A is orthogonally diagonalizable*

The next example will need our skills at the Gram Schmidt process.

Consider $A = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}$.

The characteristic polynomial is $(\lambda - 2)^2(\lambda - 4)$

There is one eigenvector with eigenvalue 4 and we can verify that we could use $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Even though the Spectral Theorem tells us we can diagonalize this matrix, to get the matrix P we will have to find an orthogonal basis for the null space for $\lambda = 2$.

We are looking for an orthonormal basis for the null space of

$$A - (-2)I = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

This matrix clearly row reduces to $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. We have two free variables. If we let $z = t$ and $y = s$ then $x + s + t = 0$ and the null space consists of the vectors

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -s - t \\ s \\ t \end{pmatrix} = t \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

So

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

is a basis for this eigenspace. Denote these vectors by $\overrightarrow{b_1}$ and $\overrightarrow{b_2}$. However we need an orthonormal basis.

We let $\overrightarrow{v_1} = \overrightarrow{b_1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

Next we need $\overrightarrow{v_2} = \overrightarrow{b_2} - \dfrac{\overrightarrow{b_2} \cdot \overrightarrow{v_1}}{\overrightarrow{v_1} \cdot \overrightarrow{v_1}} \overrightarrow{v_1}$.

Now

$$\overrightarrow{v_1} \cdot \overrightarrow{b_2} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = 1$$

Now

$$\overrightarrow{v_1} \cdot \overrightarrow{v_1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 2$$

So

$$\overrightarrow{v_2} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

For hand calculation we could make life easier by choosing $v_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}$. Explain.

So it should be clear that our orthogonal matrix could be taken to be

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}.$$