Math 1180, 2009
Notes 20

Proof of most of the spectral theorem

Here again is the theorem.

Theorem 1 *Let A be a symmetric $n \times n$ matrix. Then*
(i) All eigenvalues of A are real numbers
(ii) Two eigenvectors corresponding to different eigenvalues are orthogonal
(iii) There is an orthogonal basis for R^n which consists of eigenvectors of A .

Proof. *Parts (i) and (ii) are not hard to prove, and (iii) is not hard in the most common situation. First we prove (i) .*

To prove (i) it is useful to generalize the idea of dot product, or "inner product" as we called it later, to the case where vectors can be complex. Let's recall some facts about complex numbers.

Any complex number z can be written as

$$z = x + iy,$$

where x and y are real numbers, and $i = \sqrt{-1}$. Then x is called the "real part" of z and y is called the "imaginary part" of z. Note that what we call the imaginary part is y, not iy.

A complex vector is a vector

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_1 \\ \cdots \\ z_n \end{pmatrix},$$

where $z_1, ..., z_n$ are each complex numbers.

The complex conjugate of a complex number $z = x + iy$ is the complex number $x - iy$. This is denoted by $\bar{z}$. Note that

$$z\bar{z} = (x + iy)(x - iy) = x^2 + ixy - ixy + (iy)^2 = x^2 + y^2.$$

Thus, $z\bar{z}$ is a nonnegative real number.

If $\mathbf{z}$ and $\mathbf{w}$ are complex vectors, then we define the "inner product" of $\mathbf{z}$ and $\mathbf{w}$ to be

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1\bar{w}_1 + z_2\bar{w}_2 + \cdots + z_n\bar{w}_n. \tag{1}$$

In particular, $\langle \mathbf{z}, \mathbf{z} \rangle = z_1\bar{z}_1 + \cdots + z_n\bar{z}_n$*.This is a nonnegative real number.*

We also have the formulas

$$\langle \lambda\mathbf{z}, \mathbf{w} \rangle = \lambda\langle \mathbf{z}, \mathbf{w} \rangle$$
$$\langle \mathbf{z}, \lambda\mathbf{w} \rangle = \bar{\lambda}\langle \mathbf{z}, \mathbf{w} \rangle.$$

It should be clear from (1) *what the* $\bar{\lambda}$ *appears in the last formula.*

Now suppose that A *has a complex eigenvalue* $\lambda = \mu + i\nu$*. Then the corresponding eigenvector will also be be complex. We have*

$$A\mathbf{z} = \lambda\mathbf{z}.$$

We now consider the inner product

$$\langle A\mathbf{z}, \mathbf{z} \rangle = \langle \lambda\mathbf{z}, \mathbf{z} \rangle = \lambda\langle \mathbf{z}, \mathbf{z} \rangle.$$

But A *is symmetric, and we saw in an earlier lecture that*

$$\langle A\mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{z}, \mathbf{Az} \rangle = \langle \mathbf{z}, \lambda\mathbf{z} \rangle = \bar{\lambda}\langle \mathbf{z}, \mathbf{z} \rangle$$

In the last two formulas, $\mathbf{z}$ *is an eigenvector, and so not* 0*. Hence* $\langle \mathbf{z}, \mathbf{z} \rangle \neq 0$*. This implies that* $\lambda = \bar{\lambda}$*, or in other words,* λ *is a real number.*

For (ii), we suppose that λ_1 *and* λ_2 *are eigenvalues, with corresponding eigenvectors* $\mathbf{v}_1$ *and* $\mathbf{v}_2$*.*

Recall the general formula

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A^T\mathbf{w} \rangle,$$

which holds whether or not A *is symmetric. If* A *is symmetric, then* $A = A^T$*, so we have*

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, A\mathbf{w} \rangle$$

We now apply this in the sequence of steps below.

$$\begin{aligned}\langle A\mathbf{v}_1, \mathbf{v}_2 \rangle &= \langle \lambda_1\mathbf{v}_1, \mathbf{v}_2 \rangle = \lambda_1\langle \mathbf{v}_1, \mathbf{v}_2 \rangle, \\ &= \langle \mathbf{v}_1, A\mathbf{v}_2 \rangle = \lambda_2\langle \mathbf{v}_1, \mathbf{v}_2 \rangle.\end{aligned}$$

(Remember that λ_1 *and* λ_2 *are real numbers.) Because* $\lambda_1 \neq \lambda_2$*, we must have* $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$*, so the vectors are orthogonal.*

For (iii), the theorem is true in one case: When the eigenvalues are distinct, there is an eigenvector for each eigenvalue, and these eigenvectors are mutually orthogonal. It is then easy to normalize all the eigenvectors, giving an orthogonal matrix which diagonalizes the matrix, as desired. ■

The difficulty in the proof when there are repeated eigenvalues is to show that each has full geometric multiplicity, in order to get a basis of eigenvectors. We shall not give this proof, but it is in the text.