MTH6140 **Linear Algebra II**

Notes 8 **16th December 2010**

8 Symmetric matrices

We come to one of the most important topics of the course. In simple terms, any real symmetric matrix is diagonalisable. But there is more to be said!

8.1 Orthogonal projections and orthogonal decompositions

We say that two vectors v, w in an inner product space are *orthogonal* if $v \cdot w = 0$.

Definition 8.1 Let V be an inner product space, and U a subspace of V. The *orthogonal complement* of U is the set of all vectors which are orthogonal to everything in U:

$$U^{\perp} = \{w \in V : w \cdot u = 0 \text{ for all } u \in U\}.$$

The symbol $U^{\perp}$ is often pronounced 'U perp', where 'perp' is short for *perpendicular*, which is another word for *orthogonal*.

Proposition 8.1 If V is an inner product space and U a subspace of V, with $\dim(V) = n$ and $\dim(U) = r$, then $U^{\perp}$ is a subspace of V, and $\dim(U^{\perp}) = n - r$. Moreover, $V = U \oplus U^{\perp}$.

Proof Proving that $U^{\perp}$ is a subspace is straightforward from the properties of the inner product. If $w_1, w_2 \in U^{\perp}$, then $w_1 \cdot u = w_2 \cdot u = 0$ for all $u \in U$, so $(w_1 + w_2) \cdot u = 0$ for all $u \in U$, whence $w_1 + w_2 \in U^{\perp}$. The argument for scalar multiples is similar.

Now choose a basis for U and extend it to a basis for V. Then apply the Gram–Schmidt process to this basis (starting with the elements of the basis for U), to obtain an orthonormal basis $(v_1, \ldots, v_n)$. Since the process only modifies vectors by adding multiples of earlier vectors, the first r vectors in the resulting basis will form an orthonormal basis for U. The last $n - r$ vectors will be orthogonal to U, and so lie in $U^{\perp}$; and they are clearly linearly independent. Now suppose that $w \in U^{\perp}$ and $w = \sum_i c_i v_i$, where $(v_1,\dots,v_n)$ is the orthonormal basis we constructed. Then $c_i = w\cdot v_i = 0$ for $i = 1,\dots,r$; so w is a linear combination of the last $n-r$ basis vectors, which thus form a basis of $U^\perp$. Hence $\dim(U^\perp) = n-r$, as required.

Now the last statement of the proposition follows from the proof, since we have a basis for V which is a disjoint union of bases for U and $U^\perp$.

Recall the connection between direct sum decompositions and projections. If we have projections $P_1,\dots,P_r$ whose sum is the identity and which satisfy $P_iP_j = O$ for $i \neq j$, then the space V is the direct sum of their images. This can be refined in an inner product space as follows.

Definition 8.2 Let V be an inner product space. A linear map $P : V \to V$ is an *orthogonal projection* if

(a) P is a projection, that is, $P^2 = P$;

(b) P is self-adjoint, that is, $P^* = P$ (where $P^*(v)\cdot w = v\cdot P(w)$ for all $v,w \in V$).

Proposition 8.2 If P is an orthogonal projection, then $\mathrm{Ker}(P) = \mathrm{Im}(P)^\perp$.

Proof We know, from Proposition 5.1, that $V = \mathrm{Ker}(P)\oplus\mathrm{Im}(P)$; we only have to show that these two subspaces are orthogonal. So take $v \in \mathrm{Ker}(P)$, so that $P(v) = 0$, and $w \in \mathrm{Im}(P)$, so that $w = P(u)$ for some $u \in V$. Then

$$v\cdot w = v\cdot P(u) = P^*(v)\cdot u = P(v)\cdot u = 0,$$

as required.

Proposition 8.3 Let $P_1,\dots,P_r$ be orthogonal projections on an inner product space V satisfying $P_1+\cdots+P_r = I$ and $P_iP_j = O$ for $i \neq j$. Let $U_i = \mathrm{Im}(P_i)$ for $i = 1,\dots,r$. Then

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_r,$$

and if $u_i \in U_i$ and $u_j \in U_j$, then u_i and u_j are orthogonal.

Proof The fact that V is the direct sum of the images of the P_i follows from Proposition 5.2. We only have to prove the last part. So take u_i and u_j as in the Proposition, say $u_i = P_iv$ and $u_j = P_jw$. Then

$$u_i\cdot u_j = P_iv\cdot P_jw = P_i^*v\cdot P_jw = v\cdot P_iP_jw = 0,$$

where the second equality holds since P_i is self-adjoint and the third is the definition of the adjoint.

A direct sum decomposition satisfying the conditions of Proposition 8.3 is called an *orthogonal decomposition* of V.

Conversely, if we are given an orthogonal decomposition of V, then we can find orthogonal projections satisfying the hypotheses of Proposition 8.3.

8.2 The spectral theorem

The main theorem can be stated in two different ways. I emphasise that these two theorems are the same! Either of them can be referred to as the *spectral theorem*.

Theorem 8.4 If T is a self-adjoint linear map on a real inner product space V, then the eigenspaces of T form an orthogonal decomposition of V. Hence there is an orthonormal basis of V consisting of eigenvectors of T. Moreover, there exist orthogonal projections $P_1, \ldots, P_r$ satisfying $P_1 + \cdots + P_r = I$ and $P_i P_j = O$ for $i \neq j$, such that

$$T = \lambda_1 P_1 + \cdots + \lambda_r P_r,$$

where $\lambda_1, \ldots, \lambda_r$ are the distinct eigenvalues of T.

Theorem 8.5 Let A be a real symmetric matrix. Then there exists an orthogonal matrix P such that $P^{-1}AP$ $(= P^\top AP)$ is diagonal.

Proof The second theorem follows from the first, since the transition matrix from one orthonormal basis to another is an orthogonal matrix. So we concentrate on the first theorem. It suffices to find an orthonormal basis of eigenvectors, since all the rest follows from our remarks about projections, together with what we already know about diagonalisable maps.

The proof will be by induction on $n = \dim(V)$. There is nothing to do if $n = 1$. So we assume that the theorem holds for $(n-1)$-dimensional spaces.

The first job is to show that T has an eigenvector.

Choose an orthonormal basis; then T is represented by a real symmetric matrix A. Its characteristic polynomial has a root λ over the complex numbers, by the so-called "fundamental theorem of algebra". We temporarily enlarge the field from $\mathbb{R}$ to $\mathbb{C}$. Now we can find a column vector $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Taking the complex conjugate, remembering that A is real, we have $A\overline{v} = \overline{\lambda}\overline{v}$.

If $v = [z_1 \quad z_2 \quad \cdots \quad z_n]^\top$, then we have

$$\begin{aligned}
\overline{\lambda}(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2) &= \overline{\lambda}\overline{v}^\top v \\
&= (A\overline{v})^\top v \\
&= \overline{v}^\top A v \\
&= \overline{v}^\top (\lambda v) \\
&= \lambda(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2),
\end{aligned}$$

so $(\overline{\lambda} - \lambda)(|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2) = 0$. Since v is not the zero vector, the second factor is positive, so we must have $\overline{\lambda} = \lambda$, that is, λ is real.

Now since T has a real eigenvalue, we can choose a real eigenvector v, and (multiplying by a scalar if necessary) we can assume that $|v| = 1$.

Let U be the subspace $v^\perp = \{u \in V : v \cdot u = 0\}$. This is a subspace of V of dimension $n-1$. We claim that $T : U \to U$. For take $u \in U$. Then

$$v \cdot T(u) = T^*(v) \cdot u = T(v) \cdot u = \lambda v \cdot u = 0,$$

where we use the fact that T is self-adjoint. So $T(u) \in U$.

So T is a self-adjoint linear map on the $(n-1)$-dimensional inner product space U. By the inductive hypothesis, U has an orthonormal basis consisting of eigenvectors of T. They are all orthogonal to the unit vector v; so, adding v to the basis, we get an orthonormal basis for V, and we are done.

Corollary 8.6 If T is self-adjoint, then eigenvectors of T corresponding to distinct eigenvalues are orthogonal.

Proof This follows from the theorem, but is easily proved directly. If $T(v) = \lambda v$ and $T(w) = \mu w$, then

$$\lambda v \cdot w = T(v) \cdot w = T^*(v) \cdot w = v \cdot T(w) = \mu v \cdot w,$$

so, if $\lambda \neq \mu$, then $v \cdot w = 0$.

Remark: The above result, combined with the Remark prior to Theorem 7.1, tells us that the eigenvectors of a self-adjoint map T corresponding to distinct eigenvalues are linearly independent. But we already knew this, by Lemma 5.8!

Example 8.1 Let

$$A = \begin{bmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{bmatrix}.$$

The characteristic polynomial of A is

$$\begin{vmatrix} x-10 & -2 & -2 \\ -2 & x-13 & -4 \\ -2 & -4 & x-13 \end{vmatrix} = (x-9)^2(x-18),$$

so the eigenvalues are 9 and 18.

For eigenvalue 18 the eigenvectors satisfy

$$\begin{bmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 18x \\ 18y \\ 18z \end{bmatrix},$$

so the eigenvectors are multiples of $\begin{bmatrix}1 & 2 & 2\end{bmatrix}^\top$. Normalising, we can choose a unit eigenvector $\begin{bmatrix}\frac{1}{3} & \frac{2}{3} & \frac{2}{3}\end{bmatrix}^\top$.

For the eigenvalue 9, the eigenvectors satisfy

$$\begin{bmatrix}10 & 2 & 2\\ 2 & 13 & 4\\ 2 & 4 & 13\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix} = \begin{bmatrix}9x\\ 9y\\ 9z\end{bmatrix},$$

that is, $x + 2y + 2z = 0$. (This condition says precisely that the eigenvectors are orthogonal to the eigenvector for $\lambda = 18$, as we know.) Thus the eigenspace is 2-dimensional. We need to choose an orthonormal basis for it. This can be done in many different ways: for example, we could choose $\begin{bmatrix}0 & 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}^\top$ and $\begin{bmatrix}-4/3\sqrt{2} & 1/3\sqrt{2} & 1/3\sqrt{2}\end{bmatrix}^\top$. Then we have an orthonormal basis of eigenvectors. We conclude that, if

$$P = \begin{bmatrix}1/3 & 0 & -4/3\sqrt{2}\\ 2/3 & 1/\sqrt{2} & 1/3\sqrt{2}\\ 2/3 & -1/\sqrt{2} & 1/3\sqrt{2}\end{bmatrix},$$

then P is orthogonal, and

$$P^\top AP = \begin{bmatrix}18 & 0 & 0\\ 0 & 9 & 0\\ 0 & 0 & 9\end{bmatrix}.$$

You might like to check that the orthogonal matrix in Example 7.2 (in the previous chapter) also diagonalises A.