21-241 MATRICES AND LINEAR TRANSFORMATIONS SUMMER 1 2012 COURSE NOTES JUNE 18

PAUL MCKENNEY

1. More on projections and orthogonal subspaces

Here's a question. How do you find (a basis for) the orthogonal subspace of a subspace?

Fact 1. Let A be a size $m \times n$ matrix. Then $\mathrm{col}(A)^\perp = \mathrm{null}(A^H)$.

Proof. Suppose $x \in \mathrm{col}(A)^\perp$. Let $a_1, \ldots, a_n \in \mathbb{C}^m$ be the columns of A. Then, $a_1^H, \ldots, a_n^H$ are the rows of A^H. Hence $a_1^H x, \ldots, a_n^H x$ are the entries of $A^H x$. But for each i we have

$$a_i^H x = \langle x, a_i \rangle = 0$$

Hence $x \in \mathrm{null}(A^H)$. The opposite direction follows from the same translation; if $x \in \mathrm{null}(A^H)$, then x is orthogonal to each of the column vectors of A. But then x is orthogonal to anything in their span. □

Corollary 1. *If V is a subspace of $\mathbb{C}^n$, then $\dim(V) + \dim(V^\perp) = n$.*

Proof. Let $v_1, \ldots, v_m$ be a basis for V, and let A be the $n \times m$ matrix with columns $v_1, \ldots, v_m$. Then $V = \mathrm{col}(A)$, and $V^\perp = \mathrm{col}(A) = \mathrm{null}(A^H)$. But the column rank of A and the row rank of A^H are the same (why?); so by the rank-nullity theorem,

$$\dim(V) + \dim(V^\perp) = \mathrm{rank}(A^H) + \mathrm{nullity}(A^H) = n$$

□

Example. Find a basis for the orthogonal subspace of

$$\mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ 5 \end{pmatrix} \right\}$$

Theorem 1. *Let V be a subspace of $\mathbb{C}^n$. Then,*

(1) $\mathbb{P}_V : \mathbb{C}^n \to \mathbb{C}^n$ is a well-defined linear transformation.
(2) If $v \in V$ then $\mathbb{P}_V(v) = v$.
(3) If $w \in V^\perp$ then $\mathbb{P}_V(w) = 0$.
(4) If P_V is the matrix which implements $\mathbb{P}_V$, then P_V is a projection matrix.

Proof. What we mean in (1) is, if we used two different orthonormal bases $\{u_1, \dots, u_k\}$ and $\{\hat{u}_1, \dots, \hat{u}_k\}$ for V, would we get the same output always? That is, do we have

$$\sum_{i=1}^{k} \langle x, u_i \rangle \, u_i = \sum_{i=1}^{k} \langle x, \hat{u}_i \rangle \, \hat{u}_i$$

for all $x \in \mathbb{C}^n$? The proof is not really significant so long as you understand the problem.

For (2), recall that for any $v \in V$,

$$v = \sum_{i=1}^{k} \langle v, u_i \rangle \, u_i$$

where $u_1, \dots, u_k$ is any orthonormal basis for V. The right-hand-side is our definition of $\mathbb{P}_V(v)$ (now that we know it makes sense).

For (3), let $w \in V^\perp$. Then

$$\mathbb{P}_V(w) = \sum_{i=1}^{k} \langle w, u_i \rangle \, u_i = 0 \cdot u_1 + \cdots + 0 \cdot u_k = 0$$

□

Theorem 2. *Let V be a subspace of $\mathbb{C}^n$. Then for every $x \in \mathbb{C}^n$, $x = \mathbb{P}_V(x) + (x - \mathbb{P}_V(x))$ is the unique decomposition of x into vectors in V and $V^\perp$.*

Proof. To see that this pair is the only one that works, say $v \in V$ and $w \in V^\perp$ is another pair of vectors, such that $x = v + w$. Then we have

$$\mathbb{P}_V(x) = \mathbb{P}_V(v + w) = \mathbb{P}_V(v) + \mathbb{P}_V(w) = v + 0 = v$$

But then $w = x - v = x - \mathbb{P}_V(x)$ as well. □

Proof of (4). Let $x \in \mathbb{C}^n$, and write $P = P_V$. Let $x = v + w$ be the unique decomposition of x into vectors from V and $V^\perp$ respectively. Then,

$$Px = P(v + w) = Pv + Pw = v + 0 = v$$

On the other hand,

$$P^2 x = P(Px) = Pv = v = Px$$

So $Px = P^2 x$ for all $x \in \mathbb{C}^n$, and as we've seen this implies $P = P^2$.

Now let $x, y \in \mathbb{C}^n$ be given. Say $x = v + w$ and $y = s + t$, where $v, s \in V$ and $w, t \in V^\perp$. Then,

$$\langle Px, y \rangle = \langle P(v+w), s+t \rangle = \langle v+0, s+t \rangle = \langle v, s \rangle + \langle v, t \rangle = \langle v, s \rangle$$

Similarly,

$$\left\langle P^H x, y \right\rangle = \langle x, Py \rangle = \langle v+w, P(s+t) \rangle = \langle v+w, s+0 \rangle = \langle v, s \rangle + \langle w, s \rangle = \langle v, s \rangle$$

Hence $\langle Px, y \rangle = \left\langle P^H x, y \right\rangle$ for all $x, y \in \mathbb{C}^n$. This implies $P = P^H$.

□

2. Eigenvalues and Orthogonality

Definition. A matrix $A \in M_n(\mathbb{C})$ is *Hermitian* if $A^H = A$. A is *symmetric* if $A^\top = A$. Note that if A is a real matrix then A is symmetric if and only if A is Hermitian.

Theorem 3. *Suppose A is Hermitian, and $\lambda, \mu \in \mathbb{C}$ are distinct eigenvalues for A. Then $V_\lambda \perp V_\mu$.*

Proof. Let $v \in V_\lambda$ and $w \in V_\mu$ be given. Then,

$$\langle Av, w \rangle = \langle \lambda v, w \rangle = \lambda \langle v, w \rangle$$

and also,

$$\langle Av, w \rangle = \left\langle v, A^H w \right\rangle = \langle v, Aw \rangle = \langle v, \mu w \rangle = \bar{\mu} \langle v, w \rangle$$

By a problem on HW5, λ and μ are actually real numbers; so in particular, $\bar{\mu} = \mu$, and we have

$$\lambda \langle v, w \rangle = \mu \langle v, w \rangle$$

Since $\lambda \neq \mu$, $\langle v, w \rangle = 0$. □

Definition. Let $A \in M_n(\mathbb{C})$, and let λ be an eigenvalue of A. The *geometric multiplicity* of λ (with respect to A) is $\dim(V_\lambda)$. The *algebraic multiplicity* of λ (with respect to A) is the number of times λ appears as a root in the characteristic polynomial of A.

Example. Let

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

Then $\operatorname{spec}(A) = \{2, 5\}$. The geometric multiplicity of 2, with respect to A, is 1; whereas the geometric multiplicity of 5 with respect to A is 2. ($\{e_1\}$ is a basis for the eigenspace of 2, and $\{e_2, e_3\}$ is a basis for the eigenspace of 5.) The characteristic polynomial of A is

$$p_A(\lambda) = \det(\lambda I - A) = (\lambda - 2)(\lambda - 5)^2$$

hence the algebraic multiplicities are 1 (for 2) and 2 (for 5).

Fact 2. The sum of the algebraic multiplicities of the eigenvalues of an $n \times n$ matrix A is exactly n.

Proof. We've seen that the characteristic polynomial of A has degree exactly n; since the eigenvalues of A are the roots of p_A, it follows that their multiplicities must sum up to n. $\square$

The same is not true for the geometric multiplicities. The shift matrix provides the canonical example. (As you should see on HW5.)

Example. Let

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

The characteristic polynomial of A is $\lambda^2 - (-1) \cdot 0 = \lambda^2$, hence the only eigenvalue of A is 0, with algebraic multiplicity 2. The eigenspace is exactly null(A), which has a basis of $\{e_1\}$. Hence the geometric multiplicity of 0 with respect to A is 1.