21-241 MATRICES AND LINEAR TRANSFORMATIONS
SUMMER 1 2012
COURSE NOTES
JUNE 26

PAUL MCKENNEY

Lemma 1. *If A is Hermitian, $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of A, and $V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$, then V and $V^\perp$ are both invariant for A.*

Theorem 1. *If A is Hermitian, then A is diagonalizable by a unitary matrix.*

Proof. Let's say $\lambda_1, \ldots, \lambda_k$ are the eigenvalues of A, without repeats, and their geometric multiplicities are $g_1, \ldots, g_k$. Let

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_k}$$

The lemma we proved yesterday tells us that V and $V^\perp$ are both invariant for A.

Let $v^i_1, \ldots, v^i_{g_i}$ be an orthonormal basis for V_{λ_i}; then the list

$$v^1_1, \ldots, v^1_{g_1}, v^2_1, \ldots, v^2_{g_2}, \ldots, v^k_1, \ldots, v^k_{g_k}$$

is an orthonormal basis for V. Let $m = \sum g_i$. Then $\dim(V) = m$, so $\dim(V^\perp) = n - m$.

If $m = n$, then we're done, for the above list of eigenvectors must be a basis for $\mathbb{C}^n$, and a theorem we've stated before (but not proven) says that this is equivalent to diagonalizability. In the remainder of the proof we will assume that $m < n$, and eventually get a contradiction. The work we do therein will also show how to see that A is diagonalizable when $m = n$, so if you didn't believe the theorem before, that should convince you.

Let $w_1, \ldots, w_{n-m}$ be an orthonormal basis for $V^\perp$; then

$$v^1_1, \ldots, v^1_{g_1}, v^2_1, \ldots, v^2_{g_2}, \ldots, v^k_1, \ldots, v^k_{g_k}, w_1, \ldots, w_{n-m}$$

is an orthonormal basis for $\mathbb{C}^n$. Label these vectors $u_1, \ldots, u_n$, in the order above, and let U be the unitary matrix whose columns are $u_1, \ldots, u_n$.

Now consider the matrix $U^H AU$. Its (i, j)-entry is

$$\left\langle U^H AUe_j, e_i \right\rangle = \langle AUe_j, Ue_i\rangle = \langle Au_j, u_i\rangle$$

Let's work out what these entries are in the various cases.

				u_i		
	$\langle Au_j, u_i \rangle$	v_p^1	v_p^2	$\cdots$	v_p^k	w_p
	v_q^1	λ_1	0	$\cdots$	0	0
	v_q^2	0	λ_2	$\cdots$	0	0
u_j	$\vdots$			$\ddots$		
	v_q^k	0	0	$\cdots$	λ_k	0
	w_q	0	0	$\cdots$	0	$\langle Aw_q, w_p \rangle$

It follows that

$$U^H A U = \begin{pmatrix} \lambda_1 I_{g_1} & & & & \\ & \lambda_2 I_{g_2} & & & \\ & & \ddots & & \\ & & & \lambda_k I_{g_k} & \\ & & & & \hat{A} \end{pmatrix}$$

where $\hat{A}$ is the $n-m \times n-m$ matrix with entries $\langle Aw_j, w_i \rangle$. Now as we've seen before, p_A and $p_{U^H AU}$ are the same. But clearly,

$$p_{U^H AU}(z) = (z-\lambda_1)^{g_1} \cdots (z-\lambda_k)^{g_k} \det(zI_{n-m} - \hat{A})$$

Since the roots of $p_{U^H AU}$ and p_A are the same, and $\lambda_1, \ldots, \lambda_k$ are the roots of p_A, it follows that

$$p_{\hat{A}}(z) = \det(zI_{n-m} - \hat{A}) = (z-\lambda_1)^{a_1} \cdots (z-\lambda_k)^{a_k}$$

where a_i is the algebraic multiplicity of λ_i with respect to $\hat{A}$. (Note that a_i may be 0.)

Let $\hat{v}$ be an eigenvector of $\hat{A}$. (We're using here our assumption that $m < n$, to even talk about $\hat{A}$; if $m = n$ then its size would be "0×0".) Let v be the vector

$$v = U \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \hat{v}_1 \\ \vdots \\ \hat{v}_{n-m} \end{pmatrix}$$

It follows that v is an eigenvector of A, with eigenvalue the same as that of $\hat{v}$ with respect to $\hat{A}$. But, if $p \leq m$,

$$\langle v, u_p \rangle = \left\langle U \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \hat{v}_1 \\ \vdots \\ \hat{v}_{n-m} \end{pmatrix}, U e_p \right\rangle = \left\langle \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \hat{v}_1 \\ \vdots \\ \hat{v}_{n-m} \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\rangle = 0$$

so $v \perp u_p$ for all $p \leq m$. But then v is orthogonal to every eigenspace of A (since u_p, for $p \leq m$, lists basis vectors for all the eigenspaces of A); in particular, if λ_i is the eigenvalue associated to v, then $v \perp V_{\lambda_i}$. Then $v = 0$, but this is a contradiction. $\square$