Math 201 lecture for Friday, Week 13

The Spectral theorem

Theorem (Spectral theorem). Let A be an $n\times n$ symmetric matrix over $\mathbb{R}$. Then A is diagonalizable over $\mathbb{R}$, and there exists an orthonormal basis for $\mathbb{R}^n$ (with respect to the standard inner product) consisting of eigenvectors for A.

Example. Let

$$A=\begin{pmatrix} -1 & -1 & -2\\ -1 & -1 & 2\\ -2 & 2 & 2\end{pmatrix}.$$

The characteristic polynomial of A is

$$p_A(t)=\det\begin{pmatrix} -1-t & -1 & -2\\ -1 & -1-t & 2\\ -2 & 2 & 2-t\end{pmatrix}=-t^3+12t+16=(4-t)(-2-t)^2.$$

So the eigenvalues are $4,-2,-2$. We next compute bases for the eigenspaces. For $\lambda=4$,

$$A-4I_4=\begin{pmatrix} -5 & -1 & -2\\ -1 & -5 & 2\\ -2 & 2 & -2\end{pmatrix}\rightsquigarrow\begin{pmatrix} 1 & 0 & \frac{1}{2}\\ 0 & 1 & -\frac{1}{2}\\ 0 & 0 & 0\end{pmatrix}.$$

So the eigenspace for $\lambda=4$ is $E_4\{(-\frac{1}{2}t,\frac{1}{2}t,t):t\in\mathbb{R}\}$. One basis is $\{(-1,1,2)\}$. Normalizing gives the basis vector

$$v_1=\frac{1}{\sqrt{6}}(-1,1,2).$$

For the eigenvalue $\lambda=-2$, we have

$$A+2I_4=\begin{pmatrix} 1 & -1 & -2\\ -1 & 1 & 2\\ -2 & 2 & 4\end{pmatrix}\rightsquigarrow\begin{pmatrix} 1 & -1 & -2\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

and the eigenspace is $E_{-2}=\{(s+2t,s,t):s,t\in\mathbb{R}\}$. A basis is $\{(1,1,0),(2,0,1)\}$. Applying Gram-Schmidt to these two vectors yields an orthonormal basis for E_{-2} consisting of

$$v_2=\frac{1}{\sqrt{2}}(1,1,0),\quad v_3=\frac{1}{\sqrt{3}}(1,-1,1).$$

Now note that something surprising has happened: these vectors are orthogonal to v_1. We arrive at an orthonormal basis $\{v_1,v_2,v_3\}$ for $\mathbb{R}^3$ consisting of eigenvectors for A. Letting P be the 3×3 matrix whose columns are v_1,v_2,v_3, we have

$$P^{-1}AP=\mathrm{diag}(4,-2,2).$$

Since the v_i form an orthonormal set, it turns out that $P^{-1}=P^t$, the transpose of P:

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

Definition. A matrix $P \in M_{n\times n}(\mathbb{R})$ is *orthogonal* if its columns form an orthonormal set in $\mathbb{R}^n$.

Lemma. $P \in M_{n\times n}(\mathbb{R})$ is orthogonal if and only if $P^{-1} = P^t$.

Proof. Note that $(P^tP)_{ij} = v_i \cdot v_j$. So $P^tP = I_n$ if and only if the columns of P form and orthonormal set. $\square$

Restatement of the spectral theorem. If A is a real $n \times n$ symmetric matrix, then there exists a real diagonal matrix D and an orthogonal matrix P such that

$$A = PDP^t.$$

Proof of the spectral theorem. We first prove that the characteristic polynomial of A splits over $\mathbb{R}$. By the Fundamental Theorem of Algebra, it splits over $\mathbb{C}$. So $p_A(t) = \prod_{k=1}^{n}(\lambda_k - t)$ for some $\lambda_k \in \mathbb{C}$. We must show that $\lambda_k \in \mathbb{R}$ for all k. So let $\lambda = \lambda_k$ for some k. Then there exists a nonzero $v \in \mathbb{C}^n$ such that $Av = \lambda v$. Recall the standard inner product on $\mathbb{C}^n$: for $y, z \in \mathbb{C}^n$, we have $\langle y, z\rangle = y \cdot \bar{z}$. Thinking of y and z as column vectors, we have $\langle y, z\rangle = z^*y$ where $(\ )^*$ denotes the conjugate transpose:

$$\langle y, z\rangle = y \cdot \bar{z} = \sum_{k=1}^{n} y_i \bar{z}_i = \begin{pmatrix} \bar{z}_1 & \cdots & \bar{z}_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = z^*y.$$

Therefore, for an arbitrary $n \times n$ complex matrix B, we have

$$\langle y, B^*z\rangle = (B^*z)^*y = z^*(B^*)^*y = z^*By = \langle By, z\rangle.$$

Our matrix, A, is real and symmetric; so $A^* = \bar{A}^t = A^t = A$. Therefore,

$$\langle y, Az\rangle = \langle Ay, z\rangle.$$

Going back to $Av = \lambda v$, we have

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \bar{\lambda}\langle v, v\rangle.$$

Since $v \neq 0$ and inner products are positive-definite, it follows that $\lambda = \bar{\lambda}$, and hence $\lambda \in \mathbb{R}$.

We now prove the theorem by induction on n, the case $n = 1$ being trivial. Suppose $n > 1$ and let $\lambda_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^n$ be an eigenvalue-eigenvector pair for A. Next, complete and apply Gram-Schmidt to construct and ordered orthonormal basis $\langle v_1, \cdots, v_n\rangle$ for $\mathbb{R}^n$. Let Q be the $n \times n$ matrix whose columns are the v_i. Then Q is orthogonal. Define

$$\tilde{A} := Q^{-1}AQ = Q^tAQ.$$

Then $\tilde{A}$ is symmetric:

$$\tilde{A}^t = \left(Q^tAQ\right)^t = Q^tA^t\left(Q^t\right)^t = Q^tA^tQ = Q^tAQ = \tilde{A}.$$

We would like to investigate the structure of $\tilde{A}$ further. To find its first column, we use the fact that $Av_1 = \lambda_1 v_1$. Let e_1 be the first standard basis vector of $\mathbb{R}^n$. Then the first column of $\tilde{A}$ is

$$\tilde{A}e_1 = Q^t A Q e_1 = Q^t A v_1 = Q^t \lambda_1 v_1 = \lambda_1 Q^t v_1.$$

The rows of Q^t are the orthonormal set $v_1, \ldots, v_n$. Therefore,

$$(Q^t v_1)_i = v_i \cdot v_1 = \begin{cases} 1 & \text{if } i = 1 \\ 0 & \text{otherwise.} \end{cases}$$

So the first column of $\tilde{A}$ is the vector $(\lambda_1, 0, \cdots, 0)$. Since $\tilde{A}$ is symmetric, its first column and first row are the same vector. Therefore, $\tilde{A}$ has the form

$$\left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$

where B is an $n \times n$ matrix. Since $\tilde{A}$ is symmetric, so is B. So we can apply induction to find an $(n-1) \times (n-1)$ orthogonal matrix T and a real diagonal matrix E such that $B = TET^t$. We then have

$$\tilde{A} = \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T & \\ 0 & & & \end{array}\right)}_{S} \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & E & \\ 0 & & & \end{array}\right)}_{D} \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T^t & \\ 0 & & & \end{array}\right)}_{S^t},$$

where the matrices S and T are defined as shown. Since T is orthogonal, so is S. Finally, define $P = QS$. Since Q and S are orthogonal, so is P (check: $(QS)^t(QS) = S^t(Q^tQ)S = S^t I_n S = I_n$). We have

$$A = Q\tilde{A}Q^t = Q(SDS^t)Q^t = (QS)D(QS)^t = PDP^t,$$

as desired. □

We now discuss a more general version of the spectral theorem.

Definition. A matrix $A \in M_{n\times n}(\mathbb{C})$ is *Hermitian* if $A^* = A$ (so $A = \bar{A}^t$). A matrix $U \in M_{n\times n}(\mathbb{C})$ is *unitary* if its columns are orthonormal, or equivalently, if U is invertible with $U^{-1} = U^*$.

Theorem (Spectral theorem) Let A be an $n \times n$ Hermitian matrix. Then $A = UDU^*$ where U is unitary and D is a real diagonal matrix.