Lecture 4-30

We learned last time that given any real symmetric matrix A we have $U^T AU = U^{-1}AU = D$ for some orthogonal matrix U and diagonal matrix D; the (real) diagonal entries of D are the eigenvalues of A. (Note that conversely any matrix of the form $U^{-1}DU$ is symmetric if U is orthogonal and D is diagonal.) Previously we had shown by a completely different argument that we have $P^T AP = D'$ for some invertible matrix P and diagonal matrix D', where P is *not* necessarily orthogonal. We also know that A is positive definite if and only if the diagonal entries of D are positive, or if and only if the diagonal entries of D' are positive. We now want to recall that we obtained the matrix D' by performing row and column operations on A in such a way that throughout multiples of higher rows in A were added to lower ones and multiples of columns to the left were added to columns to the right. These operations preserve not only the determinant of A but also the determinant of the submatrix $A^{(i)}$ consisting of the entries of the first i rows and columns of A, for all i between 1 and n (since every such submatrix is either unaffected by any such operation, or else has the same operation performed on it; this is where it is important that multiples of lower-numbered rows and columns are added to higher-numbered ones rather than the reverse). Moreover, if A is positive definite, then none of the pivots occurring in the ii-entries of any matrix B arising during the row reduction can be zero, since A is positive definite if and only if B is, and if the ii-th entry b_{ii} of B is 0, then $\vec{e}_i B \vec{e}_i^T = 0$, where as usual $\vec{e}_i$ is the i-th unit coordinate vector in $\mathbb{R}^n$, written as a row vector. We conclude that *a symmetric matrix A is positive definite if and only if all submatrices $A^{(i)}$ as defined above have positive determinant*; we had previously observed this same necessary and sufficient condition for positive definiteness of 2×2 matrices. Similarly, *a symmetric matrix A is negative definite if and only if* $\det A^{(i)}$ *is positive for i even and negative for i odd*, since this is the criterion for a diagonal matrix to be negative definite. Finally, *if all determinants* $\det A^{(i)}$ *are nonzero, then any other pattern of signs of the determinants $A^{(i)}$ implies that A is indefinite*. (Note however that even if some $\det A^{(i)}$ is 0, this does not necessarily imply that some eigenvalue of A is 0, as the example $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ shows.)

More generally, if A is Hermitian, then we have $\bar{U}^T AU = U^{-1}AU = D$ for some (usually complex) matrix U (called *unitary*, since $\bar{U}^T = U^{-1}$) and real diagonal D. This last result is called the *spectral theorem for Hermitian matrices*; actually, there are a number of spectral theorems, all of them asserting that matrices satisfying a certain condition on their conjugate transposes are always diagonalizable. In the special case of real symmetric matrices (where the matrix U is orthogonal), the spectral theorem is sometimes called the *principal axis theorem*. Note that *complex* symmetric matrices, in contrast to real ones or complex Hermitian ones, are nothing special; in particular, they need not be diagonalizable.

There is a small modification of the dot product which turns out to be much more useful than the dot product itself for complex vector spaces, which we should introduce explicitly; it is called the *Hermitian inner product* and is denoted $(\vec{v}, \vec{w})$ rather than $\vec{v} \cdot \vec{w}$ for $\vec{v}, \vec{w} \in \mathbb{C}^n$. If $\vec{v} = (v_1, \ldots, v_n), \vec{w} = (w_1, \ldots, w_n)$, then $(\vec{v}, \vec{w})$ is defined to be $\sum_i v_i \bar{w}_i$; note that this quantity depends linearly on $\vec{v}$ but *conjugate*-linearly on $\vec{w}$, so that $(\vec{v}, \alpha\vec{w}) =$

$\bar{\alpha}(\vec{v}, \vec{w})$ for $\alpha \in \mathbb{C}$. Also we have $(\vec{w}, \vec{v}) = \overline{(\vec{v}, \vec{w})}$ rather than $(\vec{v}, \vec{w})$; this product is conjugate symmetric rather than symmetric. If is more useful than the dot product since it is *positive definite*: we have $(\vec{v}, \vec{v}) \in \mathbb{R}$ and $(\vec{v}, \vec{v}) > 0$ if $\vec{v} \neq \vec{0}$; by contrast, $\vec{v} \cdot \vec{v}$ need not be real for $\vec{v} \in \mathbb{C}^n$ and even if it is real is need not be positive for $\vec{v} \neq \vec{0}$.

An $n \times n$ real matrix U is orthogonal if and only if its columns form an orthonormal basis of $\mathbb{R}^n$ (one consisting of pairwise orthogonal unit vectors), or if and only if its P*rows* form an orthonormal basis of $\mathbb{R}^n$, or if and only if $Uv \cdot Uw = v \cdot w$ for all $v, w \in \mathbb{R}^n$. Indeed, the columns of U are on orthonormal basis if and only if $U^T U = I$, or if and only if $UU^T = I$, or if and only if the rows of U are an orthonormal basis, or if and only if $Uv \cdot Uw = v^t U^T U w = v^T w = v \cdot w$ for all $v, w \in \mathbb{R}^n$, writing v and w as column vectors. The corresponding condition on unitary matrices V is that the columns of V, or its rows, form an orthonormal basis of $\mathbb{C}^n$ with respect to the Hermitian inner product.

More generally, for any real $n \times n$ matrix A and any $v, w \in \mathbb{R}^n$, we have $Av \cdot w = w^T Av = v^T A^T w = v \cdot A^T w$ (writing v, w as column vectors). This last equation gives the precise relationship between the linear transformations of multiplication by A and multiplication by A^T; note that we need the dot product to formulate this relationship. It is expressed in words by saying that A^T is the *adjoint* of A. A similar remark holds for $\bar{A}^T$ and the Hermitian inner product; indeed, another name for the conjugate transpose of A is its adjoint.