Lecture 5-28

We now return (for the last time) to real symmetric matrices. We have seen that any such matrix A may be diagonalized by an orthogonal matrix U, so that $A = U^T D U = U^{-1} D U$ and the matrices D and U are real. If the diagonal entries of D are $d_1, \ldots, d_n$, we now rewrite D as $E^T D' E$, where $E^T = E = \begin{pmatrix} e_1 & & & \\ & e_2 & & \\ & & \ddots & \\ & & & e_n \end{pmatrix}$, with $e_i = \sqrt{|d_i|}$ if $d_i \neq 0$ and $e_i = 1$ if $d_i = 0$; here D' is diagonal with i-th diagonal entry sgn d_i (equal to 1 if d_i is positive, -1 if d_i is negative, and 0 if $d_i = 0$). Thus we have $A = Q^T D' Q$, where the matrix Q is invertible (though not necessarily orthogonal) and the diagonal matrix D' has all entries equal to ± 1 or 0. We say that P is *congruent* (but not in general similar) to D'. If the matrix D' has p diagonal entries equal to 1 and q entries equal to -1 then we saw that P has *signature* (p, q).

This terminology seems to indicate that the signature of a symmetric matrix is uniquely determined by the (congruence class of) the matrix, so that two congruent symmetric matrices have the same signature. This is indeed the case; the theorem asserting it is called *Sylvester's law of inertia* in the LADW text. If the $n \times n$ symmetric matrices A and A' are congruent, then first of all they must have the same rank, since multiplication by an invertible matrix on the left or right preserves rank, whence the number t of zeroes along the diagonal of any diagonal matrix congruent to A is uniquely determined (as n minus the rank of A). Now if $A = Q^T DQ$ is congruent to the diagonal matrix D with all entries ± 1 or 0 and signature (p, q), then there is an obvious subspace S of $\mathbb{R}^n$ of dimension $p + t$ consisting entirely of vectors v such that $v^T DV \geq 0$; here S is the span of the columns of D whose diagonal entries are 1 or 0. Similarly there is a subspace T of $\mathbb{R}^n$ of dimension q consisting of vectors w with $w^T Dw < 0$ if $w \neq 0$, spanned by the columns of D whose diagonal entries are -1. The matrix Q^{-1} then takes the subspaces S, T (called *positive semidefinite* and *negative definite* for D) to subspaces S', T' with the same dimensions that are respectively positive semidefinite and negative definite for A.

Now whenever two subspaces U, W of a finite-dimensional vector space V have dimensions adding to more than the dimension of V, we must have $U \cap W \neq 0$, since the union of a basis for U and a basis for W *cannot* be a basis for V (having too many vectors), so that some nonzero combination of vectors in the first basis must coincide with another such combination for the second basis. It follows at once that two congruent symmetric matrices cannot have different signatures. We already know that two such matrices have the same rank and thus the same number of zero diagonal entries in any diagonalization. If they had different signatures then the positive semidefinite subspace attached by the preceding paragraph to one of them would have to have nonzero intersection with the negative definite subspace attached to the other, which is impossible.

We have seen that there is a natural bijection between quadratic forms (homogeneous quadratic polynomials p in n variables $x_1, \ldots, x_n$) and symmetric matrices A, any such p sending the column vector $v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ to $v^T A v$ for a unique A. Replacing the matrix A by a congruent one $Q^T A Q$ in effect just changes the variables, replacing v by $Qv = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$.

We deduce from the above analysis that *any quadratic form in n variables can be written as the difference of two sums of squares of independent linear functions of the variables, with the total number of sums involved at most n.* The linear functions are not unique, but the number of such functions whose squares appear with coefficient 1 is uniquely determined by the form, as is the number of such functions whose squares appear with coefficient -1. The form is positive definite if and only if it is the sum of the squares of n independent linear functions and positive semidefinite if and only if it is the sum of the squares of $m \leq n$ such functions. In general we can take the linear functions appearing in the difference of the two sums of squares to correspond to orthogonal (though not necessarily orthonormal) vectors in $\mathbb{R}^n$. Quadratic forms have well-defined signatures which measure in a precise way the extent to which they fail to be positive or negative definite. In general polynomials in n variables of degree more than 2 taking only nonnegative values need *not* be sums of squares; such polynomials continue to be intensively studied in current research.