Lecture 4-28

Continuing from last time, we now present a simple example of a non-diagonalizable matrix, namely $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Computing $p(\lambda) = \det(A - \lambda I)$, we get $(1-\lambda)^2$, so A has 1 as its unique eigenvalue. (This is no accident; we will see below that any matrix whose characteristic polynomial factors into *distinct* linear factors is diagonalizable.) The multiplicity of 1 as a root of $p(\lambda)$ is 2, but in solving the system $(A - I)X = 0$ we find that there is only one free variable. The 1-eigenspace of A, or equivalently the nullspace of $A - I$, is spanned by the single vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Since this is the only eigenvector of A up to scalar multiple, there is no basis of $\mathbb{R}^2$ consisting of eigenvectors of A and A is not diagonalizable, as claimed.

Now if $\lambda_1, \ldots, \lambda_m$ are *distinct* eigenvalues of a matrix B, with corresponding eigenvectors $v_1, \ldots, v_m$, then the v_i are independent. To prove this, suppose contrarily that we had a dependence relation $\sum a_i v_i = 0$ among the v_i in this situation, with m as small as possible; then it follows that $a_i \neq 0$ for all i (lest there be a dependence relation with fewer than m terms). Taking the product $A \sum a_i v_i = \sum a_i \lambda_i v_i = 0$ and subtracting from this $\lambda_1 \sum a_i v_i$, we get a dependence relation in which v_1 does not appear but $v_2, \ldots, v_m$ do, each with nonzero coefficient, since $\lambda_1 \neq \lambda_i$ for $i \neq 1$. This dependence relation has fewer than m terms, a contradiction, so the v_i must be independent, as claimed. In particular, if the characteristic polynomial of an $n \times n$ matrix A has n distinct roots $\lambda_1, \ldots, \lambda_n$, each with corresponding eigenvector $v_1, \ldots, v_n$, then the v_i are independent and so must form a basis of $\mathbb{R}^n$. Thus A is diagonalizable, as claimed. More generally, a square matrix A fails to be diagonalizable if and only if the dimension of its λ-eigenspace (called its *geometric multiplicty* is less than the multiplicity of λ as a root of its characteristic polynomial (called its *algebraic multiplicity*, as with the matrix A in the first paragraph).

Given two similar square matrices A and $P^{-1}AP$, we have $\det(P^{-1}AP - \lambda I) = \det(P^{-1}(A - \lambda I)P) = (\det P)^{-1} \det(A - \lambda I) \det P = \det(A - \lambda I)$, by the product rule for determinants. Thus *similar matrices have the same characteristic polynomial* (and thus also the same eigenvalues, as we observed earlier). We also see from this that the characteristic polynomial of a linear transformation $f : V \to V$ is well defined for any finite-dimensional vector space V, being the characteristic polynomial of the matrix of f with respect to any basis of V. We can now derive two important links between numbers attached to a matrix A and its characteristic polynomial $p(\lambda)$. First of all, plugging in $\lambda = 0$, we get $\det A$ as the constant term of $p(\lambda)$; but now recall that the constant term of any polynomial of degree n whose leading coefficient (like that of $p(\lambda)$) is $(-1)^n$, is the product of the roots. Hence *the determinant of any square matrix is the product of its eigenvalues, counting each with its multiplicity as a root of the characteristic polynomial.* Similarly, but a bit more subtly, one shows that *the trace of a square matrix is the sum of its eigenvalues, counting each with multiplicity in the same way.* To do this, note from the definition of $p(\lambda) = \det(A - \lambda I)$ that the coefficient of λ^n in this polynomial is $(-1)^n$, since the only contribution to this power of λ arises from taking the $-\lambda$ from each term of the product P of the $d_i - \lambda$, where the d_i are the diagonal entries of the matrix, while the coefficient of λ^{n-1} in this polynomial is $(-1)^{n-1}$ times the sum of the d_i, since the only contribution to this power of λ arises from taking d_i from the term $d_i - \lambda$ in P for a unique i and $-\lambda$ from all other terms. Then one concludes by recalling that the next to leading coefficient of any polynomial divided by its leading coefficient equals the negative of the sum of its roots. In particular, if two 2×2 matrices have the same trace and determinant, then they have the same characteristic polynomial (though they need not be similar: observe that the matrix A in the first paragraph and the 2×2 identity matrix I both have characteristic polynomial $(1 - \lambda)^2$).

In general, it is quite difficult to compute eigenvalues of large square matrices; this is fundamentally more complicated than solving large linear systems. There are two special cases, however, for which eigenvalues are easy to read off, namely those of an upper or lower triangular matrix (with either all entries below the main diagonal, or all entries above this diagonal, equal to 0; we call a matrix *triangular* if it is either upper or lower triangular). We have already observed that the determinant of a triangular matrix is the product of its diagonal entries; it follows at once that the eigenvalues of a triangular matrix are its diagonal entries (though the eigenvectors are not necessarily the unit coordinate vectors, as they would be for a diagonal matrix; note also that triangular matrices need not be diagonalizable).