Lecture 5-18

We now combine our earlier work with infinite series with linear algebra, considering a particular infinite series of matrices. Let A be an $n \times n$ matrix. The powers A^k of A all make sense and are $n \times n$ matrices for $k \geq 0$ (we take A^0 to be the identity matrix I), so we can form the infinite series $e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$. This series converges for any A (that is, the series for each of its entries converges), since if the entries of A are bounded in absolute value by M, then an easy induction shows that the entries of A^k are bounded by $n^{k-1}M^k$ and $\sum \frac{n^{k-1}M^k}{k!}$ converges for any M, so that the series for any entry of e^A converges absolutely (by comparison with the series just mentioned) and so converges. Letting t be a real variable we see by differentiating the matrix-valued power series $e^{At} = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$ term by term (and entry by entry) that $(e^{At})' = Ae^{At}$; here we differentiate a matrix-valued function by differentiating each of its column vectors (and thus each of its entries) in turn.

Hence every column of e^{At}, regarded as a vector-valued function $\vec{v}(t)$ of t, satisfies the system of equations $\vec{v}'(t) = A\vec{v}(t)$ (note the analogy between this and a linear system of equations). Moreover, given any two $n \times n$ matrices A, B, we have the law of exponents $e^{A+B} = e^A e^B$, *provided* that $AB = BA$; indeed, we can prove this law for real numbers directly from the power series for e^x, using the binomial theorem, and this theorem carries over to give the obvious formula for any power $(A + B)^k$ whenever A and B commute. In particular, we have $e^{As}e^{At} = e^{A(s+t)}$ for any real numbers s, t; this shows that the matrix e^{At} is always invertible with inverse e^{-At}. Thus the columns of the matrix e^{At} provide n linearly independent vector-valued solutions to the system $\vec{v}'(t) = A\vec{v}(t)$, whence (analogously to any n linearly independent solutions to a linear n-th order homogeneous differential equation) the general solution to $\vec{v}'(t) = A\vec{v}(t)$ is a constant linear combination of the columns of e^{At}. We call e^{At} a *fundamental matrix* for the system $\vec{v}'(t) = A\vec{v}(t)$.

How can the exponential e^{At} be computed? The clue comes from our earlier observation that if a matrix A is diagonalizable, so that $A = P^{-1}DP$ for some diagonal matrix D, then $A^k = P^{-1}D^kP$ for all $k \geq 0$. It follows at once that $e^{At} = P^{-1}e^{Dt}P$ in this case; moreover, an easy direct computation shows that if $d_1, \ldots, d_n$ are the diagonal entries of D, then e^D is again diagonal, with diagonal entries $e^{d_1}, \ldots, e^{d_n}$ (and similarly of course for e^{Dt}). In particular, *the determinant* $\det e^A = \det e^D = e^t$, where t is the trace of A. In fact it is not difficult to show that this last formula holds for any square matrix A, diagonalizable or not, and in fact the eigenvalues of e^A are exactly the e^λ as λ runs through the eigenvalues of A, each occurring with its algebraic multiplicity.

How can we compute e^{At} if A is not diagonalizable? We still have $e^{At} = P^{-1}e^{Bt}P$ if $A = P^{-1}BP$, so we look for a matrix B similar to A such that e^{Bt} can be computed directly. Here a tool called the *Jordan (canonical) form* (which we will state but not prove) comes into play. It turns out that any A is similar to what is called a *block diagonal* matrix J with square matrices $J_1, \ldots, J_k$ marching down the main diagonal from left to right and all other entries zero, where $J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}$ has λ_i as its unique diagonal entry, ones above the diagonal, and zeroes everywhere else.

The J_i are called *Jordan blocks*; A is diagonalizable if and only if all of its Jordan blocks has size 1×1 (so that there is no diagonal above their main diagonals). Any Jordan block J_i may be written as the sum $\lambda_i I + N_i$ of the scalar matrix $\lambda_i I$ (where λ_i is the unique diagonal entry of J_i) and N is a matrix with zeroes on the diagonal, ones above it, and zeroes everywhere else. The matrices $\lambda_i I$ and N_i commute and an easy computation shows that $N_i^{n_i} = 0$ if N_i has size $n_i \times n_i$. Hence the binomial theorem applies in computing the powers $(\lambda_i I + N_i)^k$ and only finitely many powers N_i^k are nonzero. Thus it is not too hard to compute the exponentials $e^{J_i t}$; moreover, it is easy to see that e^{Jt} is again block diagonal with the $e^{J_i t}$ as its blocks.

Note that the nondiagonalizable matrix $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ that we saw three weeks ago is the simplest example of a Jordan block (having size 2×2). One remarkable consequence of the Jordan form is that *up to similarity there are only finitely many $n \times n$ matrices having only one fixed eigenvalue α.* This is because the Jordan form of any such matrix has α as its only diagonal entry and the numbers of rows of its Jordan blocks have to add to n, so there are only finitely many possibilities for its Jordan blocks.