MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 12

Computing powers of matrices

It is worth remarking that Jordan normal forms are mainly useful for "theoretical" applications. The problem is that the corresponding result is too sensitive to "small perturbations": if we only know entries of a matrix approximately (which is always the case in real life), then we cannot really make any conclusions about its Jordan normal form: for example, as we just saw in the previous lecture, for every matrix we can alter its entries arbitrarily small so that the eigenvalues of the resulting matrix are all distinct, so it can be diagonalised, and the Jordan decomposition is essentially trivial: all threads are of length one.

The main theoretical application of Jordan normal forms is to computing functions of matrices. Suppose that A is a given matrix, that J is the Jordan form of A, and that C is the matrix whose columns form a Jordan basis for A. Then, of course, C is the transition matrix to the Jordan basis, so $C^{-1}AC = J$, and $A = CJC^{-1}$. This implies that $A^m = CJ^mC^{-1}$. To compute J^m, we need to compute powers of individual Jordan blocks. It turns out that for a Jordan block

$$J_k(\lambda) = \begin{pmatrix} \lambda & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & \lambda & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \lambda & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 & \lambda \end{pmatrix}$$

we have

$$J_k(\lambda)^m = \begin{pmatrix} \lambda^m & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ m\lambda^{m-1} & \lambda^m & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ \binom{m}{2}\lambda^{m-2} & m\lambda^{m-1} & \lambda^m & 0 & \dots & \dots & \dots & \dots & 0 \\ \binom{m}{3}\lambda^{m-3} & \binom{m}{2}\lambda^{m-2} & m\lambda^{m-1} & \lambda^m & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \binom{m}{k}\lambda^{m-k} & \binom{m}{k-1}\lambda^{m-k+1} & \binom{m}{k-2}\lambda^{m-k+2} & \dots & \dots & \lambda^m & \dots & \dots & 0 \\ 0 & \binom{m}{k}\lambda^{m-k} & \binom{m}{k-1}\lambda^{m-k+1} & \dots & \dots & \dots & \lambda^m & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \dots & \dots & m\lambda^{m-1} & \lambda^m & 0 \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & m\lambda^{m-1} & \lambda^m \end{pmatrix}.$$

(For $m > k$, there will be no zeros below the diagonal, etc., so this formula should be used carefully).

Indeed, if we let $J_k(\lambda) = \lambda I + N$, where $N = J_k(0)$ is the Jordan block corresponding to one thread for a nilpotent transformation, then

$$J_k(\lambda)^m = (\lambda I + N)^m = (\lambda I)^m + \binom{m}{1}(\lambda I)^{m-1}N + \binom{m}{2}(\lambda I)^{m-2}N^2 + \ldots,$$

because the binomial formula for computing $(a+b)^m$ works whenever $ab = ba$, and it remains to compute powers of N. Since N moves vectors along the thread, N^2 moves vectors two steps ahead, N^3 moves vectors three steps ahead, etc. But moving vectors p steps ahead is done by a matrix looking almost like N, but ones on the p-th diagonal below the main diagonal, which corresponds precisely to what we claim.

This gives an easy method for computing powers of matrices. Let A be a matrix, viewed as a linear transformation of $\mathbb{C}^n$, let J be its Jordan normal form, and let C be the matrix whose columns are made of coordinates of vectors of some Jordan basis. Then C is the transition matrix from the standard basis of $\mathbb{C}^n$ to the given Jordan basis, and so $J = C^{-1}AC$ because of the general formulas for matrices representing the given linear transformation relative to different bases. Hence $A = CJC^{-1}$, and the matrix A^m is equal to CJ^mC^{-1}. Computing J^n amounts to computing powers of individual Jordan blocks, which we already know how to do.

Examples of computations

One instance where it is useful to compute powers of matrices is when dealing with recurrent sequences. An example of that sort was considered before in the first semester when we dealt with Fibonacci numbers. We shall now consider a similar question where however Jordan decomposition will be important.

Let us consider the sequence defined as follows: $x_0 = 7$, $x_1 = 3$, $x_{n+2} = -10x_{n+1} - 25x_n$ for $n \geqslant 0$. In order to find a closed formula for x_n (that is, a formula that expresses it in terms of n, without the need to compute all the previous terms of the sequence), we consider the vectors $v_n = \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix}$, for which we have

$$v_{n+1} = \begin{pmatrix} x_{n+1} \\ x_{n+2} \end{pmatrix} = \begin{pmatrix} x_{n+1} \\ -10x_{n+1} - 25x_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -25 & -10 \end{pmatrix}\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = Av_n,$$

where $A = \begin{pmatrix} 0 & 1 \\ -25 & -10 \end{pmatrix}$. Thus, $v_n = A^n v_0$, so in order to compute x_n, it is enough to find a formula for A^n.

We have $\chi_A(t) = \det(A - tI) = t^2 + 10t + 25 = (t+5)^2$, so -5 is the only eigenvalue. The kernel of $A + 5I$ is spanned by the vector $\begin{pmatrix} 1 \\ -5 \end{pmatrix}$. We take the vector $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ outside the kernel of $A + 5I$ that would compensate for the missing pivot; we have $(A + 5I)v = \begin{pmatrix} 1 \\ -5 \end{pmatrix}$, so the columns of $C = \begin{pmatrix} 0 & 1 \\ 1 & -5 \end{pmatrix}$ form a Jordan basis. Thus, $C^{-1}AC = \begin{pmatrix} -5 & 0 \\ 1 & -5 \end{pmatrix}$, and

$$A^n = C\begin{pmatrix} -5 & 0 \\ 1 & -5 \end{pmatrix}^n C^{-1} = \begin{pmatrix} (-5)^n - n(-5)^n & n(-5)^{n-1} \\ -n(-5)^{n+1} & (-5)^n + n(-5)^n \end{pmatrix}.$$

Finally, $v_n = A^n \begin{pmatrix} 7 \\ 3 \end{pmatrix} = \begin{pmatrix} (-5)^{n-1}(38n - 35) \\ (-5)^n(38n + 3) \end{pmatrix}$, so $x_n = (-5)^{n-1}(38n - 35)$.

Another example where computing powers of matrices is important is suggested by probabilistic models known as Markov chains. Let us mention a simple example. Suppose that a particle can be in two states, that we call 1 and 2. Suppose that if it is in the state 1, then with probability p_{11} it remains in that state in one second, and with probability p_{12} changes to the state 2, and similarly, if it is in the state 2, then with probability p_{21} it changes to the state 1, and with probability p_{22} remains in the same state (of course, $p_{11} + p_{12} = 1$ and $p_{21} + p_{22} = 1$). Then, if in the beginning we only know that the particle is in the state 1 with probability p and in the state 2 with probability $q = 1 - p$, then in one second the probabilities change to $p' = pp_{11} + qp_{21}$ and $q' = pp_{12} + qp_{22}$, in other words,

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix},$$

and the probabilities after n seconds are computed using the n-th power of the "transfer matrix" $\begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix}$.

Finally, another instance when powers of matrices will be really useful is the case of ordinary differential equations with constant coefficients. Suppose that $\mathbf{x}(t)$ is a vector function satisfying the differential equation $\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t)$, where A is a given matrix that does not depend on t. Then, it is easy to prove that we have $\mathbf{x}(t) = e^{tA}\mathbf{x}(0)$, where

$$e^{tA} := I + tA + \frac{t^2A^2}{2} + \cdots + \frac{t^nA^n}{n!} + \cdots$$

Thus, using Jordan forms, one can solve all differential equations of this kind.