1212: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 22

Jordan decomposition theorem

Let us state once again the result that we proved over the course of two weeks, and emphasize various related subtleties. First and foremost, it is important to recall that for this result to hold in full generality we have to assume that scalars are complex numbers (so that every linear transformation has eigenvectors).

Let V be a finite-dimensional vector space. For a linear transformation $\varphi\colon V \to V$, there exist

- a decomposition of V
$$V = V_1 \oplus V_2 \oplus \ldots \oplus V_p$$
into a direct sum of invariant subspaces of A;
- a basis $e_1^{(i)}$, ..., $e_{n_i}^{(i)}$ of V_i for each $i = 1, \ldots, p$ such that
$$(\varphi - \lambda_i I)e_1^{(i)} = e_2^{(i)}; (\varphi - \lambda_i I)e_2^{(i)} \qquad = e_3^{(i)}; \ldots (\varphi - \lambda_i I)e_{n_i-1}^{(i)} = e_{n_i}^{(i)} (\varphi - \lambda_i I)e_{n_i}^{(i)} \qquad = 0$$
for some λ_i (which may coincide or be different for different i).

Dimensions of these subspaces and numbers λ_i are determined uniquely up to re-ordering. Relative to this basis the matrix representing φ is a block-diagonal matrix made of Jordan blocks

$$J_{n_i}(\lambda_i) = \begin{pmatrix} \lambda_i & 0 & 0 & 0 & \ldots & 0 & 0 \\ 1 & \lambda_i & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \lambda_i & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & \lambda_i & 0 \\ 0 & 0 & 0 & 0 & \ldots & 1 & \lambda_i \end{pmatrix}$$

placed along the diagonal. Each block corresponds to a thread constructed above. This matrix is called the Jordan normal form of φ, and a basis relative to which φ has this matrix is referred to as a Jordan basis for φ.

The uniqueness of lengths of threads as well as the numbers λ_i follows from our previous results. First, the numbers λ_i are eigenvalues of φ. Furthermore, for each eigenvalue λ, we consider the subspaces $\mathrm{Ker}((\varphi - \lambda I)^k)$, search for where those subspaces stabilise, consider the corresponding subspace of V where $\varphi - \lambda I$ is automatically nilpotent, and use our results on nilpotent transformations. When dealing with nilpotent transformations, we established that for each m the number of threads of length m is determined from dimensions of kernels of powers, which is what we need.

Applications of Jordan forms

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, 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 1.

The main theoretical application of Jordan normal forms is to computing functions of matrices. Let us give two illustrations.

Computing powers of matrices

Let us note 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

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

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.

Statement of Cayley–Hamilton theorem

The celebrated Cayley–Hamilton theorem states that "every matrix is a root of its own characteristic polynomial", that is that if we consider the characteristic polynomial $\chi_A(t) = \det(A - tI) = a_0 + a_1 t + \cdots + a_n t^n$ for the given matrix A, then we have

$$\chi_A(A) = a_0 I + a_1 A + \cdots + a_n A^n = 0.$$

Of course, it is tempting to say that this theorem is obvious, because $\chi_A(A) = \det(A - A \cdot I) = \det(0) = 0$. However, $A - tI$ is a matrix whose entries depend on t, and we cannot simply substitute $t = A$ in those entries! Another way to see the problem is to note that our "proof" would be equally applicable to $\operatorname{tr}(A - tI) = \operatorname{tr}(A) - t \operatorname{tr}(I) = \operatorname{tr}(A) - nt$, but substituting $t = A$ in that polynomial yields $\operatorname{tr}(A)I - nA$, which only is equal to zero for matrices A proportional to I. Next time we shall discuss a rigorous proof of this result.