1212: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 21

Finishing the case of a general linear transformation

Recall that last time we proved that for the sequence of subspaces $N_1 = \ker(\varphi)$, $N_2 = \mathrm{Ker}(\varphi^2)$, …, $N_m = \ker(\varphi^m)$, …, once we have $N_k = N_{k+1}$ for some k, we have $N_{k+l} = N_k$ for all $l > 0$, and moreover $\mathrm{Ker}(\varphi^k) \cap \mathrm{Im}(\varphi^k) = \{0\}$ and $V = \mathrm{Ker}(\varphi^k) \oplus \mathrm{Im}(\varphi^k)$.

Note that the latter result explains the difference between the case $\varphi^2 = \varphi$ and $\varphi^2 = 0$ that we discussed last week. In the case $\varphi^2 = \varphi$ we of course have $\mathrm{Ker}(\varphi) = \mathrm{Ker}(\varphi^2)$, so $V = \mathrm{Ker}(\varphi) \oplus \mathrm{Im}(\varphi)$, while in the case $\varphi^2 = 0$, usually $\mathrm{Ker}(\varphi) \neq \mathrm{Ker}(\varphi^2)$ but $\mathrm{Ker}(\varphi^2) = \mathrm{Ker}(\varphi^3)$ always, so we cannot expect that $V = \mathrm{Ker}(\varphi) \oplus \mathrm{Im}(\varphi)$.

Lemma 1. *1. Both $\mathrm{Ker}(\varphi^k)$ and $\mathrm{Im}(\varphi^k)$ are invariant subspaces of φ.*

2. On the first subspace, the linear transformation φ has just the zero eigenvalue.

3. On the second subspace, all eigenvalues of φ are different from zero.

Proof. 1. The invariance is straightforward: if $v \in \mathrm{Ker}(\varphi^k)$, so that $\varphi^k(v) = 0$, then of course

$$\varphi^k(\varphi(v)) = \varphi^{k+1}(v) = 0,$$

so $\varphi(v) \in \mathrm{Ker}(\varphi^k)$, and similarly, if $v \in \mathrm{Im}(\varphi^k)$, so that $v = \varphi^k(w)$, then of course

$$\varphi(v) = \varphi(\varphi^k(w)) = \varphi^{k+1}(w) = \varphi^k(\varphi(w)),$$

so $\varphi(v) \in \mathrm{Im}(\varphi^k)$.

2. If $\varphi(v) = \mu v$ for some $0 \neq v \in \mathrm{Ker}(\varphi^k)$, then $0 = \varphi^k(v) = \mu^k v$, so $\mu = 0$.

3. If $\varphi(v) = 0$ for some $0 \neq v \in \mathrm{Im}(\varphi^k)$, then $\varphi^k(v) = 0$, but we know that $\mathrm{Im}(\varphi^k) \cap \mathrm{Ker}(\varphi^k) = \{0\}$, which is a contradiction. □

The end of the proof utilises these results to proceed by induction, namely by induction by the number of distinct eigenvalues of φ.

We shall decompose V into a direct sum of invariant subspaces for each of which φ has only one eigenvalue, proving the following theorem.

Theorem 1. *For every linear transformation $\varphi\colon V \to V$ whose (different) eigenvalues are $\lambda_1, \ldots, \lambda_k$, there exist integers $m_1, \ldots, m_k$ such that*

$$V = \mathrm{Ker}(\varphi - \lambda_1 I)^{m_1} \oplus \ldots \oplus \mathrm{Ker}(\varphi - \lambda_k I)^{m_k}.$$

Proof. We shall prove this result by induction on the number of distinct eigenvalues of φ.

Let λ be an eigenvalue of φ, and let us consider the transformation $B_\lambda = \varphi - \lambda I$. Considering kernels of its powers, we find the first place k where they stabilise, so that $\mathrm{Ker}(B_\lambda^k) = \mathrm{Ker}(B_\lambda^{k+1}) = \ldots$.

Note that the subspaces $\mathrm{Ker}(B_\lambda^k)$ and $\mathrm{Im}(B_\lambda^k)$ are invariant subspaces of φ. (Indeed, we already know that these are invariant subspaces of B_λ, and $\varphi = B_\lambda + \lambda I$). Note also that we have $V = \mathrm{Ker}(B_\lambda^k) \oplus \mathrm{Im}(B_\lambda^k)$.

On the invariant subspace $\mathrm{Ker}(B_\lambda^k)$, B_λ has only the eigenvalue 0, so $\varphi = B_\lambda + \lambda I$ has only the eigenvalue λ. Also, on the invariant subspace $\mathrm{Im}(B_\lambda^k)$, B_λ has no zero eigenvalues, hence φ has no eigenvalues equal to λ. Hence, we may apply the induction hypothesis to the linear transformation φ on the vector space $V' = \mathrm{Im}(B_\lambda^k)$ where it has fewer eigenvalues. □

Let us see what happens for each individual subspace $\operatorname{Ker}(\varphi-\lambda I)^m$. Naturally, the linear transformation $B_\lambda = \varphi - \lambda I$ is nilpotent when restricted to that subspace. Therefore, the results of last week allow us to find a basis of threads for this linear transformation, and its matrix is block-diagonal made of blocks

$$J_l = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix},$$

one block J_l for each thread of length l. Recalling that $\varphi = B_\lambda + \lambda I$, we see that relative to the same basis of threads that we found, the linear transformation φ has a block-diagonal matrix made of blocks

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

one block $J_l(\lambda)$ for each thread of length l.

Summing up, we obtain the following theorem (which is usually called Jordan normal form theorem, or Jordan decomposition theorem):

Jordan normal form theorem. Let V be a finite-dimesional 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 φ;
- a basis $e_1^{(i)}, \ldots, e_{n_i}^{(i)}$ of V_i for each $i = 1, \ldots, p$ such that

$$\begin{aligned} (\varphi - \lambda_i I)e_1^{(i)} &= e_2^{(i)}; \\ (\varphi - \lambda_i I)e_2^{(i)} &= e_3^{(i)}; \\ &\dots \\ (\varphi - \lambda_i I)e_{n_i}^{(i)} &= 0 \end{aligned}$$

  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.

Examples

From our proof, one sees that for computing the Jordan normal form and a Jordan basis of a linear transformation φ on a vector space V, one can use the following plan:

- Find all eigenvalues of φ (that is, compute the characteristic polynomial $\det(A - cI)$ of the corresponding matrix A, and determine its roots $\lambda_1, \ldots, \lambda_k$).

- For each eigenvalue λ, form the linear transformation $B_\lambda = \varphi - \lambda I$ and consider the increasing sequence of subspaces
$$\mathrm{Ker}\, B_\lambda \subset \mathrm{Ker}\, B_\lambda^2 \subset \dots$$
and determine where it stabilizes, that is find k which is the smallest number such that $\mathrm{Ker}\, B_\lambda^k = \mathrm{Ker}\, B_\lambda^{k+1}$. Let $U_\lambda = \mathrm{Ker}\, B_\lambda^k$. The subspace U_λ is an invariant subspace of B_λ (and φ), and B_λ is nilpotent on U_λ, so it is possible to find a basis consisting of several "threads" of the form $f, B_\lambda f, B_\lambda^2 f, \dots$, where B_λ shifts vectors along each thread (as in the previous homework).
- Joining all the threads (for different λ) together, we get a Jordan basis for A. A thread of length p for an eigenvalue λ contributes a Jordan block $J_p(\lambda)$ to the Jordan normal form.

Example 1. Let $V = \mathbb{R}^3$, and $A = \begin{pmatrix} -2 & 2 & 1 \\ -7 & 4 & 2 \\ 5 & 0 & 0 \end{pmatrix}$.

The characteristic polynomial of A is $-t + 2t^2 - t^3 = -t(1-t)^2$, so the eigenvalues of A are 0 and 1.

Furthermore, $\mathrm{rk}(A) = 2$, $\mathrm{rk}(A^2) = 2$, $\mathrm{rk}(A - I) = 2$, $\mathrm{rk}(A - I)^2 = 1$. Thus, the kernels of powers of A stabilise instantly, so we should expect a thread of length 1 for the eigenvalue 0, whereas the kernels of powers of $A - I$ do not stabilise for at least two steps, so that would give a thread of length at least 2, hence a thread of length 2 because our space is 3-dimensional.

To determine the basis of $\mathrm{Ker}(A)$, we solve the system $Av = 0$ and obtain a vector $f = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$.

To deal with the eigenvalue 1, we see that the kernel of $A - I$ is spanned by the vector $\begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}$, the kernel of $(A-I)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 10 & -5 & -3 \\ -20 & 10 & 6 \end{pmatrix}$ is spanned by the vectors $\begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 3/10 \\ 0 \\ 1 \end{pmatrix}$. Reducing the latter vectors using the former one, we end up with the vector $e = \begin{pmatrix} 0 \\ 3 \\ -5 \end{pmatrix}$, which gives rise to a thread $e, (A-I)e = \begin{pmatrix} 1 \\ -1 \\ 5 \end{pmatrix}$. Overall, a Jordan basis is given by $f, e, (A - I)e$, and the Jordan normal form has a block of size 1 with 0 on the diagonal, and a block of size 2 with 1 on the diagonal:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Example 2. Let $V = \mathbb{R}^4$, and $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 11 & 6 & -4 & -4 \\ 22 & 15 & -8 & -9 \\ -3 & -2 & 1 & 2 \end{pmatrix}$.

The characteristic polynomial of A is $1 - 2t^2 + t^4 = (1+t)^2(1-t)^2$, so the eigenvalues of A are -1 and 1. To avoid unnecessary calculations (similar to avoiding computing $(A - I)^3$ in the previous example), let us compute the ranks for both eigenvalues simultaneously. For $\lambda = -1$ we have $A + I = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 11 & 7 & -4 & -4 \\ 22 & 15 & -7 & -9 \\ -3 & -2 & 1 & 3 \end{pmatrix}$, $\mathrm{rk}(A + I) = 3$, $(A + I)^2 = \begin{pmatrix} 12 & 8 & -4 & -4 \\ 12 & 8 & -4 & -4 \\ 60 & 40 & -20 & -24 \\ -12 & -8 & 4 & 8 \end{pmatrix}$, $\mathrm{rk}((A + I)^2) = 2$.
For $\lambda = 1$ we have $A - I = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 11 & 5 & -4 & -4 \\ 22 & 15 & -9 & -9 \\ -3 & -2 & 1 & 1 \end{pmatrix}$, $\mathrm{rk}(A - I) = 3$, $(A - I)^2 = \begin{pmatrix} 12 & 4 & -4 & -4 \\ -32 & -16 & 12 & 12 \\ -28 & -20 & 12 & 12 \\ 0 & 0 & 0 & 0 \end{pmatrix}$,

$\mathrm{rk}((A-I)^2) = 2$. Thus, each of these eigenvalues gives rise to a thread of length at least 2, and since our vector space is 4-dimensional, each of the threads should be of length 2, and in each case the stabilisation happens on the second step.

In the case of the eigenvalue -1, we first determine the kernel of $A+I$, solving the system $(A+I)v = 0$; this gives us a vector $\begin{pmatrix} -1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$. The equations that determine the kernel of $(A+I)^2$ are $t = 0, 3x + 2y = z$ so y and z are free variables, and for the basis vectors of that kernel we can take $\begin{pmatrix} 1/3 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} -2/3 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. Reducing the basis vectors of $\mathrm{Ker}(A+I)^2$ using the basis vector of $\mathrm{Ker}(A+I)$, we end up with a relative basis vector $e = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}$, and a thread $e, (A+I)e = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$.

In the case of the eigenvalue 1, we first determine the kernel of $A-I$, solving the system $(A-I)v = 0$; this gives us a vector $\begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$. The equations that determine the kernel of $(A-I)^2$ are $4x = z + t, 4y = z + t$ so z and t are free variables, and for the basis vectors of that kernel we can take $\begin{pmatrix} 1/4 \\ 1/4 \\ 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1/4 \\ 1/4 \\ 0 \\ 1 \end{pmatrix}$. Reducing the basis vectors of $\mathrm{Ker}(A-I)^2$ using the basis vector of $\mathrm{Ker}(A+I)$, we end up with a relative basis vector $f = \begin{pmatrix} 1/4 \\ 1/4 \\ 0 \\ 1 \end{pmatrix}$, and a thread $e, (A-I)e = \begin{pmatrix} 0 \\ 0 \\ 1/4 \\ -1/4 \end{pmatrix}$.

Finally, the vectors $e, (A+I)e, f, (A-I)f$ form a Jordan basis for A; the Jordan normal form of A is $\begin{pmatrix} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$.