MA1112: Linear Algebra II

Dr. Vladimir Dotsenko (Vlad)

Lecture 9

Case of a general linear transformation

Now, suppose that φ is an arbitrary linear transformation of V (no assumption $\varphi^k = 0$ anymore). Let us nevertheless consider the sequence of subspaces $N_1 = \ker(\varphi)$, $N_2 = \ker(\varphi^2)$, ..., $N_m = \ker(\varphi^m)$, ...

Note that this sequence is increasing:

$$N_1 \subset N_2 \subset \ldots N_m \subset \ldots$$

Indeed, if $v \in N_s$, that is $\varphi^s(v) = 0$, then we have $\varphi^{s+1}(v) = \varphi(\varphi^s(v)) = 0$.

Since we only work with finite-dimensional vector spaces, this sequence of subspaces cannot be strictly increasing; if $N_i \neq N_{i+1}$, then, obviously, $\dim N_{i+1} \geqslant 1 + \dim N_i$. It follows that for some k we have $N_k = N_{k+1}$.

Lemma 1. *In this case we have* $N_{k+l} = N_k$ *for all* $l > 0$.

Proof. We shall prove that $N_{k+l} = N_{k+l-1}$ by induction on l. The induction basis (case $l = 1$) follows immediately from our notation. Suppose that $N_{k+l} = N_{k+l-1}$; let us prove that $N_{k+l+1} = N_{k+l}$. Let us take a vector $v \in N_{k+l+1}$, so $\varphi^{k+l+1}(v) = 0$. We have $\varphi^{k+l+1}(v) = \varphi^{k+l}(\varphi(v))$, so $\varphi(v) \in N_{k+l}$. But by the induction hypothesis $N_{k+l} = N_{k+l-1}$, so $\varphi^{k+l-1}(\varphi(v)) = 0$, or $\varphi^{k+l}(v) = 0$, so $v \in N_{k+l}$, as required. □

Lemma 2. *For the index* k *that we found, we have* $V = \ker(\varphi^k) \oplus \mathrm{Im}(\varphi^k)$.

Proof. Let us first show that $\ker(\varphi^k) \cap \mathrm{Im}(\varphi^k) = \{0\}$. Indeed, suppose that $v \in \ker(\varphi^k) \cap \mathrm{Im}(\varphi^k)$. This means that $\varphi^k(v) = 0$ and that $v = \varphi^k(w)$ for some vector w. It follows that $\varphi^{2k}(w) = 0$, so $w \in N_{2k}$. But from the previous lemma we know that $N_{2k} = N_k$, so $w \in N_k$. Thus, $v = \varphi^k(w) = 0$, which is what we claimed.

Now we consider the sum of these two subspaces which we just proved to be direct. It is a subspace of V of dimension $\dim\ker(\varphi^k) + \dim\mathrm{Im}(\varphi^k) = \dim(V)$, so it has to coincide with V. □

Note that the result we just proved explains the difference between the case $\varphi^2 = \varphi$ and $\varphi^2 = 0$. 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 $V = \mathrm{Ker}(\varphi) \oplus \mathrm{Im}(\varphi)$, but we only have the trivial decomposition $V = V \oplus 0 = \mathrm{Ker}(\varphi^2) \oplus \mathrm{Im}(\varphi^2)$.

Lemma 3. *For the index* k *that we found,*

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. □

We shall conclude the proof by induction, using the results we obtained. The induction parameter would be slightly unconventional: the number of distinct eigenvalues of φ. Our strategy would be to decompose V into a direct sum of invariant subspaces for each of which φ has only one eigenvalue, proving the following result.

Theorem 1. *For every linear transformation $\varphi\colon V \to V$ whose (different) eigenvalues are $\lambda_1, \ldots, \lambda_k$, there exist invariant subsspaces $U_1, \ldots, U_k$ such that on each subspace U_i the only eigenvalue of φ is λ_i, and*

$$V = U_1 \oplus U_2 \oplus \cdots \oplus U_k.$$

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

Let us consider the transformation $\varphi_{\lambda_1} = \varphi - \lambda_1 I$. Considering kernels of its powers, we find the first place k_1 where they stabilise, so that $\mathrm{Ker}(\varphi_{\lambda_1}^{k_1}) = \mathrm{Ker}(\varphi_{\lambda_1}^{k_1+1}) = \ldots$.

Note that the subspaces $\mathrm{Ker}(\varphi_{\lambda_1}^{k_1})$ and $\mathrm{Im}(\varphi_{\lambda_1}^{k_1})$ are invariant subspaces of φ. (Indeed, we already know that these are invariant subspaces of φ_{λ_1}, and $\varphi = \varphi_{\lambda_1} + \lambda_1 I$). Note also that we have $V = \mathrm{Ker}(\varphi_{\lambda_1}^{k_1}) \oplus \mathrm{Im}(\varphi_{\lambda_1}^{k_1})$.

On the invariant subspace $\mathrm{Ker}(\varphi_{\lambda_1}^{k_1})$, the transformation φ_{λ_1} has only the eigenvalue 0, so $\varphi = \varphi_{\lambda_1} + \lambda_1 I$ has only the eigenvalue λ_1. Also, on the invariant subspace $\mathrm{Im}(\varphi_{\lambda_1}^{k_1})$, φ_{λ_1} has no zero eigenvalues, hence φ has no eigenvalues equal to λ_1. Hence, we may put $U_1 := \mathrm{Ker}(\varphi_{\lambda_1}^{k})$ and then apply the induction hypothesis to the linear transformation φ on the subspace $V' = \mathrm{Im}(\varphi_{\lambda_1}^{k})$ where it has fewer eigenvalues.

Let us remark that the claim that the eigenvalues of φ on the subspace $V' = \mathrm{Im}(\varphi_{\lambda_1}^{k})$ will be $\lambda_2, \ldots, \lambda_k$. This follows from the fact that if $V = V_1 \oplus V_2$ is a decomposition of V into a sum of two invariant subspaces, the characteristic polynomial of φ on V is the product of the characteristic polynomials on V_1 and V_2: the determinant of a block matrix is equal to the product of determinant of blocks. □