Lecture 33

Andrei Antonenko

May 2, 2003

1 Direct sum of vector spaces

In the first part of this lecture we will consider some more concepts from the theory of vector spaces.

Definition 1.1. ***Sum*** *of two vector spaces* U *and* V $U+V$ *is a vector space which consists of all vectors* $u+v$*, where* $u \in U$ *and* $v \in V$*.*

Definition 1.2. *Vector spaces* $U_1, U_2, \ldots, U_n$ *are called* ***linearly independent*** *if from* $u_1 + \cdots + u_n = 0$*, where* $u_i \in U_i$ *it follows that* $u_i = 0$ *for all* i*.*

Sum of the linearly independent vector spaces is called a **direct sum** of these vector spaces and is denoted by $U_1 \oplus \cdots \oplus U_n$.

Definition 1.3. *The vector space* V *is said to be equal to a* ***direct sum*** *of vector spaces* $U_1, \ldots, U_n$

$$V = U_1 \oplus \cdots \oplus U_n$$

if any vector v *from* V *can be represented as*

$$v = u_1 + \cdots + u_n, \quad u_i \in U_i$$

uniquely.

For example, the plane $\mathbb{R}^2$ is equal to a direct sum of $x-$ and $y-$axes.

Example 1.4. *The space of all matrices is equal to a direct sum of the space of all symmetric matrices and all skewsymmetric matrices, since any matrix* A *can be uniquely represented as*

$$A = \frac{A + A^\top}{2} + \frac{A - A^\top}{2},$$

and one can check that $\frac{A+A^\top}{2}$ *is always symmetric, and* $\frac{A-A^\top}{2}$ *is always skewsymmetric. Moreover, the sum is direct, since if a matrix is both symmetric and skewsymmetric, it is equal to* 0*-matrix.*

2 Invariant spaces

Definition 2.1. *Let $\mathcal{A}$ be an operator in vector space V. The subspace $U \subset V$ is called an **invariant subspace** with respect to operator $\mathcal{A}$ if*

$$\mathcal{A}U \subset U.$$

This definition means, that the vectors from invariant subspace remain in this subspace after application of the operator $\mathcal{A}$.

Example 2.2. *Considering the operator of rotation in the 3-dimensional space around some axes, we can see, that all the planes, perpendicular to the axes of rotation are invariant. Moreover, the axes of rotation is invariant itself.*

If the basis $\{e_1, \dots e_n\}$ of V is such that first k vectors $\{e_1, \dots, e_k\}$ is a basis of U, then the matrix of the operator $\mathcal{A}$ in this basis has the following form:

$$\begin{pmatrix} B & D \\ 0 & C \end{pmatrix}.$$

Moreover, if the space V is equal to a direct sum of two subspaces $V = U \oplus W$, and $\{e_1, \dots, e_k\}$ is a basis of U, and $\{e_{k+1}, \dots, e_n\}$ is a basis of W, then the matrix of $\mathcal{A}$ has the following form:

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}.$$

Example 2.3. *Consider the rotation in the 3-dimensional space about some axes be an angle α. In the basis $\{e_1, e_2, e_3\}$, if the vector e_3 is directed along the axes of rotation, the matrix of this operator has the following form:*

$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This matrix is consistent with the decomposition of $\mathbb{R}^3$ into a direct sum of two invariant subspaces:

$$\mathbb{R}^3 = \langle e_1, e_2 \rangle \oplus \langle e_3 \rangle.$$

3 Jordan canonical form

As we saw in previous lectures, some of the operators are not diagonalizable. But we are still able to simplify the matrix of the operator to some extent.

Definition 3.1. *The vector $v \in V$ is called a* ***root vector*** *of the operator $\mathcal{A}$ corresponding to λ if*

$$(\mathcal{A} - \lambda \mathfrak{I})^m v = 0 \tag{1}$$

for some natural number m. The minimal m is called the ***height*** *of the root vector v.*

We see that it is a generalization of the concept of the eigenvector, since eigenvectors are root vectors, for which $m = 1$: vector v is called an eigenvector if

$$(\mathcal{A} - \lambda \mathfrak{I})v = 0$$

Example 3.2. *For the operator of taking a derivative in the space of polynomials root vectors of the height m corresponding to $\lambda = 0$ are the polynomials of degree $m-1$, since, their m-th derivative is equal to 0.*

If the vector v is a root vector of the height m, then the vector

$$u = (\mathcal{A} - \lambda \mathfrak{I})^{m-1} v \tag{2}$$

is an eigenvector with the eigenvalue λ. Thus, λ's should be roots of the characteristic polynomial.

The root vectors corresponding to some particular λ form a subspace. This subspace is called a **root space** and it is denoted by $V^\lambda(\mathcal{A})$. If v is a root vector of the height m, then vector $(\mathcal{A} - \lambda \mathfrak{I})v$ is a root vector of the height $m-1$. Thus the root space $V^\lambda(\mathcal{A})$ is invariant with respect to $(\mathcal{A} - \lambda \mathfrak{I})$, and thus with respect to $\mathcal{A}$.

The set of the root vectors of the height $\leq m$ is a kernel of the operator $(\mathcal{A} - \lambda \mathfrak{I})$. Thus the root space $V^\lambda(\mathcal{A})$ is a union of the following chain of subspaces:

$$\operatorname{Ker}(\mathcal{A} - \lambda \mathfrak{I}) \subset \operatorname{Ker}(\mathcal{A} - \lambda \mathfrak{I})^2 \subset \ldots$$

Since we consider the finite-dimensional spaces, this chain will stabilize, and thus $V^\lambda(\mathcal{A}) = \operatorname{Ker}(\mathcal{A} - \lambda \mathfrak{I})^m$ for some m. The matrix of the operator $(\mathcal{A} - \lambda \mathfrak{I})$ in the basis of $V^\lambda(\mathcal{A})$ consistent with this chain of subspaces is triangular with zeros on the diagonal, and thus the matrix of the operator $\mathcal{A}$ is triangular with λ's on the diagonal. From here we have the following property: the characteristic polynomial of the operator $\mathcal{A}$ on the space $V^\lambda(\mathcal{A})$ is equal to $(t-\lambda)^k$, where k is a dimension of $V^\lambda(\mathcal{A})$. No we have the following proposition:

Proposition 3.3. *The dimension of the root space is equal to the multiplicity of the corresponding root of the characteristic polynomial.*

Proof. If $\{e_1, e_2, \ldots, e_n\}$ is a basis of V, and first k vectors of it are a basis of $V^\lambda(\mathcal{A})$, then the matrix of $\mathcal{A}$ has the following form:

$$\begin{pmatrix} B & D \\ 0 & C \end{pmatrix}$$

Thus

$$p_{\mathcal{A}}(t) = (t-\lambda)^k \det(tI - C).$$

Now let $\mathcal{C}$ be an operator in the space $W = \langle e_{k+1}, \dots, e_n \rangle$ with the matrix C. We need to prove, that λ is not a root of the polynomial $\det(tI - C)$, i.e. it is not an eigenvalue of $\mathcal{C}$. Let's assume the contrary. Then there exists $v \in W$, such that $\mathcal{C}v = \lambda v$. Then

$$\mathcal{A}v = \lambda v + u, \quad u \in V^{\lambda}(\mathcal{A}),$$

and thus $(\mathcal{A} - \lambda\mathcal{I})v = u$ is a root vector, but in this case v is also a root vector, which contradicts the definition of $V^{\lambda}(\mathcal{A})$. □

Proposition 3.4. *The root spaces, corresponding to different λ_i's are linearly independent.*

Proof. Assume

$$c_1e_1 + c_2e_2 + \dots + c_ke_k = 0, \quad e_i \in V^{\lambda_i}(\mathcal{A}). \tag{3}$$

Let's apply to this equality the operator $(\mathcal{A} - \lambda_k\mathcal{I})^m$, where m is the height of e_k. We obtain:

$$(\mathcal{A} - \lambda_k\mathcal{I})^m c_1e_1 + \dots + (\mathcal{A} - \lambda_k\mathcal{I})^m c_{k-1}e_{k-1} = 0. \tag{4}$$

Using induction we have

$$(\mathcal{A} - \lambda_k\mathcal{I})^m c_1e_1 = \dots = (\mathcal{A} - \lambda_k\mathcal{I})^m c_{k-1}e_{k-1} = 0. \tag{5}$$

But since the operator $(\mathcal{A} - \lambda_k\mathcal{I})$ is not degenerate on any of $V^{\lambda_1}(\mathcal{A}), \dots, V^{\lambda_{k-1}}(\mathcal{A})$ (i.e., not equal to 0 for nonzero vectors), we have $c_1 = \dots = c_{k-1} = 0$, and thus $c_k = 0$ also. □

These two propositions lead to the following theorem:

Theorem 3.5. *If the characteristic polynomial $p_{\mathcal{A}}(t)$ can be factored into linear terms, then*

$$V = \bigoplus_{i=1}^{s} V^{\lambda_i}(\mathcal{A}), \tag{6}$$

where λ_i's are different roots of $p_{\mathcal{A}}(t)$.

Now we will discuss the action of the operator $\mathcal{A}$ on any of the root spaces.

Definition 3.6. *The linear operator $\mathcal{N}$ is called **nilpotent** if there exists integer $m \geq 0$ such that $\mathcal{N}^m = 0$. The minimal m is called the **height** of the nilpotent operator $\mathcal{N}$.*

Example 3.7. *The operator of taking derivative in the space of polynomials of bounded degree $P_n(t)$ is a nilpotent operator of the height $n+1$.*

Since $V^\lambda(\mathcal{A}) = \operatorname{Ker}(\mathcal{A} - \lambda\mathcal{I})^m$ for some m, the operator $\mathcal{N} = (\mathcal{A} - \lambda\mathcal{I})$ is nilpotent on $V^\lambda(\mathcal{A})$. Thus we need to study nilpotent operators.

Let $\mathcal{N}$ be a nilpotent operator in the space V. The **height** of the vector v with respect to $\mathcal{N}$ is the minimal number m, such that $\mathcal{N}^m v = 0$. Obviously, the height of any vector is less than or equal to the height of the nilpotent operator. We will denote the height of the vector v as $\operatorname{ht} v$.

Lemma 3.8. *If v is a vector of the height m, then vectors*

$$v, \mathcal{N}v, \mathcal{N}^2 v, \dots, \mathcal{N}^{m-1} v$$

are linearly independent.

Proof. Assume

$$\lambda_0 v + \lambda_1 \mathcal{N} v + \lambda_2 \mathcal{N}^2 v + \cdots + \lambda_{m-1} \mathcal{N}^{m-1} v = 0. \tag{7}$$

Let λ_k is the first nonzero coefficient. Then applying the operator $\mathcal{N}^{m-k-1}$ we obtain incorrect equality

$$\lambda_k \mathcal{N}^{m-1} v = 0. \tag{8}$$

□

Definition 3.9. *The subspace $\langle v, \mathcal{N}v, \mathcal{N}^2 v, \dots, \mathcal{N}^{m-1} v\rangle$ ($m = \operatorname{ht} v$) is called a **cyclic subspace** of the nilpotent operator $\mathcal{N}$, generated by vector v.*

Obviously the cyclic subspace is invariant with respect to $\mathcal{N}$. The operator $\mathcal{N}$ on the cyclic subspace $\langle v, \mathcal{N}v, \mathcal{N}^2 v, \dots, \mathcal{N}^{m-1} v\rangle$ has height m and in the basis $\{v, \mathcal{N}v, \mathcal{N}^2 v, \dots, \mathcal{N}^{m-1} v\}$ has a matrix

$$J(0) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \tag{9}$$

which is called a **nilpotent Jordan block**.

Theorem 3.10. *The space V can be decomposed into a direct sum of the cyclic subspaces of the operator $\mathcal{N}$. The number of the spaces in such decomposition is equal to* $\dim \operatorname{Ker} \mathcal{N}$.

Proof. The proof is done by induction over $n = \dim V$. If $n = 1$ the theorem is obvious. If $n > 1$ let $U \subset V$ be an arbitrary $(n-1)$-dimensional subspace, containing $\operatorname{Im} \mathcal{N}$. Obviously, U is invariant with respect to $\mathcal{N}$. By the induction hypothesis,

$$U = U_1 \oplus \cdots \oplus U_k,$$

where U_i's are invariant subspaces. Let's take any vector $v \in V \setminus U$. We have

$$\mathcal{N}v = u_1 + \dots + u_k.$$

If for some i $u_i = \mathcal{N}v_i \in \mathcal{N}U_i$, then we can substitute v with $v - v_i$, and get that $u_i = 0$. Thus we can assume that either $u_i = 0$, or $u_i \notin \mathcal{N}U_i$.

If $u_i = 0$ for all i's, i.e. $\mathcal{N}v = 0$, then

$$V = \langle v \rangle \oplus U_1 \oplus \dots \oplus U_k$$

is the decomposition of V into a direct sum of cyclic subspaces.

Now let $\mathcal{N}v \neq 0$. Then

$$\operatorname{ht} \mathcal{N}v = \max_i \operatorname{ht} u_i.$$

Let's assume that

$$\operatorname{ht} \mathcal{N}v = \operatorname{ht} u_1 = m.$$

Then $\operatorname{ht} v = m + 1$. We will prove that in this case

$$V = \langle v, \mathcal{N}v, \mathcal{N}^2 v, \dots, \mathcal{N}^m v \rangle \oplus U_1 \oplus \dots \oplus U_k.$$

Since $u_1 \notin \mathcal{N}U_1$, then $\dim U_1 = \operatorname{ht} u_1 = m$, and thus

$$\dim V = \dim U + 1 = (m + 1) + \dim U_2 + \dots + \dim U_k.$$

Thus it is enough to prove that

$$\langle v, \mathcal{N}v, \mathcal{N}^2 v, \dots, \mathcal{N}^m v \rangle \cap (U_2 \oplus \dots \oplus U_k) = 0.$$

Assume that

$$\lambda_0 v + \lambda_1 \mathcal{N}v + \dots + \lambda_m \mathcal{N}^m v \in U_2 \oplus \dots \oplus U_k.$$

Since $v \notin U$, $\lambda_0 = 0$. Taking projection of other summands onto U_1, we get

$$\lambda_1 u_1 + \lambda_2 \mathcal{N}u_1 + \dots + \lambda_m \mathcal{N}^{n-1} u_1 = 0,$$

and thus $\lambda_1 = \dots \lambda_m = 0$.

Now we will prove the second assertion of the theorem. Let

$$V = V_1 \oplus \dots \oplus V_k$$

is a decomposition of V into a direct sum of the cyclic subspaces of $\mathcal{N}$. Obviously the kernel of $\mathcal{N}$ can be decomposed into a direct sum of kernels of the operator $\mathcal{N}$, restricted to summands $V_1, \dots, V_k$. But since the dimension of the kernel of the restricted operator is equal to 1, then $\dim \operatorname{Ker} \mathcal{N} = k$. □

Now getting back to any linear operator $\mathcal{A}$, we can see that in the cyclic subspace of the nilpotent operator $(\mathcal{A}-\lambda\mathfrak{I})$, restricted to $V^{\lambda}(\mathcal{A})$, the operator $\mathcal{A}$ has the matrix of the following form:

$$J(\lambda) = J(0) + \lambda I = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \tag{10}$$

This matrix is called a **Jordan block with the eigenvalue** λ.

Definition 3.11. *The **Jordan matrix** is a matrix with Jordan blocks over diagonal, and zeros everywhere else.*

Combining the previous results, we obtain the following most important result from the theory of linear operators:

Theorem 3.12. *If the characteristic polynomial of the operator can be factored in linear terms, then there exists a basis, in which the matrix of the operator is Jordan.*

Corollary 3.13. *Matrix of any operator can be transposed to Jordan canonical form over the complex numbers.*