LECTURE 7: LINEAR OPERATORS

NIKOLAY BOGACHEV

Contents

§ 1. Definition, coordinates and preliminaries

Let V be a vector space over field $\mathbb{F} = \mathbb{R}, \mathbb{C}$, ..., etc.

Definition 1.1. *A linear map $\mathcal{A}\colon V \to V$ is called a linear operator.*

The matrix of an operator $\mathcal{A}$ in the basis $\{e_1, \ldots, e_n\}$ is $A = (a_{ij})$, where the j-th column of A is $\mathcal{A}(e_j) = \mathcal{A}e_j = \sum_{i=1}^{n} a_{ij} e_i$. That is,

$$(\mathcal{A}e_1, \ldots, \mathcal{A}e_n) = (e_1, \ldots, e_n)A.$$

We write $A = \mathrm{Mat}(\mathcal{A})$. If $y = \mathcal{A}x$, then in the matrix form one can write $Y = AX$.

Let $(e'_1, \ldots, e'_n) = (e_1, \ldots, e_n)C$ be another basis of V. Then

$$(\mathcal{A}e'_1, \ldots, \mathcal{A}e'_n) = (\mathcal{A}e_1, \ldots, \mathcal{A}e_n)C = (e_1, \ldots, e_n)AC = (e'_1, \ldots, e'_n)C^{-1}AC.$$

Thus,

$$A' = C^{-1}AC.$$

Main Question: How can we change a basis in such a way that the operator matrix has a "simple" form?

Date: January 2020.

§ 2. Invariant subspaces, eigenvectors and eigenvalues

2.1. Invariant subspaces.

Definition 2.1. *A subspace $U \subset V$ is invariant with respect to $\mathcal{A}\colon V \to V$ if $\mathcal{A}U \subset U$, i.e. $\mathcal{A}u \in U$ for any $u \in U$.*

The restriction on an invariant subspace is a well-defined linear operator: $\mathcal{A}|_U \colon U \to U$.

In the basis of V that agrees with U the matrix of $\mathcal{A}$ has the following form:

$$\begin{pmatrix} A_U & B \\ 0 & C \end{pmatrix},$$

where $A_U = \mathrm{Mat}(\mathcal{A}|_U)$.

If $V = V_1 \oplus \ldots \oplus V_k$, where all V_j are invariant, then

$$A = \mathrm{diag}(A_1, \ldots, A_k) = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_k \end{pmatrix},$$

where $A_j = \mathrm{Mat}(\mathcal{A}|_{V_j})$.

Example 1. $A = \mathrm{diag}(a_1, a_2)$*, where* $V = \mathbb{R}^2 = \langle e_1 \rangle \oplus \langle e_1 \rangle$.

2.2. Eigenvectors and eigenvalues.

Definition 2.2. *A vector $v \in V$ is called an eigenvector for an operator $\mathcal{A}\colon V \to V$ if $\mathcal{A}v = \lambda v$ for some number $\lambda \in \mathbb{F}$. The corresponding number $\lambda \in \mathbb{F}$ is called an eigenvalue.*

If $\mathcal{A}v = \lambda v$, then $\langle v \rangle$ is an invariant subspace for $\mathcal{A}$. It is easy to verify that in the basis $\{v_1, \ldots, v_n\}$ of eigenvectors of $\mathcal{A}$ we have

$$A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

Geometrically, eigenvectors are exactly the directions, where an operator acts by stretching of a space by the corresponding eigenvalues.

The natural question arises here: how can we calculate eigenvectors and eigenvalues?

The answer is the following: $\mathcal{A}v = \lambda v$ if and only if the operator $\mathcal{A} - \lambda\mathcal{I}$ is singular (degenerate), where I is the identical operator on V: $\mathcal{I}x := \mathrm{Id}(x) \equiv x$.

The last is equivalent to the fact that

$$\det(A - \lambda E) = 0.$$

Definition 2.3. *The space $V_\lambda(\mathcal{A}) := \mathrm{Ker}(\mathcal{A} - \lambda\mathcal{I})$ is called the eigenspace of $\mathcal{A}$ associated with the eigenvalue λ.*

Definition 2.4. *The characteristic polynomial of $\mathcal{A}$ is*

$$f_{\mathcal{A}}(\lambda) = (-1)^n \det(A - \lambda E).$$

Eigenvalues are exactly the roots of the characteristic polynomial. When the eigenvalues are already known, one can calculate the eigenvalues in the following way: it remains just to find all non-zero solution of a system of linear equations: $(\mathcal{A} - \lambda\mathcal{I})x = 0$.

Theorem 2.1. *The following holds:*

1. $\dim V_\lambda(\mathcal{A}) \leq$ *the multiplicity of λ in $f_{\mathcal{A}}$,*
2. $V_{\lambda_1}(\mathcal{A}), \ldots, V_{\lambda_s}(\mathcal{A})$ *are linearly independent for different lambda's.*

Proof. Sketch.

Part (1): It is enough to consider the matrix of $\mathcal{A}$ in the basis of V that agrees with $\dim V_\lambda(\mathcal{A})$. If $\dim V_\lambda(\mathcal{A}) = k$, then $f_\mathcal{A}(t) = (t-\lambda)^k h(t)$.

Part (2): Induction by s. If they are linearly independent, then there exist such vectors v_j that $v_1 + \ldots + v_s = 0$. Taking $\mathcal{A}$ of it, we obtain $\lambda_1 v_1 + \ldots + \lambda_s v_s = 0$. After that it remains to take the difference of this equality and the previous one multiplied by one of the non-zero numbers λ_j. After that we can use the induction hypothesis. ∎

Corollary 1. *If $f_\mathcal{A}$ has n different roots, then there exists the diagonal basis for $\mathcal{A}$ consisting of its eigenvectors.*

§ 3. Existence of a 1-dim or 2-dim invariant subspace for an operator in a real vector space

Suppose V is a real vector space, then its complexification is

$$V(\mathbb{C}) := \{u + iv \mid u, v \in V\}.$$

It is clear that $V(\mathbb{C})$ is also a vector space, $V(\mathbb{C}) \supset V = \{v + i \cdot 0 \mid v \in V\}$. It also clear that the basis of V is a basis for $V(\mathbb{C})$, i.e. $\dim V_\mathbb{R} = \dim V(\mathbb{C})_\mathbb{C}$.

Every linear operator in V can be uniqely extended to an operator in $V(\mathbb{C})$: $\mathcal{A}_\mathbb{C}(u+iv) = \mathcal{A}u + i \cdot \mathcal{A}v$ (with the same matrix in the basis of V).

Theorem 3.1. *For every linear operator in a real vector space there exist a 1-dim or a 2-dim invariant subspace.*

Proof. If $f_\mathcal{A}$ has a real root, then $\mathcal{A}$ has 1-dim invariant subspace.

Suppose now that $f_\mathcal{A}$ has a complex root $\lambda + \mu i$, and $u + iv \in V(\mathbb{C})$ is the corresponding eigenvector. That is, $\mathcal{A}u + i\mathcal{A}v = (\lambda + \mu i)(u + iv)$, which follows that

$$\begin{cases} \mathcal{A}u = \lambda u - \mu v \\ \mathcal{A}v = \mu u - \lambda v. \end{cases}$$

Thus, $\langle u, v\rangle$ is an invariant subspace. ∎

§ 4. Linear operators in Euclidean and Hermitian spaces

4.1. Euclidean spaces. Let V be a real Euclidean space with an inner product $(\cdot,\cdot)$. Then any operator in V naturally corresponds to a bilinear form $\varphi_\mathcal{A}(x, y) = (x, \mathcal{A}y)$.

In the orthonormal basis of V we have $\mathrm{Mat}(\varphi_\mathcal{A}) = \mathrm{Mat}(\mathcal{A})$: $\varphi_\mathcal{A}(e_i, e_j) = (e_i, \mathcal{A}e_j) = a_{ij}$. Recall that the basis of a Euclidean space V is orthonormal if and only if $\mathrm{Mat}((\cdot,\cdot)) = E$ in this basis.

A map $\mathcal{A} \mapsto \varphi_\mathcal{A}$ is an isomorphism of the space $\mathcal{L}(V)$ of linear operators to the space of bilinear forms on V.

One can also define a transposed bilinear form $\varphi^T_\mathcal{A}(x, y) = \varphi_\mathcal{A}(y, x)$. Clearly, $\mathrm{Mat}(\varphi^T_\mathcal{A}) = \mathrm{Mat}(\varphi_\mathcal{A})^T$. We can also define the corresponding adjoint operator $\mathcal{A}^*$: $(x, \mathcal{A}^* y) = (y, \mathcal{A}x) = (\mathcal{A}x, y)$. In the orthonormal basis $\mathrm{Mat}(\mathcal{A}^*) = \mathrm{Mat}(\mathcal{A})^T$.

Definition 4.1. *A linear operator $\mathcal{A}$ is called symmetric (or self-adjoint) if $\mathcal{A}^* = \mathcal{A}$ (i.e. $(x, \mathcal{A}y) = (y, \mathcal{A}x)$).*

Definition 4.2. *A linear operator $\mathcal{A}$ is called skew-symmetric if $\mathcal{A}^* = -\mathcal{A}$.*

Definition 4.3. *A linear operator $\mathcal{A}$ is called orthogonal if $\mathcal{A}^*\mathcal{A} = \mathcal{I}$ (i.e. $\mathcal{A}$ preserves the inner product in V: $(\mathcal{A}x, \mathcal{A}y) = (\mathcal{A}^*\mathcal{A}x, y) = (x, y)$).*

It is clear that non-degenerate operators correspond to non-degenerate matrices ($\det A \neq 0$) and orthogonal operators correspond to orthogonal matrices ($AA^T = A^T A = E$).

Thus, $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{O}(n, \mathbb{R})$ denote respectively the groups of non-degenerate and orthogonal operators on $\mathbb{R}^n$. One can also write $\mathrm{GL}(V)$ and $\mathrm{O}(V)$.

4.2. Hermitian spaces. Let V be a Hermitian space. Recall that in V (which is defined over $\mathbb{C}$) we have the inner product with following property: $(x, y) = \overline{(y, x)}$.

In this case we can also define a form $\varphi_{\mathcal{A}}$ and its conjugate form $\varphi^*_{\mathcal{A}}$ in the way similar to Euclidean spaces.

That is, $(x, \mathcal{A}^* y) = \overline{(y, \mathcal{A}x)} = (\mathcal{A}x, y)$.

Definition 4.4. *An operator $\mathcal{A}$ is called hermitian, skew-hermitian, and unitary, if $\mathcal{A}^* = \mathcal{A}$, $\mathcal{A}^* = -\mathcal{A}$, and $\mathcal{A}^* = \mathcal{A}^{-1}$, respectively.*

§ 5. Orthonormal eigenbasis for a symmetric operator

An eigenbasis is a basis of eigenvectors.

Theorem 5.1. *Let $\mathcal{A}$ be either a symmetric, or a skew-symmetric, or an orthogonal operator in a Euclidean space V, and $U \subset V$ be its invariant subspace. Then $U^\perp$ also is invariant for $\mathcal{A}$.*

Proof. Let $\mathcal{A}$ be a symmetric operator. If $x \in U$, $y \in U^\perp$, then $(x, \mathcal{A}y) = (\mathcal{A}x, y) = 0$, since $\mathcal{A}x \in U$, $y \in U^\perp$. The same works for a skew-symmetric operator.

Suppose $\mathcal{A}$ is orthogonal. Then $\mathcal{A}\mid_U$ also is orthogonal and non-degenerate. Let $y \in U^\perp$ and $x \in U$. We need $\mathcal{A}y \in U^\perp$, i.e. $(x, \mathcal{A}y) = 0$. But since $\mathcal{A}\mid_U$ is non-degenerate, then there exists $z \in U$, such that $x = \mathcal{A}z$. Then $(x, \mathcal{A}y) = (\mathcal{A}z, \mathcal{A}y) = (z, y) = 0$. ∎

Theorem 5.2. *For any symmetric operator in a Euclidean space, there exists an orthonormal eigenbasis.*

Proof. Induction by $n = \dim V$. Case $n = 1$ is trivial.

If $n = 2$, then $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. We have

$$f_{\mathcal{A}}(t) = \det \begin{pmatrix} t-a & -b \\ -b & t-c \end{pmatrix} = t^2 - (a+c)t + ac - b^2.$$

If $f_{\mathcal{A}}(t) = 0$, then $t_{1,2} = \frac{a+c\pm\sqrt{(a-c)^2+4b^2}}{2} \in \mathbb{R}$, which implies that there exists 1-dim invariant subspace U (generated by one of the eigenvectors), and in this case $V = U \oplus U^\perp = \langle e_1 \rangle \oplus \langle e_1 \rangle$ (see Theorem 5.1) is the sum of 1-dim invariant subspaces, where e_1, e_2 are orthonormal eigenvectors.

For $n > 2$ we choose an invariant (1- or 2-dim) subspace U (see Theorem 3.1) and $V = U \oplus U^\perp$ (see Theorem 5.1), where $\dim U, \dim U^\perp < n$ and we can use the induction hypothesis. ∎

Corollary 2. *If $\mathcal{A}$ is a symmetric operator in V, then $V = \oplus_\lambda V_\lambda(\mathcal{A})$, where $V_\lambda(\mathcal{A}) \perp V_\mu(\mathcal{A})$ if $\lambda \neq \mu$.*

Proof. Let $(e_1, \dots, e_n)$ be an orthonormal eigenbasis from Theorem 5.2, $\mathcal{A}e_j = \lambda_j e_j$. Then $V_\lambda(\mathcal{A}) = \langle e_i \mid \lambda_i = \lambda \rangle$ is orthogonal to all other $V_\mu(\mathcal{A})$. ∎

Theorem 5.3. *For any quadratic form $q(x)$ in a Euclidean space, there exists an orthonormal basis, where $q(x) = \lambda_1 x_1^2 + \ldots + \lambda_n x_n^2$, where λ_j are the eigenvalues of $\mathrm{Mat}(q)$ in any orthonormal basis and are defined up to permutation.*

Proof. Indeed, $q(x) = (\mathcal{A}x, x)$ for a symmetric operator $\mathcal{A}$, which has the same matrix in some orthonormal basis as a form q. Then in any orthonormal basis $\mathrm{Mat}(q) = \mathrm{Mat}(\mathcal{A})$. It remains to use the eigenbasis from Theorem 5.2. ∎

§ 6. Canonical form of an orthogonal operator

Theorem 6.1. *For any orthogonal operator in a Euclidean space, there exists an orthonormal basis, where*

$$A = \begin{pmatrix} \Pi(\alpha_1) & & & & \\ & \ddots & & & \\ & & \Pi(\alpha_k) & & \\ & & & \mathrm{diag}(-1,\dots,-1) & \\ & & & & \mathrm{diag}(1,\dots,1) \end{pmatrix},$$

where (see Exercise 1)

$$\Pi(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Proof. Induction by $n = \dim V$. Case $n = 1$ is trivial: $A = (\pm 1)$.

Let $n = 2$ and (e_1, e_2) be an orthonormal basis. Suppose $\angle(\mathcal{A}e_1, e_1) = \alpha$. Since $\mathcal{A}e_1 \perp \mathcal{A}e_2$, then either $\mathcal{A}$ is a rotation on α (and $A = \Pi(\alpha)$) or $\mathcal{A}$ is a reflection with respect to the bisector of the angle between e_1 and $\mathcal{A}e_1$, and in this case $A = \mathrm{diag}(-1, 1)$ in a suitable basis.

For $n > 2$ we can choose again an invariant (1- or 2-dim) subspace U (see Theorem 3.1) and $V = U \oplus U^\perp$ (see Theorem 5.1), where $\dim U, \dim U^\perp < n$ and we can use the induction hypothesis. ■

§ 7. Orthonormal basis for a Hermitian and a unitary operator

Theorem 7.1. *Eigenvalues of a hermitian operator are real numbers, and eigenvalues of a unitary operator have absolute values equal 1.*

Proof. If $\mathcal{A}$ is Hermitian, then $\lambda(e,e) = (\mathcal{A}e, e) = (e, \mathcal{A}e) = \overline{\lambda}(e,e)$, i.e. $\lambda = \overline{\lambda} \in \mathbb{R}$.

If $\mathcal{A}$ is unitary, then $(\mathcal{A}e, \mathcal{A}e) = \lambda\overline{\lambda}(e,e) = (e,e)$, that is $|\lambda| = 1$. ■

Theorem 7.2. *For any Hermitian or unitary operator $\mathcal{A}$ the subspace $U^\perp$ is invariant if U is invariant.*

For any Hermitian or unitary operator $\mathcal{A}$ there exists an orthonormal eigenbasis.

Proof. Similar to the Euclidean case. ■

§ 8. Polar decomposition

Definition 8.1. *An operator $\mathcal{A}$ is called positive definite ($\mathcal{A} > 0$) if the corresponding quadratic form $q(x) = (\mathcal{A}x, x) > 0$ (is positive definite). It is equivalent to the fact that $\lambda_1, \dots, \lambda_n > 0$.*

Lemma 8.1. *Prove that for any positive definite symmetric linear operator $\mathcal{A}$ there is a unique positive definite symmetric linear operator $\mathcal{B}$ such that $\mathcal{A} = \mathcal{B}^2$.*

Proof. In some orthonormal basis $\mathrm{Mat}(\mathcal{A}) = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. We take in this basis $\mathrm{Mat}(\mathcal{B}) = \mathrm{diag}(\sqrt{\lambda}_1, \dots, \sqrt{\lambda}_n)$. Since all $\sqrt{\lambda_j} > 0$, then $\mathcal{B} > 0$.

This operator $\mathcal{B}$ is unique, since we can consider different eigenvalues $\mu_1, \dots, \mu_m$ for $\mathcal{B}$, and $V = V_{\mu_1}(\mathcal{B}) \oplus \dots \oplus V_{\mu_m}(\mathcal{B})$, where the summands are pairwise orthogonal. The operator $\mathcal{B}$ acts on each $V_\mu(\mathcal{B})$ as a multiplication by μ^2. Thus, $V_{\mu_j}(\mathcal{B}) = V_{\mu_j^2}(\mathcal{A})$. It means that μ_j and $V_{\mu_j}(\mathcal{B})$ are uniquely determined. ■

Theorem 8.1. (Polar Decomposition.) *Prove that each invertible operator $\mathcal{A}$ in a Euclidean space can be decomposed in so called 'polar decomposition'*

$$\mathcal{A} = \mathcal{S}_1\mathcal{O}_1 = \mathcal{O}_2\mathcal{S}_2,$$

where operators $\mathcal{S}_j$ are unique positive definite symmetric operators and $\mathcal{O}_j$ are unique orthogonal operators.

Proof. Let us consider $\mathcal{A}\mathcal{A}^*$. Clearly, it is symmetric and $\mathcal{A}\mathcal{A}^* > 0$. Then there exists $\mathcal{S} > 0$, such that $\mathcal{S}^2 = \mathcal{A}\mathcal{A}^*$, and it remains to take $\mathcal{O} = \mathcal{S}^{-1}\mathcal{A}$.

(We can check: if $\mathcal{A} = \mathcal{S}\mathcal{O}$, then $\mathcal{A}\mathcal{A}^* = \mathcal{S}(\mathcal{O}\mathcal{O}^*)\mathcal{S}^* = \mathcal{S}^2$.) ∎

§ 9. Exercises

1. Find the matrix of a linear operator of a 2-dimensional rotation on some angle α. We will denote it by $\Pi(\alpha)$.

2. Find the matrix of an operator of a rotation on an angle $\alpha = 2\pi/3$ around the line $\ell = \{x \in \mathbb{R}^3 \mid x_1 = x_2 = x_3\}$.

3. Suppose $\mathcal{A}$ is a real operator. Is it true that $f_{\mathcal{A}}(t) \equiv f_{\mathcal{A}^*}(t)$?

4. Prove that if $\mathcal{A}$ and $\mathcal{B}$ are similar to each other (that is, there exists C, such that $C^{-1}AC = B$), then $f_{\mathcal{A}}(t) = f_{\mathcal{B}}(t)$.

5. What are the invariants of an operator $\mathcal{A}$ with respect to changing of a basis? Find at least two of them.