Тема II. Линейные операторы

§ 5. Нормальные операторы

М.В.Волков

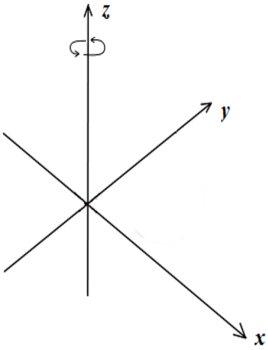
Уральский федеральный университет
Институт естественных наук и математики
кафедра алгебры и фундаментальной информатики

2021/2022 учебный год

Определение

Пусть V – векторное пространство, $\mathcal{A}\colon V \to V$ – линейный оператор. Подпространство $S \subseteq V$ называется *инвариантным относительно* $\mathcal{A}$ или $\mathcal{A}$-*инвариантным*, если $\mathcal{A}\mathbf{x} \in S$ для любого $\mathbf{x} \in S$.

Примеры. 1) Если $\mathcal{A}$ – поворот обычного трехмерного пространства относительно оси Qz на какой-то угол θ, то плоскость Oxy и прямая Oz будут $\mathcal{A}$-инвариантными подпространствами.



Инвариантные подпространства и матрицы операторов

Пусть V – векторное пространство, $\mathcal{A}\colon V \to V$ – линейный оператор, $S \subset V$ – ненулевое подпространство, инвариантное относительно $\mathcal{A}$. Обозначим $n := \dim V$, $k := \dim S$; тогда $1 \le k < n$. Выберем в S базис $\mathbf{e}_1, \dots, \mathbf{e}_k$ и дополним его векторами $\mathbf{e}_{k+1}, \dots, \mathbf{e}_n$ до базиса V.
Как выглядит матрица A оператора $\mathcal{A}$ в базисе $\mathbf{e}_1, \dots, \mathbf{e}_n$?
Поскольку подпространство S инвариантно, $\mathcal{A}\mathbf{e}_i \in S$ при $i = 1, \dots, k$.
Поэтому при $i = 1, \dots, k$ в разложении вектора $\mathcal{A}\mathbf{e}_i$ по базису $\mathbf{e}_1, \dots, \mathbf{e}_n$ ненулевые коэффициенты могут быть только у векторов $\mathbf{e}_1, \dots, \mathbf{e}_k$.
Это означает, что матрица A будет *верхней полураспавшейся*:
$A = \begin{pmatrix} B & C \\ O & D \end{pmatrix}$; у нее будет $k \times k$-блок B, отвечающий векторам $\mathbf{e}_1, \dots, \mathbf{e}_k$, под которым будет идти нулевая $(n-k) \times k$-матрица O.
Матрица B есть не что иное как матрица *ограничения оператора* $\mathcal{A}$ на подпространство S в базисе $\mathbf{e}_1, \dots, \mathbf{e}_k$.

Допустим теперь, что пространство V является *прямой суммой* ненулевых $\mathcal{A}$-инвариантных подпространств $S_1, \ldots, S_t$.
Выберем в каждом S_i базис; объединение этих базисов есть базис V.
Как выглядит матрица A оператора $\mathcal{A}$ в устроенном так базисе?

Определение

Квадратная матрица называется *блочно-диагональной*, если ее можно разбить на блоки A_{ij} так, что все блоки A_{ij} при $i \neq j$ нулевые матрицы, а все диагональные блоки A_{ii} – квадратные матрицы:

$$\begin{pmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{tt} \end{pmatrix}.$$

Понятно, что матрица A будет блочно-диагональной, причем i-й диагональный блок будет матрицей ограничения оператора $\mathcal{A}$ на подпространство S_i в выбранном в этом подпространстве базисе.

Пусть V – пространство со скалярным произведением, а S – подпространство в V. Вспомним, что множество $S^{\perp}$ всех векторов, ортогональных к произвольному вектору из S, называется *ортогональным дополнением* подпространства S. Ортогональное дополнение подпространства само является подпространством и $V = S \oplus S^{\perp}$.

Лемма 1

Если V – пространство со скалярным произведением, а S – его подпространство, инвариантное относительно линейного оператора $\mathcal{A}\colon V \to V$, то подпространство $S^{\perp}$ инвариантно относительно сопряженного оператора $\mathcal{A}^*\colon V \to V$.

Доказательство. Возьмем произвольные вектора $\mathbf{x} \in S$ и $\mathbf{y} \in S^{\perp}$. Имеем

$$\begin{aligned} \mathbf{x}\mathcal{A}^*\mathbf{y} &= \mathcal{A}\mathbf{x}\mathbf{y} && \text{свойство сопряженного оператора} \\ &= 0 && \text{так как } \mathcal{A}\mathbf{x} \in S, \text{ а } \mathbf{y} \in S^{\perp}. \end{aligned}$$

Итак, вектор $\mathcal{A}^*\mathbf{y}$ ортогонален произвольному вектору $\mathbf{x} \in S$, откуда $\mathcal{A}^*\mathbf{y} \in S^{\perp}$. □

Нормальный оператор

Определение

Линейный оператор $\mathcal{A}\colon V \to V$ называется *нормальным*, если он перестановочен со своим сопряженным, т.е. если $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$.

Примерами нормальных операторов служат *самосопряженные* операторы (когда $\mathcal{A}^* = \mathcal{A}$) и *унитарные/ортогональные* операторы (когда $\mathcal{A}^* = \mathcal{A}^{-1}$). Эти типы нормальных операторов будут рассмотрены позднее, а сейчас займемся произвольными нормальными операторами.

Лемма 2

Пусть $\mathbf{x}$ – собственный вектор нормального оператора $\mathcal{A}$, принадлежащий собственному значению λ. Тогда $\mathbf{x}$ является собственным вектором сопряженного оператора $\mathcal{A}^*$, принадлежащим собственному значению $\overline{\lambda}$.

Доказательство. Положим $\mathcal{B} := \mathcal{A} - \lambda\mathcal{E}$, где $\mathcal{E}$ – тождественный оператор. Тогда $\mathcal{B}^* = \mathcal{A}^* - \overline{\lambda}\mathcal{E}$ и из $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$ следует, что

$$\mathcal{B}\mathcal{B}^* = (\mathcal{A} - \lambda\mathcal{E})(\mathcal{A}^* - \overline{\lambda}\mathcal{E}) = (\mathcal{A}^* - \overline{\lambda}\mathcal{E})(\mathcal{A} - \lambda\mathcal{E}) = \mathcal{B}^*\mathcal{B}.$$

Хотим доказать, что $\mathcal{A}^*\mathbf{x} = \overline{\lambda}\mathbf{x}$, т.е. что $(\mathcal{A}^* - \overline{\lambda}\mathcal{E})\mathbf{x} = \mathcal{B}^*\mathbf{x} = \mathbf{0}$. Для этого достаточно убедиться, что $\mathcal{B}^*\mathbf{x}\mathcal{B}^*\mathbf{x} = 0$.

Имеем

$$\begin{aligned} \mathcal{B}^*\mathbf{x}\mathcal{B}^*\mathbf{x} &= \mathbf{x}\mathcal{B}\mathcal{B}^*\mathbf{x} && \text{свойство сопряженного оператора} \\ &= \mathbf{x}\mathcal{B}^*\mathcal{B}\mathbf{x} && \text{так как } \mathcal{B}\mathcal{B}^* = \mathcal{B}^*\mathcal{B} \\ &= \mathcal{B}\mathbf{x}\mathcal{B}\mathbf{x} && \text{свойство сопряженного оператора} \\ &= 0 && \text{так как } \mathcal{B}\mathbf{x} = (\mathcal{A} - \lambda\mathcal{E})\mathbf{x} = \mathbf{0}. \quad \square \end{aligned}$$

Мы знаем, что собственные вектора, принадлежащие попарно различным собственным значениям, линейно независимы.
Для нормальных операторов верно более сильное свойство:

Следствие

Собственные вектора нормального оператора, принадлежащие различным собственным значениям, ортогональны.

Доказательство. Пусть $\mathbf{x}$ и $\mathbf{y}$ – собственные вектора нормального оператора $\mathcal{A}$, принадлежащие соответственно λ и μ. Имеем

$$\lambda\mathbf{x}\mathbf{y} = \mathcal{A}\mathbf{x}\mathbf{y} = \mathbf{x}\mathcal{A}^*\mathbf{y} = \mathbf{x}\overline{\mu}\mathbf{y} = \mu\mathbf{x}\mathbf{y} \quad \text{т.е.} \quad (\lambda - \mu)\mathbf{x}\mathbf{y} = 0.$$

При $\lambda \neq \mu$ отсюда следует $\mathbf{x}\mathbf{y} = 0$. $\square$

Теорема 1

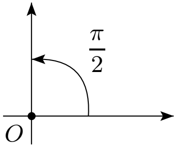
Линейный оператор $\mathcal{A}$ на унитарном пространстве V нормален тогда и только тогда, когда в V существует ортонормированный базис из собственных векторов $\mathcal{A}$.

Доказательство. Необходимость. Индукция по $\dim V$ с очевидной базой.

При $\dim V > 1$ возьмем собственный вектор $\mathbf{x}$ оператора $\mathcal{A}$. Его орт $\mathbf{e}_1$ также будет собственным вектором для $\mathcal{A}$, а по лемме 2 $\mathbf{e}_1$ будет собственным вектором и для сопряженного оператора $\mathcal{A}^*$. Подпространство S, натянутое на $\mathbf{e}_1$, инвариантно относительно $\mathcal{A}$ и $\mathcal{A}^*$. По лемме 1 ортогональное дополнение $S^\perp$ также инвариантно относительно $\mathcal{A}$ и $\mathcal{A}^*$. Понятно, что ограничения операторов $\mathcal{A}$ и $\mathcal{A}^*$ на $S^\perp$ перестановочны между собой, в силу чего ограничение $\mathcal{A}$ на $S^\perp$ – нормальный оператор. Поскольку $\dim S^\perp = \dim V - \dim S < \dim V$, по предположению индукции в $S^\perp$ существует ортонормированный базис $\mathbf{e}_2, \ldots, \mathbf{e}_{\dim V}$ из собственных векторов ограничения $\mathcal{A}$ на $S^\perp$. Добавив к нему вектор $\mathbf{e}_1$, получим ортонормированный базис всего пространства V, состоящий из собственных векторов оператора $\mathcal{A}$.

Достаточность. Матрица A оператора $\mathcal{A}$ в базисе из собственных векторов этого оператора диагональна. Раз базис ортонормированный, то в этом базисе матрица сопряженного оператора $\mathcal{A}^*$ равна эрмитово сопряженной к A матрице $A^* = \overline{A^T}$ и, следовательно, тоже диагональна. Поэтому $AA^* = A^*A$, так как диагональные матрицы перестановочны. Отсюда $\mathcal{A}\mathcal{A}^* = \mathcal{A}^*\mathcal{A}$, т.е. $\mathcal{A}$ – нормальный оператор. □

Для нормального оператора евклидова пространства аналог теоремы 1 неверен, поскольку такой оператор может не иметь собственных векторов. Примером может служить оператор $\mathcal{R}_{\frac{\pi}{2}}$ поворота плоскости на угол $\frac{\pi}{2}$, матрица которого равна $R := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.



Ее эрмитово сопряженная матрица есть попросту $R^T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ и легко подсчитать, что $RR^T = R^TR = E$. Поэтому оператор $\mathcal{R}_{\frac{\pi}{2}}$ нормален, но собственных векторов у $\mathcal{R}_{\frac{\pi}{2}}$ нет.

Несмотря на указанную трудность, структуру нормального оператора на евклидовом пространстве удается полностью прояснить.

Теорема 2

Линейный оператор $\mathcal{A}$ на евклидовом пространстве V нормален тогда и только тогда, когда в V существует ортонормированный базис, в котором матрица оператора $\mathcal{A}$ блочно-диагональна с диагональными блоками либо размера 1, либо размера 2 и вида $\rho\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$.

Вот «развернутый» вид матрицы из формулировки теоремы 2:

$$\begin{pmatrix} \alpha_1 & & & & & & & \\ & \alpha_2 & & & & & & \\ & & \dots & & & & & \\ & & & \alpha_k & & & & \\ & & & & \rho_1 \begin{pmatrix} \cos\varphi_1 & \sin\varphi_1 \\ -\sin\varphi_1 & \cos\varphi_1 \end{pmatrix} & & & \\ & & & & & \rho_2 \begin{pmatrix} \cos\varphi_2 & \sin\varphi_2 \\ -\sin\varphi_2 & \cos\varphi_2 \end{pmatrix} & & \\ & & & & & & \dots & \\ & & & & & & & \rho_m \begin{pmatrix} \cos\varphi_m & \sin\varphi_m \\ -\sin\varphi_m & \cos\varphi_m \end{pmatrix} \end{pmatrix}.$$

Необходимость. Индукция по $\dim V$ с очевидной базой.
Пусть $\dim V > 1$. Рассмотрим два случая.

Случай 1.

У оператора $\mathcal{A}$ есть действительное собственное значение.

В этом случае проходят рассуждения из доказательства теоремы 1.

Возьмем собственный вектор $\mathbf{x}$ оператора $\mathcal{A}$. Его орт $\mathbf{e}_1$ также будет собственным вектором для $\mathcal{A}$, а по лемме 2 $\mathbf{e}_1$ будет собственным вектором и для сопряженного оператора $\mathcal{A}^*$. Подпространство S, натянутое на $\mathbf{e}_1$, инвариантно относительно $\mathcal{A}$ и $\mathcal{A}^*$. По лемме 1 ортогональное дополнение $S^\perp$ также инвариантно относительно $\mathcal{A}$ и $\mathcal{A}^*$. Понятно, что ограничения операторов $\mathcal{A}$ и $\mathcal{A}^*$ на $S^\perp$ перестановочны между собой, в силу чего ограничение $\mathcal{A}$ на $S^\perp$ – нормальный оператор. Поскольку $\dim S^\perp = \dim V - \dim S < \dim V$, по предположению индукции в $S^\perp$ существует ортонормированный базис $\mathbf{e}_2, \dots, \mathbf{e}_{\dim V}$, в котором матрица A' ограничения $\mathcal{A}$ на $S^\perp$ имеет указанный в теореме блочно-диагональный вид. Добавив к нему вектор $\mathbf{e}_1$, получим ортонормированный базис всего пространства V. Матрица оператора $\mathcal{A}$ в этом базисе получается из A' добавлением одного блока размера 1.

Случай 2.

У оператора $\mathcal{A}$ нет действительных собственных значений.

В этом случае характеристический многочлен оператора $\mathcal{A}$ разлагается в произведение квадратных трехчленов с отрицательным дискриминантом. Возьмём один из них и пусть $\alpha = \sigma + \tau i$ и $\overline{\alpha} = \sigma - \tau i$ – его корни. Зафиксируем некоторый ортонормированный базис пространства V и запишем в нём матрицу A оператора $\mathcal{A}$. Возьмем теперь унитарное пространство U размерности $\dim V$, зафиксируем в нём некоторый ортонормированный базис и определим на U линейный оператор $\mathcal{A}_{\mathbb{C}}$ с матрицей A (*комплексификация* $\mathcal{A}$). Так как A – действительная матрица, $A^* = A^T$, а так как $\mathcal{A}$ – нормальный оператор, $AA^T = A^T A$. Заключаем, что и $\mathcal{A}_{\mathbb{C}}$ – нормальный оператор. Отметим еще, что отождествляя зафиксированные базисы в V и U, можно считать, что $V \subset U$. При таком отождествлении $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ действуют на V одинаково. Характеристические многочлены операторов $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ совпадают, так что α и $\overline{\alpha}$ – собственные значения оператора $\mathcal{A}_{\mathbb{C}}$. Если $\mathbf{x}$ – собственный вектор оператора $\mathcal{A}_{\mathbb{C}}$, принадлежащий α, то $A[\mathbf{x}] = \alpha[\mathbf{x}]$. Сопрягая это равенство, с учетом того, что A – действительная матрица, получаем $A[\overline{\mathbf{x}}] = \overline{\alpha}[\overline{\mathbf{x}}]$. Видим, что $\overline{\mathbf{x}}$ – собственный вектор оператора $\mathcal{A}_{\mathbb{C}}$, принадлежащий $\overline{\alpha}$.

Запишем $\mathbf{x} = \mathbf{a} + \mathbf{b}i$, где $\mathbf{a}, \mathbf{b} \in V$. Тогда $\overline{\mathbf{x}} = \mathbf{a} - \mathbf{b}i$. Выразив отсюда вектора $\mathbf{a}$ и $\mathbf{b}$, получим: $\mathbf{a} = \dfrac{\mathbf{x} + \overline{\mathbf{x}}}{2}$ и $\mathbf{b} = \dfrac{\mathbf{x} - \overline{\mathbf{x}}}{2i}$. Учитывая, что $\mathcal{A}$ и $\mathcal{A}_{\mathbb{C}}$ действуют на V одинаково и $\alpha = \sigma + \tau i$ и $\overline{\alpha} = \sigma - \tau i$, имеем

$$\mathcal{A}\mathbf{a} = \mathcal{A}_{\mathbb{C}}\mathbf{a} = \frac{1}{2}(\mathcal{A}_{\mathbb{C}}\mathbf{x} + \mathcal{A}_{\mathbb{C}}\overline{\mathbf{x}}) = \frac{1}{2}(\alpha\mathbf{x} + \overline{\alpha}\,\overline{\mathbf{x}}) =$$
$$= \frac{1}{2}((\sigma + \tau i)\mathbf{x} + (\sigma - \tau i)\overline{\mathbf{x}}) = \frac{1}{2}\sigma(\mathbf{x} + \overline{\mathbf{x}}) + \frac{1}{2}\tau i(\mathbf{x} - \overline{\mathbf{x}}) = \sigma\mathbf{a} - \tau\mathbf{b}.$$

Аналогично,

$$\mathcal{A}\mathbf{b} = \mathcal{A}_{\mathbb{C}}\mathbf{b} = \frac{1}{2i}(\mathcal{A}_{\mathbb{C}}\mathbf{x} - \mathcal{A}_{\mathbb{C}}\overline{\mathbf{x}}) = \frac{1}{2i}(\alpha\mathbf{x} - \overline{\alpha}\,\overline{\mathbf{x}}) =$$
$$= \frac{1}{2i}((\sigma + \tau i)\mathbf{x} - (\sigma - \tau i)\overline{\mathbf{x}}) = \frac{1}{2i}\sigma(\mathbf{x} - \overline{\mathbf{x}}) + \frac{1}{2}\tau(\mathbf{x} + \overline{\mathbf{x}}) = \tau\mathbf{a} + \sigma\mathbf{b}.$$

Итак, подпространство S в V, натянутое на вектора $\mathbf{a}$ и $\mathbf{b}$, инвариантно относительно оператора $\mathcal{A}$. Из леммы 2 вытекает, что вектора $\mathbf{x}$ и $\overline{\mathbf{x}}$ – собственные для оператора $\mathcal{A}_{\mathbb{C}}^*$ и принадлежат собственным значениям $\overline{\alpha}$ и α соответственно. Пользуясь этим легко проверить, что подпространство S инвариантно и относительно оператора $\mathcal{A}^*$.

По лемме 1 ортогональное дополнение $S^{\perp}$ также инвариантно относительно $\mathcal{A}$ и $\mathcal{A}^*$. Понятно, что ограничения операторов $\mathcal{A}$ и $\mathcal{A}^*$ на $S^{\perp}$ перестановочны между собой, в силу чего ограничение $\mathcal{A}$ на $S^{\perp}$ – нормальный оператор. Поскольку $\dim S^{\perp} = \dim V - \dim S < \dim V$, по предположению индукции в $S^{\perp}$ существует ортонормированный базис $\mathbf{e}_3, \dots, \mathbf{e}_{\dim V}$, в котором матрица ограничения $\mathcal{A}$ на $S^{\perp}$ имеет указанный в теореме блочно-диагональный вид. Остается построить «правильный» ортонормированный базис в подпространстве S.

По следствию леммы 2 вектора $\mathbf{x}$ и $\overline{\mathbf{x}}$ ортогональны как собственные вектора нормального оператора $\mathcal{A}_{\mathbb{C}}$, принадлежащие его различным собственным значениям α и $\overline{\alpha}$. Отсюда

$$0 = \mathbf{x}\overline{\mathbf{x}} = (\mathbf{a} + \mathbf{b}i)(\mathbf{a} - \mathbf{b}i) = \mathbf{aa} - \mathbf{a}(\mathbf{b}i) + (\mathbf{b}i)\mathbf{a} - (\mathbf{b}i)(\mathbf{b}i) =$$
$$= \mathbf{aa} + i\mathbf{ab} + i\mathbf{ba} - \mathbf{bb} = |\mathbf{a}|^2 - |\mathbf{b}|^2 + 2i\mathbf{ab}.$$

Заключаем, что $|\mathbf{a}|^2 - |\mathbf{b}|^2 = 0$ и $\mathbf{ab} = 0$, т.е. $|\mathbf{a}| = |\mathbf{b}|$ и $\mathbf{a} \perp \mathbf{b}$. Поэтому орты $\mathbf{e}_1 = \dfrac{\mathbf{a}}{|\mathbf{a}|}$ и $\mathbf{e}_2 = \dfrac{\mathbf{b}}{|\mathbf{a}|}$ образуют ортонормированный базис в S.

Выше подсчитано, что $\mathcal{A}\mathbf{a} = \sigma\mathbf{a} - \tau\mathbf{b}$, а $\mathcal{A}\mathbf{b} = \tau\mathbf{a} + \sigma\mathbf{b}$. Разделив эти равенства на $|\mathbf{a}| = |\mathbf{b}|$, получим действие оператора $\mathcal{A}$ на базис $\mathbf{e}_1, \mathbf{e}_2$:
$$\begin{cases} \mathcal{A}\mathbf{e}_1 = \sigma\mathbf{e}_1 - \tau\mathbf{e}_2, \\ \mathcal{A}\mathbf{e}_2 = \tau\mathbf{e}_1 + \sigma\mathbf{e}_2. \end{cases}$$
Итак, матрица ограничения оператора $\mathcal{A}$ на подпространство S в базисе $\mathbf{e}_1, \mathbf{e}_2$ равна $\begin{pmatrix} \sigma & \tau \\ -\tau & \sigma \end{pmatrix}$. Если записать комплексное число $\alpha = \sigma + \tau i$ в тригонометрической форме $\alpha = \rho(\cos\varphi + i\sin\varphi)$, то эта матрица запишется как $\rho\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}$, т.е. в точности в том виде, который указан в формулировке теоремы 2. Итак, добавив к базису $\mathbf{e}_3, \dots, \mathbf{e}_{\dim V}$ подпространства $S^\perp$ вектора $\mathbf{e}_1$ и $\mathbf{e}_2$, получим ортонормированный базис всего пространства V, в котором матрица оператора $\mathcal{A}$ имеет требуемый вид.

Достаточность. Для действительных матриц эрмитово сопряжение сводится к транспонированию. Легко проверяется, что каждый блок блочно-диагональной матрицы из формулировки теоремы 2 перестановочен со своей транспонированной матрицей (*проверьте!*). $\square$

Теорема о строении нормального оператора на евклидовом пространстве – еще один пример «100% действительного» факта, для **формулировки** которого комплексные числа не нужны, но **доказательство** которого использует комплексные числа.

Сравнение формулировок и особенно доказательств теорем 1 и 2 еще раз показывает, насколько комплексные числа лучше действительных!