

Полярное разложение линейных операторов

М.В.Волков

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

2023/2024 учебный год

На прошлой лекции мы обсуждали два важных типа линейных операторов на пространстве со скалярным произведением: *изометрические* операторы и *самосопряженные* операторы.

Мы определяли через некоторые формальные условия, но из обсуждения стал ясен их содержательный геометрический смысл.

Изометрические операторы – это в точности движения, сохраняющие расстояния между точками. Они не меняют форму и размеры «твердых тел», а только поворачивают пространство вокруг начала координат или отражают его относительно каких-то подпространств.

Самосопряженные операторы – это растяжения/сжатия или проектирования вдоль каких-то взаимно перпендикулярных осей, возможно в сочетании с отражениями относительно подпространств.

Сегодня мы покажем, что произвольный линейный оператор на пространстве со скалярным произведением является произведением некоторого самосопряженного и некоторого изометрического операторов.

Итак, на пространстве со скалярным произведением любой линейный оператор сначала растягивает/сжимает или проектирует пространство вдоль взаимно перпендикулярных осей, а затем поворачивает его вокруг начала координат или отражает относительно какого-то подпространства.

Хотим представить оператор \mathcal{A} в виде $\mathcal{A} = \mathcal{B}\mathcal{U}$, где $\mathcal{B}^* = \mathcal{B}$, $\mathcal{U}^* = \mathcal{U}^{-1}$.
 Откуда взять такие операторы \mathcal{B} и \mathcal{U} ? Заметим, что $\mathcal{A}^* = \mathcal{U}^*\mathcal{B}^* = \mathcal{U}^{-1}\mathcal{B}$.
 Перемножая, получаем $\mathcal{A}\mathcal{A}^* = \mathcal{B}\mathcal{U}\mathcal{U}^{-1}\mathcal{B} = \mathcal{B}^2$.
 Итак, оператор \mathcal{B} необходимо должен быть «квадратным корнем» из $\mathcal{A}\mathcal{A}^*$.
 Но мы не извлекали квадратные корни из операторов! Придется научиться.

Определение

Самосопряженный оператор $\mathcal{A}: V \rightarrow V$ называется *неотрицательным*, если $(\mathcal{A}\mathbf{x}, \mathbf{x}) \geq 0$ для любого вектора $\mathbf{x} \in V$. Если $(\mathcal{A}\mathbf{x}, \mathbf{x}) > 0$ для любого ненулевого вектора $\mathbf{x} \in V$, оператор называется *положительным*.

$\nabla 1$. Для любого линейного оператора \mathcal{A} оператор $\mathcal{A}\mathcal{A}^*$ неотрицательный.

Доказательство. Имеем $(\mathcal{A}\mathcal{A}^*)^* = (\mathcal{A}^*)^*\mathcal{A}^* = \mathcal{A}\mathcal{A}^*$, поэтому $\mathcal{A}\mathcal{A}^*$ – самосопряженный оператор. Для любого вектора \mathbf{x} имеем

$$((\mathcal{A}\mathcal{A}^*)\mathbf{x}, \mathbf{x}) = (\mathcal{A}^*(\mathcal{A}\mathbf{x}), \mathbf{x}) = (\mathcal{A}\mathbf{x}, \mathcal{A}\mathbf{x}) \geq 0. \quad \square$$

∇2. Самосопряженный оператор неотрицателен тогда и только тогда, когда все его собственные значения неотрицательны.

Доказательство. Необходимость. Пусть \mathcal{A} – неотрицательный оператор, λ – его собственное значение, а \mathbf{x} – собственный вектор, принадлежащий λ . Тогда $\lambda(\mathbf{x}, \mathbf{x}) = (\lambda\mathbf{x}, \mathbf{x}) = (\mathcal{A}\mathbf{x}, \mathbf{x}) \geq 0$. Поскольку $(\mathbf{x}, \mathbf{x}) > 0$, имеем $\lambda \geq 0$.

Достаточность. Пусть \mathcal{A} – самосопряженный оператор, $\lambda_1, \dots, \lambda_n$ – все его собственные значения, а $\mathbf{e}_1, \dots, \mathbf{e}_n$ – ортонормированный базис из принадлежащих этим собственным значениям собственных векторов. Разложим произвольный вектор \mathbf{x} по этому базису: $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$. Тогда $(\mathcal{A}\mathbf{x}, \mathbf{x}) = (\mathcal{A}(x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n), x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) = (\lambda_1x_1\mathbf{e}_1 + \dots + \lambda_nx_n\mathbf{e}_n, x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n) = \lambda_1x_1\overline{x_1} + \dots + \lambda_nx_n\overline{x_n} \geq 0$. \square

∇3. Неотрицательный оператор положителен тогда и только тогда, когда он невырожденный.

Это сразу следует из доказательства **∇2**, если учесть, что оператор невырожденный тогда и только тогда, когда все его собственные значения отличны от нуля (объясните, почему).

Теорема (существование и единственность квадратного корня)

Для любого неотрицательного оператора \mathcal{A} существует единственный неотрицательный оператор \mathcal{B} такой, что $\mathcal{B}^2 = \mathcal{A}$.

Доказательство. Существование. Возьмем тот ортонормированный базис e_1, \dots, e_n , в котором оператор \mathcal{A} имеет диагональную матрицу

$A = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$. Числа, стоящие на диагонали, суть собственные значения оператора \mathcal{A} , и по $\nabla 2$ все они неотрицательны. Оператор \mathcal{B} ,

имеющий в том же базисе матрицу $B := \begin{pmatrix} \sqrt{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\lambda_n} \end{pmatrix}$, неотрицателен

по $\nabla 2$. По построению имеем матричное равенство $B^2 = A$, откуда следует равенство операторов $\mathcal{B}^2 = \mathcal{A}$.

Единственность. Пусть $\alpha_1, \alpha_2, \dots, \alpha_s$ – все попарно различные собственные значения \mathcal{A} , т.е. $\lambda_1 = \dots = \lambda_{k_1} = \alpha_1$, $\lambda_{k_1+1} = \dots = \lambda_{k_2} = \alpha_2$, \dots , $\lambda_{k_{s-1}+1} = \dots = \lambda_n = \alpha_s$.

Квадратные корни из неотрицательных операторов (2)

Тогда матрицу A можно представить как блочную матрицу

$$A = \begin{pmatrix} \alpha_1 E_1 & O & \dots & O \\ O & \alpha_2 E_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & \alpha_s E_s \end{pmatrix}, \text{ где } E_1, E_2, \dots, E_s \text{ – единичные матрицы}$$

размеров $k_1, k_2 - k_1, \dots, n - k_{s-1}$, а через O обозначены нулевые блоки соответствующих размеров. Соответственно, матрица B имеет вид

$$B = \begin{pmatrix} \sqrt{\alpha_1} E_1 & O & \dots & O \\ O & \sqrt{\alpha_2} E_2 & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & \sqrt{\alpha_s} E_s \end{pmatrix}. \text{ Возьмем любой неотрицательный}$$

оператор C такой, что $C^2 = A$, и покажем, что $C = B$. Пусть C – матрица оператора C в том же базисе; тогда $C^2 = A$. Разбив C на блоки тех же

размеров, что и у матрицы A , запишем $C = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1s} \\ C_{21} & C_{22} & \dots & C_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ C_{s1} & C_{s2} & \dots & C_{ss} \end{pmatrix}$. Тогда

у матрицы AC блок в позиции (i, j) равен $\alpha_i C_{ij}$, а у матрицы CA в той же позиции стоит блок $C_{ij} \alpha_j = \alpha_j C_{ij}$. Поскольку $AC = C^3 = CA$, имеем $\alpha_i C_{ij} = \alpha_j C_{ij}$, т.е. $(\alpha_i - \alpha_j) C_{ij} = O$ для всех i, j , но при $i \neq j$ такое равенство возможно только, если $C_{ij} = O$.

Итак, C имеет вид $\begin{pmatrix} C_{11} & O & \dots & O \\ O & C_{22} & \dots & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \dots & C_{ss} \end{pmatrix}$. В частности, подпространство V_1 ,

натянутое на вектора $\mathbf{e}_1, \dots, \mathbf{e}_{k_1}$, инвариантно относительно оператора C . Заметим, что для любого вектора $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_{k_1}\mathbf{e}_{k_1} \in V_1$

$$\mathcal{A}\mathbf{x} = \mathcal{A}(x_1\mathbf{e}_1 + \dots + x_{k_1}\mathbf{e}_{k_1}) = x_1\alpha_1\mathbf{e}_1 + \dots + x_{k_1}\alpha_1\mathbf{e}_{k_1} = \alpha_1\mathbf{x}.$$

Ограничение оператора C на подпространство V_1 – неотрицательный оператор с матрицей C_{11} . Выберем в V_1 ортонормированный базис $\mathbf{f}_1, \dots, \mathbf{f}_{k_1}$ из собственных векторов этого оператора, принадлежащих собственным значениям $\gamma_1, \dots, \gamma_{k_1}$. Для каждого $j = 1, \dots, k_1$ имеем $\alpha_1\mathbf{f}_j = \mathcal{A}\mathbf{f}_j = C^2\mathbf{f}_j = C(C\mathbf{f}_j) = C(\gamma_j\mathbf{f}_j) = \gamma_j^2\mathbf{f}_j$. Отсюда $\gamma_j^2 = \alpha_1$, и так как $\gamma_j \geq 0$, имеем $\gamma_j = \sqrt{\alpha_1}$. Итак, $\gamma_1 = \dots = \gamma_{k_1} = \sqrt{\alpha_1}$. Вектора $\mathbf{e}_1, \dots, \mathbf{e}_{k_1}$ выражаются через $\mathbf{f}_1, \dots, \mathbf{f}_{k_1}$, откуда $C\mathbf{e}_i = \sqrt{\alpha_1}\mathbf{e}_i$ для всех $i = 1, \dots, k_1$. Блок $C_{11} = \sqrt{\alpha_1}E_1$ равен первому диагональному блоку матрицы B . Тот же аргумент работает для любого диагонального блока матрицы C , откуда $C = B$ и $\mathcal{C} = \mathcal{B}$. □

Теорема (полярное разложение невырожденного оператора)

Для любого невырожденного линейного оператора A на унитарном (евклидовом) пространстве существует единственная пара операторов B и U такая, что $A = BU$, причем B – положительный оператор, а U – унитарный (ортогональный) оператор.

Доказательство. Существование. По $\nabla 1$ и $\nabla 3$ оператор AA^* положительный. Квадратный корень B из AA^* также будет положительным оператором. Пусть $U := B^{-1}A$; тогда ясно, что $A = BU$, и остается проверить, что U – унитарный (ортогональный) оператор. Действительно, $U^* = A^*B^{-1}$, так как B^{-1} – самосопряженный оператор и $UU^* = B^{-1}AA^*B^{-1} = B^{-1}B^2B^{-1} = \mathcal{E}$. Итак, $U^* = U^{-1}$.

Единственность. Мы уже отмечали, что если $A = BU$, где B – неотрицательный, а U – изометрический операторы, то B необходимо будет квадратным корнем из AA^* . Если B – положительный оператор, то U однозначно определяется из равенства $A = BU$ как $B^{-1}A$. □

Следствие

Любая невырожденная матрица над полем \mathbb{C} однозначно представима как произведение эрмитовой и унитарной матриц.

Используя те же рассуждения, можно разложить любой невырожденный линейный оператор в произведение изометрического и положительного:

Теорема (дуальный вариант полярного разложения)

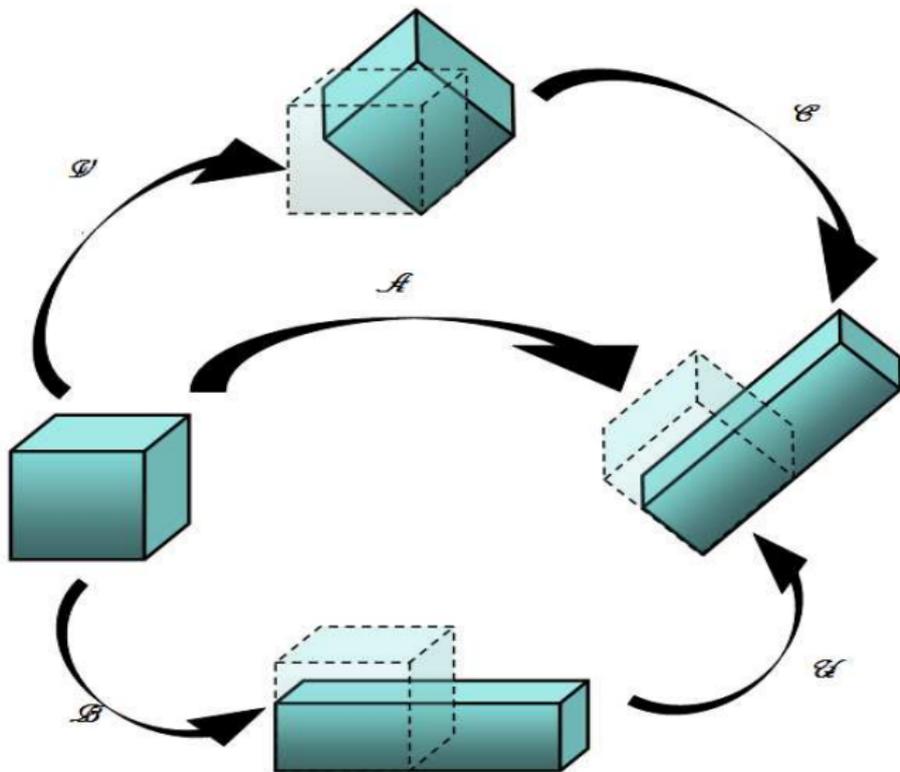
Для любого невырожденного линейного оператора A на унитарном (евклидовом) пространстве существует единственная пара операторов C и V такая, что $A = VC$, причем C – положительный оператор, а V – унитарный (ортогональный) оператор.

Докажите этот вариант самостоятельно. Подсказка: примените reverse engineering и поймите, какой оператор необходимо взять в роли C .

Замечание 1. Из того, что $A = BU = VC$, не следует, вообще говоря, что $B = C$ и $U = V$.

Замечание 2. Полярное разложение невырожденного оператора напоминает представление ненулевого комплексного числа в тригонометрической форме – как произведение положительного действительного числа и комплексного числа с модулем 1.

Замечание 3. Полярное разложение раскрывает геометрический смысл произвольного невырожденного линейного преобразования.



Теорема (полярное разложение произвольного оператора)

Для любого линейного оператора A на унитарном (евклидовом) пространстве существует пара операторов B и U такая, что $A = BU$, где B – однозначно определяемый неотрицательный оператор, а U – унитарный (ортогональный) оператор.

Доказательство опирается на следующую лемму.

Лемма (о метрически равных операторах)

Пусть A и B – линейные операторы на пространстве V со скалярным произведением и $(Ax, Ay) = (Bx, By)$ для всех $x, y \in V$. Тогда существует изометрический оператор U такой, что $A = BU$.

Вывод теоремы из леммы. По $\nabla 1$ оператор AA^* неотрицательный. Пусть B – квадратный корень из AA^* . Для всех x, y имеем

$$(Ax, Ay) = (x, A^*(Ay)) = (x, (AA^*)y) = (x, B^2y) = (Bx, By).$$

По лемме $A = BU$ для некоторого изометрического оператора U . □

Лемма (о метрически равных операторах)

Пусть \mathcal{A} и \mathcal{B} – линейные операторы на пространстве V со скалярным произведением и $(\mathcal{A}x, \mathcal{A}y) = (\mathcal{B}x, \mathcal{B}y)$ для всех $x, y \in V$. Тогда существует изометрический оператор \mathcal{U} такой, что $\mathcal{A} = \mathcal{B}\mathcal{U}$.

Доказательство. Заметим, что из равенства $(\mathcal{A}x, \mathcal{A}x) = (\mathcal{B}x, \mathcal{B}x)$ следует, что $\mathcal{A}x = \mathbf{0}$ тогда и только тогда, когда $\mathcal{B}x = \mathbf{0}$, т.е. $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}$. По теореме о сумме ранга и дефекта ранги \mathcal{A} и \mathcal{B} равны. Выберем в $\text{Im } \mathcal{A}$ некоторый ортонормированный базис $\mathbf{e}_1, \dots, \mathbf{e}_r$, где r – ранг \mathcal{A} . Пусть $\mathbf{h}_1, \dots, \mathbf{h}_r \in V$ – какие-то вектора со свойством $\mathcal{A}\mathbf{h}_i = \mathbf{e}_i$ для каждого $i = 1, \dots, r$. Положим $\mathbf{f}_i := \mathcal{B}\mathbf{h}_i$; тогда $(\mathbf{f}_i, \mathbf{f}_j) = (\mathbf{e}_i, \mathbf{e}_j) = \begin{cases} 1 & \text{при } i = j, \\ 0 & \text{при } i \neq j. \end{cases}$

Поэтому вектора $\mathbf{f}_1, \dots, \mathbf{f}_r$ образуют ортонормированный базис в $\text{Im } \mathcal{B}$. Выберем в ортогональном дополнении $(\text{Im } \mathcal{A})^\perp$ подпространства $\text{Im } \mathcal{A}$ какой-то ортонормированный базис $\mathbf{e}_{r+1}, \dots, \mathbf{e}_n$, а в ортогональном дополнении $(\text{Im } \mathcal{B})^\perp$ – какой-то ортонормированный базис $\mathbf{f}_{r+1}, \dots, \mathbf{f}_n$. Тогда $\mathbf{e}_1, \dots, \mathbf{e}_r, \mathbf{e}_{r+1}, \dots, \mathbf{e}_n$ и $\mathbf{f}_1, \dots, \mathbf{f}_r, \mathbf{f}_{r+1}, \dots, \mathbf{f}_n$ – два ортонормированных базиса пространства V . Оператор \mathcal{U} , переводящий второй из них в первый, будет изометрическим. Покажем, что $\mathcal{A} = \mathcal{B}\mathcal{U}$.

Напомним, что мы зафиксировали вектора $\mathbf{h}_1, \dots, \mathbf{h}_r$ со свойством $\mathcal{A}\mathbf{h}_i = \mathbf{e}_i$ для каждого $i = 1, \dots, r$. Эти вектора линейно независимы, поскольку линейно независимы их образы $\mathbf{e}_1, \dots, \mathbf{e}_r$. Дополним систему $\mathbf{h}_1, \dots, \mathbf{h}_r$ базисом $\text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}$ до базиса $\mathbf{h}_1, \dots, \mathbf{h}_r, \mathbf{h}_{r+1}, \dots, \mathbf{h}_n$ всего пространства V (это конструкция из доказательства теоремы о сумме ранга и дефекта). Разложим произвольный вектор $\mathbf{x} \in V$ по этому базису:

$$\mathbf{x} = x_1\mathbf{h}_1 + \dots + x_r\mathbf{h}_r + x_{r+1}\mathbf{h}_{r+1} + \dots + x_n\mathbf{h}_n$$

и подсчитаем $\mathcal{A}\mathbf{x}$ и $(\mathcal{B}\mathcal{U})\mathbf{x}$.

$$\begin{aligned} \mathcal{A}\mathbf{x} &= \mathcal{A}(x_1\mathbf{h}_1 + \dots + x_n\mathbf{h}_n) = \text{(так как } \mathbf{h}_{r+1}, \dots, \mathbf{h}_n \in \text{Ker } \mathcal{A} = \text{Ker } \mathcal{B}) \\ &= x_1\mathcal{A}\mathbf{h}_1 + \dots + x_r\mathcal{A}\mathbf{h}_r = x_1\mathbf{e}_1 + \dots + x_r\mathbf{e}_r. \end{aligned}$$

$$\begin{aligned} (\mathcal{B}\mathcal{U})\mathbf{x} &= \mathcal{U}(\mathcal{B}\mathbf{x}) = \mathcal{U}(\mathcal{B}(x_1\mathbf{h}_1 + \dots + x_n\mathbf{h}_n)) = \\ &= \mathcal{U}(x_1\mathcal{B}\mathbf{h}_1 + \dots + x_r\mathcal{B}\mathbf{h}_r) = \mathcal{U}(x_1\mathbf{f}_1 + \dots + x_r\mathbf{f}_r) = \\ &= x_1\mathcal{U}\mathbf{f}_1 + \dots + x_r\mathcal{U}\mathbf{f}_r = x_1\mathbf{e}_1 + \dots + x_r\mathbf{e}_r. \end{aligned}$$

Итак, $\mathcal{A}\mathbf{x} = (\mathcal{B}\mathcal{U})\mathbf{x}$ для любого $\mathbf{x} \in V$, т.е. $\mathcal{A} = \mathcal{B}\mathcal{U}$. □