ЗАДАЧА О ПОДОБИИ ДЛЯ ЛИНЕЙНЫХ ОПЕРАТОРОВ

Методическая разработка для студентов 1 курса ММФ (черновик)

В.А.Чуркин

Во второй половине 19 века усилиями, в основном, Жордана, Сильвестра и Фробениуса была получена классификация линейных операторов конечномерных пространств с точностью до подобия. Ее можно получить либо «алгебраическим» методом, основанном на приведении полиномиальных матриц к канонической форме элементарными преобразованиями, либо «геометрическим», связанным с действием оператора в пространстве. В методичке избран геометрический метод приведения к жордановой нормальной форме, поскольку именно такой подход с его понятиями и теоремами востребован в приложениях теории линейных операторов. Геометрический метод в свою очередь допускает различные реализации. Здесь изложен с доказательствами новый относительно быстрый метод вычисления жордановой нормальной формы и матрицы перехода к ней, активно использующий, во-первых, одновременное вычисление ядра и образа оператора и, во вторых, сведение к инвариантным подпространствам с целью понижения размерности. Арифметика поля скаляров предполагается точной, а спектр оператора, как правило, известным. Краткое изложение ранее: В.А. Чуркин, Жорданова классификация конечномерных линейных операторов, НГУ, 1991.

Отметим, что жорданова нормальная форма — хороший теоретический инструмент с точки зрения классической математики. С другой стороны нужно понимать, что она неустойчива относительно малых изменений элементов матрицы и потому плоха с точки зрения вычислительной математики. Вопросы устойчивости спектра матрицы — отдельная большая тема, по сути относящаяся к матричному анализу. Наша цель — введение в одну из глав теории линейных операторов, жорданову нормальную форму, с акцентом на самостоятельное чтение, решение задач и использование на семинарах. Большое внимание уделено исключительно важной идее собственных векторов и диагонализируемых операторов. Часть приведенных примеров иллюстрирует доказательства теорем. Примеры другого типа призваны показать прикладной характер линейной алгебры и ее связь с геометрией, анализом и дифференциальными уравнениями. Без понимания этой связи многие задачи линейной алгебры решить трудно или даже невозможно.

Методичку следует рассматривать как вспомогательное средство при начальном изучении теории линейных операторов. Она не заменит серьезные объемные учебники и задачники.

Рекомендуемые учебники: 1) *Э.Б. Винберг*, Курс алгебры, М., МЦНМО, 2011;
2) *А.И. Кострикин*, Введение в алгебру 2. Линейная алгебра, М., ФМЛ, 2000;
3) *И.Р. Шафаревич, А.О. Ремизов*, Линейная алгебра и геометрия, М., ФМЛ, 2009.

Рекомендуемый задачник: Сборник задач по алгебре (под ред. *А.И. Кострикина*), т. 1. М., ФМЛ, 2007.

Для самостоятельной работы можно использовать следующую подборку задач: *В.А. Чуркин*, Задания по алгебре для 1 курса ММФ, НГУ, 1994 (обновленный вариант: http://www.mmfd.nsu.ru/ , Russian version, Кафедры, Кафедра алгебры и математической логики, Russian version, Учебные пособия кафедры, Месячные задания по алгебре (В.А.Чуркин, pdf)).

МАТРИЦА ЛИНЕЙНОГО ОТОБРАЖЕНИЯ

Пусть V и W — векторные пространства над полем K. Отображение $\mathcal{A} : V \to W$ называется *линейным*, если оно аддитивно и однородно:

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

Отсюда следуют свойства

$$\mathcal{A}0 = 0, \quad \mathcal{A}\left(\sum \lambda_i v_i\right) = \sum \lambda_i \mathcal{A} v_i.$$

При $W = V$ линейное отображение называется линейным преобразованием или линейным оператором, при $W = K$ — линейной функцией или линейным функционалом. Отображение $\mathcal{O} : v \mapsto 0,\ v \in V$, называется *нулевым*, оператор $\mathcal{E} : v \mapsto v,\ v \in V$, называется *единичным* или *тождественным*, оператор $\lambda\mathcal{E} : v \mapsto \lambda v,\ v \in V$, при заданном $\lambda \in K$ называется *скалярным.* Скалярные операторы играют особую роль в алгебре операторов: во-первых, они характеризуются как операторы, коммутирующие с любым линейным оператором, и, во-вторых, все скалярные операторы образуют алгебру, изоморфную полю скаляров.

Линейные отображения и операторы естественно возникают не только в алгебре, но и в геометрии и анализе, в механике и физике, в экономике и других приложениях.

Примеры. 1) *Геометрические.* Преобразование евклидова пространства, сохраняющее расстояния между любыми точками, называется движением или изометрией. В механике и физике такие преобразования имитируют движение твердого тела. Если движение имеет неподвижную точку O, то оно индуцирует линейное преобразование в пространстве радиус-векторов с началом O, поскольку отображает параллелограммы в параллелограммы и сохраняет коэффициент подобия отрезков на общей прямой (см. определение). В частности, можно считать линейными отражения пространства относительно точки, прямой или плоскости, повороты плоскости вокруг точки, повороты пространства вокруг неподвижной прямой и т. д.

Проецирование пространства геометрических векторов на плоскость параллельно некомпланарной прямой линейно.

Скалярное, векторное или смешанное произведение геометрических векторов линейно по каждому аргументу при заданных других.

2) *Алгебраические.* Пусть $V = K^n$, $W = K^s$ — пространства столбцов над полем K. Для любой матрицы A размера $s \times n$ над K отображение $\mathcal{A} : x \mapsto Ax,\ x \in K^n$, «умножение столбца на матрицу», всегда линейно.

Пусть $V = M_n(K)$ — алгебра матриц и задана матрица $A \in M_n(K)$. Тогда преобразования

$$X \mapsto X^\top, \quad L_A : X \mapsto AX, \quad R_A : X \mapsto XA,$$

линейны. Они называются соответственно транспонирование матриц, левый и правый сдвиги на матрицу A.

3) *Функциональные.* Пусть $V = \mathbb{R}[t]$ — алгебра многочленов, $a(t)$ — заданный многочлен. Тогда сдвиг $f(t) \mapsto a(t)f(t)$, дифференцирование $d/dt : f \mapsto df/dt,\ f \in \mathbb{R}[t]$, и интегрирование $f(t) \mapsto \int_a^t f(s)\,ds$ или $f(t) \mapsto \int_a^b K(s,t)f(s)\,ds$ — линейные операторы в алгебре многочленов или в алгебре непрерывных функций на отрезке $[a,\ b] \subset \mathbb{R}$. Здесь $K(s,t)$ — непрерывная функция в квадрате $[a,b] \times [a,b] \subset \mathbb{R}^2$.

В анализе доказывается, что *произвольное* дифференцируемое отображение можно локально приблизить суммой постоянного и линейного. Это линейное отображение называется дифференциалом отображения в данной точке.

Легко проверить, что сумма линейных отображений $V \to W$ линейна, после умножения линейного отображения на скаляр получается линейное отображение, композиция линейных отображений $V \to W$ и $W \to U$ линейна. В частности, линейны преобразования $X \mapsto AXB$, $X \mapsto AX - XA$ в пространстве матриц, $\mathcal{D} : f(t) \mapsto \sum_{k\geqslant 0} a_k(t)(d^k f/dt^k)$ в алгебре многочленов.

Легко проверить, что множество $L(V, W)$ всех линейных отображений $V \to W$ относительно двух первых операций образует векторное пространство, а множество $L(V)$ всех линейных преобразований пространства V относительно всех трех операций образует алгебру.

Итак, линейные отображения встречаются в разнообразной форме. Можно ли навести порядок в этом классе, представить его единообразно, распознавать подобие и существенное различие? В случае конечномерных пространств ключ к решению — алгебра матриц.

Заметим сначала, что любое линейное отображение задается образом базиса отображаемого пространства, причем образом базиса мощности n может быть любой набор из n векторов во втором пространстве.

Пусть $e_1, \ldots, e_n$ — базис пространства V, $f_1, \ldots, f_s$ — базис пространства W. Пусть

$$\mathcal{A}e_j = \sum_i a_{ij}\, f_i, \quad a_{ij} \in K.$$

Тогда матрица $A = (a_{ij})$ размера $s \times n$ называется *матрицей линейного отображения* $\mathcal{A} : V \to W$ в заданной паре базисов. Мы видим, что она составлена из координатных столбцов векторов $\mathcal{A}e_1, \ldots, \mathcal{A}e_n$ в базисе $f_1, \ldots, f_s$. Поэтому матрица однозначно задает линейное отображение при заданных базисах и может быть любой $s\times n$-матрицей над полем K. В случае линейных операторов, когда пространство только одно, матрица обычно задается выбором одного базиса и определяется по формуле $\mathcal{A}e_j = \sum_i a_{ij}\, e_i$, $a_{ij} \in K$. Если мы желаем подчеркнуть, что эта матрица вычислялась в базисе $e = (e_1, \ldots, e_n)$, то будем обозначать ее через A_e.

Примеры. 1) Пусть $\lambda = a + bi = r(\cos\varphi + i \sin\varphi)$ — комплексное число в алгебраической и тригонометрической форме. Пусть $L_\lambda : z \mapsto \lambda z$, $z \in \mathbb{C}$. Очевидно, что L_λ — линейный оператор в поле $\mathbb{C}$ как одномерном векторном пространстве над полем $\mathbb{C}$. В то же время L_λ — линейный оператор в поле $\mathbb{C}$ как двумерном векторном пространстве над $\mathbb{R}$. Найдем матрицы оператора L_λ в базисах 1 и 1, i соответственно. Имеем

$$L_\lambda 1 = \lambda \cdot 1, \quad \left\{ \begin{array}{rr} \lambda 1 = & a + bi, \\ \lambda i = & -b + ai \end{array} \right. .$$

Следовательно, матрицы оператора L_λ соответственно имеют вид

$$(\lambda), \quad \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = r \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}.$$

Отсюда следуют два вывода. Во-первых, отображение

$$a + bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

является изоморфным вложением поля комплексных чисел в алгебру вещественных матриц порядка 2. Во-вторых, мы видим, что последняя матрица задает поворот плоскости на угол φ вокруг начала координат с последующим растяжением (сжатием) с коэффициентом $r = |\lambda|$ в ортонормированном базисе $1,\ i$, поскольку именно такое геометрическое описание имеет оператор L_λ.

2) Пусть V — пространство вещественных многочленов степени $\leqslant n$ от переменной t. Выберем $c \in \mathbb{R}$. Отображение $f(t) \mapsto f(c),\quad f \in V$, линейно. Найдем его матрицу в паре базисов $1,\ t,\ t^2, \ldots, t^n$ и 1 пространств V и $\mathbb{R}$. Поскольку $t^j \mapsto c^j$, то матрица отображения имеет вид строки

$$(1\ \ c\ \ c^2\ \ \ldots\ \ c^n).$$

Упр. 1. Вычислить матрицы линейных отображений, указанных в приведенных геометрических, алгебраических и функциональных примерах, используя «естественные» базисы. В частности, докажите, что данная квадратная матрица A совпадает с матрицей оператора умножения столбцов на матрицу A в стандартном базисе пространства столбцов.

Таким образом, любое линейное отображение конечномерных пространств в заданных базисах задается конечным набором чисел-скаляров. Легко проверяется, что соответствие $\mathcal{A} \leftrightarrow A$ является изоморфизмом пространства $L(V, W)$ всех линейных отображений и пространства $K^{s\times n}$ всех прямоугольных матриц, а также алгебр $L(V)$ и $M_n(K)$, и это очень хорошо. С другой стороны, данное линейное отображение может задаваться *разными* матрицами, если менять базисы пространств. Возникает задача извлечения из матрицы тех свойств соответствующего линейного отображения, которые не зависят от выбора базиса — инвариантов отображения.

Известны формулы замены матрицы отображения при замене базисов и действия отображения в координатах.

1) Если A и A' — матрицы линейного оператора $\mathcal{A}$ в базисах $e = (e_1, \ldots, e_n)$ и $e' = (e'_1, \ldots, e'_n)$ пространства V и $C : e \to e'$ — матрица перехода, то $A' = C^{-1}AC$. Такие матрицы A и A' называются *подобными*.

2) Если A и A' — матрицы линейного отображения $\mathcal{A}$ в парах базисов $e,\ f$ и $e',\ f'$ пространства V и W и если $C : e \to e'$, $D : f \to f'$ — матрицы перехода, то $A' = D^{-1}AC$. Такие матрицы A и A' называются *эквивалентными*.

3) Координатный столбец вектора $\mathcal{A}x$ в базисе f получается умножением матрицы A отображения $\mathcal{A}$ в паре базисов $e,\ f$ на координатный столбец вектора x в базисе e. Другими словами, действие линейного отображения в координатах сводится к умножению столбца на матрицу слева.

Задача эквивалентности матриц решается легко за счет свободы в выборе двух базисов: матрицы A и B одного размера над одним полем эквивалентны в том и только том случае, когда $\operatorname{rk} A = \operatorname{rk} B$ (проверьте это). Она становится содержательной, если ограничиться выбором только ортонормированных базисов над полями вещественных или комплексных чисел. При этом условии эквивалентность означает равенство сингулярных чисел матриц. Эта тема важна с точки зрения вычислительной математики и рассматривается при классификации линейных отображений евклидовых или эрмитовых пространств.

Задача подобия для квадратных матриц является трудной и мы будем заниматься далее именно этой задачей. Отметим, что она допускает примитивный подход, который почти ничего не объясняет. Матрицы A и A' порядка n подобны над данным полем K в том и только том случае, когда однородная система n^2 линейных уравнений $AX = XA'$ имеет решение $C \in M_n(K)$, для которого $\det C \neq 0$.

Вернемся к задаче вычисления матрицы линейного отображения. Отождествляя векторы с их координатными столбцами, будем считать, что $V = K^n$, $e_1, \dots, e_n$ — столбцы единичной матрицы E порядка n, $W = K^s$ и что $\mathcal{A} : x \mapsto Ax$, $x \in K^n$. Тогда полная информация о линейном отображении задается матрицей

$$\begin{pmatrix} e_1 & \dots & e_n \\ Ae_1 & \dots & Ae_n \end{pmatrix} = \begin{pmatrix} E \\ A \end{pmatrix}.$$

Такую матрицу, элементы которой — векторы, назовем $\mathcal{A}$-*слойной*. Утверждается, что элементарные преобразования столбцов такой матрицы сохраняют $\mathcal{A}$-слойность:

$$\begin{pmatrix} u & v \\ Au & Av \end{pmatrix} \xrightarrow{\mathrm{I}} \begin{pmatrix} u+v & v \\ Au+Av & Av \end{pmatrix}, \quad \begin{pmatrix} u & v \\ Au & Av \end{pmatrix} \xrightarrow{\mathrm{II}} \begin{pmatrix} \lambda u & v \\ \lambda(Au) & Av \end{pmatrix}.$$

Упр. 2. Пусть $v_1, \dots, v_n$ — базис K^n, $w_1, \dots, w_n$ — произвольный набор векторов из K^s. Пусть $\mathcal{A}$ — линейное отображение $K^n \to K^s$ со свойством $\mathcal{A}v_j = w_j$ при всех j. Составим из столбцов $v_1, \dots, v_n$ и $w_1, \dots, w_n$ матрицы M и N соответственно и приведем матрицу $\begin{pmatrix} M \\ N \end{pmatrix}$ элементарными преобразованиями системы столбцов к виду $\begin{pmatrix} E \\ A \end{pmatrix}$. Доказать, что матрица A является матрицей отображения $\mathcal{A}$ в паре стандартных базисов пространств K^n и K^s.

Поскольку приходится часто вычислять образ вектора при умножении его на матрицу, то напомним матричную и векторную запись системы $Ax = b$ линейных уравнений:

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{s1} & \dots & a_{sn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_j a_{1j}x_j \\ \vdots \\ \sum_j a_{sj}x_j \end{pmatrix} = \sum_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{sj} \end{pmatrix} x_j = \begin{pmatrix} b_1 \\ \vdots \\ b_s \end{pmatrix}.$$

Она означает, что i-я координата вектора Ax является скалярным произведением i-й строки матрицы A и вектора x. В целом, весь *вектор Ax является линейной комбинацией столбцов матрицы A с коэффициентами, равными координатам вектора x.*

Пример. Пусть $A = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}$. Тогда

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \xrightarrow{A} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \xrightarrow{A} \begin{pmatrix} 6 \\ 7 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \xrightarrow{A-E} \begin{pmatrix} 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \xrightarrow{A-E} \begin{pmatrix} 4 \\ 4 \end{pmatrix}.$$

ЯДРО И ОБРАЗ

Пусть $\mathcal{A} : V \to W$ — линейное отображение векторных пространств над полем K. Множества $\operatorname{Ker} \mathcal{A} = \{v \in V \,|\, \mathcal{A}v = 0\}$ и $\operatorname{Im} \mathcal{A} = \{\mathcal{A}v \,|\, v \in V\} = \mathcal{A}V$ и называются соответственно *ядром* и *образом* отображения $\mathcal{A}$.

Примеры. 1) Пусть $\mathcal{P}$ — ортогональное проецирование пространства геометрических векторов V на плоскость W. Тогда $\operatorname{Ker}\mathcal{P} = U$ — прямая, ортогональная плоскости W, $\operatorname{Im}\mathcal{P} = W$.

2) Пусть $\Sigma : X \mapsto (X + X^\top)/2$ — линейный оператор в пространстве матриц $M_n(\mathbb{R})$. Тогда $\operatorname{Ker}\Sigma$ — подпространство кососимметрических матриц, а $\operatorname{Im}\Sigma$ — подпространство симметрических матриц.

3) Пусть $\mathcal{D} = d/dt$ — оператор дифференцирования в пространстве V всех вещественных многочленов от переменной t. Тогда $\operatorname{Ker}\mathcal{D}$ — подпространство постоянных многочленов (констант), а $\operatorname{Im}\mathcal{D} = V$.

Здесь все равенства легко проверить. Но в общем случае найти ядро и образ бывает совсем не просто — решение следующих задач потребуют бóльших усилий.

Упр. 1. Найти ядро и образ линейного отображения $\mathcal{D}_\lambda : f \mapsto \lambda f - df/dt$ пространства вещественных дифференцируемых функций, заданных на всей вещественной прямой. Здесь $\lambda \in \mathbb{R}$ — заданное число.

Упр. 2. Найти ядро и образ для оператора Лапласа $\Delta : f \mapsto \partial^2 f/\partial x^2 + \partial^2 f/\partial y^2$ на пространстве вещественных многочленов от переменных x, y. Ядро состоит из многочленов, которые называются гармоническими.

Вернемся к задаче вычисления ядра и образа линейного отображения в конечномерном случае. Легко проверить, что ядро и образ линейного отображения $\mathcal{A} : V \to W$ — подпространства из V и W соответственно, потому каждое из них вполне задается выбором базиса. Размерности ядра и образа называются соответственно *дефектом* и *рангом* отображения $\mathcal{A}$ и обозначаются $\operatorname{def}\mathcal{A}$ и $\operatorname{rk}\mathcal{A}$. Дефект называется также корангом ввиду следующей теоремы.

Теорема. Сумма ранга и дефекта линейного отображения конечномерного пространства равна размерности отображаемого пространства.

Доказательство. Используем предыдущие обозначения. Пусть $u_1, \ldots, u_d$ — базис ядра, $w_1 = \mathcal{A}v_1, \ldots, w_r = \mathcal{A}v_r$ — базис образа отображения $\mathcal{A}$. Достаточно доказать, что $u_1, \ldots, u_d, v_1, \ldots, v_r$ — базис пространства V.

Линейная независимость. Пусть $\sum \alpha_i u_i + \sum \beta_j v_j = 0$, где α_i, β_j из K. Действуя на обе части этого равенства отображением $\mathcal{A}$, получим, что $\sum \beta_j w_j = 0$. Но $w_1, \ldots, w_r$ — базис, значит, $\beta_1 = \ldots = \beta_r = 0$ и, отсюда, $\sum \alpha_i u_i = 0$. Так как $u_1, \ldots, u_d$ — базис, то $\alpha_1 = \ldots = \alpha_d = 0$.

Максимальность. Пусть v — произвольный вектор из V. Тогда $\mathcal{A}v$ из $\operatorname{Im}\mathcal{A}$, $\mathcal{A}v = \sum \beta_j w_j$ при некоторых β_j из K и $u = v - \sum \beta_j v_j \in \operatorname{Ker}\mathcal{A}$. Поэтому $u = \sum \alpha_i u_i$ и, следовательно, $v = \sum \alpha_i u_i + \sum \beta_j v_j$. Теорема доказана.

Упр. 3. Пусть U и W — подпространства конечномерного пространства V с единственным ограничением $\dim U + \dim W = \dim V$. Доказать, что тогда найдется линейный оператор $\mathcal{A}$ в пространстве V, для которого $\operatorname{Ker}\mathcal{A} = U$, $\operatorname{Im}\mathcal{A} = W$.

Упр. 4. Доказать, что $V = \operatorname{Im}\mathcal{A} \oplus \operatorname{Ker}\mathcal{A}$ в том и только том случае, когда $\mathcal{A}^2 V = \mathcal{A} V$.

Чтобы найти ядро и образ для линейного отображения $\mathcal{A} : x \mapsto Ax$ координатных пространств, заданного матрицей A, достаточно решить следующие стандартные задачи о системах линейных уравнений.

1) Найти фундаментальную систему решений однородной системы линейных уравнений $Ax = 0$. Получится базис ядра.

2) Найти базис системы столбцов (или линейной оболочки системы столбцов) матрицы A. Получится базис образа.

Укажем способ *одновременного вычисления базисов ядра и образа* линейного отображения $\mathcal{A} : V \to W$, заданного матрицей в паре базисов. Приведем элементарными преобразованиями слоев матрицу $\begin{pmatrix} E \\ A \end{pmatrix}$ к виду $\begin{pmatrix} B \\ C \end{pmatrix}$, где B — матрица порядка n, C — матрица размера $s \times n$, ступенчатая по столбцам.

Утверждается, что система ненулевых столбцов $w_1, \ldots, w_r$ матрицы C образует базис $\operatorname{Im} \mathcal{A}$, а система столбцов $u_1, \ldots, u_d$ матрицы B, имеющих нулевое продолжение в матрице C, образует базис $\operatorname{Ker} \mathcal{A}$.

$$\begin{pmatrix} E \\ A \end{pmatrix} \xrightarrow{\text{эл.пр.столбцов}} \begin{pmatrix} B \\ C \end{pmatrix} = \begin{pmatrix} u_1 & .. & u_d & v_1 & .. & v_r \\ 0 & .. & 0 & w_1 & .. & w_r \end{pmatrix}.$$

Действительно, матрица $\begin{pmatrix} B \\ C \end{pmatrix}$ является $\mathcal{A}$-слойной, поэтому $Au_1 = 0, \ldots, Au_d = 0,\ Av_1 = w_1, \ldots, Av_r = w_r$. Столбцы матрицы B образуют базис пространства K^n, так как система столбцов B получилась элементарными преобразованиями системы столбцов невырожденной матрицы E. Поэтому

$$\mathcal{A}V = \mathcal{A}\langle u_1, \ldots, u_d,\ v_1, \ldots, v_r\rangle = \langle Au_1, \ldots, Au_d,\ Av_1, \ldots, Av_r\rangle = \langle w_1, \ldots, w_r\rangle.$$

Но система векторов $w_1, \ldots, w_r$ линейно независима ввиду ступенчатости матрицы C. Следовательно, она образует базис образа. С другой стороны, система $u_1, \ldots, u_d$ линейно независима как часть базиса пространства столбцов K^n, содержится в ядре и является его базисом, если учесть, что размерность ядра равна $n - r = d$ по теореме о сумме ранга и дефекта.

Пример. Ведущие элементы при элементарных преобразованиях системы строк в матрицах будем выделять рамкой. Пусть $\mathcal{A} : x \mapsto Ax,\ x \in \mathbb{R}^4$ и

$$A = \begin{pmatrix} 2 & 3 & 0 & 5 \\ 1 & 0 & -3 & 1 \\ -1 & 1 & 5 & 0 \\ 1 & 2 & 1 & 3 \end{pmatrix}.$$

Тогда

$$\begin{pmatrix} E \\ A \end{pmatrix} = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 2 & 3 & 0 & 5 \\ 1 & 0 & -3 & 1 \\ -1 & 1 & 5 & 0 \\ \boxed{1} & 2 & 1 & 3 \end{array}\right) \sim \left(\begin{array}{cccc} 1 & -2 & -1 & -3 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 2 & -1 & -2 & -1 \\ 1 & -2 & -4 & -2 \\ -1 & \boxed{3} & 6 & 3 \\ 1 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & -2 & 3 & -1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right).$$

В последней матрице столбцы, расположенные внизу и слева, образуют базис образа $\operatorname{Im} \mathcal{A}$, а столбцы вверху и справа — базис ядра $\operatorname{Ker} \mathcal{A}$.

Упр. 5. Найти ядро линейного оператора $\mathcal{D} = \partial f/\partial x - \partial f/\partial y$ в пространстве вещественных многочленов степени $\leqslant 2$ от переменных $x,\ y$. (*Указание.* Применить указанный метод, не используя координаты.)

СОБСТВЕННЫЕ ВЕКТОРЫ И СОБСТВЕННЫЕ ЗНАЧЕНИЯ

Задача отыскания собственных векторов и собственных значений линейного оператора считается одной из основных задач линейной алгебры. Причина в том, что довольно часто пространство может иметь базис, состоящий из собственных векторов оператора. Такой оператор называется *диагонализируемым* или *полупростым*. Для него есть хорошее геометрическое описание, можно найти все инвариантные подпространства, прогнозировать поведение итераций оператора, вычислять значения скалярных функций и т. д.

Напомним, что для линейного оператора $\mathcal{A}$ в пространстве V над полем K вектор $v \in V$ называется *собственным*, отвечающим *собственному значению* $\lambda \in K$, если

$$\mathcal{A}v = \lambda v, \quad v \neq 0.$$

Примеры. 1) Для поворота трехмерного евклидова пространства вокруг некоторой оси на некоторый нетривиальный угол собственные векторы с собственным значением 1 — это направляющие векторы неподвижной оси, других собственных значений и собственных векторов нет.

Для проецирования пространства на плоскость W параллельно некомпланарной прямой U ненулевые векторы плоскости W — собственные с собственным значением 1, а ненулевые векторы прямой U — собственные с собственным значением 0, других собственных значений и собственных векторов нет.

2) Для линейного оператора транспонирования матрицы $X \mapsto X^\top$ в алгебре $M_n(\mathbb{R})$ собственные значения только 1 и -1, а собственные векторы — соответственно ненулевые симметрические и кососимметрические матрицы, других собственных значений и собственных векторов нет.

Для линейного оператора коммутирования $\mathcal{C} : X \mapsto AX - XA$ в пространстве матриц $M_n(\mathbb{R})$ ненулевые матрицы, коммутирующие с матрицей A являются собственными с собственным значением 0. (*Замечание.* Если известны собственные значения для оператора $x \mapsto Ax, \quad x \in \mathbb{R}^n$, то можно найти собственные вектора и собственные значения для оператора коммутирования $\mathcal{C}$. Это хорошая задача — используйте матрицы ранга 1!)

3) Для оператора дифференцирования в пространстве вещественных многочленов собственное значение только 0, а собственные векторы — ненулевые постоянные многочлены. Для того же оператора в пространстве гладких вещественных функций на прямой для любого $\lambda \in \mathbb{R}$ функция $f(t) = c\,e^{\lambda t}$ при $c \neq 0$ является собственной с собственным значением λ.

4) Для линейного оператора сдвига вправо счетных последовательностей комплексных чисел

$$\mathcal{R} : (x_1, x_2, x_3, \ldots) \mapsto (0, x_1, x_2, x_3, \ldots)$$

в пространстве $\mathbb{C}^\infty$ вообще *нет собственных значений* и собственных векторов.

Как найти собственные векторы и собственные значения линейного оператора в конечномерном случае?

Отметим, что нередко хотя бы часть собственных векторов и собственных значений можно угадать, исходя из контекста задачи, если она имеет естественный характер — см. приведенные примеры. Такую попытку следует сделать обязательно. Но в общем случае приходится переходить к вычислению собственных значений и собственных векторов через матрицу оператора.

Если A – матрица оператора $\mathcal{A}$ в произвольном базисе пространства V конечной размерности n, то все собственные значения — в точности решения характеристического уравнения $\det(A - \lambda E) = 0$, принадлежащие полю скаляров. Если λ считать неизвестным, то левая часть уравнения — многочлен степени n. Он называется *характеристическим многочленом* линейного оператора $\mathcal{A}$ и не зависит от выбора базиса пространства.

Таким образом, собственные значения — корни характеристического многочлена, принадлежащие полю скаляров. Конечно, в данном поле многочлен может не иметь корней, но в подходящем расширении поля скаляров содержатся все его корни, а поле комплексных чисел даже не нужно расширять: всякий многочлен степени $n \geqslant 1$ с комплексными коэффициентами обязательно имеет комплексный корень. Следовательно, в ненулевом *конечномерном* комплексном пространстве любой линейный оператор имеет собственные векторы. Множество корней характеристического многочлена линейного оператора $\mathcal{A}$ называется *спектром оператора* и обозначается $\operatorname{Sp}\mathcal{A}$. По традиции буква λ будет обозначать и переменную, и конкретное значение собственное значение в зависимости от контекста. Это удобно потому, что в прикладных задачах как правило буквы t и x заняты — t для времени, а $x = x(t)$ — для набора координат, задающих положение точки в фазовом пространстве. В общем случае будем писать

$$\det(A - \lambda E) = \pm \prod_{j=1}^{s} (\lambda - \lambda_j)^{k_j}, \quad \operatorname{Sp}\mathcal{A} = \{\lambda_1, \ldots, \lambda_s\}.$$

Здесь предполагается, что $\lambda_i \neq \lambda_j$ при $i \neq j$.

Наметим теперь, как искать, хотя бы теоретически, собственные значения и собственные векторы.

Сначала два простых случая. Если матрица *треугольна* с элементами главной диагонали $\lambda_1, \ldots, \lambda_n$, то легко проверить, что это и есть корни характеристического многочлена и собственные значения оператора. Если матрица оказалась *полураспавшейся*, то характеристический многочлен распадается на множители меньшей степени — характеристические многочлены клеток главной диагонали, поэтому достаточно найти только их корни.

В общем случае первая вычислительная трудность — подсчитать коэффициенты характеристического многочлена.

Для быстрого вычисления всех его коэффициентов развиты специальные методы, но для нас достаточна следующая формула

$$\det(A - \lambda E) = (-\lambda)^n + c_1(-\lambda)^{n-1} + \ldots + c_k(-\lambda)^{n-k} + \ldots + c_n,$$

где c_k — сумма всех главных миноров порядка k из матрицы A, в частности, $c_1 = \operatorname{tr} \mathrm{A}$, $c_n = \det A$. Напомним, что минор главный, если номера задающих его строк и столбцов, совпадают.

Упр. 1. Докажите эту формулу, используя свойство полилинейности определителя.

Отметим, что все коэффициенты характеристического многочлена — *инварианты оператора*, т. е. зависят только от оператора и не зависят от выбора базиса пространства.

Задача отыскания собственных векторов и собственных значений (еще говорят — *спектральная задача*) расщепляется на две части: 1) отыскание всех корней характеристического многочлена, принадлежащих полю скаляров, 2) при известных собственных значениях надо найти собственные векторы, отвечающие этим значениям.

Мы уже знаем, как трудна первая задача в общем случае. Что касается второй, то достаточно для каждого собственного значения λ_j, $j = 1, \ldots, s$, решить систему линейных уравнений $(A - \lambda_j E)x = 0$, т. е. найти фундаментальную систему решений нескольких систем однородных линейных уравнений.

Предположим, что известны все собственные значения. Покажем, как ускорить решение второй задачи при точной арифметике, используя одновременный поиск ядра и образа и инвариантность подпространств.

Отметим сначала, что в пространстве V множество всех собственных векторов, отвечающих собственному значению λ линейного оператора $\mathcal{A}$, вместе с нулем образует подпространство V_λ, совпадающее с ядром оператора $\mathcal{A} - \lambda\mathcal{E}$. Оно называется *собственным подпространством линейного оператора $\mathcal{A}$, отвечающим собственному значению* λ. Оказывается, что собственные векторы с *другими* собственными значениями содержатся в образе оператора $\mathcal{A} - \lambda\mathcal{E}$.

Действительно, равенство $\mathcal{A}u = \lambda u$ равносильно равенству $(\mathcal{A} - \lambda\mathcal{E})u = 0$. Если же $\mathcal{A}v = \lambda' v$, $\lambda' \neq \lambda$, то

$$\mathcal{A}v - \lambda v = (\lambda' - \lambda)v, \quad v = (\mathcal{A} - \lambda\mathcal{E})(\tfrac{1}{\lambda' - \lambda}v).$$

Поскольку подпространство $V' = (\mathcal{A} - \lambda\mathcal{E})V$ отображается операторами $\mathcal{A}$ и $\mathcal{A} - \lambda'\mathcal{E}$ в себя, то можно искать собственное подпространство $V_{\lambda'}$ пространства V, отвечающее другому собственному значению $\lambda' \neq \lambda$, как ядро оператора $\mathcal{A} - \lambda'\mathcal{E}$ в пространстве V'. При этом собственное подпространство $V_{\lambda''}$ пространства V, отвечающее третьему собственному значению $\lambda'' \neq \lambda,\ \lambda'$, содержится в подпространстве $V'' = (\mathcal{A} - \lambda'\mathcal{E})V'$ и его можно снова выделить как ядро оператора $\mathcal{A} - \lambda''\mathcal{E}$ в этом подпространстве V'', и так далее.

Отметим, что для начала нам достаточно знать только одно собственное значение λ, при этом поиск собственных значений и собственных векторов оператора $\mathcal{A}$ в пространстве V редуцируется к поиску собственных значений и собственных векторов оператора $\mathcal{A}$ в пространстве $(\mathcal{A} - \lambda\mathcal{E})V$ меньшей размерности.

Пример 1. Пусть $V = \mathbb{R}^2$ и $\mathcal{A} : x \mapsto Ax$, $x \in \mathbb{R}^2$, — линейный оператор умножения столбца на матрицу

$$A = \begin{pmatrix} 4 & 1 \\ 3 & 2 \end{pmatrix}.$$

Тогда характеристический многочлен $|A - \lambda E| = (-\lambda)^2 + \operatorname{tr} A\,(-\lambda) + \det A = \lambda^2 - 6\lambda + 5 = (\lambda - 5)(\lambda - 1)$. Теперь найдем собственные векторы:

$$\begin{pmatrix} E \\ A - 5E \end{pmatrix} = \left(\begin{array}{rr} 1 & 0 \\ 0 & 1 \\ \hline -1 & 1 \\ \boxed{3} & -3 \end{array}\right) \sim \left(\begin{array}{rr} 1 & 1 \\ 0 & 1 \\ \hline -1 & 0 \\ 3 & 0 \end{array}\right).$$

Следовательно, вектор $f_1 = (1, 1)^\top$ — базис пространства $\operatorname{Ker}(\mathcal{A} - 5\mathcal{E})$ собственных векторов оператора $\mathcal{A}$ с собственным значением 5, вектор $f_2 = (-1, 3)^\top$ — базис пространства $\operatorname{Ker}(\mathcal{A} - 5\mathcal{E})$ собственных векторов $\mathcal{A}$ с собственным значением 1 или $\mathcal{A}$-неподвижных векторов.

Теперь легко увидеть, что оператор $\mathcal{A}$ растягивает плоскость в 5 раз вдоль каждой прямой, параллельной прямой $\mathbb{R}f_1$, от неподвижной точки пересечения этой прямой с прямой $\mathbb{R}f_2$. Действительно, если $x = \alpha_1 f_1 + \alpha_2 f_2$, то $Ax = 5\alpha_1 f_1 + \alpha_2 f_2$.

Пример 2. Заменим матрицу в предыдущем примере на следующую матрицу

$$A = \frac{1}{3}\begin{pmatrix} 4 & 5 \\ 3 & 6 \end{pmatrix}.$$

Тогда

$$|A - \lambda E| = (-\lambda)^2 + \operatorname{tr} A\,(-\lambda) + \det A = \lambda^2 - (10/3)\lambda + 1 = (\lambda - 3)(\lambda - (1/3)).$$

Найдем собственные векторы:

$$\left(\begin{array}{c} E \\ A - 3E \end{array}\right) \sim \left(\begin{array}{rr} 3 & 0 \\ 0 & 3 \\ \hline -5 & 5 \\ \boxed{3} & -3 \end{array}\right) \sim \left(\begin{array}{rr} 3 & 3 \\ 0 & 3 \\ \hline -5 & 0 \\ 3 & 0 \end{array}\right).$$

Теперь собственные прямые оператора, отвечающие собственным значениям 3 и $1/3$, имеют соответственно базисы $f_1 = (3,3)^\top$ и $f_2 = (-5,3)^\top$. В новой системе координат плоскости с базисом f_1, f_2 легко дать геометрическое описание оператора. Очевидно, плоскость должна растягиваться в 3 раза вдоль первой координатной прямой $\mathbb{R}f_1$ и сжиматься в 3 раза вдоль второй координатной прямой $\mathbb{R}f_2$. При этом если $x = \alpha_1 f_1 + \alpha_2 f_2$, то $Ax = 3\alpha_1 f_1 + (1/3)\alpha_2 f_2$. Следовательно, произведение координат остается постоянным. Поскольку условие $\alpha_1\alpha_2 = \mathrm{const}$ задает либо гиперболу с асимптотами, равными координатным прямым, либо пару пересекающихся координатных прямых, то каждая точка плоскости вне координатных прямых под действием оператора «скользит» вдоль такой гиперболы все ближе к «притягивающей» первой координатной прямой. Как действует оператор на точки координатных прямых — совершенно очевидно. Такой оператор называется *гиперболическим поворотом*.

Пример 3. Покажем, как собственные векторы и собственные значения линейных операторов используются для прогнозирования. Предположим, что некая фирма осуществляет перевозки между городами X, Y и Z, располагая M грузовыми машинами. Пусть в начальный момент времени в этих городах находились соответственно x_0, y_0 и z_0 машин. Контроль через неделю показал, что треть машин из города X осталась там же, а две трети оказалась в городе Y; треть машин из Y оказалась в городе X, а две трети — в городе Z; треть машин из Z осталась в Z, а две трети перебралось в Y. Предполагая, что такой закон сохраняется достаточно долго, найдем предельное распределение машин по городам.

Если обозначить x_1, y_1, z_1 соответствующее количество машин в городах через неделю, то

$$\begin{cases} x_1 = x_0/3 + y_0/3 \\ y_1 = 2x_0/3 + + 2z_0/3 \\ z_1 = 2y_0/3 + z_0/3. \end{cases}$$

Это равенство можно записать в матричной форме $u_1 = Au_0$, где $u_k = (x_k,\ y_k,\ z_k)^\top$ — вектор-столбец распределения машин по городам через k недель и

$$A = \frac{1}{3}\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

Из нашего предположения следует, что $u_2 = Au_1 = A^2u_0$ и вообще $u_k = A^ku_0$. По сути дела нам нужно найти $\lim_{k\to\infty} A^ku_0$. Так как возводить матрицу в степень — дело трудоемкое, то естественно разобраться, как действует на пространстве столбцов $\mathbb{R}^3$ линейный оператор $\mathcal{A} : u \mapsto Au$ умножения столбца u на матрицу A. С большой долей вероятности пространство имеет базис, состоящий из собственных векторов оператора. Если v_1, v_2, v_3 — такой базис и $Av_i = \lambda_i v_i,\ i = 1, 2, 3$, то разложим вектор u_0 по базису $u_0 = \sum \alpha_i v_i$. Тогда $Au_0 = \sum \alpha_i Av_i = \sum \alpha_i \lambda_i v_i$ и вообще $A^ku_0 = \sum \alpha_i \lambda_i^k v_i$ при всех k. Теперь ясно, что предел, если он существует, определяется пределами $\lim_{k\to\infty} \lambda_i^k$, а также собственными векторами матрицы.

В нашем случае $|A - \lambda E| = -\lambda^3 + 2\lambda^2/3 + 5\lambda/9 - 2/9 = -(\lambda - 1)(\lambda + (2/3))(\lambda - (1/3))$. Отсюда заключаем, что существует базис $\mathbb{R}^3$, состоящий из собственных векторов. Найдем его. Имеем

$$\begin{pmatrix} E \\ A - E \end{pmatrix} \sim \left(\begin{array}{rrr} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \\ \hline -2 & 1 & 0 \\ 2 & -3 & 2 \\ 0 & 2 & \boxed{-2} \end{array}\right) \sim \left(\begin{array}{rrr} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 3 & 3 \\ \hline -2 & 1 & 0 \\ 2 & \boxed{-1} & 2 \\ 0 & 0 & -2 \end{array}\right) \sim \left(\begin{array}{rrr} 3 & 0 & 0 \\ 6 & 3 & 0 \\ 6 & 3 & 3 \\ \hline 0 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -2 \end{array}\right).$$

Отсюда $Av_1 = v_1$, если $v_1 = (1, 2, 2)^\top$. Собственные векторы с собственными значениями 1/3 и -2/3 содержатся в линейной оболочке векторов $f_1 = (1, -1, 0)^\top,\ f_2 = (0, 1, -1)^\top$. Найдем их таким же способом.

$$\begin{pmatrix} f_1 & f_2 \\ (3A - E)f_1 & (3A - E)f_2 \end{pmatrix} = \left(\begin{array}{rr} 1 & 0 \\ -1 & 1 \\ 0 & -1 \\ \hline -1 & 1 \\ \boxed{3} & -3 \\ -2 & 2 \end{array}\right) \sim \left(\begin{array}{rr} 1 & 1 \\ -1 & 0 \\ 0 & -1 \\ \hline -1 & 0 \\ 3 & 0 \\ -2 & 0 \end{array}\right).$$

Отсюда $Av_2 = (1/3)v_2,\ Av_3 = (-2/3)v_3$, если $v_2 = (1, 0, -1)^\top,\ v_3 = (1, -3, 2)^\top$. Следовательно, $\lim_{k\to\infty} A^ku_0 = \alpha_1 v_1 = (\alpha_1, 2\alpha_1, 2\alpha_1)$, если $u_0 = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$. Конечно, конкретный начальный вектор u_0 всегда можно разложить по собственным векторам v_1, v_2, v_3 и узнать коэффициент α_1, но мы поступим иначе. Заметим, что общее количество M всех машин в процессе не менялось, поэтому и в пределе $\alpha_1 + 2\alpha_1 + 2\alpha_1 = M$. Тогда $\alpha_1 = M/5$ и, следовательно, предельное распределение грузовых машин по городам X, Y, Z имеет вид $\lim_{k\to\infty} A^ku_0 = (M/5, 2M/5, 2M/5)^\top$. Теперь ясно, что векторы v_2, v_3 можно было и не искать, — итоговый ответ определяется неподвижным вектором v_1, общим числом M машин и даже не зависит от исходного распределения машин по городам.

Замечание. Конечномерное комплексное пространство относительно случайно выбранного линейного оператора как правило содержит базис, состоящий из собственных векторов оператора. Действительно, ввиду алгебраической замкнутости поля $\mathbb{C}$ все корни характеристического многочлена являются собственными значениями, условие существования кратного корня для характеристического многочлена равносильно равенству нулю дискриминанта многочлена, что дает нетривиальное полиномиальное условие на коэффициенты матрицы оператора. Но для случайно выбранной матрицы это условие как правило не выполняется. Если оператор $\mathcal{A}$ имеет n различных собственных значений $\lambda_1, \ldots, \lambda_n$, то соответствующие n собственных векторов $v_1, \ldots,\ v_n$ линейно независимы и образуют базис пространства размерности n. Действие оператора $\mathcal{A}$ в базисе $v_1, \ldots,\ v_n$ становится предсказуемо простым

$$\mathcal{A} : x_1 v_1 + \ldots + x_n v_n \mapsto \lambda_1 x_1 v_1 + \ldots + \lambda_n x_n v_n.$$

Поскольку его матрица в данном базисе диагональна, то такой оператор называется *диагонализируемым.*

Упр. 2. Если $f(\lambda)$ — комплексный многочлен, то в случае диагонализируемого оператора $\mathcal{A}$ в предыдущих обозначениях

$$f(\mathcal{A}) : x_1 v_1 + \ldots + x_n v_n \mapsto f(\lambda_1) x_1 v_1 + \ldots + f(\lambda_n) x_n v_n.$$

Это утверждение верно и для рядов $f(\lambda)$, сходящихся во всех точках спектра диагонализируемого оператора.

Если поле скаляров не является алгебраически замкнутым, то доля конечномерных диагонализируемых операторов среди всех линейных операторов может быть мала. Главная причина — доля многочленов, не имеющих корней в данном поле K, среди всех многочленов данной степени над полем K может быть велика.

Пример 4. Найдем числа $x,\ y \in \mathbb{R}$, при которых матрица

$$A = \begin{pmatrix} 2x & 1+y^2 \\ -1 & 0 \end{pmatrix}$$

подобна вещественной диагональной матрице. Оценим также вероятность этого события, то есть долю площади множества таких точек (x, y) в круге $x^2+y^2 \leqslant r^2$ при $r \to \infty$.

Если матрица A подобна вещественной диагональной матрице $D = Diag\,(\lambda_1,\ \lambda_2)$, то тогда $\operatorname{Sp} A = \{\lambda_1,\ \lambda_2\} \subset \mathbb{R}$ и характеристический многочлен $|A - \lambda E| = \lambda^2 - 2x\,\lambda + 1 + y^2$ имеет вещественные корни. Поэтому дискриминант неотрицателен, $x^2 - (1+y^2) \geqslant 0$. Если он равен нулю, то корни совпадают, матрица D скалярна, может быть подобна только себе и, значит, матрица A не диагонализируема. Следовательно, дискриминант положителен, $y^2 < x^2 - 1,\ |y| < \sqrt{x^2-1}$. Легко видеть, что множество таких точек образует «внутренность» гиперболы $x^2 - y^2 = 1$ с асимптотами $x^2 - y^2 = 0,\ y = \pm x$. Если радиус круга достаточно большой, то доля площади таких точек в круге приближается к доле площади в данном круге точек $|y| < |x|$ между асимптотами поскольку гипербола приближается к асимптотам. Следовательно, искомая вероятность равна $1/2$.

Пример 5. Покажем, как идея собственных векторов может помочь с оценкой скорости роста рекуррентных последовательностей, $x_0,\ x_1,\ x_2, \ldots,$ для которых каждый член последовательности, начиная с номера n, выражается через n предыдущих членов последовательности по одному и тому же заданному правилу.

Пусть даны комплексные числа $a_1, \ldots, a_n$. Зададим последовательность по следующему «линейному» правилу: $x_0,\ x_1,\ x_2, \ldots, x_{n-1}$ — произвольные комплексные числа,

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_n x_0,$$

и вообще $x_{n+k} = a_1 x_{n-1+k} + a_2 x_{n-2+k} + \cdots + a_n x_k$ при $k \geqslant 0$.

Дополним условие $x_n = a_1x_{n-1} + a_2x_{n-2} + \dots + a_nx_0$ равенствами $x_i = x_i$ при $i < n$ и запишем полученную систему уравнений в матричном виде $u_1 = Au_0$, где вектор $u_k = (x_{n-1+k},\ x_{n-2+k}, \dots, x_k)^\top$, а матрица A имеет вид

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_{n-1} & a_n \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ .. & .. & \dots & .. & .. \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

По условию $u_k = Au_{k-1} = A^2u_{k-2} = \ldots = A^ku_0$. Получается, что скорость роста последовательности задается в случае диагонализируемости оператора $\mathcal{A} : x \mapsto Ax$ его спектром.

Утверждается, что

$$|A - \lambda E| = (-1)^n(\lambda^n - a_1\lambda^{n-1} - a_2\lambda^{n-2} - \ldots - a_n).$$

Это равенство легко установить индукцией по n с помощью разложения определителя $D_n(\lambda) = |A - \lambda E|$ по последнему столбцу. Мы видим, в частности, что *любой* многочлен степени n со старшим коэффициентом $(-1)^n$ является характеристическим для подходящей матрицы порядка n.

Предположим, что $\lambda_1, \dots, \lambda_n$ — корни характеристического многочлена $|A - \lambda E|$. Легко проверяется, что вектор $v_j = (\lambda_j^{n-1},\ \lambda_j^{n-2}, \dots, \lambda_j,\ 1)^\top$ является собственным для $\mathcal{A}$ с собственным значением λ_j. Если все корни различны, то векторы $v_1, \dots,\ v_n$ образуют базис пространства $\mathbb{C}^n$ и оператор $\mathcal{A}$ диагонализируем.

Упр. 3. Вычислить скорость роста последовательности Фибоначчи

$$F_0 = 0,\ F_1 = 1,\ F_k = F_{k-1} + F_{k-2} \text{ при } k \geqslant 2,$$

то есть $\lim_{k\to\infty} \dfrac{F_k}{F_{k-1}}$.

Пример 6. Покажем, как с помощью собственных векторов можно расщепить и решить многие системы линейных дифференциальных уравнений с постоянными коэффициентами. Предположим, что положение изменяющейся во времени системы (материальной, экономической и т. д.) в момент времени t задается координатами $x(t) = (x_1(t), \dots, x_n(t))^\top \in \mathbb{R}^n$ и что ее вектор скорости зависит от положения линейно:

$$\frac{dx_i}{dt} = \sum_{j=1}^n a_{ij}x_j(t), \quad i = 1, \dots, n.$$

Коэффициенты матрицы $A = (a_{ij})$ системы уравнений предполагаются постоянными. Эта система уравнений имеет краткую матричную запись $\dfrac{dx}{dt} = Ax$. Сложность решения системы уравнений определяется сложностью матрицы A. Разберемся, как матрица изменится при замене системы координат. Допустим, что старая декартова система координат x заменена на новую декартову систему координат y с постоянной обратимой матрицей замены C. Тогда $x = Cy$ и

$$\frac{dx}{dt} = C\frac{dy}{dt} = ACy, \quad \frac{dy}{dt} = C^{-1}ACy,$$

матрица A исходной системы уравнений заменилась на подобную матрицу $C^{-1}AC$ новой системы уравнений.

Если матрица A диагонализируема над $\mathbb{R}$, то можно считать $C^{-1}AC$ диагональной матрицей $Diag(\lambda_1, \ldots, \lambda_n)$. После замены получится, что новая система уравнений расщепилась на уравнения для отдельных координат $\dfrac{dy_i}{dt} = \lambda_i y_i, \quad i = 1, \ldots, n$. Ее решение очевидно: $y_i(t) = c_i e^{\lambda_i t}$. Неопределенные константы $c_i = y_i(0)$ задаются однозначно, если известно начальное положение системы в нулевой момент времени:

$$x(0) = x_0 = Cy(0), \quad y(0) = C^{-1}x_0.$$

Упр. 4. Решить систему уравнений

$$\begin{cases} dx_1/dt &= 3x_1 - \frac{1}{2}x_2 \\ dx_2/dt &= 5x_1 - \frac{1}{2}x_2 \end{cases}$$

при начальном условии $x_1(0) = 3,\ x_2(0) = 12$. Нарисовать траекторию решения $x(t)$ при $t \to +\infty$.

Упр. 5. Предположим, что линейный оператор $\mathcal{A}$ в $\mathbb{R}^2$ диагонализируем. Разбить $\mathbb{R}^2$ в объединение семейства таких кривых, что дискретная траектория $x,\ \mathcal{A}x,\ \mathcal{A}^2x,\ \mathcal{A}^3x,\ \ldots$ с началом в любой точке $x \in \mathbb{R}^2$ принадлежит ровно одной кривой семейства. Указать уравнения этих кривых в зависимости от спектра $\mathcal{A}$.

ИНВАРИАНТНЫЕ ПОДПРОСТРАНСТВА

Линейный оператор векторного пространства на каждом подпространстве сужается до линейного отображения. Если это подпространство отображается *в себя*, то оно называется *инвариантным относительно оператора*. Сужение оператора на инвариантном подпространстве — снова линейный оператор. В формулах: подпространство U пространства V инвариантно относительно линейного оператора $\mathcal{A} : V \to V$, если

$$u \in U \Longrightarrow \mathcal{A}u \in U.$$

Последнее отображение называется *сужением* (или *ограничением*) $\mathcal{A}$ на U и обозначается $\mathcal{A}|_U$.

Всегда инвариантны нулевое подпространство и все пространство. Остальные инвариантные подпространства назовем *нетривиальными*. Если v — собственный вектор для линейного оператора $\mathcal{A}$, то его линейная оболочка — одномерное инвариантное подпространство, на котором оператор действует как умножение на скаляр. Верно и обратное утверждение — одномерное инвариантное подпространство порождается собственным вектором оператора. Таким образом, понятие инвариантного подпространства в некотором роде обобщает понятие собственного вектора. Существование нетривиального инвариантного подпространства облегчает описание действия оператора в пространстве. Такое пространство называются *приводимым* относительно данного оператора. Пространство называется *неприводимым* относительно линейного оператора, если оно содержит ровно два инвариантных подпространства — нулевое и все пространство. Таковы нетривиальные минимальные инвариантные подпространства.

Примеры. 1) Плоскость неприводима относительно поворота на угол φ, некратный π. При $\varphi = k\pi,\ k \in \mathbb{Z}$, инвариантны любые прямые (= одномерные подпространства).

Относительно ортогонального проецирования пространства геометрических векторов на плоскость U нетривиальные инвариантные подпространства — это а) все прямые плоскости U, б) прямая L, ортогональная плоскости U, в) сама плоскость U и г) все плоскости, содержащие прямую L. Других нетривиальных инвариантных подпространств нет.

2) Ядро и образ линейного оператора всегда инвариантны относительно оператора.

Пространство матриц $M_n(K)$ распадается в прямую подпространств V_j, состоящих из матриц, в которых равны нулю все столбцы с номерами $\neq j$. Эти попространства инвариантны относительно оператора левого сдвига $L_A : X \mapsto AX$ на матрицу A.

3) Подпространство многочленов степени $\leqslant k$ инвариантно относительно дифференцирования.

Теорема 1. 1) Пространство приводимо относительно линейного оператора в том и только том случае, когда в подходящем базисе пространства матрица оператора является полураспавшейся.

2) Характеристический многочлен сужения оператора на нетривиальном инвариантном подпространстве делит характеристический многочлен оператора.

3) Если пространство распадается в прямую сумму ненулевых подпространств, инвариантных относительно оператора, то его характеристический многочлен распадается в произведение характеристических многочленов сужений оператора на слагаемых.

Доказательство. 1) Пусть $\mathcal{A} : V \to V,\quad \mathcal{A}U \subset U$. Пусть $e_1, \ldots, e_k$ — базис U. Дополним его до базиса пространства. Пусть $e_1, \ldots, e_k,\ e_{k+1}, \ldots, e_n$ — базис V. В этом базисе матрица A оператора $\mathcal{A}$ является клеточно треугольной с клетками B и C порядков k и $n-k$ по главной диагонали, поскольку ввиду инвариантности

$$\mathcal{A}e_j = \sum_{i \leqslant k} b_{ij} e_i \text{ при } j \leqslant k.$$

Отметим, что при этом матрица B является матрицей сужения $\mathcal{B} = \mathcal{A}|_U$ в базисе $e_1, \ldots, e_k$.

Наоборот, предположим, что матрица A оператора $\mathcal{A}$ в базисе $e_1, \ldots, e_k,\ e_{k+1}, \ldots, e_n$ является клеточно треугольной с клетками B и C порядков k и $n-k$ по главной диагонали. Тогда линейная оболочка U векторов $e_1, \ldots, e_k$ инвариантна относительно $\mathcal{A}$ и имеет размерность k, где $0 < k < n$.

2) В предыдущих обозначениях по свойствам определителей полураспавшихся матриц $|A - \lambda E| = |B - \lambda E| \cdot |C - \lambda E|$. В частности, спектр оператора $\mathcal{B}$ содержится в спектре $\mathcal{A}$.

3) В предыдущих обозначениях достаточно выбрать в качестве $e_{k+1}, \ldots, e_n$ базис W и обозначить $\mathcal{C} = \mathcal{A}|_W$. В этом случае спектр оператора $\mathcal{A}$ является объединением спектров $\mathcal{B}$ и $\mathcal{C}$.

Теорема доказана.

Следствие 1. Если характеристический многочлен оператора неразложим над полем скаляров, то пространство неприводимо относительно оператора.

Следствие 2. Если спектр оператора содержится в поле скаляров, то всякое ненулевое инвариантное подпространство содержит собственный вектор оператора.

Следствие 3. Если спектр оператора содержится в поле скаляров, то пространство содержит неуплотняемую цепочку инвариантных подпространств:

$$0 = V_0 < V_1 < V_2 < \ldots < V_{n-1} < V_n = V, \ \ \dim V_k/V_{k-1} = 1, \ \ \mathcal{A}V_k \subset V_k, \ k = 1, \ldots, n.$$

Более того, такую цепочку можно провести через любое заданное подпространство, инвариантное относительно $\mathcal{A}$.

Доказательство первой части следствия 3 вытекает из существования инвариантного подпространства коразмерности 1. Можно также использовать существование собственных векторов и конструкцию фактор-оператора на фактор-пространстве пространства V по инвариантному подпространству U:

$$\mathcal{A}U \subset U, \ \ \tilde{\mathcal{A}} : \tilde{v} \mapsto \widetilde{\mathcal{A}v}, \ \tilde{x} = x + U.$$

Второе утверждение доказывается с использованием обеих подходов.

Следствие 4. Всякая квадратная матрица подобна треугольной над полем, содержащим спектр матрицы.

Доказательство. Свяжем с матрицей $A \in M_n(K)$ оператор $\mathcal{A} : x \mapsto Ax$ в пространстве столбцов $V = K^n$. Его матрица в стандартном базисе совпадает с A. По следствию 3 существует неуплотняемая цепочка инвариантных подпространств. Выберем базис $f_1, \ldots, f_n$, согласованный с этой цепочкой так, что $f_1, \ldots, f_k$ — базис V_k при всех k. Ввиду инвариантности подпространств матрица оператора $\mathcal{A}$ в базисе $f_1, \ldots, f_n$ является верхне-треугольной.

Итак, выбор нетривиального инвариантного подпространства позволяет понизить размерность спектральной задачи.

Следующая теорема помогает найти инвариантные подпространства.

Теорема 2. Если два линейных оператора коммутируют, то ядро и образ одного оператора всегда инвариантны относительно второго:

$$\mathcal{P}\mathcal{Q} = \mathcal{Q}\mathcal{P} \Longrightarrow \mathcal{P}(\operatorname{Ker}\mathcal{Q}) \subseteq \operatorname{Ker}\mathcal{Q}, \ \ \mathcal{P}(\operatorname{Im}\mathcal{Q}) \subseteq \operatorname{Im}\mathcal{Q}.$$

В частности, это верно, когда $\mathcal{P}$ и $\mathcal{Q}$ — многочлены от оператора $\mathcal{A}$.

Доказательство почти очевидно.

1) Если $v \in \operatorname{Ker}\mathcal{Q}$, то $\mathcal{Q}(\mathcal{P}v) = \mathcal{P}(\mathcal{Q}v) = \mathcal{P}0 = 0, \ \mathcal{P}v \in \operatorname{Ker}\mathcal{Q}$.

2) Если $\mathcal{Q}v \in \operatorname{Im}\mathcal{Q}$, то $\mathcal{P}(\mathcal{Q}v) = \mathcal{Q}(\mathcal{P}v) \in \operatorname{Im}\mathcal{Q}$ для всех $v \in V$.

Теорема 3. Пусть в конечномерном пространстве задано коммутативное семейство линейных операторов. Если каждый оператор диагонализируем, то в пространстве существует базис, векторы которого собственные для каждого оператора семейства.

Доказательство. Используем индукцию по размерности пространства V. Если пространство одномерно, то утверждение очевидно. Если каждый оператор скалярен, то утверждение также очевидно, любой базис пространства подходит. Отметим, что мы можем при необходимости изменить семейство, прибавляя к некоторым операторам скалярный, что не влияет на посылку и заключение теоремы.

В общем случае относительно любого диагонализируемого оператора пространство распадается в прямую сумму собственных подпространств. Если существует нескалярный оператор $\mathcal{A}$, то собственных значений и соответствующих слагаемых по крайней мере два — иначе оператор скалярен. Пусть

$$\operatorname{Sp}\mathcal{A} = \{\lambda_1, \ldots, \lambda_s\}, \ V = V_1 \oplus \ldots \oplus V_s, \ V_j = \operatorname{Ker}(\mathcal{A} - \lambda_j\mathcal{E}), \ s \geqslant 2, \ \dim V_j < \dim V.$$

Любой оператор $\mathcal{B}$, коммутирующий с $\mathcal{A}$, коммутирует и с $\mathcal{A} - \lambda_j \mathcal{E}$. По теореме 2 подпространства V_j инвариантны относительно $\mathcal{B}$.

Покажем, что V_j имеет базис, состоящий из собственных векторов оператора $\mathcal{B}$. Пусть

$$v = v_1 + \ldots + v_s \neq 0,\ v_j \in V_j,\ \mathcal{B}v = \lambda v = \mathcal{B}v_1 + \ldots + \mathcal{B}v_s.$$

Применим оператор $\mathcal{A}$. Тогда

$$\mathcal{A}\mathcal{B}v = \mathcal{A}(\lambda v) = \lambda(\mathcal{A}v) = \lambda\lambda_1 v_1 + \ldots + \lambda\lambda_s v_s.$$

С другой стороны,

$$\mathcal{A}\mathcal{B}v = \mathcal{B}\mathcal{A}v = \mathcal{B}(\lambda_1 v_1 + \ldots + \lambda_s v_s) = \lambda_1 \mathcal{B}v_1 + \ldots + \lambda_s \mathcal{B}v_s.$$

Поскольку сумма подпространств V_j прямая, то вектор $\mathcal{A}\mathcal{B}v$ единственным образом разлагается в проекции на V_j. Поэтому

$$\lambda_j \mathcal{B}v_j = \lambda\lambda_j v_j, \quad j = 1, \ldots, s.$$

Если $v_j \neq 0$, то можно считать, что $\lambda_j \neq 0$, иначе заменим $\mathcal{A}$ на $\mathcal{A} + \mathcal{E}$. Сокращая на λ_j, получим, что

$$\mathcal{B}v_j = \lambda v_j.$$

Это доказывает, что проекция собственного вектора оператора $\mathcal{B}$ на подпространство V_j либо равна нулю, либо является собственным вектором. Поскольку пространство V имеет базис, состоящий из собственных оператора $\mathcal{B}$, и проекция базиса V порождает V_j, то V_j имеет базис, состоящий из собственных векторов оператора $\mathcal{B}$. Итак, сужение $\mathcal{B}$ на V_j диагонализируемо. Поскольку $\dim V_j < \dim V$, то по предположению индукции каждое подпространство V_j имеет базис, векторы которого собственные для каждого оператора семейства. Объединяя по j, получим требуемый базис пространства V.

Теорема доказана.

Упр. 1. Найти все конечномерные подпространства в алгебре многочленов $\mathbb{R}[t]$, инвариантные относительно дифференцирования d/dt.

Упр. 2. Пусть оператор $\mathcal{A}$ задается матрицей A в базисе $e_1, \ldots, e_n$ векторного пространства V над полем K. Любое подпространство U *коразмерности 1* из пространства V в координатах в базисе $e_1, \ldots, e_n$ задается одном ненулевым линейным уравнением $a^\top x = a_1 x_1 + \cdots + a_n x_n = 0$. Доказать, что подпространство U инвариантно относительно оператора $\mathcal{A}$ в том и только том случае, когда вектор a собственный для матрицы $A^\top$, то есть

$$A^\top a = \lambda a, \quad a \neq 0.$$

Упр. 3. Доказать, что коммутирующие операторы в конечномерном комплексном пространстве всегда имеют общий собственный вектор. Более того, существует базис пространства, в котором матрицы всех операторов верхнетреугольны.

Упр. 4. Используя разложение вещественного многочлена в произведение линейных и квадратичных неприводимых множителей, докажите, что ненулевое конечномерное вещественное векторное пространство относительно любого линейного оператора содержит либо инвариантную прямую, либо инвариантную плоскость.

Упр. 5. Доказать обратное утверждение к следствию 1 теоремы 1.

НИЛЬПОТЕНТНЫЕ ОПЕРАТОРЫ

Линейный оператор $\mathcal{N}$ векторного пространства V над полем K называется *нильпотентным*, если $\mathcal{N}^k = O$ для некоторого натурального числа k. Такой оператор может быть ненулевым, как показывает пример оператора дифференцирования в пространстве многочленов степени $\leqslant n$, если $n \geqslant 1$. Другие примеры нильпотентных операторов — это разностные операторы

$$f(t) \mapsto \frac{f(t+h) - f(t)}{h} \text{ или } f(t) \mapsto \frac{f(t+h) - f(t-h)}{2h}$$

в пространстве многочленов ограниченной степени от переменной t, которыми заменяют приближенно взятие производной в вычислительной математике. Нильпотентность следует из того, что они понижают степень многочлена.

Предложение. 1) Все корни характеристического многочлена нильпотентного оператора — нулевые.

2) Если нильпотентный оператор диагонализируем, то он нулевой.

Доказательство. 1) Действительно, если N — матрица нильпотентного оператора, то $N^k = 0$. С другой стороны, если $\lambda \in \operatorname{Sp} N \subset L$ для подходящего расширения L поля K, то в пространстве L^n существует вектор-столбец $u \neq 0$, такой, что $Nu = \lambda u$. Отсюда

$$N^2 u = \lambda^2 u, \ldots, N^k u = \lambda^k u = 0,\ \lambda^k = 0,\ \lambda = 0.$$

2) Если бы в пространстве нашелся базис, состоящий из собственных векторов нильпотентного оператора, то оператор был бы нулевым, поскольку все собственные значения нулевые. Предложение доказано.

Итак, ненулевые нильпотентные операторы недиагонализируемы. Можно ли описать нильпотентные операторы с точностью до подобия?

Покажем, что конечномерное векторное пространство относительно нильпотентного оператора распадается в прямую сумму инвариантных подпространств, на каждом из которых оператор действует подобно оператору дифференцирования на пространстве многочленов ограниченной степени от одной переменной.

Последовательность векторов $v,\ \mathcal{N}v,\ \mathcal{N}^2 v, \ldots, \mathcal{N}^{h-1} v$ назовем *ниль-слоем высоты h с порождающим вектором v* относительно линейного оператора $\mathcal{N}$, если $\mathcal{N}^h v = 0$. Отметим, что, возможно $\mathcal{N}^m v = 0$ при $m < h$.

Таблица, столбцы которой — ниль-слои относительно $\mathcal{N}$ с общей нижней горизонтальной границей и порождающими векторами сверху, назовем *ниль-таблицей* относительно $\mathcal{N}$.

Таким образом, элементы ниль-таблицы — это векторы; верхний край ниль-таблицы может быть «рваным», если ниль-слои имеют разную высоту. Оператор опускает каждый вектор на этаж ниже, векторы первого этажа переходят в нуль.

Следующие преобразования ниль-таблиц назовем *элементарными*.

1) Прибавление к слою высоты h нижнего отрезка высоты h из другого слоя высоты $\geqslant h$, умноженного на некоторый скаляр.

2) Перестановка слоёв.

3) Умножение слоя на ненулевой скаляр.

4) Исключение нулевых векторов (элементов ниль-таблицы) сдвигом слоя вниз и обратное действие.

Очевидно, такие преобразования переводят ниль-таблицу в ниль-таблицу.

Лемма 1. Элементарные преобразования ниль-таблиц сохраняют линейную оболочку системы векторов ниль-таблицы.

Доказательство. Системы векторов преобразованной и исходной таблицы линейно эквивалентны, так как получаются друг из друга несколькими обычными элементарными преобразованиями систем векторов. Линейные оболочки эквивалентных систем совпадают.

Лемма 2. Если первый этаж ниль-таблицы — линейно независимая система векторов, то и все векторы ниль-таблицы — линейно независимая система.

Доказательство. Допустим, что векторы ниль-таблицы в совокупности линейно зависимы. Тогда в нетривиальной линейной комбинации этих векторов, равной нулю, присутствуют с ненулевыми коэффициентами векторы ниль-таблицы на максимальной высоте h и $h > 1$. Применяя к комбинации оператор $\mathcal{N}^{h-1}$, получим линейную зависимость векторов первого этажа. Полученное противоречие доказывает лемму.

Отметим, что обратное утверждение к утверждению леммы 2 также верно.

Базис пространства назовем *жордановым* относительно нильпотентного оператора N, если его можно представить в виде ниль-таблицы. Другими словами, такой базис является объединением непересекающихся ниль-слоев относительно $\mathcal{N}$.

Теорема 2. Конечномерное векторное пространство V относительно любого нильпотентного оператора $\mathcal{N}$ имеет жорданов базис. Число s_h максимальных ниль-слоев высоты h в любом таком базисе равно $r_{h-1} - 2r_h + r_{h+1}$, где $r_k = \dim \mathcal{N}^k V$ и потому зависит только от пространства V и оператора $\mathcal{N}$.

Доказательство. Составим ниль-таблицу из ниль-слоёв, порождаемых векторами некоторого базиса пространства, а затем элементарными преобразованиями перестроим её в базис пространства. Если векторы первого этажа таблицы линейно зависимы, то среди них найдется вектор, который выражается линейно через векторы, принадлежащие, вообще говоря, более высоким ниль-слоям, чем исходный. Поэтому элементарным преобразованием 1) на его месте можно получить нулевой вектор, а затем исключить и его преобразованием 4). Общее число векторов в таблице уменьшается и через конечное число шагов получится таблица с независимым первым этажом. По лемме 2 все векторы последней ниль-таблицы линейно независимы. По лемме 1 их линейная оболочка совпадает с линейной оболочкой исходной системы, которая включала базис пространства, и, потому равна всему пространству. Следовательно, получен базис пространства, являющийся объединением непересекающихся ниль-слоёв, т. е. жорданов базис.

Пусть дан произвольный жорданов базис. Тогда легко найти базисы подпространств $\mathcal{N}V$, $\mathcal{N}^2V, \ldots$ и подсчитать их размерности. Действительно, образ пространства — линейная оболочка образа базиса, но образ жорданова базиса получается сдвигом векторов соответствующей ниль-таблицы вниз на один этаж, кроме векторов первого этажа, переходящих в нуль. Поскольку ниль-слои под действием оператора сокращают высоту на единицу, то

$$\left\{\begin{array}{lllll} r_0 & = & \dim V & = & s_1 + 2s_2 + 3s_3 + 4s_4 + \ldots \\ r_1 & = & \dim \mathcal{N}V & = & s_2 + 2s_3 + 3s_4 + \ldots \\ r_2 & = & \dim \mathcal{N}^2V & = & s_3 + 3s_4 + \ldots \\ \ldots & \ldots & \ldots\ldots\ldots & \ldots & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right.$$

Вычитая из каждого уравнения следующее, а затем из каждой разности следующую, получим, что $s_h = r_{h-1} - 2r_h + r_{h+1}$. Теорема доказана.

Следствие. В подходящем базисе пространства матрица нильпотентного оператора принимает клеточно-диагональный вид с диагональными клетками

$$J_h = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ . & . & . & \dots & . & . \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

порядка h. Этот вид однозначен с точностью до порядка клеток J_h по диагонали. Характеристический многочлен нильпотентного оператора равен $(-\lambda)^n$, где n — размерность пространства, а спектр состоит только из нуля.

Доказательство. Матрица оператора имеет такой вид, если и только если базис составлен из ниль-слоев оператора и вектора в ниль-таблице нумеруются *по слоям снизу вверх, начиная с первого этажа.* Когда закончится один слой, можно перейти к любому другому и снова нумеровать снизу вверх. (Если нумеровать по слоям сверху вниз, то матрица транспонируется и получится нижне-треугольный вид клеток.) При этом клетка порядка h отвечает столбцу ниль-таблицы высоты h.

Упр. 1. По данному жорданову базису укажите базисы ядер $\operatorname{Ker}\mathcal{N}^k$ степеней нильпотентного оператора $\mathcal{N}$, а также базисы фактор-пространств $\operatorname{Ker}\mathcal{N}^k/\operatorname{Ker}\mathcal{N}^{k-1}$.

Предположим, что в пространстве столбцов $V = K^n$ над полем K задан линейный оператор $\mathcal{N} : x \mapsto Nx$ умножения столбцов на нильпотентную матрицу N. Следующий *алгоритм* выстраивает жорданов базис V относительно $\mathcal{N}$, моделируя доказательство теоремы с одним изменением — не следует сразу вычислять *все* ниль-слои, порождаемые базисными векторами.

Пусть $e_1, \dots, e_n$ — некоторый базис $V = K^n$, например, стандартный.

1. Вычислить ниль-слой, порождаемый e_1. Пусть

$$e_1, Ne_1, N^2e_1, \dots, N^{h-1}e_1 \neq 0$$

и $N^he_1 = 0$. Этот ниль-слой является ниль-таблицей с ненулевым последним вектором и по лемме 2 все его векторы линейно независимы. Перейти к шагу 2.

2. Если число векторов в полученной таблице равно размерности пространства, то они и образуют жорданов базис пространства, вычисления закончить. Иначе перейти к шагу 3.

3. Дополнить полученную ниль-таблицу ниль-слоем, порождаемым вектором e_j, который ранее не использовался. Перейти к шагу 4.

4. Перестроить ниль-таблицу элементарными преобразованиями так, чтобы система векторов её первого этажа была линейно независима, например, ступенчата и без нулевых векторов. Перейти к шагу 2.

Пример 1. $V = \mathbb{R}^3$, $\mathcal{N} : x \mapsto Nx$, $x \in V$,

$$N = \begin{pmatrix} 2 & 4 & 8 \\ 1 & 2 & 4 \\ -1 & -2 & -4 \end{pmatrix}, \quad \begin{array}{l} e_1 = (1,0,0)^\top, \quad Ne_1 = (2,1,-1)^\top, \quad N^2e_1 = 0, \quad 2 < 3 = \dim V, \\ e_2 = (0,1,0)^\top, \quad Ne_2 = (4,2,-2)^\top, \quad N^2e_2 = 0. \end{array}$$

$$\left(\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \hline 2 & 4 \\ 1 & 2 \\ \boxed{-1} & -2 \end{array}\right) \sim \left(\begin{array}{cc} 1 & -2 \\ 0 & 1 \\ 0 & 0 \\ \hline 2 & 0 \\ 1 & 0 \\ -1 & 0 \end{array}\right) \sim \left(\begin{array}{cc} 1 & \\ 0 & \\ 0 & \\ \hline 2 & -2 \\ 1 & 1 \\ -1 & 0 \end{array}\right).$$

Жорданов базис V относительно $\mathcal{N}$ — это

$$f_1 = (2, 1, -1)^\top,\ f_2 = (1, 0, 0)^\top,\ f_3 = (-2, 1, 0)^\top.$$

Пример 2. $V = \mathbb{R}^4,\ \ \mathcal{N} : x \mapsto Nx,\ x \in V,$

$$N = \begin{pmatrix} 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{array}{rr} e_4 = (0, 0, 0, 1)^\top, & Ne_4 = (4, 3, 2, 0)^\top, \\ N^2 e_4 = (6, 4, 0, 0)^\top, & N^3 e_4 = (8, 0, 0, 0)^\top. \end{array}$$

Жорданов базис V относительно $\mathcal{N}$ — это

$$f_1 = (8, 0, 0, 0)^\top,\ f_2 = (6, 4, 0, 0)^\top,\ f_3 = (4, 3, 2, 0)^\top,\ f_4 = (0, 0, 0, 1)^\top.$$

Пример 3. Пусть V — пространство вещественных многочленов степени не выше двух от переменных x, y с базисом $x^2,\ xy,\ y^2,\ x,\ y,\ 1$ и пусть $\mathcal{D} : f \mapsto \partial f/\partial x + \partial f/\partial y$ — понижающий общую степень, а потому нильпотентный линейный оператор на пространстве V.

Преобразуем ниль-таблицу с порождающими многочленами $x^2,\ xy,\ y^2$ относительно $\mathcal{D}$ (очевидно, что линейная оболочка векторов ниль-таблицы с порождающими многочленами $x^2,\ xy$ не совпадает с пространством V):

$$\begin{pmatrix} x^2 & xy & y^2 \\ 2x & x+y & 2y \\ \boxed{2} & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} x^2 & xy - x^2 & y^2 - x^2 \\ 2x & \boxed{y - x} & 2(y-x) \\ 2 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} x^2 & & \\ 2x & xy - x^2 & \\ 2 & y - x & (y-x)^2 \end{pmatrix}.$$

Столбцы последней таблицы образуют жорданов базис V относительно $\mathcal{D}$, её последняя строка — базис ядра $\mathcal{D}$, в частности, если $f \in V$, то

$$\partial f/\partial x + \partial f/\partial y = 0 \Leftrightarrow f = a + b(y-x) + c(y-x)^2$$

для некоторых $a,\ b,\ c \in \mathbb{R}$.

Упр. 2. Пусть

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}.$$

Прверить, что оператор коммутирования $\mathcal{C} : X \mapsto AX - XA$ в пространстве вещественных матриц порядка 2 нильпотентен. Найти жорданов базис пространства матриц относительно $\mathcal{C}$.

Упр. 3. Допустим, что известен жорданов базис пространства относительно нильпотентного оператора. Попробуйте описать все инвариантные подпространства. При каких условиях число инвариантных подпространств конечно?

КОРНЕВОЕ РАЗЛОЖЕНИЕ

Предположим, что *все корни* характеристического многочлена линейного оператора $\mathcal{A}$ в конечномерном векторном пространстве V принадлежат полю скаляров K. Случай кратных корней не исключается, различных корней s.

Предполагая, что спектр расщепляется, т. е. характеристический многочлен имеет по крайней мере два корня, разложим пространство в прямую сумму s ненулевых инвариантных подпространств. Разделим решение на следующие шаги.

1) *Поиск инвариантных слагаемых.* Выберем собственное значение λ. Тогда оператор $\mathcal{A}-\lambda\mathcal{E}$ вырожден и, следовательно, $V > (\mathcal{A}-\lambda\mathcal{E})V$ по теореме о сумме ранга и дефекта. Применяя $\mathcal{A}-\lambda\mathcal{E}$ несколько раз, получим убывающую *цепочку образов*

$$V > (\mathcal{A}-\lambda\mathcal{E})V > \ldots > (\mathcal{A}-\lambda\mathcal{E})^h V = (\mathcal{A}-\lambda\mathcal{E})^{h+1} V = \ldots,$$

которая выравнивается на конечном шаге $h = h(\lambda)$ ввиду конечномерности пространства V. Соответствующая *цепочка ядер*

$$0 < \operatorname{Ker}(\mathcal{A}-\lambda\mathcal{E}) < \ldots < \operatorname{Ker}(\mathcal{A}-\lambda\mathcal{E})^h = \operatorname{Ker}(\mathcal{A}-\lambda\mathcal{E})^{h+1} = \ldots$$

выравнивается ввиду теоремы о сумме ранга и дефекта.

Обозначим $\mathcal{P} = (\mathcal{A}-\lambda\mathcal{E})^h$. Подпространство $V^\lambda = \operatorname{Ker}\mathcal{P} = \operatorname{Ker}(\mathcal{A}-\lambda\mathcal{E})^{h(\lambda)}$, на котором выравнялась цепочка ядер, состоит из всевозможных решений уравнений $(\mathcal{A}-\lambda\mathcal{E})^k x = 0$, $k = 1, 2, \ldots$, в пространстве V, и называется *корневым подпространством*, отвечающим собственному значению λ оператора $\mathcal{A}$ в пространстве V.

Обозначим через $W = \operatorname{Im}\mathcal{P} = (\mathcal{A}-\lambda\mathcal{E})^h V$ подпространство, на котором выравнялась цепочка образов. Ввиду леммы 1 оба подпространства V^λ и W инвариантны относительно $\mathcal{A}$.

2) *Отщепление корневого подпространства.* Покажем, что пространство V — прямая сумма V^λ и W, а оператор $\mathcal{P}$ — «*проецирует*» V на W параллельно V^λ, т. е.

$$\mathcal{P}V^\lambda = 0, \quad \mathcal{P}V = W, \quad \mathcal{P}W = W, \quad V = V^\lambda \oplus W.$$

Действительно, первое и второе равенства сразу следуют из определений, третье следует из выравнивания цепочки образов. Поскольку $\mathcal{P}$ невырожден на W, то $(\operatorname{Ker}\mathcal{P}) \cap W = V^\lambda \cap W = 0$. Теперь четвертое равенство следует из сравнения размерностей V и $V^\lambda \oplus W$ с учетом теоремы о сумме ранга и дефекта для оператора $\mathcal{P}$ в пространстве V.

3) *Расположение других корневых подпространств.* Покажем, что остальные корневые подпространства $V^{\lambda'}$ при $\lambda' \neq \lambda$ содержатся во втором слагаемом W. Достаточно доказать, что $(\mathcal{A}-\lambda\mathcal{E})V^{\lambda'} = V^{\lambda'}$, поскольку тогда $(\mathcal{A}-\lambda\mathcal{E})^h V^{\lambda'} = V^{\lambda'}$, $V^{\lambda'} \subset \operatorname{Im}\mathcal{P} = W$. Включение $(\mathcal{A}-\lambda\mathcal{E})V^{\lambda'} \subset V^{\lambda'}$ вытекает из инвариантности ввиду леммы 1 всех корневых подпространств относительно многочленов от $\mathcal{A}$. Равенство следует из того, что ядро оператора $\mathcal{A}-\lambda\mathcal{E}$ на $V^{\lambda'}$ нулевое. Действительно, пусть $\mathcal{A}v = \lambda v$ и $(\mathcal{A}-\lambda'\mathcal{E})^k v = 0$. Тогда $\mathcal{A}v - \lambda' v = \lambda v - \lambda' v$, $(\mathcal{A}-\lambda'\mathcal{E})v = (\lambda-\lambda')v$, $(\mathcal{A}-\lambda'\mathcal{E})^k v = (\lambda-\lambda')^k v = 0$, $\lambda' \neq \lambda$, поэтому $v = 0$.

Отсюда вытекает, что спектр сужения $\mathcal{A}$ на V^λ состоит только из λ: ввиду леммы 2 он содержится в спектре $\mathcal{A}$, но собственные векторы, отвечающие другим собственным значениям, содержатся в W.

4) *Отщепление других корневых подпространств.* Выбирая $\lambda' \neq \lambda$, можно продолжить аналогичный процесс отщепления корневых подпространств уже от пространства W или использовать индукцию по размерности пространства:

$$W = W^{\lambda'} \oplus W', \quad W^{\lambda'} = V^{\lambda'}, \quad V = V^{\lambda} \oplus V^{\lambda'} \oplus W',$$

и так далее.

5) *Подведение итога.* Предположим, что

$$\operatorname{Sp} \mathcal{A} = \{\lambda_1, \ldots, \lambda_s\} \subseteq K, \quad \lambda_i \neq \lambda_j \text{ при } i \neq j.$$

Тогда

$$V = V_1 \oplus \ldots \oplus V_s, \quad V_i = V^{\lambda_i} = \operatorname{Ker}(A - \lambda_i E)^{h_i}.$$

Если $\mathcal{A}_i$ — сужение $\mathcal{A}$ на V_i, то $\operatorname{Sp} \mathcal{A}_i = \{\lambda_i\}$, $|\mathcal{A}_i - \lambda\mathcal{E}| = \pm(\lambda - \lambda_i)^{k_i}$, где $k_i = \dim V_i$. С другой стороны, $|\mathcal{A} - \lambda\mathcal{E}| = \prod |\mathcal{A}_i - \lambda\mathcal{E}| = \pm \prod (\lambda - \lambda_i)^{k_i}$ по лемме 2. Следовательно, размерность корневого подпространства совпадает с кратностью соответствующего собственного значения как корня характеристического многочлена.

Отметим еще, что $k_i = \dim V_i \geqslant h_i$. Следовательно, *можно определить корневое подпространство* V_i как ядро оператора $(A - \lambda_i E)^{k_i}$ или, равносильно, как пространство решений одной системы линейных уравнений $(A - \lambda_i E)^{k_i} x = 0$, где k_i — кратность собственного значения λ_i как корня характеристического многочлена $|A - \lambda E|$.

Таким образом, доказана

Теорема (о корневом разложении). Пусть $\mathcal{A}$ — линейный оператор в конечномерном векторном пространстве V на полем K. Предположим, что характеристический многочлен оператора разлагается над полем K на множители степени один:

$$|\mathcal{A} - \lambda\mathcal{E}| = \pm \prod_{i=1}^{s} (\lambda - \lambda_i)^{k_i}, \quad \lambda_i \in K, \quad \lambda_i \neq \lambda_j \text{ при } i \neq j,\ 1 \leqslant i, j \leqslant s.$$

Тогда

1) пространство разлагается в прямую сумму корневых подпространств

$$V = V_1 \oplus \ldots \oplus V_s, \quad V_i = V^{\lambda_i} = \operatorname{Ker}(\mathcal{A} - \lambda_i\mathcal{E})^{k_i}, \quad 1 \leqslant i \leqslant s,$$

которые инвариантны относительно оператора $\mathcal{A}$,

2) размерность корневого подпространства совпадает с кратность соответствующего собственного значения как корня характеристического многочлена

$$\dim V_i = k_i,\ 1 \leqslant i \leqslant s.$$

Пример 1. Найдем спектр и корневые подпространства оператора $\mathcal{A} : x \mapsto Ax,\ x \in V = \mathbb{R}^3$, где

$$A = \begin{pmatrix} 1 & 4 & 6 \\ -3 & -7 & -7 \\ 4 & 8 & 7 \end{pmatrix}.$$

Легко найти $\det(A - \lambda E) = -(\lambda - 3)(\lambda + 1)^2,\ \operatorname{Sp} A = \{3,\ -1\} \subset \mathbb{R}$. Поэтому

$$V = V_1 \oplus V_2, \quad V_1 = \operatorname{Ker}(\mathcal{A} - 3\mathcal{E}),\ V_2 = \operatorname{Ker}(\mathcal{A} + \mathcal{E})^2, \quad \dim V_1 = 1,\ \dim V_2 = 2.$$

Мы можем найти базисы этих подпространств, отыскивая фундаментальные системы решений для однородных систем линейных уравнений $(A-3E)x=0$, $(A+E)^2x=0$. Поскольку возведение матрицы в степень — трудоемкая операция, то лучше воспользуемся доказательством теоремы о корневом разложении, согласно которому при отщеплении корневого подпространства $V_1=\operatorname{Ker}(\mathcal{A}-3\mathcal{E})$ другое корневое подпространство $V_2=\operatorname{Ker}(\mathcal{A}+\mathcal{E})^2$ попадает в образ $\operatorname{Im}(\mathcal{A}-3\mathcal{E})$, а так как третьего корневого подпространства нет, то и совпадает с этим образом. Применим одновременный поиск ядра и образа для оператора $\mathcal{A}-3\mathcal{E}$.

$$\begin{pmatrix} E \\ A-3E \end{pmatrix} = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 2 & 4 & 6 \\ -3 & -10 & -7 \\ \boxed{4} & 8 & 4 \end{array}\right) \sim \left(\begin{array}{ccc} 1 & -2 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 2 & 8 & 8 \\ -3 & -4 & -4 \\ 4 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{ccc} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \\ \hline 2 & 8 & 0 \\ -3 & -4 & 0 \\ 4 & 0 & 0 \end{array}\right).$$

В итоге получаем, что вектор $f_1=(1,-1,1)^\top$ образует базис корневого подпространства V_1, а векторы $f_2=(2,-3,4)^\top$, $f_3=(8,-4,0)^\top$ — базис V_2.

В предыдущем примере всего два корневых подпространства, причем одно из них одномерно. Как избежать трудоемкой операции возведения матрицы в степень, если все корни характеристического многочлена кратные?

Предположим, что в пространстве столбцов $V=K^n$ над полем K задан линейный оператор $\mathcal{A}: x\mapsto Ax$ умножения столбцов на матрицу A и что известны его собственные значения. Следующий *алгоритм* вычисляет базисы корневых подпространств. Он вытекает из теоремы о корневом разложении и использует многократное одновременное вычисление ядер и образов относительно операторов $\mathcal{N}=\mathcal{A}-\lambda\mathcal{E}$.

Пусть B_0 — единичная или любая невырожденная матрица порядка n над полем K. Важно только, что столбцы B_0 образуют базис пространства K^n. Пусть $N=A-\lambda E$, где λ — собственное значение матрицы A.

1) Элементарными преобразованиями слоев $\mathcal{N}$-слойную матрицу $\begin{pmatrix} B_0 \\ NB_0 \end{pmatrix}$ привести к виду $\begin{pmatrix} B_1 \\ C_1 \end{pmatrix}$, где C_1 — n-строчная матрица, ступенчатая по столбцам. Здесь важно только, что ненулевые столбцы матрицы C_1 линейно независимы.

2) Вычислить произведение NC_1 и элементарными преобразованиями слоев N^2-слойную матрицу $\begin{pmatrix} B_1 \\ NC_1 \end{pmatrix}$ привести к виду $\begin{pmatrix} B_2 \\ C_2 \end{pmatrix}$, где C_2 — n-строчная матрица, ступенчатая по столбцам.

3) Вычислить произведение NC_2 и т. д.

Вычисления закончить на матрице $\begin{pmatrix} B_h \\ C_h \end{pmatrix}$, если ранги матриц C_h и NC_h совпадают, или, равносильно, если число нулевых столбцов в C_h равно кратности $k(\lambda)$ корня λ в характеристическом многочлене матрицы A. В этом случае система столбцов матрицы B_h, имеющих нулевое продолжение в матрице C_h, образует базис корневого подпространства $V^\lambda(\mathcal{A})$, а ненулевые столбцы матрицы C_h, образуют базис пространства $W=\operatorname{Im}(\mathcal{A}-\lambda\mathcal{E})^{h(\lambda)}$, содержащего остальные корневые подпространства.

Теперь пусть новая матрица B_0 получается из матрицы C_h удалением нулевых столбцов, и пусть $N := A - \lambda' E$, где λ' — другое собственное значение матрицы A. Вернемся к началу алгоритма и продолжим вычисления в соответствии с указанными правилами. Тогда получим базис корневого подпространства $V^{\lambda'}(\mathcal{A})$ и базис пространства $W' = \mathrm{Im}\,(\mathcal{A} - \lambda'\mathcal{E})^{h(\lambda')}|_W$, содержащего все корневые подпространства, кроме $V^{\lambda}(\mathcal{A})$ и $V^{\lambda'}(\mathcal{A})$. Продолжая вычисления, найдем в итоге базисы всех корневых подпространств.

Отметим, что в предложенном алгоритме активно используется инвариантность подпространств, за счет которой происходит понижение размерности, матрицы вида B_0 содержат все меньше столбцов и произведения вида NB_0 вычисляются с шагом алгоритма все быстрее.

Пример 2. Пусть $\mathcal{A} : x \mapsto Ax,\ x \in V = \mathbb{R}^4$, где

$$A = \begin{pmatrix} 2 & 2 & 0 & -1 \\ -1 & -2 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Тогда

$$\det(A - \lambda E) = (\lambda-1)^2(\lambda+1)^2,\ \ V = \mathrm{Ker}\,(\mathcal{A}-\mathcal{E})^2 \oplus \mathrm{Ker}\,(\mathcal{A}+\mathcal{E})^2,\ \ \mathrm{Ker}\,(\mathcal{A}+\mathcal{E})^2 = \mathrm{Im}\,(\mathcal{A}-\mathcal{E})^2.$$

Чтобы отыскать базисы этих подпространств и не вычислять квадрат $N = A - E$, найдем одновременно базисы ядра и образа относительно $\mathcal{N} = \mathcal{A} - \mathcal{E}$, а потом применим оператор $\mathcal{N}$ еще раз к базису образа. В соответствии с алгоритмом

$$\begin{pmatrix} E \\ N \end{pmatrix} = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 2 & 0 & -1 \\ -1 & -3 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & \boxed{-1} \end{array}\right) \sim \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 0 & 1 & -1 & -1 \\ 0 & -2 & 1 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array}\right) = \begin{pmatrix} B_1 \\ C_1 \end{pmatrix},$$

$$\begin{pmatrix} B_1 \\ NC_1 \end{pmatrix} = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 0 & -3 & 1 & 2 \\ 0 & 5 & -2 & -3 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & \boxed{1} \end{array}\right) \sim \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 1 & 3 & 2 & 1 \\ \hline 0 & 0 & 3 & 2 \\ 0 & 0 & -5 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right) = \begin{pmatrix} B_2 \\ C_2 \end{pmatrix}.$$

Поскольку кратность корня 1 равна двум, то вычисления следует закончить, векторы $f_1 = (0,1,1,1)^\top,\ f_2 = (0,1,2,3)^\top$ образуют базис корневого подпространства $\mathrm{Ker}\,(\mathcal{A}-\mathcal{E})^2$, а векторы $f_3 = (3,-5,1,0)^\top,\ f_4 = (2,-3,0,1)^\top$ — базис корневого подпространства $\mathrm{Ker}\,(\mathcal{A}+\mathcal{E})^2$.

Упр. 1. Пусть $\mathcal{A}$ — линейный оператор конечномерного векторного пространства V над полем K и $\operatorname{Sp}\mathcal{A} \subset K$. Доказать, что $\mathcal{A}$ диагонализируем тогда и только тогда, когда $\operatorname{rk}(\mathcal{A} - \lambda\mathcal{E}) = \operatorname{rk}(\mathcal{A} - \lambda\mathcal{E})^2$ для всех $\lambda \in \operatorname{Sp}\mathcal{A}$.

Упр. 2. Доказать, что если пространство V относительно линейного оператора $\mathcal{A}$ распадается в прямую сумму корневых подпространств V_i, то всякое подпространство $U \subset V$, инвариантное относительно $\mathcal{A}$, распадается в прямую сумму подпространств $U_i \subset V_i$, инвариантных относительно $\mathcal{A}$.

МИНИМАЛЬНЫЙ АННУЛИРУЮЩИЙ МНОГОЧЛЕН И ТЕОРЕМА ГАМИЛЬТОНА-КЭЛИ

Пусть $\mathcal{A}$ — линейный оператор векторного пространства V над полем K. Многочлен $m(\lambda) = m_{\mathcal{A}}(\lambda)$ из $K[\lambda]$ называется *минимальным аннулирующим* для оператора $\mathcal{A}$, если он имеет минимальную степень среди всех нормированных многочленов, аннулирующих оператор $\mathcal{A}$. Аналогично определяется минимальный аннулирующий многочлен для матрицы порядка n над полем K.

Лемма. Минимальный аннулирующий многочлен делит любой другой аннулирующий многочлен для данного оператора.

Доказательство. Пусть $f(\lambda) \in K[\lambda]$ и $f(\mathcal{A}) = 0$. Разделим $f(\lambda)$ на $m(\lambda)$ с остатком:

$$f(\lambda) = m(\lambda)q(\lambda) + r(\lambda), \quad \deg r(\lambda) < \deg m(\lambda).$$

Тогда $0 = f(\mathcal{A}) = m(\mathcal{A})q(\mathcal{A}) + r(\mathcal{A}) = r(\mathcal{A})$. Если $r(\lambda) \neq 0$, то разделив его на старший коэффициент, получим нормированный аннулирующий многочлен меньшей степени по сравнению с $m(\lambda)$. Противоречие показывает, что $r(\lambda) = 0$ и, следовательно, $m(\lambda)$ делит $f(\lambda)$.

Теорема. Пусть $\mathcal{A}$ — линейный оператор в конечномерном векторном пространстве V над полем K. Предположим, что $\operatorname{Sp}\mathcal{A} = \{\lambda_1, \ldots, \lambda_s\} \subset K$. Тогда минимальный аннулирующий многочлен $m(\lambda)$ для оператора $\mathcal{A}$ существует и только один. Он равен $\prod_{i=1}^{s}(\lambda - \lambda_i)^{h_i}$, где h_i — максимальная высота корневого вектора, отвечающего собственному значению λ_i оператора $\mathcal{A}$.

Доказательство. Существование минимального аннулирующего многочлена следует из конечномерности алгебры линейных операторов: степени оператора $\mathcal{A}$ с показателями от нуля до n^2, $n = \dim V$, линейно зависимы. Точнее, если перейти к от оператора к его матрице A порядка n над полем K, то существует наименьшее натуральное число k и элементы $c_1, \ldots, c_k$ из K, для которых матрицы A^{k-1}, $A^{k-2}, \ldots, E$ линейно независимы над полем K, но

$$A^k = c_1A^{k-1} + c_2A^{k-2} + \ldots + c_kE.$$

Коэффициенты линейной зависимости всегда можно выбрать из наименьшего поля, содержащего коэффициенты матрицы A. Поскольку независимость векторов из K^m сохраняется при расширении поля K, то такого рода зависимость между степенями матрицы A единственна и минимальный аннулирующий многочлен матрицы однозначно задается матрицей независимо от поля, объемлющего все коэффициенты матрицы. Для заданного поля единственность следует из леммы: два минимальных аннулирующих многочлена делят друг друга, нормированы и потому равны.

Теперь установим вид минимального аннулирующего многочлена. Из доказательства теоремы о корневом разложении

$$V = V_1 \oplus \ldots \oplus V_s, \quad V_i = \operatorname{Ker}(\mathcal{A} - \lambda_i\mathcal{E})^{h_i},$$

где $h_i = \min\{h \mid (\mathcal{A}-\lambda_i\mathcal{E})^h V = (\mathcal{A}-\lambda_i\mathcal{E})^{h+1}V\}$, т. е. h_i — максимальная высота корневого вектора, отвечающего собственному значению λ_i оператора $\mathcal{A}$, $h_i \leqslant k_i$. Отсюда очевидно, что оператор $\prod_{i=1}^{s}(\mathcal{A}-\lambda_i\mathcal{E})^{h_i}$ аннулирует каждое корневое подпространство V_i, их сумму V и, значит, равен нулю. По лемме многочлен $\prod_{i=1}^{s}(\lambda - \lambda_i)^{h_i}$ делится на минимальный аннулирующий многочлен $m(\lambda)$ оператора $\mathcal{A}$.

Пусть $\mathcal{A}_i$ — сужение оператора $\mathcal{A}$ на V_i и $m_i(\lambda)$ — минимальный аннулирующий многочлен для оператора $\mathcal{A}_i$. Из определения подпространств V_i получаем $(\mathcal{A}_i-\lambda_i\mathcal{E})^{h_i} = 0$. По лемме многочлен $m_i(\lambda)$ делит $(\lambda - \lambda_i)^{h_i}$. Поэтому $m_i(\lambda) = (\lambda - \lambda_i)^{l_i}$, $\; l_i \leqslant h_i$. Если $l_i < h_i$, то $(\mathcal{A}_i-\lambda_i\mathcal{E})^{l_i}v \neq 0$, где v — вектор высоты h_i из V_i. Следовательно, $l_i = h_i$. Снова ввиду леммы $m_i(\lambda) = (\lambda-\lambda_i)^{h_i}$ делит $m(\lambda)$. По свойствам взаимно простых многочленов многочлен $\prod_{i=1}^{s}(\lambda - \lambda_i)^{h_i}$ делит $m(\lambda)$.

Теорема доказана.

Следствие (теорема Гамильтона–Кэли). Линейный оператор аннулируется своим характеристическим многочленом.

Доказательство достаточно провести для матриц. Свяжем с матрицей A линейный оператор умножения столбцов на матрицу $\mathcal{A} : x \mapsto Ax$, $x \in K^n$. Разложим характеристический многочлен матрицы A на множители степени один над расширением L поля K как в теореме о корневом разложении. Тогда ввиду теоремы о корневом разложении оператор $\prod_{i=1}^{s}(\mathcal{A} - \lambda_i\mathcal{E})^{k_i}$ аннулирует каждое корневое подпространство в L^n, т. е. все пространство L^n, и, следовательно, равен нулевому.

Упр. 1. Линейный оператор в конечномерном векторном пространстве, спектр которого содержится в поле скаляров, диагонализируем в том и только том случае, когда его минимальный аннулирующий многочлен не имеет кратных корней.

Упр. 2. Покажите, что любой линейный оператор $\mathcal{P}$ со свойством $\mathcal{P}^2 = \mathcal{P}$ проецирует пространство на подпространство $W = \operatorname{Im}\mathcal{P}$ параллельно подпространству $U = \operatorname{Ker}\mathcal{P}$.

ЖОРДАНОВА КЛАССИФИКАЦИЯ ОПЕРАТОРОВ

Жорданова клетка порядка h, отвечающая собственному значению λ — это матрица порядка h

$$J_h(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ . & . & . & \ldots & . & . \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}.$$

Матрица имеет жорданову нормальную форму, если она клеточно-диагональна с жордановыми клетками по главной диагонали.

Пример 1. Жордановы нормальные формы порядка 2 — это матрицы

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 \\ 0 & \lambda' \end{pmatrix}.$$

Теорема. Пусть $\mathcal{A}$ — линейный оператор ненулевого конечномерного векторного пространства V над полем K и пусть $\mathrm{Sp}\,\mathcal{A} \subseteq K$. Тогда

1) в подходящем базисе V матрица оператора $\mathcal{A}$ имеет жорданову нормальную форму J;

2) число $s_h(\lambda)$ диагональных жордановых клеток $J_h(\lambda)$ в матрице J зависит только от $\mathcal{A}$ и равно $r_{h-1}(\lambda) - 2r_h(\lambda) + r_{h+1}(\lambda)$, где $r_k(\lambda) = \mathrm{rk}\,(\mathcal{A} - \lambda\mathcal{E})^k$.

Доказательство. 1) По теореме о корневом разложении $V = \oplus_{\lambda\in\mathrm{Sp}\,\mathcal{A}} V^\lambda$, где $V^\lambda = \mathrm{Ker}\,(\mathcal{A} - \lambda\mathcal{E})^{h(\lambda)}$. Сужение $\mathcal{A} - \lambda\mathcal{E}$ на V^λ нильпотентно и по основной теореме о нильпотентных операторах пространство V^λ имеет базис, составленный из непересекающихся ниль-слоев $f_1, \ldots, f_h$ относительно $\mathcal{A} - \lambda\mathcal{E}$. Но

$$\left\{\begin{array}{lcl} (\mathcal{A} - \lambda\mathcal{E})f_1 & = & 0 \\ (\mathcal{A} - \lambda\mathcal{E})f_2 & = & f_1 \\ \quad\ldots & . & . \\ (\mathcal{A} - \lambda\mathcal{E})f_h & = & f_{h-1} \end{array}\right. \iff \left\{\begin{array}{lcl} \mathcal{A}f_1 & = & \lambda f_1 \\ \mathcal{A}f_2 & = & f_1 + \lambda f_2 \\ \ \ldots & . & \ldots \\ \mathcal{A}f_h & = & f_{h-1} + \lambda f_h \end{array}\right. \iff \mathcal{A}_f = J_h(\lambda).$$

Таким образом, при вычислении матрицы оператора ниль-слою высоты h относительно $\mathcal{A} - \lambda\mathcal{E}$ соответствует жорданова клетка порядка h, отвечающая собственному значению λ, расположенная по диагонали. Следовательно, в базисе V, составленном из непересекающихся ниль-слоев относительно $\mathcal{A} - \lambda\mathcal{E}$, $\lambda \in \mathrm{Sp}\,\mathcal{A}$, и только таком базисе, матрица оператора $\mathcal{A}$ имеет жорданову нормальную форму J. Этот базис будем называть *жордановым*.

2) Ввиду предыдущего можно отождествить число $s_h(\lambda)$ с числом максимальных по включению ниль-слоев высоты h относительно $\mathcal{A} - \lambda E$ в некотором базисе V. Все такие векторы лежат в корневом подпространстве V^λ и линейно независимы. Если их число меньше, чем $\dim V^\lambda$ для некоторого λ, то их общее число по всем λ меньше, чем $\sum_{\lambda\in\mathrm{Sp}\,\mathcal{A}} \dim V^\lambda = \dim V$, что невозможно. Следовательно, такие векторы всегда образуют базис V^λ. По теореме о нильпотентных операторах $s_h(\lambda) = r_{h-1} - 2r_h + r_{h+1}$, где $r_k = \dim(\mathcal{A} - \lambda\mathcal{E})^k V^\lambda$. Ввиду теоремы о корневом разложении

$$V = V^\lambda \oplus W, \quad (\mathcal{A} - \lambda\mathcal{E})W = W.$$

Поэтому $(\mathcal{A} - \lambda\mathcal{E})^k V = (\mathcal{A} - \lambda\mathcal{E})^k V^\lambda \oplus W$. Если $d = \dim W$, то отсюда $r_k(\lambda) = r_k + d$, $r_k = r_k(\lambda) - d$, и тогда $s_h(\lambda) = (r_{h-1}(\lambda) - d) - 2(r_h(\lambda) - d) + (r_{h+1}(\lambda) - d) = r_{h-1}(\lambda) - 2r_h(\lambda) + r_{h+1}(\lambda)$. Теорема доказана.

Следствие. Всякая квадратная матрица над полем, содержащим её спектр, подобна матрице в жордановой нормальной форме. Она задается исходной матрицей однозначно с точностью до порядка жордановых клеток по диагонали.

Упр. 1. Число клеток в жордановой нормальной форме матрицы оператора равно сумме размерностей собственных подпространств оператора.

Пример 2. Для матрицы A найдем её жорданову нормальную форму J и матрицу C перехода к ней, так что $J = C^{-1}AC$, если

$$A = \begin{pmatrix} 0 & 6 & -2 & 2 \\ -1 & 5 & -1 & 1 \\ -1 & 1 & -1 & -1 \\ 2 & -4 & 4 & 2 \end{pmatrix}.$$

Пусть $V = \mathbb{R}^4$, $\mathcal{A} : x \mapsto Ax$, $x \in \mathbb{R}^4$. Так как $\det(A - \lambda E) = \lambda(\lambda - 2)^3$, то можно найти жорданов базис $\mathbb{R}^4$ относительно оператора $\mathcal{A}$. По теореме о корневом разложении $V = V^0 \oplus V^2$, $V^0 = \operatorname{Ker} \mathcal{A}$, $\dim V^0 = 1$, $V^2 = \operatorname{Ker}(\mathcal{A} - 2\mathcal{E})^3$, $\dim V^2 = 3$. Выгоднее сначала искать ниль-слои, связанные с корнем большой кратности, так как ранг матрицы $A - \lambda E$ может быть мал. При $\lambda = 2$ получаем

$$\begin{pmatrix} E \\ A - 2E \end{pmatrix} = \left(\begin{array}{rrrr} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 6 & -2 & 2 \\ -1 & 3 & -1 & 1 \\ -1 & 1 & -3 & -1 \\ \boxed{2} & -4 & 4 & 0 \end{array}\right) \sim \left(\begin{array}{rrrr} 1 & 2 & -2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 2 & 2 & 2 \\ -1 & 1 & 1 & 1 \\ -1 & \boxed{-1} & -1 & -1 \\ 2 & 0 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{rrrr} 1 & 2 & -4 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 2 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{array}\right)$$

Отсюда $\dim \operatorname{Ker}(\mathcal{A} - 2\mathcal{E}) = 2$, но $\dim V^2 = 3$. Поэтому применим к базису $\operatorname{Im}(\mathcal{A} - 2\mathcal{E})$ оператор $\mathcal{A} - 2\mathcal{E}$ ещё раз:

$$\left(\begin{array}{rrrr} 1 & 2 & -4 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 2 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ \hline 4 & 4 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ \boxed{-4} & -4 & 0 & 0 \end{array}\right) \sim \left(\begin{array}{rrrr} 1 & 1 & -4 & -2 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline -2 & 4 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 2 & -2 & 0 & 0 \\ \hline 4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ -4 & 0 & 0 & 0 \end{array}\right).$$

Отсюда $V^2 = \operatorname{Ker}(\mathcal{A} - 2\mathcal{E})^2$, $V^0 = \operatorname{Ker} \mathcal{A} = \operatorname{Im}(\mathcal{A} - 2\mathcal{E})^2$, вектор $f_1 = (4, 2, 2, -4)^\top$ образует базис V^0, а начальные векторы последних трех слоёв образуют базис V^2. Теперь найдем базисы V^0 и V^2, составленные из ниль-слоев относительно операторов $\mathcal{A} - 0\mathcal{E}$ и $\mathcal{A} - 2\mathcal{E}$ соответственно. Так как $Af_1 = 0$, то для V^0 задача решена. Ниль-слои относительно $\mathcal{A} - 2\mathcal{E}$ уже есть. Второй столбец матрицы содержит ненулевой слой высоты 2. Так как $\dim V^2 = 3$, то дополним его ненулевым вектором третьего столбца и запишем результат в виде ниль-таблицы

$$\left(\begin{array}{rr} 1 & \\ 1 & \\ 0 & \\ 0 & \\ \hline 4 & -4 \\ 2 & -1 \\ 0 & 1 \\ -2 & 0 \end{array}\right).$$

Векторы нижнего этажа линейно независимы, поэтому все три вектора независимы и образуют жорданов базис для V^2. Обозначая $f_2 = (4, 2, 0, -2)^\top$, $f_3 = (1, 1, 0, 0)^\top$, $f_4 = (-4, -1, 1, 0)^\top$, получим

$$\begin{cases} Af_1 & = & 0 \\ (A-2E)f_2 & = & 0 \\ (A-2E)f_3 & = & f_2 \\ (A-2E)f_4 & = & 0 \end{cases}, \quad \begin{cases} Af_1 & = & 0 \\ Af_2 & = & 2f_2 \\ Af_3 & = & f_2 + 2f_3 \\ Af_4 & = & 2f_4 \end{cases}, \quad J = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = A_f.$$

Если $e_1, \ldots, e_4$ — стандартный базис $\mathbb{R}^4$, то $A_e = A$. Пусть C — матрица перехода от стандартного базиса к жорданову. Тогда

$$C = (f_1 f_2 f_3 f_4) = \begin{pmatrix} 4 & 4 & 1 & -4 \\ 2 & 2 & 1 & -1 \\ 2 & 0 & 0 & 1 \\ -4 & -2 & 0 & 0 \end{pmatrix}, \quad J = \mathcal{A}_f = C^{-1}\mathcal{A}_e C = C^{-1}AC.$$

Пример 3. В трехмерном евклидовом пространстве задана декартова система координат, в которой координаты движущейся материальной точки $x(t) = (x_1(t), x_2(t), x_3(t))^\top$ и компоненты вектора скорости $dx/dt = (dx_1/dt, dx_2/dt, dx_3/dt)^\top$ связаны системой уравнений

$$\begin{cases} dx_1/dt & = & -x_1 + 0x_2 + x_3, \\ dx_2/dt & = & -3x_1 + 2x_2 + 7x_3, \\ dx_3/dt & = & x_1 - x_2 - 3x_3. \end{cases}$$

В начальный момент времени $t = 0$ точка имеет координаты $x_0 = (1, 2, -2)^\top$. Требуется найти параметрические уравнения движения точки.

Представим систему в матричном виде $dx/dt = Ax$, где

$$A = \begin{pmatrix} -1 & 0 & 1 \\ -3 & 2 & 7 \\ 1 & -1 & -3 \end{pmatrix}.$$

Переход в другую декартову систему координат $x = Cy$ приведет систему уравнений к виду $dy/dt = C^{-1}ACy$. Выберем ее так, чтобы матрица $J = C^{-1}AC$ стала достаточно простой, например, жордановой нормальной формой.

Свяжем с матрицей A оператор $\mathcal{A} : x \mapsto Ax$ в пространстве $V = \mathbb{C}^3$.

Легко вычислить $|A - \lambda E| = -\lambda(\lambda + 1)^2$.

Найдем базисы корневых подпространств $V_1 = \operatorname{Ker} \mathcal{A}$ и $V_2 = \operatorname{Ker}(\mathcal{A} + \mathcal{E})^2 = \operatorname{Im} \mathcal{A}$:

$$\begin{pmatrix} E \\ A \end{pmatrix} = \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline -1 & 0 & \boxed{1} \\ -3 & 2 & 7 \\ 1 & -1 & -3 \end{array}\right) \sim \left(\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 4 & \boxed{2} & 7 \\ -2 & -1 & -3 \end{array}\right) \sim \left(\begin{array}{ccc} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 0 & 2 & 7 \\ 0 & -1 & -3 \end{array}\right).$$

Отсюда $u_1 = (1, -2, 1)^\top$ — базис V_1, $u_2 = (0, 2, -1)^\top$, $u_3 = (1, 7, -3)^\top$ — базис V_2.

Найдем жордановы базисы корневых подпространств.

Для V_1 — это вектор $f_1 = u_1$.

В корневом подпространстве V_2 вытягиваем ниль-слой относительно $\mathcal{A} + \mathcal{E}$, порождаемый u_2:

$$\begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \xrightarrow{A+E} \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}.$$

Отсюда $f_2 = (-1, -1, 0)^\top$, $f_3 = u_2 = (0, 2, -1)^\top$ — жорданов базис пространства V_2.

Объединяя жордановы базисы V_1 и V_2, получим жорданов базис f_1, f_2, f_3 пространства V. Матрица перехода от стандартного базиса к жорданову — это матрица $C = (f_1\ f_2\ f_3)$. Кроме того,

$$C^{-1}AC = J = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

— жорданова форма матрицы A.

В координатах в базисе f_1, f_2, f_3 система уравнений $dy/dt = Jy$ расщепляется и принимает вид

$$\begin{cases} dy_1/dt &= 0, \\ dy_2/dt &= -y_2 + y_3, \\ dy_3/dt &= -y_3. \end{cases}$$

Очевидное решение

$$y_1(t) = c_1, \quad y_2(t) = (c_2 + c_3 t)e^{-t}, \quad y_3(t) = c_3 e^{-t}.$$

Константы c_1, c_2, c_3 можно найти, решая систему линейных уравнений $x(0) = Cy(0)$. Нетрудно подсчитать, что $c_1 = 3$, $c_2 = 2$, $c_3 = 5$.

Мы видим, что движение точки происходит в плоскости, заданной уравнением $y_1 = 3$, параллельной линейной оболочке f_2 и f_3. Кроме того, движение устойчиво: $x(t) \to 3f_1 = (3, -6, 3)^\top$ при $t \to +\infty$.

Замечание. Систему дифференциальных уравнений $dy/dt = Jy$, матрица которой — жорданова клетка $J = J_h(\lambda)$, можно легко решить, разбивая J в сумму $J = \lambda E + N$. Общее решение системы $dy/dt = Ny$ имеет вид $y_1 = c_1$, $y_2 = c_2 + c_1 t$, $y_3 = c_3 + c_2 t + c_1 t^2/2, \ldots,$ $y_h = c_h + c_{h-1}t + \ldots + c_1 t^{h-1}/(h-1)!$. Умножая его на $e^{\lambda t}$, получим общее решение исходной системы уравнений.

Пример 4. Решим в алгебре $M_2(\mathbb{R})$ с помощью нормальной жордановой формы уравнение $\sin X = A$, где

$$A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

Поскольку $\sin X = \sum_{k=0}^{\infty} (-1)^k X^{2k+1}/(2k+1)!$, то $XA = X(\sin X) = (\sin X)X = AX$. Это линейное уравнение можно решить, но вряд ли оно поможет решить исходное уравнение.

Попытаемся упростить матрицу A с помощью преобразования подобия. Пусть $J = C^{-1}AC$ — жорданова форма матрицы A. Нетрудно вычислить, что в данном конкретном случае

$$J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}.$$

Тогда
$$J = C^{-1}AC = C^{-1}(\sin X)C = \sin(C^{-1}XC),$$
поскольку синус — сумма ряда. Если обозначить $Y = C^{-1}XC$, то получаем уравнение $\sin Y = J$. Снова $YJ = JY$ и это уравнение решить проще: если
$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$
то $\gamma = 0,\ \delta = \alpha$. Треугольная матрица Y легко возводится в степень:
$$Y^k = \begin{pmatrix} \alpha^k & \alpha^{k-1}\beta \\ 0 & \alpha^k \end{pmatrix}.$$
Отсюда
$$\sin Y = \sum_{k=0}^{\infty}(-1)^k\frac{Y^{2k+1}}{(2k+1)!} = \begin{pmatrix} \sin\alpha & \beta\cos\alpha \\ 0 & \sin\alpha \end{pmatrix}.$$
Если $\sin Y = J$, то $\sin\alpha = 0,\quad \beta\cos\alpha = 1,\ \ \alpha = k\pi,\ \beta = (-1)^k$. Окончательно
$$X = CYC^{-1} = C(k\pi E + (-1)^k J)C^{-1} = k\pi E + (-1)^k A.$$

ФУНКЦИИ ОТ МАТРИЦ И ЖОРДАНОВА НОРМАЛЬНАЯ ФОРМА

Вычисление многочленов от матриц кажется простой операцией:
$$f(\lambda) = \sum_{k=0}^{N} c_k\lambda^k \Longrightarrow f(A) = \sum_{k=0}^{N} c_k A^k.$$
Проблема в том, что при перемножении двух матриц порядка n требуется выполнить n^3 операций умножения элементов матриц, значит, для вычисления значения многочлена нужно выполнить порядка Nn^3 операций умножения.

Уменьшить трудоемкость может помочь знание *минимального аннулирующего* многочлена $m(\lambda)$ матрицы A: если $r(\lambda)$ — остаток от деления $f(\lambda)$ на $m(\lambda)$ и $h = \deg m(\lambda)$, то $f(A) = r(A)$, число операций умножения скаляров для вычисления $f(A)$ сокращается до порядка $(h-1)n^3$.

Допустим, что известна жорданова нормальная форма J матрицы A и матрица C перехода к ней: $J = C^{-1}AC$. Тогда легко проверяется, что
$$f(J) = C^{-1}f(A)C,\quad f(A) = Cf(J)C^{-1}.$$

Пусть $J = Diag(J_1, \ldots, J_m)$, где J_k — жордановы клетки. Тогда легко проверяется, что
$$f(J) = Diag(f(J_1), \ldots, f(J_m)).$$
Осталось вычислить значение многочлена от жордановой клетки.

Лемма.

$$f(J_h(\lambda_j)) = \begin{pmatrix} f(\lambda_j) & f'(\lambda_j)/1! & f''(\lambda_j)/2! & ... & ... & f^{(h-1)}(\lambda_j)/(h-1)! \\ 0 & f(\lambda_j) & f'(\lambda_j)/1! & f''(\lambda_j)/2! & ... & ... \\ 0 & 0 & f(\lambda_j) & f'(\lambda_j)/1! & f''(\lambda_j)/2! & ... \\ ... & ... & ... & ... & ... & ... \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & 0 & f(\lambda_j) \end{pmatrix}$$

Доказательство. Разложим многочлен по формуле Тейлора в точке λ_j:

$$f(\lambda) = \sum_{k\geqslant 0} \frac{f^{(k)}(\lambda_j)}{k!}(\lambda-\lambda_j)^k.$$

Тогда

$$f(J_h(\lambda_j)) = \sum_{k\geqslant 0} \frac{f^{(k)}(\lambda_j)}{k!}(J_h(\lambda_j)-\lambda_j E)^k.$$

Матрица $N = J_h(\lambda_j) - \lambda_j E$ является матрицей оператора сдвига ниль-слоя высоты h на один этаж вниз. Тогда N^k — матрица сдвига ниль-слоя на k этажей ниже. Она отличается от нулевой единицами на диагонали выше главной на $k-1$ шагов. Умножая ее на скаляры $\dfrac{f^{(k)}(\lambda_j)}{k!}$ и суммируя, получим требуемую матрицу. Лемма доказана.

Упр. 1. Оцените трудоемкость вычисления многочлена от матрицы с известными жордановой нормальной формой и матрицей перехода к ней.

Важный момент — подход через жорданову нормальную форму позволяет определить *значение произвольной функции от матрицы*, если определены значения функции и достаточного количества ее производных в точках спектра матрицы — нужно использовать схему, описанную для многочлена. Это определение не совсем корректно, результат вычислений может зависеть от выбора жордановой нормальной формы и матрицы перехода к ней, которые задаются исходной матрицей неоднозначно.

Теорема 1. Пусть h_j — максимальный порядок жордановой клетки $J_h(\lambda_j)$ из жордановой формы J матрицы A и пусть $h = \sum_{j=1}^{s} h_j$ — степень минимального аннулирующего многочлена матрицы A. Если функция f определена на спектре матрицы A, то существует единственный многочлен $p(\lambda) = p_A(\lambda)$ степени $< h$, для которого

$$p^{(k)}(\lambda_j) = f^{(k)}(\lambda_j), \quad j = 1, \ldots, s, \quad k = 0, \ldots, h_j - 1.$$

В частности, $p(A) = f(A)$ и значение $f(A)$ *не зависит* от порядка жордановых клеток в жордановой нормальной форме матрицы A и матрицы перехода к ней.

Замечание. Значения $f^{(k)}(\lambda_j)$ в системе уравнений для $p(\lambda)$ из теоремы могут быть произвольными.

Многочлен $p_A(\lambda)$ называется *интерполяционным многочленом Лагранжа-Сильвестра* для функции f на спектре матрицы A.

Доказательство. Пусть V — пространство многочленов степени $< h$ и $W = K^h$. Отображение $\varphi : V \to W$ сопоставляет многочлену $p(\lambda)$ набор значений его производных $(p^{(k)}(\lambda_j) \mid j = 1, \ldots, s,\ k = 0, \ldots, h_j - 1)$ в определенном порядке. Это отображение линейно. Если $p(\lambda) \in \operatorname{Ker}\varphi$, то ввиду формулы Тейлора многочлен $p(\lambda)$ делится $(\lambda - \lambda_j)^{h_j}$. Тогда по свойствам взаимно простых многочленов $p(\lambda)$ делится на произведение всех многочленов $(\lambda - \lambda_j)^{h_j}$, имеющее степень $h = \sum_{j=1}^{s} h_j$. Поскольку степень $p(\lambda)$ меньше h, то $p(\lambda) = 0$.

Следовательно, $\operatorname{Ker}\varphi = 0$ и отображение φ — вложение, поэтому $\dim \operatorname{Im}\varphi = \dim V = h$. С другой стороны, $\operatorname{Im}\varphi \leqslant W$, $\dim W = h$. Поэтому $\operatorname{Im}\varphi = W$. Итак, φ — взаимно однозначное соответствие между V и W. Это доказывает теорему.

Пример 1. Для любой комплексной матрицы корректно определяется экспонента, косинус и синус, обычные и гиперболические, поскольку эти функции определены и бесконечно дифференцируемы в любой точке из $\mathbb{C}$.

Сохраняются ли при этом тождества для экспоненты суммы аргументов или косинуса суммы аргументов? Известно, что эти тождества определяются свойствами степенных рядов, в которые разлагаются функции. Надо изучить вопрос о сходимости степенных рядов на матрице. Для этого можно использовать либо матричные нормы, либо поточечную сходимость функций. Напомним, что последовательность функций $f_k(x)$ в области определения X с комплексными значениями сходится к функции $f(x)$ поточечно, если $\lim_{k\to\infty} |f_k(x) - f(x)| = 0$ для всех $x \in X$. Это определение можно распространить на комплексные матрицы порядка n, поскольку такая матрица может рассматриваться как комплексная функция на множестве пар индексов $X = \{(i, j) \mid 1 \leqslant i,\ j \leqslant n\}$.

Если $f(\lambda) = \sum_{k=0}^{\infty} c_k\lambda^k$ — степенной ряд, то он сходится на матрице A, если сходится последовательность матриц $f_N(A)$, где $f_N(\lambda) = \sum_{k=0}^{N} c_k\lambda^k$ — частичная сумма ряда. Предельная матрица обозначается через $f(A) = \sum_{k=0}^{\infty} c_k A^k$ и называется суммой матричного ряда.

Теорема 2. Степенной ряд сходится на матрице A, если он и его формальные производные до порядка $h_j - 1$ сходятся во всех точках λ_j из спектра матрицы A. Здесь h_j — максимальный порядок жордановых клеток $J_h(\lambda_j)$ из жордановой нормальной формы матрицы A.

Доказательство легко следует из формул

$$J = C^{-1}A \Longrightarrow f_N(J) = C^{-1}f_N(A)C, \quad f_N(A) = Cf_N(J)C^{-1},$$
$$J = Diag(J_1, \ldots, J_m) \Longrightarrow f_N(A) = Diag(f_N(J_1), \ldots, f_N(J_m))$$

и леммы.

Следствие. Если формальные ряды сходятся на матрице и удовлетворяют формальному тождеству, то суммы рядов тоже удовлетворяют этому тождеству.

Упр. 2. Проверить следующие тождества
1) $\cos^2 A + \sin^2 A = E$,
2) $\exp(A + B) = \exp A \exp B$, если $AB = BA$,
3) $\cos(A + B) = \cos A \cos B - \sin A \sin B$, если $AB = BA$.

Важную роль в матричном анализе, теории групп Ли и алгебр Ли играет матричная экспонента.

Пример 2. Для вектор-функции $x(t)$ и постоянной матрицы A найдем решение уравнения $dx/dt = Ax$ при условии $x(0) = x_0$.

Допустим, что вектор-функция $x(t)$ разлагается в нуле в ряд

$$x(t) = x_0 + x_1 t + x_2 t^2 + \ldots = \sum_{k\geqslant 0} t^k x_k, \quad x_k \in \mathbb{C}^n.$$

Игнорируя сходимость, продифференцируем ряд:

$$\frac{dx}{dt} = \sum_{k\geqslant 1} kt^{k-1}x_k = Ax(t) = \sum_{k\geqslant 0} t^k Ax_k.$$

Сравнивая коэффициенты при t^{k-1}, получаем

$$kx_k = Ax_{k-1}, \quad x_k = \frac{1}{k}Ax_{k-1} = \frac{1}{k(k-1)}A^2x_{k-2} = \ldots = \frac{1}{k!}A^kx_0.$$

Отсюда

$$x(t) = \sum_{k\geqslant 0} t^k x_k = \sum_{k\geqslant 0} \frac{1}{k!}t^kA^kx_0 = (\exp(tA))x_0.$$

Теперь легко доказать, что это действительно решение уравнения:

$$\frac{d}{dt}(\exp(tA))x_0 = \lim_{h\to 0}\frac{\exp((h+t)A) - \exp(tA)}{h}x_0 = \lim_{h\to 0}\frac{\exp(hA) - E}{h}(\exp(tA))x_0 =$$

$$= A\ (\exp(tA))x_0, \quad (\exp(tA))x_0|_{t=0} = Ex_0 = x_0.$$

Отсюда следует, что траектория $x(t)$ устойчива (т. е. имеет конечный предел) при $t \to +\infty$, если спектр матрицы A расположен в левой полуплоскости $Re\, z < 0$ плоскости комплексных чисел $\mathbb{C}$. Действительно, если $\lambda = a + bi$ — собственное значение матрицы A, где $a,\ b \in \mathbb{R}$, то матрица $\exp(tA)$ имеет собственные значения вида

$$e^{t(a+bi)} = e^{ta}(\cos\,(tb) + i\sin\,(tb)).$$

Если $a < 0$ для всех собственных значений, то предел при $t \to +\infty$ существует. Если $a = 0$ для всех собственных значений, то траектория $x(t)$ ограничена при $b \neq 0$, более того, в диагонализируемом случае она может рассматриваться как траектория в декартовом произведении n окружностей, т. е. n-мерном торе. Если $a > 0$ хотя бы для одного собственного значения, то траектория может уходить на бесконечность в зависимости от выбора начальных данных. Если $a = b = 0$, то соответствующая направлению собственного вектора координата постоянна.

Пример 3. Используя экспоненту, решим систему уравнений $dx/dt = Ax$, где

$$A = \begin{pmatrix} -1 & 0 & 1 \\ -3 & 2 & 7 \\ 1 & -1 & -3 \end{pmatrix}.$$

при начальном условии $x(0) = x_0 = (1, 2, -2)^\top$.

Эту систему мы уже решали с помощью перехода к жордановой форме матрицы системы.

Имеем $|A - \lambda E| = -\lambda(\lambda+1)^2$, $\mathrm{Sp}\, A = \{0, -1, -1\}$. Поэтому $\mathrm{Sp}\,(tA) = \{0, -t, -t\}$.

Многочлен Лагранжа-Сильвестра ищем в форме $p(\lambda) = a + b\lambda + c\lambda^2$. Система уравнений из теоремы 1 для коэффициентов $p(\lambda)$ от экспоненты на матрице tA имеет вид

$$p(0) = e^0, \quad p(-t) = e^{-t}, \quad p'(-t) = e^{-t}.$$

Тогда получаем систему линейных уравнений

$$a = 1,\ a - bt + ct^2 = e^{-t},\quad b - 2ct = e^{-t}.$$

Прибавим ко второму уравнению третье, умноженное на t, и найдем c, а потом b. В итоге получим

$$a = 1,\quad c = (1 - (t+1)e^{-t})/t^2,\quad b = e^{-t} + (2(1 - (t+1)e^{-t})/t).$$

Решение системы дифференциальных уравнений имеет вид

$$p(tA)x_0 = (E + (2 - (t+2)e^{-t})A + (1 - (t+1)e^{-t})A^2)x_0 =$$

$$= (E + A)^2 x_0 - e^{-t}[(t+2)A + (t+1)A^2]x_0.$$

По свойствам экспоненты при $t \to \infty$ оно сходится к $(E + A)^2 x_0 = (3,\ -6,\ 3)^\top$. Окончательный подсчет дает решение в форме

$$x(t) = (3,\ -6,\ 3)^\top + e^{-t}(2,\ -8,\ -1)^\top + te^{-t}(4,\ 3,\ -4)^\top.$$

Пример 4. Вычислим матрицу $\cos A$, если

$$A = \begin{pmatrix} \pi/2 & 1 & 1 \\ 0 & \pi/2 & 1 \\ 0 & 0 & \pi/2 \end{pmatrix}.$$

Удобно разложить косинус в ряд Тейлора в точке $\pi/2$:

$$\cos\lambda = \sum_{k\geqslant 0} \frac{\cos^{(k)}(\pi/2)}{k!}(\lambda - (\pi/2))^k = \sum_{l\geqslant 0} (-1)^l \frac{1}{2l+1)!}(\lambda - (\pi/2))^{2l+1}.$$

Тогда

$$\cos A = \sum_{l\geqslant 0} (-1)^l \frac{1}{(2l+1)!}(A - (\pi/2)E))^{2l+1}.$$

Поскольку по теореме Гамильтона-Кэли $(A - (\pi/2)E)^3 = 0$, то

$$\cos\ A = -(A - (\pi/2)E)) = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Проверим результат вычисления, используя многочлен Лагранжа-Сильвестра.

В данном случае жорданова нормальная форма является жордановой клеткой, поэтому многочлен Лагранжа-Сильвестра имеет степень $\leqslant 2$.

Пусть $p(\lambda) = a + b\lambda + c\lambda^2$. Система уравнений из теоремы 1 имеет вид

$$\left\{ \begin{array}{ccccccc} p(\frac{\pi}{2}) & = & a + b\frac{\pi}{2} + c\frac{\pi^2}{4} & = & \cos\frac{\pi}{2} & = & 0, \\ p'(\frac{\pi}{2}) & = & b + c\pi & = & -\sin\frac{\pi}{2} & = & -1, \\ p''(\frac{\pi}{2}) & = & 2c & = & -\cos\frac{\pi}{2} & = & 0. \end{array} \right.$$

Отсюда $c = 0,\ b = -1,\ a = \pi/2$. Следовательно, $p(\lambda) = (\pi/2) - \lambda$ и $\cos\ A = (\pi/2)E - A$.

Сравним результат с вычислением по определению функции от матрицы.

Если обозначить $N = A - (\pi/2)E$, то N нильпотентна, последовательность векторов

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \xrightarrow{N} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \xrightarrow{N} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

образует ниль-базис. Отсюда $J = C^{-1}AC$, где

$$J = \begin{pmatrix} \pi/2 & 1 & 0 \\ 0 & \pi/2 & 1 \\ 0 & 0 & \pi/2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

По лемме находим

$$\cos J = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} = (\pi/2)E - J.$$

Отсюда $\cos A = C((\pi/2)E - J)C^{-1} = (\pi/2)E - A$.

Упр. 3. Докажите, что любая матрица A порядка n над полем K допускает единственное «аддитивное разложение Жордана» $A = P + N$, где $PN = NP$, N нильпотентна, P полупроста (диагонализируема над подходящим расширением поля K). Более того, P и N являются многочленами от A.