КОНЕЧНОМЕРНЫЕ ЛИНЕЙНЫЕ ОПЕРАТОРЫ

(задача о подобии)

Чуркин В.А.

ПРЕДИСЛОВИЕ

Конечномерный линейный оператор однозначно определяется своей матрицей в некотором базисе векторного пространства. Таким образом, в n-мерном пространстве он задается всего лишь n^2 числами. Это позволяет подключить к изучению конечномерных операторов методы теории чисел, алгебры и анализа. С другой стороны, матрица оператора почти всегда существенно зависит от выбора базиса — как известно, при смене базиса она заменяется на подобную матрицу. При случайном выборе базиса матрица оператора может быть "сложной". Как подобрать базис так, чтобы матрица оператора стала "простой" настолько, что прояснилась бы "механика" действия оператора на пространстве? Как распознать, задают ли две данные матрицы один и тот же оператор (в разных базисах) или эти операторы существенно различны? Расширяя при необходимости поле скаляров, можно считать, что все корни характеристического многочлена лежат в поле скаляров. Если все они различны, т.е. нет кратных корней, то существует базис пространства, состоящий из собственных векторов оператора. Разбору этой важной и, главное, типичной ситуации посвящены начальные параграфы этой методички. Случай кратных корней — достаточно редкая ситуация — разбирается далее и составляет содержание теоремы Жордана.

Сразу отметим трудности при практической реализации. Во-первых, задача классификации разбивается на две подзадачи. Сначала надо найти корни характеристического многочлена — и здесь нет общего удовлетворительного решения. Во вторых, надо построить базис из собственных векторов (или жорданов базис) пространства относительно оператора и жорданову форму для матрицы оператора — эта задача имеет вполне ясное алгоритмическое решение, но *только при точной арифметике в поле скаляров*, т.е. если мы умеем складывать и умножать элементы поля точно, а не приближенно как часто бывает на практике.

При решении задачи о подобии теорема Фробениуса иногда позволяет обойти задачу отыскания корней многочлена, заменяя ее задачей разложения многочлена на неразложимые множители над данным полем скаляров. При неразрешимой первой задаче вторая может быть алгоритмически разрешима или может решаться проще. Например, такова ситуация над важным полем рациональных чисел или над конечными полями. Мы объясним здесь как после разложения на множители можно построить фробениусов базис пространства и фробениусову форму матрицы оператора.

Задача приведения матриц к удобной для работы форме с помощью замены координат возникает кроме алгебры во многих областях — в геометрии, анализе, теории дифференциальных уравнений. В конце работы на примерах показаны типичные применения классификации операторов по Жордану и Фробениусу.

Несколько слов об отличии нашего подхода от других. В настоящее время известны различные доказательства теоремы Жордана. На их основе выработаны разные способы практического решения задачи о подобии. Наиболее громоздкие опираются на теорию λ-матриц (см., например, [3], гл. 13, [4], гл. 4) или связаны с решением больших систем линейных уравнений ([2], с. 71). Часто употребим геометрический метод, основанный на вычислении ядер операторов ([6], задача 1529, [8], с. 333-341) или на вычислении образов ([9], [5], с. 153-162). Разработаны и двухэтапные способы, когда сначала стремятся триангулировать матрицу, а потом разными приемами привести к жордановой форме уже треугольную матрицу ([1], [10], с. 148-158). В этой работе описывается новый способ отыскания жорданова базиса и жордановой формы для линейного оператора, основанный на параллельном вычислении ядер и образов с помощью элементарных преобразований, а также на широком использовании инвариантных подпространств, позволяющих понизить размерность задачи. Мы можем утверждать, что он существенно проще указанных выше, легче усваивается и программируется. Аналогичный подход описывается и для теоремы Фробениуса.

В дальнейшем предполагаются известными некоторые простейшие сведения о матрицах и операторах: вычисление ранга матрицы и базиса линейной оболочки её векторов-столбцов посредством приведения матрицы к ступенчатому виду элементарными преобразованиями столбцов, понятия матрицы оператора, связь матриц оператора в разных базисах, определения характеристического многочлена, характеристических корней, собственных чисел и собственных векторов оператора, инвариантного подпространства и сужения оператора на нем, прямой суммы подпространств.

ОБРАЗ И ЯДРО

Пусть A — линейный оператор векторного пространства V над полем K. Множества $\operatorname{Ker} A = \{v \in V \,|\, Av = 0\}$ и $\operatorname{Im} A = AV = \{Av \,|\, v \in V\}$ и называются соответственно *ядром* и *образом* оператора A. Легко проверить, что ядро и образ — подпространства из V и потому каждое из них вполне задается выбором базиса. Размерности ядра и образа называются соответственно *дефектом* и *рангом* оператора A и обозначаются $\operatorname{df} A$ и $\operatorname{rk} A$.

Теорема 1. Сумма дефекта и ранга линейного оператора равна размерности пространства.

Доказательство. Пусть $u_1, \ldots, u_d$ — базис ядра, $w_1 = Av_1, \ldots, w_r = Av_r$ — базис образа оператора 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 оператором 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 Av из $\operatorname{Im} A$, $Av = \sum \beta_j w_j$ при некоторых β_j из K и $u = v - \sum \beta_j v_j \in \operatorname{Ker} A$. Поэтому $u = \sum \alpha_i u_i$ и, следовательно, $v = \sum \alpha_i u_i + \sum \beta_j v_j$. Теорема доказана.

Укажем способ *одновременного вычисления базисов ядра и образа* из доказательства теоремы 1. Пусть линейный оператор A векторного пространства V над полем K задан в базисе $e_1, \ldots, e_n$ матрицей $A = (a_{ij})$, т. е. $Ae_j = \sum_i a_{ij}\, e_i,\ j = 1, \ldots, n$. Можно считать, отождествляя векторы с их координатными столбцами, что $V = K^n$, $e_1, \ldots, e_n$ — столбцы единичной матрицы I и что $A : x \mapsto Ax,\ x \in K^n$. Тогда полная информация об операторе задается матрицей

$$\begin{pmatrix} e_1 & \ldots & e_n \\ Ae_1 & \ldots & Ae_n \end{pmatrix} = \begin{pmatrix} I \\ A \end{pmatrix}.$$

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

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

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

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum a_{1j}x_j \\ \vdots \\ \sum a_{nj}x_j \end{pmatrix} = \sum \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} x_j = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

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 из соответствующих координат вектора x.

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

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

Вернемся к задаче одновременного вычисления ядра и образа. Приведем элементарными преобразованиями слоев матрицу $\begin{pmatrix} I \\ A \end{pmatrix}$ к виду $\begin{pmatrix} B \\ C \end{pmatrix}$, где B и C — матрицы порядка n, причем C ступенчатая по столбцам.

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

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

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

$$AV = A\langle u_1, \dots, u_d,\ v_1, \dots, v_r \rangle = \langle Au_1, \dots, Au_d,\ Av_1, \dots, Av_r \rangle = \langle w_1, \dots, w_r \rangle.$$

Но система векторов $w_1, \dots, w_r$ линейно независима ввиду ступенчатости матрицы C. Следовательно, она образует базис образа. С другой стороны, система $u_1, \dots, u_d$ 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 1.

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

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

Тогда

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

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

СОБСТВЕННЫЕ ВЕКТОРЫ И СОБСТВЕННЫЕ ЗНАЧЕНИЯ

Задача отыскания собственных векторов и собственных значений линейного оператора — одна из основных задач линейной алгебры. Дело в том, что характеристический многочлен наугад выбранной матрицы как правило не имеет кратных корней — условие равенства нулю дискриминанта обычно не выполняется. Если еще предположить, что все корни содержатся в поле скаляров, то тогда для каждого из них найдется свой собственный вектор. Поскольку собственные векторы с разными собственными значениями линейно независимы, то пространство имеет базис, состоящий из собственных векторов оператора. Такой оператор называется *диагонализируемым* или оператором *простой структуры*. Для него есть хорошее геометрическое описание, можно найти все его инвариантные подпространства, прогнозировать поведение его итераций, вычислять скалярные функции от него и т. д.

Предположим, что нам известны все собственные значения. Покажем, как найти собственные векторы, используя инвариантность и одновременный поиск ядра и образа. Отметим сначала, что подпространство собственных векторов с собственным значением λ для линейного оператора A пространства V совпадает с ядром оператора $A - \lambda I$. Важно, что собственные векторы с другими собственными значениями содержатся в образе оператора $A - \lambda I$. Действительно, равенство $A u = \lambda u$ равносильно равенству $(A-\lambda I)u = 0$. Если же $Av = \lambda' v$, $\lambda' \neq \lambda$, то $Av - \lambda v = (\lambda' - \lambda)v$, $v = (A-\lambda I)((1/(\lambda'-\lambda))v)$. Поскольку подпространство $W = (A - \lambda I)V$ инвариантно относительно операторов A и $A - \lambda' I$, то можно искать собственные векторы, отвечающие другому собственному значению λ', как ядро сужения оператора $A - \lambda' I$ на W. При этом собственные векторы с третьим собственным значением λ'' содержатся в подпространстве $X = (A - \lambda' I)W$, их можно снова искать как ядро сужения $A - \lambda'' I$ на X, и так далее.

Пример. Пусть $\mathbb{R}^2$ — плоскость вещественных столбцов высоты 2 и $A : x \mapsto Ax$, $x \in \mathbb{R}^2$, — линейный оператор умножения столбца на матрицу

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

Тогда характеристический многочлен $\chi(t) = (-t)^2 + \operatorname{tr} A\,(-t) + \det A = t^2 - 6t + 5 =$

$(t-5)(t-1)$. Теперь найдем собственные векторы:

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

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

Теперь легко увидеть, что A растягивает плоскость в 5 раз вдоль каждой прямой, параллельной прямой $\mathbb{R}f_1$, от 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. Заменим матрицу в предыдущем примере на матрицу

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

Тогда характеристический многочлен $\chi(t) = (-t)^2 + \operatorname{tr} A\,(-t) + \det A = t^2 - (10/3)t + 1 = (t-3)(t-(1/3))$. Найдем собственные векторы:

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

Теперь собственные прямые с собственными значениями 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 первой координатной прямой. Как действует оператор на точки координатных прямых — совершенно очевидно.

Пример. 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 N 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

$$\left\{\begin{array}{rcrcrcr} x_1 & = & x_0/3 & + & y_0/3 & & \\ y_1 & = & 2x_0/3 & + & & + & 2z_0/3 \\ z_1 & = & & & 2y_0/3 & + & z_0/3. \end{array}\right.$$

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 $u_k = A^k u_0$. По сути дела нам нужно найти $\lim_{k\to\infty} A^k u_0$. 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 $A^k u_0 = \sum \alpha_i \lambda_i^k v_i$ 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 $\chi(t) = -t^3 + 2t^2/3 + 5t/9 - 2/9 = -(t-1)(t+(2/3))(t-(1/3))$. Отсюда заключаем, что существует базис $\mathbb{R}^3$, состоящий из собственных векторов. Найдем его. Имеем

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

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.

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

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, $\lim_{k\to\infty} A^k u_0 = \alpha_1 v_1 = (\alpha_1, 2\alpha_1, 2\alpha_1)$, 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 N всех машин в процессе не менялось, поэтому и в пределе $\alpha_1 + 2\alpha_1 + 2\alpha_1 = N$.

Тогда $\alpha_1 = N/5$ и, следовательно, предельное распределение грузовых машин по городам X, Y, Z имеет вид $\lim_{k\to\infty} A^k u_0 = (N/5, 2N/5, 2N/5)$. Теперь ясно, что векторы v_2, v_3 можно было и не искать, — итоговый ответ определяется неподвижным вектором v_1, общим числом машин N и даже не зависит от исходного распределения машин по городам.

НИЛЬПОТЕНТНЫЕ ОПЕРАТОРЫ

Линейный оператор N векторного пространства V над полем K называется *нильпотентным*, если $N^q = O$ для некоторого натурального числа q. Такой оператор может быть ненулевым, как показывает пример оператора дифференцирования на пространстве многочленов ограниченной степени.

Отметим, что все собственные значения нильпотентного оператора — нулевые:

$$N^q = O,\ Nu = \lambda u,\ u \neq 0 \Longrightarrow N^2 u = \lambda^2 u, \ldots, N^q u = \lambda^q u = 0,\ \lambda^q = 0,\ \lambda = 0.$$

Теорема. Ненулевой нильпотентный оператор недиагонализируем.

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

Покажем, что относительно нильпотентного оператора конечномерное векторное пространство распадается в прямую сумму инвариантных подпространств, на каждом из которых оператор действует подобно оператору дифференцирования на пространстве многочленов ограниченной степени от одной переменной. Это описание позволит классифицировать нильпотентные операторы с точностью до подобия.

Последовательность векторов $v,\ Nv,\ N^2v, \ldots, N^{h-1}v$ назовем *ниль-слоем высоты h с началом v* относительно линейного оператора N, если $N^h v = 0$. Отметим, что, возможно $N^k v = 0$ при $k < h$. Таблица, столбцы которой — ниль-слои относительно N с общей нижней горизонтальной границей, назовем *ниль-таблицей* относительно N. Таким образом, элементы ниль-таблицы — это векторы; верхний край таблицы, содержащий начала ниль-слоёв, может быть "рваным".

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 нижняя строка ниль-таблицы — линейно независимая система векторов, то и вся система векторов ниль-таблицы линейно независима.

Доказательство. Допустим, что $\sum \lambda_{ij} N^i v_j = 0$ и $\lambda_{pq} N^p v_q$ — "самое верхнее" ненулевое слагаемое. Если r — число строк ниль-таблицы ниже $N^p v_q$, то применим к этой комбинации оператор N^r. Получится нулевая комбинация векторов последней строки ниль-таблицы с ненулевым коэффициентом λ_{pq}. Противоречие. Лемма доказана.

Отметим, что обратное утверждение к утверждению леммы 2 очевидно.

Базис пространства назовем *жордановым* относительно нильпотентного оператора N, если он является раздельным объединением ниль-слоев относительно N. Другими словами, его можно записать в виде ниль-таблицы.

Теорема 2. Конечномерное векторное пространство V относительно любого нильпотентного оператора 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 $r_k = \dim N^k V$ и потому зависит только от пространства V и оператора N.

Доказательство. Составим ниль-таблицу из ниль-слоёв с началами из некоторого базиса пространства и элементарными преобразованиями перестроим её в базис пространства. Если векторы нижней строки таблицы линейно зависимы, то среди них найдется вектор, который выражается линейно через векторы, принадлежащие, вообще говоря, более длинным 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 независимой нижней строкой. По лемме 2 все векторы 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 раздельным объединением ниль-слоёв, т.е. жорданов базис.

Пусть дан произвольный жорданов базис. Тогда легко указать базисы подпространств NV, $N^2V, \ldots$ и подсчитать их размерности. Действительно, образ пространства — линейная оболочка образа базиса, но образ жорданова базиса получается сдвигом векторов соответствующей ниль-таблицы вниз на один шаг, кроме векторов нижней строки, переходящих в нуль. Поскольку ниль-слои под действием оператора сокращают свою длину на единицу, то

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

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

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

порядка h. Этот вид однозначен с точностью до порядка клеток J_h по диагонали. Характеристический многочлен нильпотентного оператора равен $\pm t^n$, 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 нижнего 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.

Упражнение. По данному жорданову базису найдите базисы ядер степеней нильпотентного оператора.

Предположим, что в пространстве столбцов $V = K^n$ над полем K задан линейный оператор $N : x \mapsto Nx$ умножения столбцов на нильпотентную матрицу N. Следующий *алгоритм* выстраивает жорданов базис V относительно N, моделируя доказательство теоремы с одним изменением — не следует вычислять образы всех базисных векторов сразу.

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

1. Вычислить ниль-слой с началом e_1. Пусть

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

и $N^h e_1 = 0$. 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 с началом e_j, которое ранее не использовалось. Перейти к шагу 4.

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

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

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

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

Жорданов базис V относительно N — это

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

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

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

Жорданов базис V относительно 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 $D : f \mapsto \partial f/\partial x + \partial f/\partial y$ — понижающий общую степень, а потому нильпотентный линейный оператор на пространстве V.

Преобразуем ниль-таблицу с началами $x^2,\ xy,\ y^2$ относительно $\mathcal{D}$ (очевидно, начальных векторов $x^2,\ xy$ недостаточно, чтобы линейная оболочка векторов таблицы совпадала с 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 D, её последняя строка — базис ядра 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 $a,\ b,\ c \in \mathbb{R}$.

ЯДЕРНО-ОБРАЗНОЕ И КОРНЕВОЕ РАЗЛОЖЕНИЕ

Напомним, что подпространство U векторного пространства V называется *инвариантным* относительно линейного оператора A пространства V, если $A\,U \subseteq U$. Нулевое подпространство и все пространство всегда инвариантны — они называются тривиальными. Если существует нетривиальное A-инвариантное подпространство U, то оператор сводится в некотором роде к комбинации более простых операторов на пространствах меньшей размерности U и V/U. Совсем хорошо, когда пространство является прямой суммой ненулевых инвариантных подпространств:

$$V = V_1 \oplus \ldots \oplus V_s, \quad A\,V_i \subset V_i.$$

Если $v = v_1 + \ldots + v_s,\ v_i \in V_i$, то $A\,v = A\,v_1 + \ldots + A\,v_s,\ A\,v_i \in V_i$. Ввиду однозначности разложения в сумму такого вида достаточно знать как действует сужение A на каждом V_i, чтобы восстановить действие оператора на всем пространстве. Наша цель состоит в том, чтобы получить такого типа разложения относительно произвольного линейного оператора, если это вообще возможно.

Пусть A — линейный оператор векторного пространства V на полем K. Вектор v из V называется *корневым вектором высоты* h, отвечающим собственному значению λ оператора A, если

$$(A - \lambda I)^h v = 0, \quad (A - \lambda I)^{h-1} v \neq 0.$$

Корневые векторы высоты 1 — это просто собственные векторы. Ясно, что

$$0 < \mathrm{Ker}\,(A - \lambda I) \leqslant \mathrm{Ker}\,(A - \lambda I)^2 \leqslant \mathrm{Ker}\,(A - \lambda I)^3 \leqslant \ldots\ .$$

Подпространство

$$V^\lambda(A) := \bigcup_{h>0} \mathrm{Ker}\,(A - \lambda I)^h$$

называется *корневым подпространством* пространства V, отвечающим собственному значению λ оператора A. Оно состоит из нуля и *всех* корневых векторов, отвечающих собственному значению λ.

Теорема *(о корневом разложении)*. Пусть A — линейный оператор ненулевого конечномерного векторного пространства V над полем K и пусть $\operatorname{Sp} A \subseteq K$. Тогда

1) высота корневого вектора, отвечающего собственному значению λ оператора A, не превосходит числа $h(\lambda) = \min\{h \mid (A-\lambda I)^h V = (A-\lambda I)^{h+1} V\}$, в частности, $V^\lambda(A) = \operatorname{Ker}(A-\lambda I)^{h(\lambda)}$;

2) пространство V является прямой суммой своих корневых подпространств $V^\lambda(A)$, $\lambda \in \operatorname{Sp} A$, причем все слагаемые инвариантны относительно оператора A.

Доказательство начнем с лемм.

Лемма 1. Если два линейных оператора перестановочны, то ядро и образ одного оператора всегда инвариантны относительно второго оператора:

$$AB = BA \Longrightarrow A(\operatorname{Ker} B) \subseteq \operatorname{Ker} B, \quad A(\operatorname{Im} B) \subseteq \operatorname{Im} B.$$

В частности, это верно, когда один оператор является линейной комбинацией степеней второго, т. е. многочленом или рациональной функцией от второго оператора.

Доказательство почти очевидно.

1) Если $v \in \operatorname{Ker} B$, то $B(Av) = A(Bv) = A0 = 0$, $Av \in \operatorname{Ker} B$.

2) Если $Bv \in \operatorname{Im} B$, то $A(Bv) = B(Av) \in \operatorname{Im} B$ для всех $v \in V$.

Лемма 2 *(о ядерно-образном разложении)*. Пусть A — линейный оператор векторного пространства V. Тогда $V = \operatorname{Ker} A \oplus \operatorname{Im} A$ в том и только том случае, когда $A^2V = AV$.

Доказательство. Отметим, что слагаемые этого разложения A-инвариантны по лемме 1. Если $V = \operatorname{Ker} A \oplus \operatorname{Im} A$, то $AV = A(\operatorname{Ker} A) \oplus A(\operatorname{Im} A) = A(AV) = A^2V$. Если $A(AV) = AV$, то сужение A на подпространстве AV — невырожденный оператор, откуда $\operatorname{Ker} A \cap \operatorname{Im} A = \operatorname{Ker} A \cap AV = \operatorname{Ker} A|_{AV} = 0$. Ввиду теоремы о сумме дефекта и ранга оператора получаем $V = \operatorname{Ker} A \oplus \operatorname{Im} A$.

Лемма 3. Пусть A — линейный оператор векторного пространства $V = U \oplus W$ и пусть слагаемые A-инвариантны. Обозначим B и C — сужения A на U и W соответственно. Тогда

$$\chi_A(t) = \chi_B(t) \cdot \chi_C(t), \quad \operatorname{Sp} A = \operatorname{Sp} B \cup \operatorname{Sp} C.$$

Доказательство. Пусть $e_1, \ldots, e_k$ — базис U, $e_{k+1}, \ldots, e_n$ — базис W. Тогда $e_1, \ldots, e_k$, $e_{k+1}, \ldots, e_n$ — базис V. Пусть B_e, C_e, A_e — матрицы операторов B, C, A в соответствующих базисах. Ввиду инвариантности подпространств матрица A_e клеточно диагональна с клетками B_e и C_e. Поэтому $\det(A_e - tI) = \det(B_e - tI) \cdot \det(C_e - tI)$. Это означает, что спектр оператора A является объединением спектров B и C.

Упр. Докажите лемму 3 без использования матриц.

Теперь докажем теорему о корневом разложении.

1) Сначала *ограничим высоту корневых векторов*. Если $\lambda \in \operatorname{Sp} A$, то $A - \lambda I$ — вырожденный оператор и тогда $V > (A-\lambda I)V$. Применим $A - \lambda I$ несколько раз. Ввиду конечномерности

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

где $h = h(\lambda) \geqslant 1$. Первое равенство в этой цепочке влечет за собой последующие с помощью действия $A - \lambda I$. Отсюда

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

по теореме о сумме ранга и дефекта. Тогда объединение $V^\lambda(A)$ этой цепочки подпространств совпадает с $\mathrm{Ker}\,(A-\lambda I)^h$.

2) Покажем *способ отщепления одного корневого подпространства*. Выберем $\lambda \in \mathrm{Sp}\, A$, $h = h(\lambda)$. Поскольку $(A-\lambda I)^h V = (A-\lambda I)^{2h} V$, то по лемме 2

$$V = \mathrm{Ker}\,(A-\lambda I)^h \oplus \mathrm{Im}\,(A-\lambda I)^h.$$

Обозначим слагаемые U и W соответственно и отметим, что они A-инвариантны по лемме 1.

Пусть B — сужение A на U. Тогда оператор $B-\lambda I$ нильпотентен:

$$u \in U = \mathrm{Ker}\,(A-\lambda I)^h \Longrightarrow (B-\lambda I)^h u = (A-\lambda I)^h u = 0.$$

Ввиду нильпотентности $\mathrm{Sp}\,(B-\lambda I) = \{0\}$, $\mathrm{Sp}\, B = \{\lambda\}$, $\chi_B(t) = \pm(t-\lambda)^{k(\lambda)}$, где $k(\lambda) = \dim U \geqslant h(\lambda)$.

Если $W = 0$, то всё доказано. Пусть $W \neq 0$ и пусть C — сужение A на W. По лемме 3 $\chi_A(t) = \chi_B(t)\cdot\chi_C(t) = \pm(t-\lambda)^{k(\lambda)}\cdot\chi_C(t)$, $\mathrm{Sp}\, A = \{\lambda\}\cup\mathrm{Sp}\, C$. Отметим, что $\chi_C(\lambda) \neq 0$, так как $A-\lambda I$ невырожден на W, в частности, кратность λ как корня характеристического многочлена оператора A равна размерности корневого подпространства: $k(\lambda) = \dim U$.

Утверждается, что корневое подпространство из V, отвечающее собственному значению $\lambda' \neq \lambda$ оператора A, содержится в W и потому является корневым для оператора C. В самом деле, пусть $(A-\lambda' I)^h v = 0$, $v = u + w$, $u \in U$, $w \in W$. Тогда $(A-\lambda' I)^h u + (A-\lambda' I)^h w = 0$. Ввиду инвариантности U и W и свойств прямой суммы, получаем $(A-\lambda' I)^h u = (A-\lambda' I)^h w = 0$. Но оператор $A-\lambda' I$ невырожден на подпространстве U, поскольку $\mathrm{Sp}\, A|_U = \{\lambda\}$, $\lambda' \neq \lambda$. Отсюда $u = 0$, $v = w \in W$. Действуя на W оператором $C-\lambda' I = A-\lambda' I$ аналогичным образом отщепим слагаемое $W^{\lambda'}(C) = V^{\lambda'}(A)$ и т.д. Ввиду конечномерности этот процесс оборвется и в итоге получится расщепление пространства V в прямую сумму корневых подпространств оператора A.

Следствие 1 (*о минимальном многочлене*). Многочлен $\mu_A(t) = \prod_{\lambda\in\mathrm{Sp}\, A}(t-\lambda)^{h(\lambda)}$ является приведенным многочленом наименьшей степени, аннулирующим оператор A. Этими свойствами он определяется однозначно и называется *минимальным многочленом* оператора A.

Доказательство. Так как $V^\lambda = \mathrm{Ker}\,(A-\lambda I)^{h(\lambda)}$, то $\mu(A) = 0$ на V^λ, а потому и на $V = \oplus V^\lambda$. Если $f(t)$ — приведенный многочлен наименьшей степени, аннулирующий оператор A, то $\deg f(t) \leqslant \deg \mu(t)$. Разделим с остатком: $\mu(t) = f(t)q(t) + r(t)$, $\deg r(t) < \deg f(t)$. Подставляя вместо t оператор A, получим, что $r(A) = 0$. Это противоречит выбору f, если $r(t)$ — ненулевой многочлен. Следовательно, остаток нулевой и f делит μ, $f(t) = \prod_{\lambda\in\mathrm{Sp}\, A}(t-\lambda)^{m(\lambda)}$, $m(\lambda) \leqslant h(\lambda)$. Пусть v — корневой вектор максимальной высоты $h(\lambda)$. Так как $f(A)v = 0$ и оператор $\prod_{\lambda'\neq\lambda}(A-\lambda' I)^{m(\lambda')}$ действует на V^λ невырожденно, то $(A-\lambda I)^{m(\lambda)} v = 0$, $m(\lambda) \geqslant h(\lambda)$. Следовательно, $f(t) = \mu(t)$.

Следствие 2 (*теорема Гамильтона–Кэли*). Характеристический многочлен делится на минимальный и аннулирует оператор.

Доказательство. Можно считать, что наш оператор задан как оператор умножения столбцов на матрицу над некоторым полем. Поле можно всегда расширить так, что оно содержит спектр матрицы. Условия теоремы о корневом разложении будут выполнены. Ввиду доказательства теоремы о корневом разложении можно считать, что кратность $h(\lambda)$ корня λ минимального многочлена равна максимальной высоте соответствующего корневого вектора, а та не превосходит размерности содержащего его корневого подпространства, которая совпадает в свою очередь с кратностью $k(\lambda)$ этого корня в характеристическом многочлене.

Пример 1. Покажем, что формула $P^2 = P$ характеризует в точности операторы проецирования пространства V на подпространство U параллельно прямому дополнению W к этому подпространству. Действительно, многочлен $f(t) = t^2 - t = (t-1)t$ аннулирует оператор P. Поскольку минимальный многочлен делит любой аннулирующий, то он равен либо $t-1$, либо t, либо $(t-1)t$. Тогда соответственно либо $P = I$, либо $P = 0$, либо $V = \operatorname{Ker}(P-I) \oplus \operatorname{Ker} P$, причем $\operatorname{Ker}(P-I) = \operatorname{Im} P$. Если $v = u + w,\ u \in \operatorname{Ker}(P-I) = \operatorname{Im} P,\ w \in \operatorname{Ker} P$, то $Pv = Pu + Pw = u + 0 = u$. Следовательно, P — проектор V на $U = \operatorname{Ker}(P-I) = \operatorname{Im} P$ параллельно $W = \operatorname{Ker} P$.

Пример 2. Найдем спектр и корневые подпространства оператора $A : x \mapsto Ax,\ x \in V = \mathbb{R}^3$, где

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

Легко найти $\chi_A(t) = -(t-3)(t+1)^2,\ \operatorname{Sp} A = \{3,\ -1\} \subset \mathbb{R}$. Поэтому $V = V^3 \oplus V^{-1}$, $\dim V^3 = 1,\ \dim V^{-1} = 2$. Так как высоты корневых векторов не больше кратности корня в характеристическом многочлене, то $V^3 = \operatorname{Ker}(A-3I),\ V^{-1} = \operatorname{Ker}(A+I)^2$. Мы можем найти базисы этих подпространств, отыскивая фундаментальные системы решений для однородных систем линейных уравнений $(A-3I)x = 0,\ (A+I)^2 x = 0$. Поскольку возведение матрицы в степень — трудоемкая операция, то лучше воспользуемся доказательством теоремы о корневом разложении или леммой 2, согласно которым при отщеплении корневого подпространства $V^3 = \operatorname{Ker}(A-3I) = \operatorname{Ker}(A-3I)^2$ другое корневое подпространство V^{-1} попадает в образ $\operatorname{Im}(A-3I)$, а так как третьего корневого подпространства нет, то и совпадает с этим образом. Применим способ одновременного поиска ядра и образа для оператора $A - 3I$.

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

В итоге вектор $f_1 = (-2, 1, 0)^\top$ образует базис корневого подпространства V^3, а векторы $f_2 = (2, -3, 4)^\top,\ f_3 = (2, -1, 0)^\top$ — базис V^{-1}.

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

Предположим, что в пространстве столбцов $V = K^n$ над полем K задан линейный оператор $A : x \mapsto Ax$ умножения столбцов на матрицу A и что известны его собственные значения. Следующий *алгоритм* вычисляет базисы корневых подпространств. Он опирается на лемму 2 и многократное одновременное вычисление ядер и образов относительно операторов $N = A - \lambda I$.

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 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 $\begin{pmatrix} B_h \\ C_h \end{pmatrix}$, есл и 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 $V^\lambda(A)$, а ненулевые столбцы матрицы C_h, образуют базис пространства $W = \mathrm{Im}\,(A-\lambda I)^{h(\lambda)}$, содержащего остальные корневые подпространства.

Теперь пусть новая матрица B_0 получается из матрицы C_h удалением нулевых столбцов, и пусть $N := A - \lambda' I$, где λ' — другое собственное значение матрицы A. Вернемся к началу алгоритма и продолжим вычисления в соответствии с указанными правилами. Тогда получим базис корневого подпространства $V^{\lambda'}(A)$ и базис пространства $W' = \mathrm{Im}\,(A-\lambda' I)^{h(\lambda')}|_W$, содержащего все корневые подпространства, кроме $V^\lambda(A)$ и $V^{\lambda'}(A)$. 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 3. Пусть $A : x \mapsto Ax,\ x \in V = \mathbb{R}^4$, где

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

Тогда

$$\chi_A(t) = (t-1)^2(t+1)^2, \quad V = \mathrm{Ker}\,(A-I)^2 \oplus \mathrm{Ker}\,(A+I)^2, \quad \mathrm{Ker}\,(A+I)^2 = \mathrm{Im}\,(A-I)^2.$$

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

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

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

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 $\operatorname{Ker}(A-I)^2$, а векторы $f_3 = (3,-5,1,0)^\top$, $f_4 = (2,-3,0,1)^\top$ образуют базис второго корневого подпространства $\operatorname{Ker}(A+I)^2$.

ЖОРДАНОВА КЛАССИФИКАЦИЯ ОПЕРАТОРОВ

Здесь будет дана классификация линейных операторов конечномерных пространств с точностью до подобия в случае, когда поле скаляров содержит спектр оператора. Основная идея — выбрать базис пространства, согласованный с действием оператора, опираясь на корневое разложение пространства и свойства нильпотентных операторов. По традиции эта классификация имеет матричный характер. *Жордановой клеткой порядка h, отвечающей собственному значению* λ называется матрица

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

Матрица жорданова или имеет жорданов вид, если она клеточно-диагональна с жордановыми клетками по диагонали.

Пример. Жордановы матрицы порядка 2 — это

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

Теорема *(Жордан, 1870).* Пусть A — линейный оператор ненулевого конечномерного векторного пространства V над полем K и пусть $\operatorname{Sp} A \subseteq K$. Тогда

1) в подходящем базисе V матрица оператора A имеет жорданов вид J;

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

Доказательство. 1) По теореме о корневом разложении $V=\oplus_{\lambda\in \mathrm{Sp}\,A}V^\lambda$, где $V^\lambda = \mathrm{Ker}\,(A-\lambda I)^{h(\lambda)}$. Сужение $A-\lambda I$ 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 $A-\lambda I$. Но

$$\left\{\begin{array}{rcl}(A-\lambda I)f_1 &=& 0\\ (A-\lambda I)f_2 &=& f_1\\ \cdots & . & .\\ (A-\lambda I)f_h &=& f_{h-1}\end{array}\right. \iff \left\{\begin{array}{rcl}Af_1 &=& \lambda f_1\\ Af_2 &=& f_1+\lambda f_2\\ \ldots & . & \ldots\\ Af_h &=& f_{h-1}+\lambda f_h\end{array}\right. \iff A_f=J_h(\lambda).$$

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

2) Ввиду предыдущего можно отождествить число $s_h(\lambda)$ с числом максимальных ниль-слоев высоты h относительно $A-\lambda I$ 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 $\sum_{\lambda\in\mathrm{Sp}\,A}\dim V^\lambda = \dim V$, 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 $r_k=\dim(A-\lambda I)^kV^\lambda$. Ввиду теоремы о корневом разложении $V=V^\lambda\oplus W$, $(A-\lambda I)W=W$. Поэтому $(A-\lambda I)^kV=(A-\lambda I)^kV^\lambda\oplus W$. Если $d=\dim W$, то отсюда $r_k(\lambda)=r_k+d$, $r_k=r_k(\lambda)-d$, и тогда $s_h(\lambda)=(r_{h-1}(\lambda)-d)-2(r_h(\lambda)-d)+(r_{h+1}(\lambda)-d)=r_{h-1}(\lambda)-2r_h(\lambda)+r_{h+1}(\lambda)$.

Следствие. Всякая квадратная матрица над полем, содержащим её спектр, подобна жордановой. При этом жорданов вид задается исходной матрицей однозначно с точностью до порядка жордановых клеток по диагонали.

Пример. Для матрицы A найдем её жорданов вид J и сопрягающую с ним матрицу C, т. е. такую, 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 $V=\mathbb{R}^4$, $A: x\mapsto Ax$, $x\in\mathbb{R}^4$. Так как $\chi_A(t)=t(t-2)^3$, то можно найти жорданов базис $\mathbb{R}^4$ относительно оператора A. По теореме о корневом разложении $V=V^0\oplus V^2$, $V^0=\mathrm{Ker}\,A$, $\dim V^0=1$, $V^2=\mathrm{Ker}\,(A-2I)^3$, $\dim V^2=3$. Выгоднее искать сначала ниль-слои, связанные с корнем λ большой кратности, так как ранг матрицы $A-\lambda I$ может быть мал. При $\lambda = 2$ получаем

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

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

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

Отсюда $V^2 = \operatorname{Ker}(A-2I)^2$, $V^0 = \operatorname{Ker} A = \operatorname{Im}(A-2I)^2$, 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 $A - 0I$ и $A-2I$ соответственно. Так как $Af_1 = 0$, то для V^0 задача решена. Ниль-слои относительно $A-2I$ 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

$$
\left\{\begin{array}{lcl} Af_1 & = & 0 \\ (A-2I)f_2 & = & 0 \\ (A-2I)f_3 & = & f_2 \\ (A-2I)f_4 & = & 0 \end{array}\right. , \quad \left\{\begin{array}{lcl} Af_1 & = & 0 \\ Af_2 & = & 2f_2 \\ Af_3 & = & f_2+2f_3 \\ Af_4 & = & 2f_4 \end{array}\right. , \; J = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = A_f.
$$

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

$$
C = (f_1 f_2 f_3 f_4) = \begin{pmatrix} 4 & 4 & 1 & -4 \\ 2 & 2 & 1 & -1 \\ 2 & 0 & 0 & 1 \\ -4 & -2 & 0 & 0 \end{pmatrix}, \; J = A_f = C^{-1} A_e C = C^{-1} A C.
$$