

Линейная алгебра

Дима Трушин

Семинар 6

Линейные операторы

Напоминание В этом разделе я наконец-то вам начну рассказывать о самых важных объектах в линейной алгебре – линейных операторах. Пусть V – векторное пространство, тогда линейным оператором на V называется линейное отображение $\varphi: V \rightarrow V$, то есть такое отображение, что $\varphi(v_1 + v_2) = \varphi(v_1) + \varphi(v_2)$ и $\varphi(\lambda v) = \lambda \varphi(v)$. Так как линейный оператор – это частный случай линейного отображения, то для него применимо все, о чем мы уже говорили в случае отображений. Про линейный оператор надо думать как про линейную деформацию пространства V .

Примеры

1. $\text{Id}: V \rightarrow V, v \mapsto v$. Тожественный линейный оператор, ничего не деформирует.
2. $0: V \rightarrow V, v \mapsto 0$. Нулевой линейный оператор, который все отправляет в ноль.
3. Растяжения вдоль осей: $D: \mathbb{R}^n \rightarrow \mathbb{R}^n$, заданный по правилу $x \mapsto Dx$, где

$$D = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

Тогда отображение D растягивает i -ю координату в d_i раз. Если $d_i > 1$, то это растяжение, если $0 < d_i < 1$, то это сжатие, если $-1 < d_i < 0$, то это сжатие и отражение вдоль оси, если $d_i < -1$, то это растяжение и отражение вдоль оси.

4. Поворот на плоскости: $\rho_\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, где вектор x поворачивается на угол α против часовой стрелки. Такое отображение в матричном виде задается так¹

$$\rho_\alpha \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

5. Поворот в пространстве вокруг оси OX : $\rho_{1,\alpha}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$, где вектор x поворачивается вокруг оси OX на угол α против часовой стрелки, если смотреть со стороны вектора $e_1 = (1, 0, 0)^t$ на начало координат. В матричном виде эта штука имеет вид

$$\rho_{1,\alpha} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Теперь давайте объясним некоторую специфику.

¹Строго говоря я еще не рассказывал про то, что такое движение, но этот и следующий пример можно понять и без общей науки, которая будет чуть позже.

Матрица линейного оператора Пусть в векторном пространстве V задан некоторый базис e_1, \dots, e_n и пусть $\varphi: V \rightarrow V$ – линейный оператор. Так как у оператора пространство из которого он бьет и то в которое он бьет совпадают, то мы фиксируем всего лишь один базис (пространство-то у нас одно). Тогда по определению матрица линейного оператора φ – это такая матрица $A_\varphi \in M_n(\mathbb{R})$, что выполнено $\varphi e = e A_\varphi$, где $e = (e_1, \dots, e_n)$.

Пусть теперь у нас задан другой базис e'_1, \dots, e'_n в пространстве V с матрицей перехода $C \in M_n(\mathbb{R})$, то есть $(e'_1, \dots, e'_n) = (e_1, \dots, e_n)C$. Пусть так же $e' = (e'_1, \dots, e'_n)$. Тогда матрица φ в базисе e' пусть будет A'_φ , то есть $\varphi e' = e' A'_\varphi$. В этом случае связь между матрицами следующая $A'_\varphi = C^{-1} A_\varphi C$. То есть матрица A_φ сопряжена матрице A'_φ .²

Замечания

- Отметим, что матрица линейного оператора обязательно квадратная. Таким образом, изучение линейного отображения – это изучение прямоугольной матрицы, а изучение линейного оператора – это всегда изучение только квадратной матрицы.
- Если линейное отображение $\psi: V \rightarrow U$ бьет между двумя разными пространствами одинаковой размерности, то ему тоже соответствует квадратная матрица. Но принципиальная разница с линейным оператором заключается в том, что для линейного отображения мы можем независимо менять базисы в V и U , что соответствует замене $A'_\psi = C^{-1} A_\psi D$, а для линейного оператора, так как пространство одно и то же, базисы меняются одновременно, что соответствует $A'_\varphi = C^{-1} A_\varphi C$.
- Так как линейные операторы – это линейные отображения, то задавать их можно так же как и линейные отображения, например: либо с помощью образа базисных векторов, либо с помощью матрицы в фиксированном базисе.

Например, можно фиксировать векторы $e_1, \dots, e_n \in V$, которые являются базисом и зафиксировать любой набор векторов $u_1, \dots, u_n \in V$, тогда существует единственный линейный оператор $\varphi: V \rightarrow V$ такой, что $\varphi(e_i) = u_i$. Это самый простой и эффективный способ строить новые линейные операторы на пространстве. Но может быть не всегда самый удобный, чтобы потом с такими операторами работать.

Характеристики линейных операторов

След оператора Пусть $A \in M_n(\mathbb{R})$ – квадратная матрица. Тогда *след матрицы* A – это сумма ее диагональных элементов, т.е. $\text{tr } A = a_{11} + \dots + a_{nn} = \sum_{i=1}^n a_{ii}$. Заметим важное свойство следа: $\text{tr}(AB) = \text{tr}(BA)$ (это непосредственная проверка влоб). В частности $\text{tr}(C^{-1}AC) = \text{tr}(A)$ для любых $A, C \in M_n(\mathbb{R})$ с условием, что C обратима.

Пусть теперь $\phi: V \rightarrow V$ – некоторый линейный оператор. Тогда в некотором базисе e_1, \dots, e_n он задается матрицей A . Определим *след линейного оператора* ϕ как след этой матрицы A . Это определение не зависит от базиса. Действительно, в другом базисе оператор ϕ задается матрицей $A' = C^{-1}AC$, тогда $\text{tr}(A') = \text{tr}(C^{-1}AC) = \text{tr}(A)$. След оператора ϕ также обозначается через $\text{tr } \phi$. Важно понимать, что след – это характеристика линейного оператора, а не его матрицы, т.е. эта штука не зависит от матрицы, которой задается оператор. Однако, мы не можем определить эту характеристику не пользуясь базисом. Более того, в принципе невозможно определить след без базиса!

Определитель оператора Пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда в некотором базисе он задается матрицей A . Положим $\det \phi = \det A$. Надо лишь проверить, что это определение не зависит от выбора базиса. Действительно, в другом базисе ϕ задается $C^{-1}AC$, а значит $\det(C^{-1}AC) = \det A$. Величина $\det \phi$ называется *определителем линейного оператора*. Как и в случае следа, определитель линейного оператора не зависит от базиса, но его нельзя определить не пользуясь базисом.

Многочлены от операторов Пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда определен линейный оператор $\lambda\phi$ для любого $\lambda \in \mathbb{R}$. Уточним на всякий случай, что $(\lambda\phi)(v) = \lambda\phi(v)$ по определению. Более того, определены все натуральные степени оператора ϕ , как композиция, т.е. $\phi^n(v) = \phi(\dots\phi(v)\dots)$ – где композиция берется

²Напомним, что квадратные матрицы B и D называются сопряженными, если найдется обратимая матрица C такая, что $D = C^{-1}BC$.

n раз, например, $\phi^3(v) = \phi(\phi(\phi(v)))$.³ Кроме того, линейные операторы можно складывать. Напомню, что $(\phi + \psi)(v) = \phi(v) + \psi(v)$ по определению.

Из выше сказанного следует, что можно составлять выражения вида $\sqrt{\pi}\phi - 2\phi^4$ и их результат будет линейный оператор на V . Более того, можно определить ϕ^0 как тождественный оператор Id , т.е. $\text{Id}(v) = v$.⁴ Тогда можно писать выражения вроде $2/3 + \phi^2$, где имеется в виду $2/3 \text{Id} + \phi^2$.

Таким образом для любого многочлена $p(t) = a_0 + a_1 t + \dots + a_n t^n$, где $a_i \in \mathbb{R}$, и любого линейного оператора $\phi: V \rightarrow V$ определен линейный оператор $p(\phi): V \rightarrow V$. Когда мы перейдем к базисам, оператор ϕ будет соответствовать матрице A . В этом случае $p(\phi)$ соответствует матрице $p(A)$.⁵

Основной бонус от подстановки матриц и линейных операторов в многочлены состоит вот в чем. Пусть многочлен $p(t)$ раскладывается в произведение $(t - \lambda_1) \dots (t - \lambda_n)$. Тогда оператор $p(\phi)$ раскладывается в композицию операторов $(\phi - \lambda_1 \text{Id}) \dots (\phi - \lambda_n \text{Id})$. Скоро (очень очень скоро) будет видно зачем все это нужно.

Характеристический многочлен Пусть $\phi: V \rightarrow V$ – некоторый линейный оператор. Опять же для удобства, можно считать, что после выбора базиса $V = \mathbb{R}^n$ и ϕ соответствует некоторой матрице $A \in M_n(\mathbb{R})$. Тогда выражение $\det(\phi - \lambda \text{Id}) = \det(A - \lambda E)$ является многочленом от λ степени n . Действительно,

$$\det(A - \lambda E) = \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix}$$

С усилием вспоминая явную формулу для определителя через перестановки, понимаем, что получается многочлен от λ . Еще чуть внимательнее присмотревшись к нему, можно заметить, что

$$\det(A - \lambda E) = \det(A) + \dots + (-1)^{n-1} \text{tr } A \lambda^{n-1} + (-1)^n \lambda^n$$

Данный многочлен обозначим через $\chi_\phi(\lambda)$ или $\chi_A(\lambda)$ и будем называть *характеристическим многочленом* оператора ϕ или соответствующей матрицы A (в зависимости от того, о чем идет речь).⁶ Еще полезно видеть перед глазами следующее равенство

$$(-1)^n \chi_A(\lambda) = \lambda^n - \text{tr}(A) \lambda^{n-1} + \dots + (-1)^n \det(A)$$

Утверждение (Теорема Гамильтона-Кэли). Пусть $\phi: V \rightarrow V$ – линейный оператор и $\chi(t)$ – его характеристический многочлен. Тогда $\chi(\phi) = 0$.⁷

Минимальный многочлен Если $\phi: V \rightarrow V$ – линейный оператор на n мерном пространстве, то как мы видели выше, он зануляется своим характеристическим многочленом, т.е. существуют многочлены $p(t)$ такие, что $p(\phi) = 0$. На самом деле, факт существования таких многочленов доказывается проще, чем теорема Гамильтона-Кэли. Наш ϕ соответствует матрице A . Но тогда $E, A, A^2, \dots, A^{n^2}$ не могут быть линейно независимы. А значит $a_0 + a_1 A + \dots + a_{n^2} A^{n^2} = 0$, то есть A зануляется многочленом $p(t) = a_0 + a_1 t + \dots + a_{n^2} t^{n^2}$. Суть теоремы Гамильтона-Кэли в том, чтобы понизить степень многочлена с n^2 до n .

Мы скажем, что ненулевой многочлен $p(t)$ является минимальным для ϕ , если $p(\phi) = 0$ и степень многочлена p является минимально возможной. Минимальный многочлен делит любой многочлен зануляющий ϕ , так как остаток тоже должен занулять ϕ и степени меньше. Потому, если у минимальный многочлен нормировать так, чтобы его старший коэффициент был единицей, то он определен однозначно.

Утверждение. Пусть $\phi: V \rightarrow V$ – линейный оператор. Тогда

1. Минимальный многочлен $p(t)$ для ϕ со старшим коэффициентом определен однозначно.

³На самом деле умные люди вообще не пишут скобки, ибо они только загромождают обозначения. Ведь куда приятнее смотреть на $\phi\phi\phi v$ вместо $\phi(\phi(\phi(v)))$. Но еще приятнее смотреть на $\phi^3 v$.

⁴В любом базисе тождественный оператор соответствует единичной матрице.

⁵Предлагаю вам самим восстановить детали того, что такое подставить матрицу в многочлен по аналогии с операторами.

⁶Надо отметить, что часто характеристическим многочленом называют $\det(\lambda E - A)$, так как в этом случае старший коэффициент по λ становится 1. Наш многочлен от этого отличается на $(-1)^n$. Для многих вопросов это не принципиально.

⁷Есть два доказательства этого факта: (1) кустарное, методами линейной алгебры и (2) концептуальное методами коммутативной алгебры. Первое доказательство использует теорему о Жордановой нормальной форме (по сути классификацию всех линейных операторов) и очень геморройное. Второе доказательство в одну строчку – формулы Крамера для модулей над коммутативными кольцами. Его беда в том, что надо объяснить все дурацкие слова в доказательстве. Это не сложно, но требует кучу времени и усилий, чтобы их осознать.

2. Многочлен $p(t)$ делит любой многочлен аннулирующий ϕ .
3. Существует такое число m , что характеристический многочлен $\chi_\phi(t)$ делит $p^m(t)$.

Лемма о стабилизации

Утверждение. Пусть V – векторное пространство и $\varphi: V \rightarrow V$ – линейный оператор. Тогда

1. Найдется такое число $0 \leq k \leq \dim V$, что

$$0 \subsetneq \ker \varphi \subsetneq \ker \varphi^2 \subsetneq \dots \subsetneq \ker \varphi^k = \ker \varphi^{k+1} = \dots$$

2. Найдется такое число $0 \leq k \leq \dim V$, что

$$V \supsetneq \operatorname{Im} \varphi \supsetneq \operatorname{Im} \varphi^2 \supsetneq \dots \supsetneq \operatorname{Im} \varphi^k = \operatorname{Im} \varphi^{k+1} = \dots$$

По простому эта лемма переформулируется так:

Утверждение. Пусть $A \in M_n(\mathbb{R})$ – произвольная матрица. Тогда

1. Найдется такое число $0 \leq k \leq n$, что

$$0 \subsetneq \{y \in \mathbb{R}^n \mid Ay = 0\} \subsetneq \{y \in \mathbb{R}^n \mid A^2y = 0\} \subsetneq \dots \subsetneq \{y \in \mathbb{R}^n \mid A^ky = 0\} = \{y \in \mathbb{R}^n \mid A^{k+1}y = 0\} = \dots$$

2. Найдется такое число $0 \leq k \leq \dim V$, что

$$V \supsetneq \langle A \rangle \supsetneq \langle A^2 \rangle \supsetneq \dots \supsetneq \langle A^k \rangle = \langle A^{k+1} \rangle = \dots$$

В частности это означает, что у матриц A^k начиная с $k = n$ точно одинаковый ранг и одинаковое множество решений систем $A^ky = 0$. Это позволяет быстро отвечать на задачи вида: пусть дана $A \in M_4(\mathbb{R})$ (которая задана явно), найдите $\operatorname{rk} A^{2019}$. Для этого достаточно найти A^4 и из леммы о стабилизации следует, что $\operatorname{rk} A^4 = \operatorname{rk} A^{2019}$. А A^4 находим за два умножения $(A^2)^2$. Этим же методом доказывается, что если $A^N = 0$ для некоторого большого N , то уже $A^n = 0$.

Собственные значения и вектора оператора

Пусть $\phi: V \rightarrow V$ – линейный оператор на пространстве V . Будем говорить, что вектор $v \in V$ является *собственным*, если $\phi v = \lambda v$.⁸ То есть на собственный вектор оператор ϕ действует растяжением. Если $\phi v = \lambda v$ для $v \neq 0$, число λ называется *собственным значением* оператора ϕ . При фиксированном $\lambda \in \mathbb{R}$ множество всех собственных векторов с собственным значением λ , т.е. $\{v \in V \mid \phi v = \lambda v\}$, будем обозначать через V_λ . Все V_λ обязательно будут подпространствами.⁹

Если мы выберем базис в пространстве V , то оно превратится в \mathbb{R}^n . Наш оператор ϕ будет задаваться матрицей $A \in M_n(\mathbb{R})$. В этом случае, собственный вектор задается уравнением $Ax = \lambda x$, где $x \in \mathbb{R}^n$. Беда в том, что мы пока заранее не знаем, какие λ нам подходят. Чтобы это выяснить нужно переписать уравнение так: $(A - \lambda E)x = 0$. Оно имеет решение тогда и только тогда, когда $A - \lambda E$ – вырожденная матрица. Это, в свою очередь, происходит тогда и только тогда, когда $\det(A - \lambda E) = 0$. Напомним, что характеристический многочлен ϕ (он же характеристический для A) это $\chi(\lambda) = \det(A - \lambda E)$, т.е. получаем следующее.

Утверждение. Пусть $\phi: V \rightarrow V$ – некоторый линейный оператор с матрицей $A \in M_n(\mathbb{R})$ в некотором базисе. Тогда

1. Все собственные значения оператора ϕ – это в точности корни характеристического многочлена $\chi(\lambda) = \det(A - \lambda E)$.
2. Если λ – НЕ корень характеристического многочлена, то $V_\lambda = 0$.
3. Если λ – корень характеристического многочлена, то V_λ – ненулевое подпространство V . Кроме того, $\dim V_\lambda$ не превосходит кратности корня λ у характеристического многочлена.¹⁰

⁸Нулевой вектор является собственным для любого $\lambda \in \mathbb{R}$. Обратите внимание, что в некоторых учебниках собственные вектора обязательно считаются ненулевыми, но это идейно не правильно.

⁹Заметим, что $V_\lambda = \ker(\phi - \lambda \operatorname{Id})$.

¹⁰Для многочлена $p(t)$ число λ является корнем тогда и только тогда, когда $p(t) = (t - \lambda)q(t)$. Если λ корень для $q(t)$, мы можем еще раз вынести множитель $t - \lambda$ и так далее. В итоге, можно записать $p(t) = (t - \lambda)^k h(t)$, где $h(\lambda) \neq 0$. Такое число k называется кратностью корня λ .

Привет от комплексных чисел

Заметим, что собственные значения являются корнями многочлена. С действительными числами есть беда: многочлены могут вообще не иметь корней. Например: пусть $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, тогда $\chi_A(\lambda) = \lambda^2 + 1$. У этого многочлена нет вещественных корней. Потому нет собственных значений, а значит и ненулевых собственных векторов. На самом деле, для любого многочлена можно подобрать матрицу так, что он будет ее характеристическим многочленом. Так что это не случайное явление.

Так как собственные значения и вектора хотелось бы иметь, то нам придется в этом вопросе переходить к комплексным числам. и вместо пространства \mathbb{R}^n рассматривать \mathbb{C}^n . Тогда над комплексными числами каждый многочлен имеет ровно столько корней (с учетом кратности), какова его степень. Это первое место в линейной алгебре, где появляется разница в том, какие коэффициенты использовать.

Кто такие комплексные числа

По простому, мы хотим построить множество «чисел», которые бы содержали вещественные числа и на них были определены все нужные операции: сложения, вычитания, умножения и деления на любое ненулевое число. Есть несколько конструкций, я рассмотрю две.

Классическая конструкция Рассмотрим множество картинок вида $a+bi$, где $a, b \in \mathbb{R}$, а i – просто символ. Как множество $\mathbb{C} = \{a+bi \mid a, b \in \mathbb{R}\}$. Теперь на \mathbb{C} определим следующие операции:

1. Сложение: $(a+bi) + (c+di) = (a+c) + (b+d)i$
2. Вычитание: $(a+bi) - (c+di) = (a-c) + (b-d)i$
3. Умножение: $(a+bi)(c+di) = (ac-bd) + (ad+bc)i$.
4. Сопряжение: $\overline{a+bi} = a-bi$.

В этом случае нулем будет число вида $0+0i$, единицей $1+0i$. Если $z = a+bi$, то число $z\bar{z} = a^2 + b^2$ является неотрицательным вещественным числом. Модуль комплексного числа z – это $|z| = \sqrt{z\bar{z}}$. Обратный к числу z имеет вид $\frac{\bar{z}}{|z|^2}$.

Числа вида $a+0i$ можно отождествить с вещественными числами $a \in \mathbb{R}$. Таким образом $\mathbb{R} \subseteq \mathbb{C}$. Более того, операции определены так, что это вложение с ними согласовано. Обратим внимание на новое число $i = 0+1i$. По определению $i^2 = -1$. На самом деле верно следующее.

Утверждение. Для любого многочлена $p(t) = a_0 + a_1t + \dots + a_{n-1}t^{n-1} + t^n$, где $a_i \in \mathbb{C}$ существует ровно n комплексных корней с учетом кратности.

Матричная конструкция Рассмотрим матрицы вида

$$T = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M_2(\mathbb{R}) \right\}$$

Заметим, что если сложить или перемножить любые две матрицы из T , то получим матрицу из T . Более того, все матрицы из T кроме нулевой обратимы. Множество T можно отождествить с \mathbb{C} , построенным выше, следующим образом: $a+bi \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$. То есть T и \mathbb{C} – это одно и то же. Обратим внимание, что на этом языке сопряжение – это транспонирование, а определитель равен квадрату модуля комплексного числа.¹¹

Собственный базис

Утверждение. Пусть $\phi: V \rightarrow V$ – линейный оператор и пусть $\lambda_1, \dots, \lambda_k$ – его разные собственные значения (тут не важно из \mathbb{R} или \mathbb{C}) и $v_1, \dots, v_k \in V$ – соответствующие им ненулевые собственные вектора. Тогда v_1, \dots, v_k линейно независимы.

¹¹Первая конструкция обычно рассказывается в школе и потому более привычная. Вторая хороша тем, что нам не надо проверять, что операции ведут себя хорошо, все следует из знаний о матрицах. Плюс это дает некий мостик в правильную линейную алгебру над вещественными числами.

Доказательство. Предположим противное, что $a_1v_1 + \dots + a_kv_k = 0$. Мы можем считать, что все a_i не равны нулю. Это можно записать так

$$a_1v_1 + \dots + a_kv_k = \begin{pmatrix} a_1v_1 & \dots & a_kv_k \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

Применим к этой линейной комбинации ϕ , получим новую линейную комбинацию

$$a_1\lambda_1v_1 + \dots + a_k\lambda_kv_k = \begin{pmatrix} a_1v_1 & \dots & a_kv_k \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{pmatrix} = 0$$

Продолжим применять ϕ суммарно $k-1$ раз. В результате имеем

$$\begin{pmatrix} a_1v_1 & \dots & a_kv_k \end{pmatrix} \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \dots & \lambda_k^{k-1} \end{pmatrix} = \begin{pmatrix} 0 & \dots & 0 \end{pmatrix}$$

Но определитель матрицы выше есть определитель вандермонда. Значит, матрица обратима и на нее можно поделить. Значит, все вектора $a_iv_i = 0$. Так как по предположению $v_i \neq 0$ это означает, что $a_i = 0$, противоречие. \square

Утверждение. Пусть $\phi: V \rightarrow V$ – оператор на n -мерном пространстве (не важно комплексном или вещественном), при этом его характеристический многочлен имеет n различных корней $\lambda_1, \dots, \lambda_n$. Тогда соответствующие ненулевые собственные вектора v_1, \dots, v_n образуют базис V и в этом базисе матрица ϕ диагональная с числами λ_i на диагонали.

Доказательство. Действительно, для каждого такого λ_i обязательно найдется ненулевой собственный вектор. Из предыдущего утверждения все такие собственные вектора линейно независимы, а значит образуют базис. По определению в этом базисе $\phi v_i = \lambda_i v_i$, т.е. $\phi(v_1, \dots, v_n) = (v_1, \dots, v_n) \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$. \square

На это утверждение можно смотреть так: если есть квадратная матрица $A \in M_n(\mathbb{R})$ (или $A \in M_n(\mathbb{C})$) такая, что $\det(A - \lambda E)$ имеет n различных корней, то существует такая невырожденная матрица $C \in M_n(\mathbb{R})$ (соответственно из $M_n(\mathbb{C})$), что $C^{-1}AC$ является диагональной и на диагонали стоят корни многочлена $\det(A - \lambda E)$. Комплексный случай хорош лишь тем, что корни обязательно существуют у многочлена, надо лишь чтобы они были различными. В вещественном случае существование корней не гарантировано.

Важно понимать, что если матриц взялась «из жизни» или из «непрерывных случайных данных», то с вероятностью один, характеристический многочлен такой матрицы будет иметь n различных комплексных корней. То есть над комплексными числами любая случайная матрица с вероятностью один превращается в диагональную в некотором базисе.

Поиск собственных значений и векторов

Следующий алгоритм годится как для комплексных так и для вещественных матриц. Разница лишь в том, что в вещественном случае у нас вообще говоря будет меньше собственных значений. Для определенности алгоритм рассказывается для комплексных матриц.

Дано Матрица $A \in M_n(\mathbb{C})$.

Задача Найти все собственные значения λ_i для A и для каждого λ_i найти базис пространства V_{λ_i} .

Алгоритм

1. Посчитать характеристический многочлен $\chi_A(\lambda) = \det(A - \lambda E)$.
2. Найти корни многочлена $\chi(\lambda)$. Корни $\{\lambda_1, \dots, \lambda_k\}$ будут собственными значениями A .
3. Для каждого λ_i найти ФСР системы $(A - \lambda_i E)x = 0$. Тогда ФСР будет базисом V_{λ_i} .

Отметим, что общее количество собственных векторов для всех собственных значений λ_i не превосходит n – размерности матрицы, так как $\dim V_{\lambda_i}$ не превосходит кратности корня λ_i , а сумма кратностей всех корней в точности равна степени многочлена $\chi(\lambda)$, которая есть n – размер матрицы A .

Если количество собственных векторов оказалось равно n , то матрица A приводится в диагональный вид. Пусть v_{i1}, \dots, v_{in_i} – собственные вектора с собственным значением λ_i , при этом n_i будет кратность собственного значения λ_i . Пусть C – матрица составленная из векторов v_{ij} . Пусть D – диагональная матрица с диагональю $(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_k, \dots, \lambda_k)$, где каждое λ_i повторяется n_i раз. Тогда $C^{-1}AC = D$.

Проверка на диагонализуемость

Дано Матрица $A \in M_n(\mathbb{R})$, задающая линейный оператор $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Задача Выяснить существует ли базис, в котором φ задается диагональной матрицей и если задается, то какой именно. На матричном языке: существует ли невырожденная матрица $C \in M_n(\mathbb{R})$ такая, что $C^{-1}AC$ является диагональной и найти эту диагональную матрицу.

Алгоритм

1. Найдем характеристический многочлен $\chi(t)$ для φ , он же для A по формуле $\chi(t) = \det(A - tE)$.
2. Проверим, раскладывается ли $\chi(t)$ на линейные множители, то есть представляется ли он в виде $\chi(t) = (t - \lambda_1)^{d_1} \dots (t - \lambda_k)^{d_k}$. Если не представляется, то φ (или что то же самое A) не диагонализуется.
3. Если $\chi(t) = (t - \lambda_1)^{d_1} \dots (t - \lambda_k)^{d_k}$. Найдем для каждого λ_i базис V_{λ_i} как ФСР системы $(A - \lambda_i E)x = 0$. Если для хотя бы одного i количество элементов в ФСР меньше соответствующей кратности корня d_i , то φ не диагонализуется.
4. Если для каждого i мы получили, что размер ФСР совпадает с кратностью корня, то есть $\dim V_{\lambda_i} = d_i$. То φ диагонализуется и диагональная матрица $C^{-1}AC$ на диагонали содержит числа λ_i в количестве d_i штук.

Заметим, что если задача изначально дана для комплексной матрицы $A \in M_n(\mathbb{C})$, которая задает в этом случае оператор $\varphi: \mathbb{C}^n \rightarrow \mathbb{C}^n$, то первый шаг алгоритма выполнен автоматически, а именно, над комплексными числами любой многочлен разлагается на линейные множители. Потому над комплексными числами вопрос о диагонализуемости – это лишь проверка всех равенств $\dim V_{\lambda_i} = d_i$.

Признак диагонализуемости Есть очень удобный признак диагонализуемости, который часто помогает. Если у вас есть линейный оператор $\varphi: V \rightarrow V$, то обычно в начале фиксируется какой-то базис и после этого V превращается в \mathbb{R}^n , а оператор φ в оператор умножения на матрицу $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}^n$ задан $x \mapsto Ax$, где $A \in M_n(\mathbb{R})$. В итоге нам надо понять, можно ли выбрать другой базис для φ , чтобы матрицу A заменить на диагональную.

Утверждение. Пусть $A \in M_n(\mathbb{R})$ и $g \in \mathbb{R}[x]$ – аннулирующий многочлен для A , то есть $g(A) = 0$. Если он раскладывается на линейные множители $g(x) = (x - \lambda_1) \dots (x - \lambda_k)$ и все λ_i различны, то существует обратимая матрица $C \in M_n(\mathbb{R})$ такая, что $C^{-1}AC$ будет диагональной и на диагонали будут числа из множества $\{\lambda_1, \dots, \lambda_k\}$ (но может быть не все из них и часть может повторяться).

Обратите внимание, что здесь очень важно, чтобы корни λ_i были различными! Если это условие не выполнено, то утверждение не верно. Например, возьмем матрицу $J_n(\lambda)$. У нее минимальный многочлен будет $(x - \lambda)^n$, но про нее можно доказать, что нельзя сопряжением ее сделать диагональной.

Квадратные корни из единицы Это соображение полезно, если вы хотите решать различные матричные уравнения. Например, пусть мы хотим найти все возможные $A \in M_n(\mathbb{R})$ такие, что $A^2 = E$, то есть хотим найти все квадратные корни из единицы в матрицах. Тогда это означает, что $g(x) = x^2 - 1$ зануляет матрицу A . При этом $g(x) = (x - 1)(x + 1)$. Это значит, что найдется обратимая матрица C такая, что $C^{-1}AC$ будет диагональной с числами 1 и -1 на диагонали. Кроме того, если у диагональной матрицы 1 и -1 не идут подряд, то мы ее можем сопрячь некоторой матрицей так, что 1 и -1 пойдут подряд. То есть мы показали, что если A является решением уравнения $A^2 = E$, то найдется такая невырожденная матрица C , что

$$C^{-1}AC = \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}, \text{ следовательно } A = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1}$$

Здесь подразумевается, что блоки с единичной и минус единичной матрицей могут быть пустыми (то есть только одни единицы или минус единицы допустимы).

С другой стороны. Легко видеть, что матрицы полученного вида являются решениями данного уравнения

$$A^2 = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1} C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix} C^{-1} = C \begin{pmatrix} E & 0 \\ 0 & -E \end{pmatrix}^2 C^{-1} = C E C^{-1} = E$$

Жорданова нормальная форма (ЖНФ)

Самый главный вопрос о линейных операторах: на сколько хорошим можно выбрать базис, чтобы максимально упростить матрицу оператора в этом базисе? В случае «общего положения» как в предыдущем параграфе мы можем диагонализировать матрицу. И это самый популярны в приложениях случай. Но есть и плохие матрицы, которые нельзя диагонализировать. В общем случае ответ будет чуть-чуть сложнее.

Для начала несколько определений. *Жорданова клетка* $J_n(\lambda)$ размера n с числом $\lambda \in \mathbb{C}$ – это матрица вида¹²

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \ddots & \vdots \\ \vdots & \vdots & \ddots & 1 \\ 0 & 0 & \dots & \lambda \end{pmatrix} \in M_n(\mathbb{C})$$

Будем говорить, что матрица $A \in M_n(\mathbb{C})$ имеет *Жорданову нормальную форму*, если она имеет следующий блочный вид

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix} \text{ где } A_i \in M_{n_i}(\mathbb{R}) \text{ имеет вид } \begin{pmatrix} J_{r_1}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{r_2}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_s}(\lambda_i) \end{pmatrix}$$

Следующая теорема – это полная классификация линейных операторов на векторном пространстве.

Утверждение (Теорема о Жордановой нормальной форме). Пусть V – комплексное векторное пространство и $\phi: V \rightarrow V$ – произвольный линейный оператор. Пусть $\lambda_1, \dots, \lambda_k$ – корни его характеристического многочлена с кратностями n_1, \dots, n_k . Тогда, существует базис V такой, что матрица ϕ имеет следующий вид:

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_k \end{pmatrix}$$

где $A_i \in M_{n_i}(\mathbb{C})$ (размер равен кратности собственного значения). А каждая A_i имеет вид

$$\begin{pmatrix} J_{r_{i1}}(\lambda_i) & 0 & \dots & 0 \\ 0 & J_{r_{i2}}(\lambda_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_{r_{is_i}}(\lambda_i) \end{pmatrix}$$

¹²Можно определить Жорданову клетку и для действительных чисел и для рациональных и вообще каких угодно, но я буду тут обсуждать только комплексный случай.

где λ_i – соответствующее собственное значение. При этом числа r_{i1}, \dots, r_{is_i} определены однозначно и могут отличаться только порядком.¹³

Замечания Пусть $A \in M_n(\mathbb{C})$, тогда она обязательно приводится в ЖНФ. Если λ – это собственное значение для A , тогда оно является корнем характеристического многочлена и пусть его кратность будет d .¹⁴

1. Число d – это суммарный размер жордановых клеток в ЖНФ для A с числом λ на диагонали.
2. Число $\dim V_\lambda$ – это количество жордановых клеток в ЖНФ для A с числом λ на диагонали.

¹³Существуют алгоритмы нахождения базиса, в котором матрица имеет Жорданову нормальную форму, но мы их изучать не будем.

¹⁴На самом деле можно дать полный список числовых инвариантов, которые характеризуют ЖНФ для A , но это выходит за рамки нашего обсуждения.