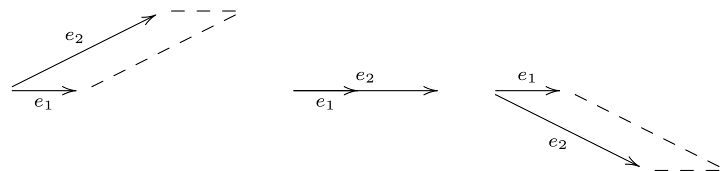
Семинар 2

Определитель (немного философии)

Сейчас я хочу обсудить «ориентированный объем» на прямой, плоскости и в пространстве.

Прямая На прямой мы можем выбрать «положительное» направление. Обычно на рисунке выбирают с лева на право. Тогда длина вектора, который смотрит с лева на право, считается положительной, а с права на лево – отрицательной.

Плоскость Здесь объем будет задаваться парой векторов, то есть некоторой квадратной матрицей размера 2, где вектора – это ее столбцы. Основная идея такая: пусть мы хотим посчитать площадь между двумя векторами на плоскости, точнее площадь параллелограмма натянутого на вектора e_1 и e_2 как на первом рисунке ниже.

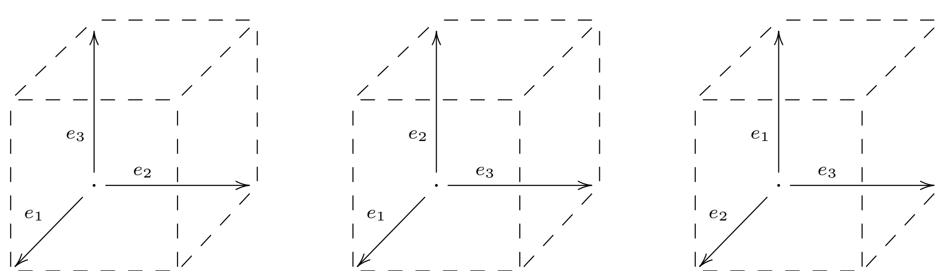


Давайте двигать вектор e_2 к вектору e_1. Тогда площадь будет уменьшаться и когда вектора лягут на одну прямую, она будет равна нулю. Однако, если мы продолжим двигать вектор e_2, то площадь между векторами опять начнет расти и картинка в конце концов станет симметрична исходной, а полученный параллелограмм равен изначальному. Однако, эта ситуация отличается от предыдущей и вот как можно понять чем. Предположим, что между векторами была натянута хорошо сжимаемая ткань, одна сторона которой красная, другая зеленая. Тогда в самом начале на нас смотрит красная сторона этой ткани, но как только e_2 прошел через e_1 на нас уже смотрит зеленая сторона. Мы бы хотели научиться отличать эти две ситуации с помощью знака, если на нас смотрит красная сторона – знак положительный, если зеленая – отрицательный.

Еще один способ думать про эту ситуацию. Представим, что плоскость – это наш стол, а параллелограмм вырезан из бумаги. Мы можем положить параллелограмм на стол двумя способами: лицевой стороной вверх или же вниз. В первом случае мы считаем площадь положительной, а во втором – отрицательной. Возможность определить лицевую сторону связана с тем, что мы знаем, где у стола верх, а где них. Это возможно, потому что наша плоскость лежит в трех мерном пространстве и мы можем глядеть на нее извне. Однако, если бы мы жили на плоскости и у нас не было бы возможности выглянуть за ее пределы, то единственный способ установить «какой стороной вверх лежит параллелограмм» был бы с помощью порядка векторов.

Еще одно важное замечание. Если мы берем два одинаковых параллелограмма на нашем столе, которые лежат лицевой стороной вверх, то мы можем передвинуть один в другой, не отрывая его от стола. А вот если один из параллелограммов имеет положительный объем, а другой отрицательный, то нельзя перевести один в другой, не отрывая от стола. То есть, если вы живете на плоскости, то вам не получится переместить положительный параллелограмм в отрицательный, не сломав или не разобрав его.

Пространство В пространстве дело с ориентацией обстоит абсолютно аналогично. Мы хотим уже считать объемы параллелепипедов натянутых на три вектора. И мы так же хотим, чтобы эти объемы показывали «с какой стороны» мы смотрим на параллелепипед.



Здесь знак объема определяется по порядку векторов.[1] На рисунке объемы первого и третьего положительные, а у второго отрицательный.[2] Если вы сделаете модельки этих кубиков из подписанных спичек, то третий кубик – это первый, но лежащий на другой грани. А вот второй кубик получить из первого вращениями не получится. Надо будет его разобрать и присобачить ребра по-другому.

Как и в случае с плоскостью, если бы мы могли выйти за пределы нашего трехмерного пространства, то у нас появилась бы лицевая и тыльная сторона, как у стола. И тогда первый и третий кубики лежали бы лицевой стороной вверх, а второй – вниз. Мы, конечно же, так сделать не сможем и никогда в жизни не увидим подобное, но думать про такое положение вещей по аналогии с плоскостью можем и эта интуиция бывает полезна.

Определитель

Существует несколько способов определить определитель.[3] Вот классическое определение через перестановки. Пусть $A \in \mathrm{M}_n(\mathbb{R})$ положим

$$\det A = \sum_{\sigma \in \mathrm{S}_n} \operatorname{sgn}(\sigma) a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$$

Именно эту формулу любят составители учебников по линейной алгебре и именно ее никто не использует, когда начинают решать задачи. А если учесть, что в этой формуле полно всяких непонятных значков, смысл которых еще надо объяснить, то это по сути означает, что данная формула, мягко говоря, не такая уж полезная и нужная.

Тем не менее, давайте я объясню, что тут написано. По-хорошему, я должен объяснить, что такое перестановки, их знак и как они тут используются. Но вместо этого я пойду коротким путем и объясню по рабоче-крестьянски.

Внутри стоят произведения вида $a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$, где первый индекс – это индекс строки и эти индексы идут по-порядку. Это значит, что мы из каждой строки выбираем по элементу. Хитрая запись для второго индекса означает, что мы выбираем элементы так, чтобы из каждого столбца выбран в точности один элемент. Например, если мы в первой строке выбрали элемент из 3-его столбца, то далее из 3-его столбца выбирать уже нельзя. То есть такое слагаемое получено так: мы выбираем n элементов из матрицы A так, чтобы в каждой строке и каждом столбце встретился только один элемент. Все эти элементы перемножаем, добавляем знак ± 1 и получаем одно слагаемое. А потом мы все подобные слагаемые складываем между собой. Всего оказывается целых $n!$ слагаемых. А это означает, что их сильно дофига и считать по этой формуле практически невозможно.

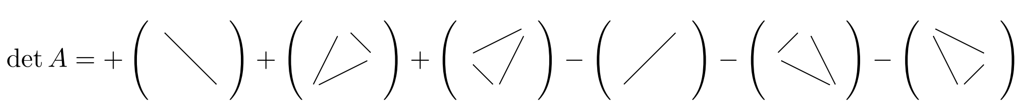
Как это определение связано с ориентированным объемом. Думать про это надо так: пусть $A = (A_1 \mid \ldots \mid A_n) \in \mathrm{M}_n(\mathbb{R})$ составлена из столбцов $A_i \in \mathbb{R}^n$. Тогда $\det A$ – это ориентированный объем n-мерного параллелепипеда натянутого на вектора $A_1, \ldots, A_n$.

Примеры

1. Если $A \in \mathrm{M}_1(\mathbb{R}) = \mathbb{R}$, то $\det A = A$.

2. Если $A \in \mathrm{M}_2(\mathbb{R})$ имеет вид $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, то $\det A = ad - bc$. Графически: главная диагональ минус побочная.

3. Если $A \in \mathrm{M}_3(\mathbb{R})$ имеет вид $A = \left(\begin{smallmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{smallmatrix}\right)$, то определитель получается из 6 слагаемых три из них с $+$ три с $-$. Графически слагаемые можно изобразить так:



Точная формула[4]

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

[1] В общем случае можно определить понятие знака для любого упорядочивания векторов, так что этот знак будет вести себя хорошим образом. Однако, мы не будем углубляться в эту ненужную нам теорию.

[2] Чтобы понять, почему именно так, надо развить теорию знака перестановки. Однако, сейчас это не важно, а ниже, я расскажу метод, который не потребует этого знания.

[3] Как бы глупо это ни звучало.

[4] Начиная с этого момента можно забыть честное определение. Оно почти никогда не нужно для использования, нам лишь пригодятся его свойства. Определитель никогда не считают по определению. Это слишком долго.

Свойства определителя

Считать определитель по явной формуле весьма проблематично. Слишком уж много слагаемых. Потому по определению его можно вычислить лишь для очень специальных матриц. Для произвольных матриц используются некоторые полезные свойства, с помощью которых их определители сводятся к определителям специальных матриц.

Пусть $A \in \mathrm{M}_n(\mathbb{R})$ – матрица, тогда на нее можно смотреть как на набор из n столбцов $A = (A_1|\ldots|A_n)$. Тогда определитель $\det(A)$ можно рассматривать, как функцию от столбцов матрицы A, то есть $\det(A) = \det(A_1|\ldots|A_n)$. Думаю таким образом, мы можем сформулировать следующие свойства:

1. $\det(A_1|\ldots|A_i + A_i'|\ldots|A_n) = \det(A_1|\ldots|A_i|\ldots|A_n) + \det(A_1|\ldots|A_i'|\ldots|A_n)$. Например,

$$\det\begin{pmatrix}1 & 3\\ 2 & 7\end{pmatrix} = \det\left(\begin{pmatrix}1\\2\end{pmatrix}\middle|\begin{pmatrix}3\\7\end{pmatrix}\right) = \det\left(\begin{pmatrix}1\\2\end{pmatrix}\middle|\begin{pmatrix}1\\3\end{pmatrix}+\begin{pmatrix}2\\4\end{pmatrix}\right) = \det\begin{pmatrix}1 & 1\\ 2 & 3\end{pmatrix} + \det\begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$$

2. $\det(A_1|\ldots|A_i|\ldots|A_j|\ldots|A_n) = -\det(A_1|\ldots|A_j|\ldots|A_i|\ldots|A_n)$. То есть если поменять местами два столбца, то определитель изменит знак, например,

$$\det\begin{pmatrix}1 & 3\\ 2 & 7\end{pmatrix} = -\det\begin{pmatrix}3 & 1\\ 7 & 2\end{pmatrix}$$

3. $\det(A_1|\ldots|A'|\ldots|A'|\ldots|A_n) = 0$, то есть если у матрицы есть два одинаковых столбца, то определитель автоматически равен нулю, например

$$\det\begin{pmatrix}1 & 0\\ 2 & 0\end{pmatrix} = 0$$

4. $\det(A_1|\ldots|\lambda A_i|\ldots|A_n) = \lambda\det(A_1|\ldots|A_i|\ldots|A_n)$. То есть, если один столбец умножить на одно и то же число, то весь определитель умножится на это число, например,

$$\det\begin{pmatrix}1 & 3\\ 2 & 9\end{pmatrix} = 3\det\begin{pmatrix}1 & 1\\ 2 & 3\end{pmatrix}$$

5. $\det(A_1|\ldots|A_i|\ldots|A_j|\ldots|A_n) = \det(A_1|\ldots|A_i|\ldots|A_j + \lambda A_i|\ldots|A_n)$. То есть, если к одному столбцу матрицы прибавить другой умноженный на коэффициент, то определитель не изменится, например

$$\det\begin{pmatrix}1 & 3\\ 2 & 7\end{pmatrix} = \det\left(\begin{pmatrix}1\\2\end{pmatrix}\middle|\begin{pmatrix}3\\7\end{pmatrix}\right) = \det\left(\begin{pmatrix}1\\2\end{pmatrix}\middle|\begin{pmatrix}3\\7\end{pmatrix}+2\begin{pmatrix}1\\2\end{pmatrix}\right) = \det\left(\begin{pmatrix}1\\2\end{pmatrix}\middle|\begin{pmatrix}5\\11\end{pmatrix}\right) = \det\begin{pmatrix}1 & 5\\ 2 & 11\end{pmatrix}$$

6. $\det A = \det A^t$. То есть определитель матрицы равен определителю транспонированной матрицы. А значит, все свойства сформулированные выше для столбцов автоматически верны и для строк.

7. Определитель треугольной матрицы

$$\det\begin{pmatrix}\lambda_1 & * & \ldots & *\\ & \lambda_2 & \ldots & *\\ & & \ddots & \vdots\\ & & & \lambda_n\end{pmatrix} = \det\begin{pmatrix}\lambda_1 & & & \\ * & \lambda_2 & & \\ \vdots & \vdots & \ddots & \\ * & * & \ldots & \lambda_n\end{pmatrix} = \lambda_1\lambda_2\ldots\lambda_n$$

То есть у треугольной матрицы определитель равен произведению ее диагональных элементов. В частности $\det E = 1$ и $\det(\lambda E) = \lambda^n$.

Вычисление определителя с помощью элементарных преобразований

Из сформулированных свойств выше следует, что определитель можно считать так: надо привести матрицу A элементарными преобразованиями к треугольному виду и по пути запоминать некоторые коэффициенты, после чего надо перемножить эти коэффициенты с определителем треугольной матрицы. Давайте продемонстрируем на примере, будем приводить матрицу к ступенчатому виду элементарными преобразованиями строк

$$\det\begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & -1 \end{pmatrix} = -\det\begin{pmatrix} 1 & 1 & -1 \\ -1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} = -\det\begin{pmatrix} 1 & 1 & -1 \\ 0 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix} =$$
$$= -3\det\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = -3\det\begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = -3$$

Связь определителя с произведением

Для определителя верны следующие формулы

1. $\det(AB) = \det(A)\det(B)$
2. $\det(A^{-1}) = \det(A)^{-1}$
3. Пусть $A \in \mathrm{M}_n(\mathbb{R})$, $C \in \mathrm{M}_m(\mathbb{R})$ и $B \in \mathrm{M}_{n\,m}(\mathbb{R})$, тогда
$$\det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \det C$$

Оказывается, что определитель является единственной функцией $\phi\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ такой, что

1. $\phi(AB) = \phi(A)\phi(B)$
2. $\phi\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{pmatrix} = \lambda$

Миноры и алгебраические дополнения

Пусть $A \in \mathrm{M}_n(\mathbb{R})$ с элементами a_{ij}. Рассмотрим матрицу $D_{ij} \in \mathrm{M}_{n-1}(\mathbb{R})$ полученную из A вычеркиванием i-ой строки и j-го столбца. Определитель матрицы D_{ij} обозначается M_{ij} и называется *минором* матрицы A или ij-минором для определенности. Число $(-1)^{i+j}M_{ij}$ называется алгебраическим дополнением элемента a_{ij} и обозначается A_{ij}.

Покажем как это все выглядит на картинках. Если мы представим матрицу A в виде

$$A = \left(\begin{array}{c|c|c} X_{ij} & \begin{matrix} * \\ \vdots \end{matrix} & Y_{ij} \\ \hline * \;\; \ldots & a_{ij} & \ldots \;\; * \\ \hline Z_{ij} & \begin{matrix} \vdots \\ * \end{matrix} & W_{ij} \end{array}\right)$$

Тогда

$$D_{ij} = \begin{pmatrix} X_{ij} & Y_{ij} \\ Z_{ij} & W_{ij} \end{pmatrix}, \quad M_{ij} = \det\begin{pmatrix} X_{ij} & Y_{ij} \\ Z_{ij} & W_{ij} \end{pmatrix} \quad \text{и} \quad A_{ij} = (-1)^{i+j}\det\begin{pmatrix} X_{ij} & Y_{ij} \\ Z_{ij} & W_{ij} \end{pmatrix}$$

Разложение определителя по строке (столбцу)

Разложение по столбцу Для матрицы $A \in \mathrm{M}_n(\mathbb{R})$ и любого k с условием $1 \leqslant k \leqslant n$ верна формула $\det A = \sum_{i=1}^n a_{ik}A_{ik}$. Здесь A_{ij} – алгебраическое дополнение a_{ij}. Например, разложим по второму столбцу

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix} = 2(-1)^{1+2}\det\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix} + 2(-1)^{2+2}\det\begin{pmatrix} 1 & 3 \\ -1 & 1 \end{pmatrix} + 0(-1)^{3+2}\det\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$$

Для разложения по строкам верны ровно те же самые формулы. Их можно получить просто перейдя к транспонированной матрице

Явная формула обратной матрицы

Пусть $A \in \mathrm{M}_n(\mathbb{R})$. Оказывается, что матрица обратима тогда и только тогда, когда $\det A \neq 0$. В этом случае мы можем написать явные формулы для обратной матрицы через определитель. Пусть A имеет вид

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

Составим матрицу A^+ из алгебраических дополнений, т.е.

$$A^+ = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}$$

Сопряженная матрица $\hat{A}$ для A определяется как $(A^+)^t$, т.е.

$$\hat{A} = \begin{pmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{pmatrix}$$

Тогда $A^{-1} = \frac{1}{\det A}\hat{A}$.

Пример Пусть $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, тогда $A^+ = \left(\begin{smallmatrix} d & -c \\ -b & a \end{smallmatrix}\right)$, $\hat{A} = \left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$, $A^{-1} = \frac{1}{ad-bc}\left(\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right)$.

Формулы Крамера

Матрица $A \in \mathrm{M}_n(\mathbb{R})$ называется *невырожденной*, если $\det A \neq 0$. В противном случае она называется вырожденной. Если матрица A не вырожденная, то из явных формул для обратной матрицы следует, что существует A^{-1}. В частности система вида $Ax = b$ имеет единственное решение для любой правой части $b \in \mathbb{R}^n$.

Пусть $A \in \mathrm{M}_n(\mathbb{R})$ – невырожденная квадратная матрица, $b \in \mathbb{R}^n$ – произвольный столбец и $Ax = b$ – система линейных уравнений с n неизвестными.

Пусть $A = (A_1|\dots|A_n)$, где $A_i \in \mathbb{R}^n$ – столбцы матрицы A. Определим матрицы $B_i = (A_1|\dots|A_{i-1}|b|A_{i+1}|\dots|A_n)$, т.е. B_i получается из A, если в A заменить i-ый столбец A_i на вектор b из правой части системы.

Так как A – невырождена, то система $Ax = b$ всегда имеет единственное решение. Это решение вычисляется по формулам $x_i = \frac{\det B_i}{\det A}$. Эти формулы и называются формулами Крамера.[5]

Пример Пусть $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ и $b = \left(\begin{smallmatrix} b_1 \\ b_2 \end{smallmatrix}\right)$. Решим систему $Ax = b$. Тогда $\det A = ad - bc$, $\det B_1 = \det\left(\begin{smallmatrix} b_1 & b \\ b_2 & d \end{smallmatrix}\right) = b_1 d - b_2 b$, $\det B_2 = \det\left(\begin{smallmatrix} a & b_1 \\ c & b_2 \end{smallmatrix}\right) = ab_2 - cb_1$. Тогда $x_1 = \frac{b_1 d - b_2 b}{ad-bc}$ и $x_2 = \frac{ab_2 - cb_1}{ad-bc}$.

Спектр матрицы

Пусть $A \in \mathrm{M}_n(\mathbb{R})$ определим спектр матрицы A следующим образом:

$$\operatorname{spec} A = \{\lambda \in \mathbb{C} \mid A - \lambda E \text{ не обратима}\}$$

Это определение спектра принадлежит функциональному анализу. На самом деле этот спектр совпадает со множеством собственных значений матрицы, о которых речь пойдет чуть позже. Обратите внимание, что хотя матрица имеет вещественные коэффициенты, спектр мы рассматриваем комплексный. Это все заговор среди математиков, которые любят комплексные числа больше вещественных за их более приличное математическое поведение.

[5] У этих формул нет особого практического смысла, в основном только теоретическое применение, за исключением малого размера.

Примеры

1. Пусть $A \in \mathrm{M}_n(\mathbb{R})$ – диагональная матрица

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

   Тогда $\operatorname{spec} A = \{\lambda_1, \ldots, \lambda_n\}$.

2. Пусть $A = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \in \mathrm{M}_2(\mathbb{R})$. Можно показать, что $\operatorname{spec} A = \{i, -i\}$.

Подстановка матриц в многочлен

Пусть $p(x) = a_0 + a_1 x + \ldots a_n x^n$ – многочлен с вещественными коэффициентами, а $A \in \mathrm{M}_n(\mathbb{R})$. Тогда можно определить $p(A) = a_0 E + a_1 A^1 + \ldots + a_n A^n$, где E – единичная матрица. Множество всех многочленов с вещественными коэффициентами я буду обозначать $\mathbb{R}[x]$.

Свойства подстановки в многочлен

1. Если $A \in \mathrm{M}_n(\mathbb{R})$ и $f \in \mathbb{R}[x]$ – многочлен, то $f(C^{-1}AC) = C^{-1}f(A)C$ для любой обратимой $C \in \mathrm{M}_n(\mathbb{R})$.

2. Если $A \in \mathrm{M}_n(\mathbb{R})$ и $f, g \in \mathbb{R}[x]$ – многочлены, то матрицы $f(A)$ и $g(A)$ коммутируют между собой.

Минимальный многочлен Пусть $A \in \mathrm{M}_n(\mathbb{R})$ – некоторая матрица. Рассмотрим множество всех ненулевых многочленов зануляющих A. Формально мы смотрим на множество

$$M = \{f \in \mathbb{R}[x] \mid f(A) = 0,\ f \neq 0\}$$

Пусть $f_{min} \in M$ – многочлен самой маленькой степени и со старшим коэффициентом 1. Тогда он называется минимальным многочленом матрицы A. Обратите внимание, что минимальный многочлен зависит от того, с какими коэффициентами мы его рассматриваем. Комплексный и вещественный минимальный многочлен могут быть разными.

Утверждение. *Пусть $A \in \mathrm{M}_n(\mathbb{R})$, тогда верны следующие утверждения:*

1. *Минимальный многочлен f_{min} существует и единственный, при этом $\deg f_{min} \leqslant n$.*

2. *Минимальный многочлен делит любой другой многочлен зануляющий A.*

3. *$\lambda \in \operatorname{spec} A$ тогда и только тогда, когда $f_{min}(\lambda) = 0$.*

4. *Для любого зануляющего многочлена $g \in \mathbb{R}[x]$, то есть $g(A) = 0$, верно $\operatorname{spec} A \subseteq$ корни g.*

Характеристический многочлен

Пусть $A \in \mathrm{M}_n(\mathbb{R})$, тогда выражение

$$\chi_A(\lambda) = \det(\lambda E - A) = (-1)^n \det(A - \lambda E)$$

является многочленом от λ степени n и называется характеристическим многочленом для A. Главная польза от характеристического многочлена – его корни это и есть спектр матрицы.

Утверждение. *Пусть $A \in \mathrm{M}_n(\mathbb{R})$ и $\chi_A(\lambda)$ – характеристический многочлен. Тогда*

1. *Легко посчитать следующие коэффициенты*[6]

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

2. *Для произвольного числа λ верно, что $\lambda \in \operatorname{spec} A$ тогда и только тогда, когда $\chi_A(\lambda) = 0$.*

3. *Многочлен $\chi_A(\lambda)$ зануляющий для A, то есть $\chi_A(A) = 0$. Это называется теорема Гамильтона-Кэли.*

[6] На самом деле есть формулы для всех коэффициентов, но в них нет смысла для нас.

Кратность корня Пусть есть многочлен $f = x^n + a_{n-1}x^{n-1} + \ldots + a_0$, где $a_i \in \mathbb{R}$ и пусть $\lambda \in \mathbb{R}$ – его корень, то есть $f(\lambda) = 0$. Это означает, что многочлен f можно представить в виде

$$f(x) = (x - \lambda)g(x)$$

где $g(x)$ уже многочлен степени $n-1$. Если λ является корнем g, то для него можно проделать ту же самую процедуру и выделить множитель вида $x - \lambda$. В итоге получим

$$f(x) = (x - \lambda)g(x) = (x - \lambda)^2 g_1(x)$$

Если λ является корнем для g_1, то опять можно продолжить эту процедуру. В какой то момент мы придем в разложению вида

$$f(x) = (x - \lambda)^k h(x)$$

где h – многочлен степени $n-k$ (он может быть равен 1 при $k = n$). Давайте подытожим написанное выше. Если λ является корнем многочлена f, то всегда можно найти разложение вида

$$f(x) = (x - \lambda)^k h(x), \quad k > 0, \quad h(\lambda) \neq 0$$

В этом случае число k называется кратностью корня λ. Обратите внимание, что если λ корень, то кратность всегда хотя бы один. То есть быть корнем – это значит уметь отщепить от многочлена линейный множитель. А кратность корня – это сколько раз мы можем этот множитель отщепить. Вот пара примеров

$$x^2 + 2x + 1 = (x + 1)^2$$
$$x^2 - 4x + 3 = (x - 1)(x - 3)$$

В первом случае один корень $\lambda = -1$ кратности 2, а во втором случае два корня кратности 1, это 3 и 1.

Рекурентные соотношения Вместо того, чтобы тут развивать супер общую теорию, я все проиллюстрирую на конкретном примере – последовательность Фибоначчи. Что это такое? Это последовательность чисел $a_n \in \mathbb{R}$, где $n \in \mathbb{Z}_{\geqslant 0}$ удовлетворяющая следующим условиям

$$\begin{cases} a_n = a_{n-1} + a_{n-2} \\ a_0 = 1 \\ a_1 = 1 \end{cases}$$

Если мы захотим по этим правилам посчитать n-ый член последовательности, то нам понадобится $O(n)$ операций, то есть последовательно посчитать все n членов последовательности, чтобы добраться до a_n. Однако, можно несколько схитрить и сделать это быстрее за $O(\log n)$ с помощью матричных операций. Для этого введем вектор

$$x_n = \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} \text{ при этом мы знаем, что } x_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Заметим, что x_n выражается через x_{n-1} следующим образом

$$\begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ a_{n-2} \end{pmatrix} \text{ то есть } x_n = Ax_{n-1} \text{ где } A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Но тогда $x_n = Ax_{n-1} = A^2 x_{n-2} = \ldots = A^{n-1}x_1$. А значит для нахождения x_n нам лишь надо возвести матрицу A в степень n. Для этого подходит хорошо известный алгоритм быстрого возведения в степень для чисел, который слово в слово работает для матриц. Давайте заведем две квадратные матрицы $X, Y \in \mathrm{M}_2(\mathbb{R})$ и число $m \in \mathbb{Z}_{\geqslant 0}$. В самом начале положим $X = E$, $Y = A$ и $m = n$. Будем поддерживать следующий инвариант $XY^m = A^n$. Алгоритм остановим тогда, когда $m = 0$, тогда X будет нашим ответом. Шаги алгоритма следующие. Если m четно, то $XY^{2m'} = X(Y^2)^{m'}$ поделим m на 2, а Y возведем в квадрат. Если m нечетно, то $XY^{2m'+1} = (XY)Y^{2m'}$ уменьшим m' на единицу и умножим X на Y.

На самом деле, можно проверить, что $A = CDC^{-1}$, где

$$D = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}, \quad C = \begin{pmatrix} \frac{2}{\sqrt{5}-1} & -\frac{2}{\sqrt{5}+1} \\ 1 & 1 \end{pmatrix} \text{ и } C^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{5-\sqrt{5}}{10} \\ -\frac{1}{\sqrt{5}} & \frac{5+\sqrt{5}}{10} \end{pmatrix}$$

Для простоты обозначений, будем считать $D = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \mu \end{smallmatrix}\right)$, тогда

$$A^n = (CDC^{-1})^n = CD^nC^{-1} = C \left(\begin{smallmatrix} \lambda^n & 0 \\ 0 & \mu^n \end{smallmatrix}\right) C^{-1}$$

То есть, зная как получить из матрицы A супер крутое разложение, мы можем еще сильнее упростить вычисления степени матрицы и свести его к вычислению степеней чисел и произведения трех матриц.