Линейная алгебра

Дима Трушин

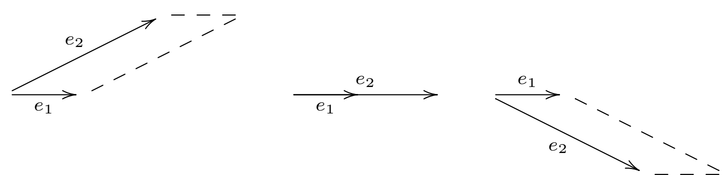
Семинар 3

Определители

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

Прямая На прямой мы можем выбрать «положительное» направление. Обычно на рисунке выбирают слева направо. Тогда длина вектора, который смотрит слева направо, считается положительной, а справа налево – 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.

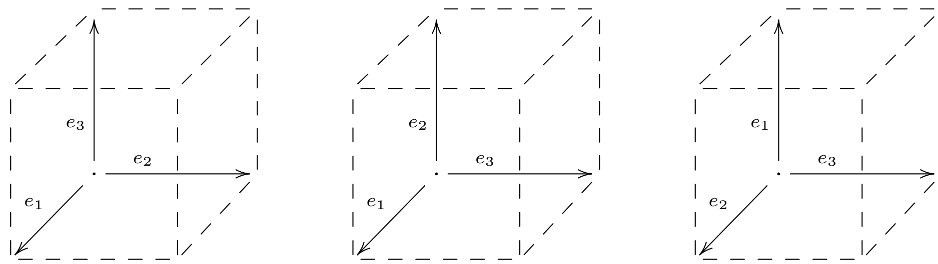


Давайте двигать вектор e_2 к вектору e_1. Тогда площадь будет уменьшаться и когда вектора совпадут, она будет равна нулю. Однако, если мы продолжим двигать вектор e_2, то площадь между векторами опять начнет расти и картинка в конце концов станет симметрична исходной, а полученный параллелограмм равен 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, как знак перестановки. На рисунке объемы первого и третьего положительные, а у второго отрицательный. 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.

Пояснение планов В текущей лекции я не собираюсь обсуждать объемы, а всего лишь хочу коснуться некоторой техники, которая используется для работы с ориентированными объемами. Чтобы начать честный рассказ про сами объемы (который обязательно будет, но позже), нам надо поговорить о том, что такое векторное пространство и как в абстрактном векторном пространстве мерить расстояния и углы. Потому, пока мы не покроем эти темы, всерьез говорить про настоящие объемы мы не сможем.

Три разных определения

Мультипликативные отображения (I) Рассмотрим множество отображений $\psi\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ удовлетворяющие следующими свойствам:

1. $\psi(AB) = \psi(A)\psi(B)$ для любых $A, B \in \mathrm{M}_n(\mathbb{R})$.
2. $\psi\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & d \end{pmatrix} = d$ для любого ненулевого $d \in \mathbb{R}$.

Всюду ниже будем упоминать отображения с такими свойствами, как отображения со свойством (I).

Полилинейные кососимметрические отображения (II) Пусть $\phi\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ – некоторое отображение и $A \in \mathrm{M}_n(\mathbb{R})$. Тогда про матрицу A можно думать, как про набор из n столбцов: $A = (A_1|\ldots|A_n)$. Тогда функцию $\phi(A) = \phi(A_1, \ldots, A_n)$ можно рассматривать как функцию от n столбцов.

В обозначениях выше рассмотрим отображения $\phi\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ удовлетворяющие следующим свойствам:

1. $\phi(A_1, \ldots, A_i + A_i', \ldots, A_n) = \phi(A_1, \ldots, A_i, \ldots, A_n) + \phi(A_1, \ldots, A_i', \ldots, A_n)$ для любого i.
2. $\phi(A_1, \ldots, \lambda A_i, \ldots, A_n) = \lambda\phi(A_1, \ldots, A_i, \ldots, A_n)$ для любого i и любого $\lambda \in \mathbb{R}$.
3. $\phi(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n) = -\phi(A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n)$ для любых различных i и j.
4. $\phi(E) = 1$.

Первые два свойства вместе называются *полилинейностью* ϕ по столбцам, т.е. это уважение суммы и умножения на скаляр. Третье свойство называется *кососимметричностью* ϕ по столбцам. Данный набор свойств можно заменить эквивалентным с переформулированным третьим свойством:

1. $\phi(A_1, \ldots, A_i + A_i', \ldots, A_n) = \phi(A_1, \ldots, A_i, \ldots, A_n) + \phi(A_1, \ldots, A_i', \ldots, A_n)$ для любого i.

2. $\phi(A_1, \ldots, \lambda A_i, \ldots, A_n) = \lambda\phi(A_1, \ldots, A_i, \ldots, A_n)$ для любого i и любого $\lambda \in \mathbb{R}$.
3. $\phi(A_1, \ldots, A', \ldots, A', \ldots, A_n) = 0$, т.е. если есть два одинаковых столбца, то значение ϕ равно нулю.
4. $\phi(E) = 1$.

Действительно, обозначим $\Phi(a, b) = \phi(A_1, \ldots, a, \ldots, b, \ldots, A_n)$. Тогда Φ полилинейная функция двух аргументов.[1] И нам надо показать, что $\Phi(a, a) = 0$ для любого $a \in \mathbb{R}^n$ тогда и только тогда, когда $\Phi(a, b) = -\Phi(b, a)$ для любых $a, b \in \mathbb{R}^n$. Для $\Rightarrow$ подставим $b = a$, получим $\Phi(a, a) = -\Phi(a, a)$. Для обратного $\Leftarrow$ подставим $a+b$, получим $\Phi(a+b, a+b) = 0$. Раскроем скобки: $\Phi(a, a)+\Phi(a, b)+\Phi(b, a)+\Phi(b, b) = 0$. Откуда следует требуемое.

Везде далее будем упоминать отображения с такими свойствами, как отображения со свойством (II).

Полилинейные кососимметрические отображения (II') Аналогично (II) можно рассмотреть полилинейные кососимметрические отображения по строкам матрицы A вместо столбцов. Тогда можно рассматривать отображения $\phi'\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ с аналогами четырех свойств выше: полилинейность, кососимметричность, значение 1 на единичной матрице. Такие отображения мы будем называть, как отображения со свойствами (II').

Определитель (III) Рассмотрим отображение $\det\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ задаваемое следующей формулой: для любой матрицы $A \in \mathrm{M}_n(\mathbb{R})$ положим

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

Данное отображение называется *определителем*, а его значение $\det A$ на матрице A называется определителем матрицы A.

Давайте неформально обсудим, как считается выражение для определителя. Как мы видим определитель состоит из суммы некоторых произведений. Каждое произведение имеет вид $a_{1\sigma(1)} \ldots a_{n\sigma(n)}$ умноженное на $\operatorname{sgn}(\sigma)$. Здесь из каждой строки матрицы A[2] выбирается по одному элементу так, что никакие два элемента не лежат в одном столбце (это гарантированно тем, что σ – перестановка и потому $\sigma(i)$ не повторяются). Заметим, что слагаемых ровно столько, сколько перестановок – $n!$ штук. Из этих слагаемых половина идет со знаком плюс, а другая – со знаком минус.

Самое важное – надо понимать, что все три подхода эквивалентны.

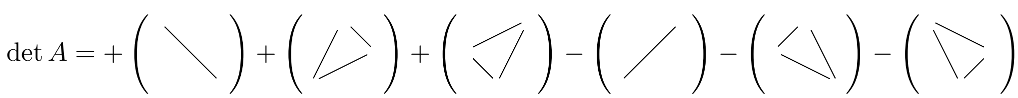
Утверждение. *Пусть $\varphi\colon \mathrm{M}_n(\mathbb{R}) \to \mathbb{R}$ – некоторое отображение, тогда эквивалентно*

1. *φ обладает свойствами определения (I).*
2. *φ обладает свойствами определения (II) (или (II')).*
3. *$\varphi = \det$.*

Явные формулы для определителя

Подсчет в малых размерностях

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[3]

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

[1] Такие отображения называются билинейными.

[2] Первый индекс – индекс строки.

[3] Для больших размерностей чем 3 на 3 явная формула не пригодна из-за слишком большого числа слагаемых. Даже с вычислительной точки зрения.

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
$$\det\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det A \det C$$

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

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

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 $A \in \mathrm{M}_n(\mathbb{R})$ 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.

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

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

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

Для любой матрицы A всегда выполнено равенство

$$\hat{A}A = A\hat{A} = \det(A)E$$

При этом матрица A обратима тогда и только тогда, когда $\det A \neq 0$. В этом случае обратную матрицу можно найти по формуле $A^{-1} = \frac{1}{\det A}\hat{A}$.

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.[4]

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 $x_2 = \frac{ab_2 - cb_1}{ad-bc}$.

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.

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

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

1. *Легко посчитать следующие коэффициенты*[5]

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

2. *Для произвольного числа λ верно, что $\lambda \in \operatorname{spec}_{\mathbb{R}} A$ тогда и только тогда, когда $\chi_A(\lambda) = 0$.*[6]

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

Принцип «продолжения по непрерывности»

Задача. Рассмотрим квадратную матрицу из $\mathrm{M}_{n+m}(\mathbb{R})$ вида $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$, где $A \in \mathrm{M}_n(\mathbb{R})$, $B \in \mathrm{M}_{n\,m}(\mathbb{R})$, $C \in \mathrm{M}_{m\,n}(\mathbb{R})$, $D \in \mathrm{M}_m(\mathbb{R})$. Покажите следующее:

1. Если A обратима, то
$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A)\det(D - CA^{-1}B)$$

2. Если $m = n$ и $AC = CA$, то[7]
$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB)$$

Решение. (1) Идея в том, чтобы применить блочные элементарные преобразования к строкам матрицы. А именно, мы берем первую строку матрицы $(A|B)$ и умножаем ее слева на «коэффициент» CA^{-1}, получим $(C|CA^{-1}B)$. После этого вычитаем эту строку из второй строки $(C|D)$ и получаем $(0|D - CA^{-1}B)$. На одной картинке мы проделали следующее

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{вычитаем из II-ой строки I-ю умноженную на } CA^{-1} \text{ слева} \quad \longrightarrow \begin{pmatrix} A & B \\ 0 & D - CA^{-1}B \end{pmatrix}$$

По аналогии с элементарными преобразованиями первого типа над строками такая процедура не должна поменять определитель. Давайте строго поймем, почему определитель не меняется. Для этого заметим, что такая процедура эквивалентна умножению слева на блочную матрицу, а именно

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

Самая левая матрица – нижнетреугольная с единицами на диагонали, а ее определитель равен 1. Потому определитель исходной матрицы равен определителю матрицы слева. А ее определитель равен требуемому, так как тут есть угол нулей.

(2) **Шаг 1** Давайте в начале рассмотрим случай A обратима. В этом случае, по первому пункту мы получаем
$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det(D - CA^{-1}B) = \det(AD - ACA^{-1}B) = \det(AD - CB)$$

В последнем равенстве мы воспользовались тем, что матрицы A и C коммутируют.

Шаг 2 А здесь я вам продемонстрирую идею, на которой держится решение многих задач. Например, если вы можете доказать что-то только для обратимых матриц, то этим способом очень легко свести доказательство необратимого случая к обратимому.

Давайте рассмотрим матрицу $A_\lambda = A - \lambda E$. Как мы знаем, спектр матрицы A – это в точности те λ, при которых $A - \lambda E$ необратима и таких λ у нас конечное число. А значит матрица A_λ необратима только для конечного числа λ и обратима для бесконечного числа λ.

Теперь рассмотрим нашу задачу для матрицы A_λ вместо матрицы A. То есть нам надо посчитать

$$\det\begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix}$$

[5] На самом деле есть формулы для всех коэффициентов, но в них нет смысла для нас.
[6] Аналогичное утверждение верно и для $\operatorname{spec}_{\mathbb{C}} A$.
[7] Обратите внимание, матрица A может не быть обратимой в этом случае.

Матрицы A и C коммутировали. Так как $A_\lambda = A - \lambda E$, то и A_λ и C коммутируют. А значит для всех $\lambda \in \mathbb{R}$ кроме конечного числа мы можем воспользоваться предыдущим шагом и написать

$$\det \begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix} = \det(A_\lambda D - CB) \quad \text{для всех } \lambda \in \mathbb{R} \text{ кроме конечного числа}$$

Теперь обратим внимание, что левая часть

$$\det \begin{pmatrix} A_\lambda & B \\ C & D \end{pmatrix}$$

является каким-то многочленом $f(\lambda)$. Действительно, если посчитать по формуле для определителя, то мы будем перемножать какие-то элементы из матриц $A - \lambda E$, B, C и D и потом складывать все вместе. Все элементы являются либо константами либо линейными многочленами от λ. А значит их произведение будет многочленом, ну и их сумма тоже. Аналогично правая часть

$$\det(A_\lambda D - CB)$$

тоже является каким-то многочленом $g(\lambda)$. И мы только что увидели, что $f(\lambda) = g(\lambda)$ для всех $\lambda \in \mathbb{R}$ кроме конечного числа. То есть многочлен $f - g$ имеет бесконечное число корней. А такое бывает лишь когда $f = g$, а значит $f(\lambda) = g(\lambda)$ вообще для любого числа $\lambda \in \mathbb{R}$. Подставим значение $\lambda = 0$ и получим желаемое. $\square$

На последний шаг можно смотреть так. Мы знаем результат для каждой матрицы A_λ в силу обратимости таких матриц при достаточно малом λ. А теперь в финальном равенстве переходим к пределу при $\lambda \to 0$. Это тоже корректное доказательство, но приведенное выше не использует анализ в явном виде. Кроме того, при правильных словах оно годится даже для экзотических полей вроде конечных (если вы понимаете о чем я говорю).