Линейная алгебра

Илья Панов

31 июля 2025 г.

Содержание

1 Введение **2**

2 Аналитическая геометрия **2**

3 Векторные пространства и матрицы **7**

4 Линейные операторы **11**

5 Определители **15**

1 Введение

Конспект в основном составлялся по лекциям Поступашек для подготовки к поступлению в ШАД и AI Masters. Читать конспект в отрыве от лекций не особо имеет смысл, потому что в лекциях всё сильно подробнее и с рисунками, но в целом можете попробовать: интернет и нейросети творят чудеса. Записи лекций у вас есть (я их пронумеровал, смотрите по порядку), домашки есть в конце каждой секции (курс рекомендует выполнять хотя бы по 10 заданий из дз), учебники и сборники задач - в репозитории.

2 Аналитическая геометрия

Сейчас мы начнём с повторения 10-11 класса школы, повыводим всякое для плоскости, потом заметим, что для пространства у нас особо ничего не меняется.

Теорема

Три вектора (a, b, c) на плоскости всегда линейно зависимы

Доказательство:
Можем просто составить систему уравнений, решить её по Гауссу (или как Вам угодно).

$$\begin{cases} x \cdot x_a + y \cdot x_b = x_c \\ x \cdot y_a + y \cdot y_b = y_c \end{cases}$$

Получим, что решения у нас есть, если вектора не коллинеарны (в конце получим $x_a \cdot y_b - x_b \cdot y_a$ в знаменателе, если это выражение равно 0, то это равносильно коллинеарности векторов a и b без ограничения общности). Если какие-то два коллинеарны (а третий не коллинеарен), то колинеарные вектора связаны каким-то коэффициентом k, а третий вектор можем взять с нулевым коэффициентом.

Определение

Метод Гаусса, также известный как метод исключения Гаусса, это алгоритм решения систем линейных алгебраических уравнений (СЛАУ) путем последовательного исключения переменных. В основе метода лежит преобразование системы уравнений к равносильной ступенчатой (треугольной) форме (то есть нолики у нас снизу выстраиваются), из которой затем последовательно находятся значения переменных.

Теорема

Угол между двумя векторами и равносильность определений скалярного произведения

Доказательство:

Если есть равносильность определений, то косинус угла выражается очевидно. Равносильность следует из теоремы косинусов: пусть хотим найти угол между векторами a и b, тогда проведем третий - c такой, что он соединяет концы двух других векторов. Пишем теорему косинусов для c, как раз получаем искомый угол и связь определений скалярного.

Теорема

Неравенство КБШ: произведение длин векторов не меньше, чем модуль их скалярного произведения: $|a| \cdot |b| \geq |(a, b)|$

Доказательство:
Рассмотрим $t \in \mathbb{R}$, теперь возьмем скалярное произведение $(x - ty, x - ty)$, оно ≥ 0 по свойствам. По линейности раскрываем, получаем квадратный трёхчлен, который ≥ 0, значит у него $D \geq 0$ - это в точности неравенство КБШ.

Теорема

Пусть даны два вектора a и b, отложенные от одной точки, тогда проекция вектора a на вектор b можно найти по формуле $a^{‘} = \frac{(a,b)}{(b,b)} \cdot b$

Доказательство:
Что такое проекция, надеюсь, все представляют (просто уронили перпендикуляр). Длина проекция очевидным образом находится из прямоугольного треугольника $|a^{‘}| = |a| \cdot \cos\phi$

Теперь попробуем выразить сам вектор $a^{‘}$, он лежит на b, тогда чтобы получить вектор проекции, мы хотим использовать направление вектора b (единичный вектор) и умножить получившийся вектор на длину проекции: $a^{‘} = \frac{b}{|b|} \cdot |a^{‘}|$. Подставляем $|a^{‘}|$, $\cos\phi$ заменяем на $\frac{(a,b)}{|a| \cdot |b|}$, получили требуемое.

Теорема

Точка (x, y) принадлежит прямой l (прямая задана точкой (x_0, y_0) и направляющим вектором (α, β)) тогда и только тогда, когда $\frac{x - x_0}{\alpha} = \frac{y - y_0}{\beta}$. Это равенство мы будем называть каноническим уравнением прямой. В эту же теорему включим вывод других способов задать прямую

Доказательство:
Очев: если у нас точка (x, y) лежит на прямой, тогда у нас вектора $(x - x_0, y - y_0)$ и (α, β) колинеарны, тогда $\exists\ k \in \mathbb{R} \mid (x - x_0, y - y_0) = k \cdot (\alpha, \beta)$. Рассмотрев это равенство покоординатно, получим требуемое отношение. В обратную сторону аналогично, просто введём k, скажем про коллинеарность, дальше принадлежность точки прямой очевидна.

Из получившегося уравнения очевным образом получаем параметрическое уравнение прямой:

$$\begin{cases} x = x_0 + t \cdot \alpha \\ y = y_0 + t \cdot \beta \end{cases}$$

По сути заменили k на t. Далее получим общее уравнение прямой $Ax+By+C=0$. Просто возьмём каноническое и крест-накрест перемножим. Получим $\beta x - \alpha y - x_o\beta + y_0\alpha = 0$. Дальше мы просто занимаемся переобозначением.

Замечание

Также заметим, что вектор с координатами (A, B) (читать как $(\beta, -\alpha)$) - это нормаль-вектор нашей прямой. Проверяется через скалярное произведение (помним, что (α, β) - это направляющий вектор нашей прямой).

Теорема

Прямая $l : Ax + By + C = 0$ разбивает плоскость на 2 полуплоскости. Если мы возьмём какие-то 2 точки I_1, I_2 из разных полуплоскостей, тогда $l(I_1) \cdot l(I_2) < 0$

Доказательство:
Зафиксируем точку $(x_0, y_0) \in l$. Теперь рассмотрим скалярное произведение нормаль-вектора и $(x_1 - x_0, y_1 - y_0)$ (если считать, что у точки I_1 координаты (x_1, y_1)). Аналогично для второй точки I_2. Тогда одно скалярное произведение будет > 0, а другое < 0 в силу свойства скалярного (если точнее, то просто пользуемся, что косинус тупого угла отрицательный).
Например, подробнее для точки I_1 из верхней полуплоскости (БОО): $A(x_1 - x_0) + B(y_1 - y_0) > 0$, раскроем скобки, обозначим $C = -Ax_0 - By_0$. Потом мы всё это перемножим с выражением для второй точки: $Ax_2 + Bx_2 + C < 0$. Получим то, что и хотели: $l(I_1) \cdot l(I_2) < 0$.

Теорема

Формула расстояния от точки (x_0, y_0) до прямой $l : Ax + By + C = 0$ - это $d(x_0, y_0) = \frac{|Ax_0 + By_0 + C|}{\sqrt{A^2+B^2}}$

Доказательство:
Рассмотрим точку $(x, y) \in l$. Теперь построим вектор $(x_0 - x, y_0 - y)$, спроецируем его на нормаль вектор (точнее мы хотим посмотреть на длину проекции): $|(A, B)| \cdot |\frac{A(x_0-x)+B(y_0-y)}{A^2+B^2}|$. $|(A, B)|$ - это внезапно $\sqrt{A^2 + B^2}$. Сокращаем, вводим обозначение C, получаем требуемое.

Замечание

Обсудим взаимное расположение прямых: совпадают, параллельны или пересекаются. В терминах коэффициентов это соответственно $\frac{A_1}{A_2} = \frac{B_1}{B_2} = \frac{C_1}{C_2}$, $\frac{A_1}{A_2} = \frac{B_1}{B_2}$ или никакое из предыдущих равенств не выполняется.

В терминах матриц совпадение это:

$$rk\begin{pmatrix} A_1 & B_1 \\ A_2 & B_2 \end{pmatrix} = 1$$

$$rk\begin{pmatrix} A_1 & B_1 & C_1 \\ A_2 & B_2 & C_2 \end{pmatrix} = 1$$

Аналогично для параллельности у нас ранг большой матрицы будет 2, а в случае пересечения у нас ранг и маленькой, и большой матрицы будет 2.

Определение

Ранг матрицы (rk) - это максимальный порядок минора матрицы, отличный от нуля. Иными словами, это число, равное максимальному количеству линейно независимых строк (или столбцов) в матрице. Ранг матрицы показывает размерность подпространства, натянутого на строки (или столбцы) матрицы.

Теорема

Площадь параллелограмма, построенного на векторах (a, c) и (b, d) - это определитель:

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix}$$

Доказательство:
$S = |a| \cdot |b| \cdot \sin\phi$. Меняем синус на косинус по ОТТ, заносим всё под корень, раскрываем скобки, там у нас получается полный квадрат: $\sqrt{(ad - bc)^2}$, а это в точности определитель.

Замечание

Из такого геометрического смысла определителя становятся очевидны всякие свойства про линейность по строке, иммутабельность при транспонировании.

Такую же формулу, кстати, можно вывести для $\mathbb{R}^3$, но это будет просто более глиномесно:

$$V = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Вот теперь мы плавно перешли к пространствам. Будем волшебным образом перетаскивать формулы из плоскости, натягивать их на пространство, также будем что-то новое вводить.

Замечание

Каноническое уравнение прямой в $\mathbb{R}^3$ - это $\frac{x-x_0}{\alpha} = \frac{y-y_0}{\beta} = \frac{z-z_0}{\gamma}$.

Уравнение плоскости можно построить по трём точкам, зафиксируем первую точку, от неё проведём 2 вектора к двум оставшимся, теперь возьмём какую-то точку (x, y, z), вектор от первой точки к новой должен быть ЛНЗ. Получили 3 вектора, которые образовали плоскость. Объём, натянутый на эти 3 вектора, равен 0, тогда мы просто пишем объём через определитель, раскрываем, получаем: $Ax + By + Cz + D = 0$.

Давайте до кучи напишем параметрическое уравнение плоскости:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \lambda \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \mu \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

где u и v - базисные векторы плоскости, а (x_0, y_0, z_0) - какая-то начальная точка.

Теорема

Формула перехода к новому базису.

Доказательство:
Рассмотрим вектор (x, y) в базисе $\{e_1, e_2\}$. Тогда наш вектор $a = xe_1 + ye_2$, а в другом базисе наш вектор - это $a = x^{'} e_1^{'} + y^{'} e_2^{'}$. Хотим узнать $(x_1^{'}, y_2^{'})$. $e_1^{'} = a_{11}e_1 + a_{21}e_2$ и $e_2^{'} = a_{12}e_1 + a_{22}e_2$. Подставим и преобразуем, потом воспользуемся единственностью представления вектора в данном базисе, тогда $x = x^{'} a_{11} + y^{'} a_{12}$, аналогично для y.

Замечание

Вопрос исследования взаимного расположения прямой и плоскости, плоскости и плоскости довольно тривиальный, просто пользуемся параметрическим уравнением плоскости и прямой, а дальше очев: сводим задачу к исследованию расположения направляющего вектора прямой и нормаль-вектора плоскости и тд и тп, просто системы уравнений.

Определение

Угол между прямой и плоскостью - это угол между прямой и проекцией прямой на данную плоскость.

Теорема

Формула расстояния от точки (x_0, y_0, z_0) до плоскости $\alpha : Ax + By + Cz + D = 0$ - это $d(x_0, y_0, z_0) = \frac{|Ax_0 + By_0 + Cz_0 + D)|}{\sqrt{A^2 + B^2 + C^2}}$

Доказательство:
Выводится также, как и для двумерного случая.

Замечание

Расстояние между плоскостями или прямыми, или прямыми и плоскостями - это длина общего перпендикуляра. Задача сводится к расстоянию от точки до плоскости/прямой.

Домашнее задание

Из Смирнова (`smirnov.pdf`) решайте все задачи из 3.1-3.4, 4.1-4.4 и 474-482. Не обяз решать всё, решайте пока не почувствуете, что прониклись. Задачи халява, все идеи есть в лекциях.

3 Векторные пространства и матрицы

Определение

Пусть у нас есть векторное пространство размерности n, тогда набор векторов $\{e_1, \ldots, e_n\}$ будет называться **базисом** этого пространства, если они все линейно независимы: $\sum \lambda_i \cdot e_i = 0 \Rightarrow \lambda_i = 0$.

Теорема

У любого конечномерного векторного пространства существует базис.

Доказательство:
Просто предъявляем алгоритм. Возьмём $e_1 = v_1 \neq 0$, потом возьмём $e_2 = v_2 \in V$ такой, что $\lambda_1 \cdot e_1 + \lambda_2 \cdot e_2 = 0 \Rightarrow \lambda_1 = 0, \lambda_2 = 0$. Если такой не нашёлся, тогда у нас базис из 1 вектора, имеем просто одномернопространство. Потом возьмём $e_3 = v_3 \in V \ \ \ldots$ и тд.

Замечание

Система векторов обладает единственным базисом только в случае 0-мерного пространства.

Определение

Векторное пространство W представимо в виде **прямой суммы** пространств U и V, если $\forall w \in W \ \exists u \in U \ \& \ \exists v \in V \mid w = u + v$. Очень важно, что U и V пересекаются только по нулевому вектору.

Теорема

а) Размерность подпространства не превосходит размерности пространства.
б) W - векторное пространство, U - его подпространство, тогда $\exists V$ такое, что $W = U + V$.

Доказательство:
а) очевидно, пытливый читатеть может самостоятельно привести доказательство этого пункта.

б) Выбираем базис в U, потом дополняем его до всего базиса W, тогда $\forall w \in W$ $w = \sum_{i=1}^{k} w_i \cdot e_i + \sum_{i=k+1}^{n} w_i \cdot e_i$, где $k = dimU, n = dimV$. Тогда первая сумма у нас лежит в U, а вот то, что осталось мы определим как V, тогда базис нового пространства - это просто те, вектора, которыми мы дополнили базис U до базиса W.

Замечание

Если V - в.п. ($dimV = n$) над полем из q элементов, тогда всего векторов у нас q^n (потому что $v = \sum q_i \cdot v_i$), а способов выбрать базис - $(q^n - 1)(q^n - q) \dots (q^n - q^{n-1})$ (первым берём любой **ненулевой** вектор, потом берем вектор, который ЛНЗ с первым, то есть $e_2 \neq \lambda_1 \cdot e_1$, на место лямбды q вариантов и тд).

Теорема

Ранг матрицы $A|B$ (*это приписывание матрицы B справа от матрицы A*) не превосходит суммы рангов матриц A и B.

Доказательство:
$A|B = A|0 + 0|B$, тогда $rk(A|B) = rk((A|0) + (0|B)) \leq rk(A|0) + rk(0|B) = rk(A) + rk(B)$.

Теорема

Всякую матрицу ранга r можно представить в виде суммы r матриц ранга 1, но нельзя представить в виде суммы меньшего числа таких матриц.

Доказательство:
$rkA = r, A = \sum A_i^{`}$. Без ограничения общности будем считать, что у нас ЛНЗ первые r строчек матрицы:

$$\begin{pmatrix} A_1 \\ A_2 \\ \dots \\ A_r \\ A_{r+1} \\ \dots \\ A_n \end{pmatrix} = \begin{pmatrix} A_1 \\ 0 \\ \dots \\ 0 \\ \lambda_1^{r+1} A_1 \\ \lambda_1^{r+2} A_1 \\ \dots \end{pmatrix} + \begin{pmatrix} 0 \\ A_2 \\ \dots \\ 0 \\ \lambda_2^{r+1} A_2 \\ \lambda_2^{r+2} A_2 \\ \dots \end{pmatrix} + \dots$$

$rkA = rk(\sum A_k) \leq \sum rk(A_k) = k$, если k < r, тогда $rkA < r$ - противоречие.

Теорема

$A^T A = A A^T \Rightarrow (A^{-1})^T = A^{-1}$.

Доказательство:
$A^{-1} A = E$, транспонируем $(A^{-1} A)^T = E^T \Leftrightarrow A^T (A^{-1})^T = E^T$, домножим слева на A^{-1}, получим $(A^{-1})^T = A^{-1}$

Далее в лекции разобраны несколько опорных задач, связанных с коммутативностью и обратимостью матрицы. Доказательства как правило проводились через **след матрицы**, потому что работать с числами куда приятнее и понятнее, чем с матрицами. Не считаю нужным конспетировать эти задачи. Может быть кто-то захочет продолжить моё дело и откроет пул реквест.

Определение

Следом квадратной матрицы A (dimA = n) мы будем называть $tr(A) = \sum_{i=1}^{n} a_{ii}$.

След становится удобным инструментом в доказательстве теорем про матрицы благодаря ряду свойств:

1. $tr(\alpha A + \beta B) = \alpha \cdot tr(A) + \beta \cdot tr(B)$
2. $tr(C^{-1}AC) = tr(A)$, в частности $tr(AB) = tr(BA)$
3. $tr(A) = tr(A^T)$
4. **След матрицы** равен сумме её собственных значений.

Замечание

Строковый и столбцовый ранг совпадают.

Определение

Симметрические матрицы: $A = A^T$, кососимметрические матрицы: $A^T = -A$. Также заметим, что в последнем случае у нас обязательно на диагонали должны быть нули, т.к. $a_{ii} = -a_{ii} \Leftrightarrow a_{ii} = 0$, а остальные элементы $a_{ij} = -a_{ji}$.

Замечание

Рассмотрми пространства симметрических матриц (U) и кососимметрических матриц (V). Размерность первого пространства - это $\frac{n^2+n}{2}$ (потому что такая матрицы задаётся с помощью n чисел на диагонали + количество чисел над диагональю, для этого нужно из всех чисел матрицы вычесть диагональ и поделить пополам - $\frac{n^2-n}{2}$. Размерность второго пространства тогда - это $\frac{n^2-n}{2}$ (раз на диагонали только нули).

Теперь сложим эти размерности $\frac{n^2+n}{2} + \frac{n^2-n}{2} = \frac{2n^2}{2} = n^2$. Получили размерность всего пространства квадратных матриц $M_n(\mathbb{R})$. Также заметим, что пространства симметрических и кососимметрических матриц пересекаются только по нулевой матрице (то есть $A = -A$). Тогда $A \in M_n(\mathbb{R}) \Rightarrow \exists U, V \; A = U \oplus V$ причём U - симметрическая матрица, а V - кососимметрическая матрица.

Теорема

$rk(A + B) \leq rk(A) + rk(B)$

Доказательство:

Идейно: $rk(A+B) = rk(A) + rk(B) - rk(A \cap B) \le rk(A) + rk(B)$

Замечание

$tr(AB) = tr(BA)$, но $rk(AB) \neq rk(BA)$

Определение

$\mathbb{R}_n[x]$ - пространство многочленов, базисом которого может быть, например, $\{1, x, x^2, \dots, x^n\}$.

Скалярное произведение двух функций f и g - это $(f, g) = \int_a^b f(x)g(x)dx$.

Замечание

А скалярное произведение матриц - это tr их произведения. Можете прогнать по свойствам скалярного произведения и убедиться в этом.

Теорема

Рассмотрим набор векторов $\{e_1, e_2, \dots, e_n\}$ со следующим свойством:

$$\begin{cases} (e_i, e_i) = 1 \\ (e_i, e_j) = 0 \end{cases}$$

Докажем, что этот набор векторов является базисом.

Доказательство:
Проверим ЛНЗ. Хотим $\sum \lambda_i e_i = 0 \Rightarrow \forall i\ \lambda_i = 0$. Рассмотрим скалярное произведение $\forall i\ (e_i, \sum \lambda_i e_i) = 0$, раскроем по свойству линейности, получим что-то такое: $\sum \lambda_j (e_i, e_j) = \lambda_i (e_i, e_i) = 0 \Rightarrow \lambda_i = 0$.

Определение

Набор векторов из теоремы выше называется **ортонормированным базисом**.

Определение

$W^{\perp} := \{w^{\perp} \in W^{\perp} \mid \forall w \in W\ (w^{\perp}, w) = 0\}$

Замечание

$V = W \oplus W^{\perp}$. Очевидно, что W и $W^{\perp}$ пересекаются только по нулю, если бы мы нашли какой-то $x \in W, W^{\perp}$, то получили бы что-то в духе $(x, x) = 0$, а отсюда по свойству скалярного произведения получаем, что $x = 0$.

Теорема

Метод Грама–Шмидта: Любой базис $\{e_1, ..., e_n\}$ евклидова пространства можно преобразовать в ортонормированный базис $\{f_1, ..., f_n\}$ следующим образом:

$$
\begin{aligned}
u_1 &= e_1,\\
u_2 &= e_2 - \frac{(e_2, u_1)}{(u_1, u_1)} u_1,\\
u_3 &= e_3 - \frac{(e_3, u_1)}{(u_1, u_1)} u_1 - \frac{(e_3, u_2)}{(u_2, u_2)} u_2,\\
&\vdots\\
f_i &= \frac{u_i}{\|u_i\|}.
\end{aligned}
$$

Теорема

Пусть A - матрица размера $n \times n$. Если для любой матрицы X размера $n \times n$ справедливо равенство $tr(AX) = 0$, то $A = 0$.

Доказательство:
Попробуем $X = A^T$. $tr(AA^T) = \sum a_{ij}^2 = 0 \Rightarrow \forall i, j \; a_{ij} = 0$. Либо можно сказать, что у нас $tr(AX)$ - это скалярное произведение на пространстве матриц, причём у нас $\forall X \; tr(AX) = 0$, то есть A перпендикулярно любому вектору, а это возможно в том случае, если A - это нулевой вектор.

Домашнее задание

Домашка есть в `2-vector-spaces-102.pdf`

4 Линейные операторы

Определение

Отображение $f : V \to V$ мы будем называть **линейным**, если оно удовлетворяет следующим свойствам:

1. $f(u + v) = f(u) + f(v)$
2. $f(\lambda v) = \lambda f(v)$

Замечание

Если отображение задаётся матрицей, то оно линейно.
Стобцы в матрице линейного отображения - это образы базисных векторов.

Определение

Движением мы будем называть отображение $f : \mathbb{R}^2 \to \mathbb{R}^2$, которое сохраняет расстояние между образами: $d(x, y) = d(f(x), f(y))$.

Теорема

Движение - это биекция и у него всегда есть обратное отображение-движение.

Доказательство:
Пусть у нас есть две точки x_1, x_2 (различные), они не могут перейти в какую-то одну точку $f(x_1) = f(x_2)$, иначе у нас нарушится свойство движения: $d(x_1, x_2) \neq d(f(x_1), f(x_2)) = 0$.

Теорема

Движение сохраняет углы.

Доказательство:
Рассмотрим треугольник. Он перейдёт в какой-то другой треугольник, причём равный изначальному по трём сторонам. Раз треугольники равны, то равны и углы.

Теорема

Теорема Шаля: Идейно - всякое движение есть поворот, симметрия или композиция симметрий и поворотов.

Доказательство:
Начнём с частного случая теоремы, разберем движения, которые сохраняют точку $(0, 0)$. Рассмотрим куда у нас перейдут базисные вектора (e_1, e_2). По сути мы хотим получить матрицу:

$$\begin{pmatrix} f(e_1) & f(e_2) \end{pmatrix}$$

Пусть у нас угол между e_1 и $f(e_1)$ будет равен ϕ. Тогда:

$$f(e_1) \to \begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}$$

Аналогично для вектора e_2. Заметим, что раз $e_1 \perp e_2$, то $f(e_1) \perp f(e_2)$, поэтому угол между $f(e_2)$ и e_2 также выражается через ϕ. С учётом ориентации получаем:

$$f(e_2) \to \begin{pmatrix} -\sin\phi \\ \cos\phi \end{pmatrix}$$

По итогу просто имеем **матрицу поворота** :)

$$\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

НО! У нас ведь $f(e_2)$ мог "смотреть"в противоположную сторону, тогда матрица для $f(e_1)$ остаётся такой же, а вот вторая немного меняется, имеем внезапно **матрицу симметрии**:

$$\begin{pmatrix} \cos\phi & \sin\phi \\ \sin\phi & -\cos\phi \end{pmatrix}$$

Если какое-то отображение не сохраняет точку ноль, то нам ничто не мешает начать откладывать базисные вектора относительно $f(0)$, таким образом, все движения, которые сохраняют 0 выражаются примерно так: $f(x) = Ax + \binom{x_0}{y_0}$, где $\binom{x_0}{y_0}$ - это координаты точки $f(0)$.

Определение

Обратным отображением в терминах матриц мы будем называть такое отображение A^{-1}, что $A \circ A^{-1} = A^{-1} \circ A = E$.

Теорема

Матрица обратима, если ее определитель отличен от нуля.

Доказательство:
Можем переформулировать $\exists A^{-1} \;\Leftrightarrow\; \{f(e_1), f(e_2)\}$ – тоже базис. Тогда из этого следует, что у нас есть обратное отображение. Почему? См. доказательство про обратимость движения. Также берём различные вектора u и v, а потом внезапно получаем $f(u) = f(v)$, тогда $u_1 f(e_1) + u_2 f(e_2) = v_1 f(e_1) + v_2 f(e_2) \Rightarrow (u_1 - v_1) f(e_1) + (u_2 - v_2) f(e_2) \Rightarrow u_1 - v_1 = u_2 - v_2 = 0$ - противоречие с различностью u и v. Стрелочка влево доказана.

Теперь в другую сторону. От противного, $f(e_1), f(e_2)$ не образуют базис. Тогда $\lambda_1 f(e_1) + \lambda_2 f(e_2) = 0$, причём коэффициенты не равны 0. По линейности также имеем $f(\lambda_1 e_1 + \lambda_2 e_2) = 0$, но $f(0) = 0$. То есть какие-то 2 различных вектора перешли в один. Противоречие.

Таким образом у нас есть обратная матрица тогда и только тогда, когда вектора $f(e_1), f(e_2)$ не коллинеарны, а это значит, что определитель не равен 0.

Замечание

У нас определитель равен 0 может быть, если он состоит из линейно зависимых векторов. Разберём на примере 2×2:

$$\begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - cb = 0 \Rightarrow \frac{a}{b} = \frac{c}{d} = k$$

Получили условие коллинеарности. Этим свойством определителя мы и пользуемся в теореме выше.

Замечание

Алгоритм поиска обратной матриц довольно прост, если хотим найти A^{-1}, тогда рассмотрим матрицу $A|E$ и с помощью элементарных преобразований пытаемся превратить нашу матрицу в $E|X$, тогда $X = A^{-1}$.

Замечание

$(AB)^{-1} = B^{-1} A^{-1}$

Теорема

$det(AB) = detA \cdot detB$

Доказательство:
С геометрической точки зрения определитель - это просто площадь параллелограмма, натянутого на базисные вектора. $C = AB$ - это же просто композиция. Сначала B как-то растянул нашу площадь, потом A. Соответственно $detA \cdot detB$ - это то же самое, сначала мы растягиваемся в A раз, потом в B раз.

Замечание

Замена базиса линейного отображения. Пусть у нас есть $f : V \to W$, $x = C_1 x^{‘}, y = C_2 y^{‘}, y = A_f x$, где C_1, C_2 - матрицы замены координат, после подстановки получим $y^{‘} = C_2^{-1} A_f C_1 x^{‘}$.

Определение

Пусть дано $f : V \to W$, тогда $Kerf = \{v \in V \mid f(v) = 0\}$ называется **ядром** f, а $Imf = \{w \in W \mid \exists v \in V \ f(v) = w\}$ называется **образом** f.

Замечание

$rk(A_f) = dim(Imf)$

Полезное замечание, выводится оно из $dim(\langle f(e_1), f(e_2), \ldots, f(e_n)\rangle)$ - линейная оболочка, состоящая из образов базисных векторов. Теперь получим ещё более интересные свойства:

1. $rkAB \leq rkA$ и $rkAB \leq rkB$, в частности $rk(C^{-1}AC) = rkA$
2. $rk(A + B) \leq rkA + rkB$

Определение

Диагональной матрицей мы будем называть матрицы вида:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ & & \ldots & & \\ 0 & 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

Определение

Хотим научиться преобразовывать нашу матрицу к диагональному виду. Помним, что столбцы у нас - это образы базисных векторов, тогда имеем:

$$\begin{cases} f(e_1) = \lambda_1 e_1 \\ f(e_2) = \lambda_2 e_2 \\ \dots \\ f(e_n) = \lambda_n e_n \end{cases}$$

В общем виде хотим научиться такие λ и v такие, что $Av = \lambda v$. Такие числа λ мы будем называть **собственными числами**, а вектора v - **собственными векторами**.

$Av = \lambda v \Leftrightarrow Av = \lambda Ev \Leftrightarrow (A - \lambda E)v = 0$, также будем считать, что $v \neq 0$, иначе уж совсем неинтересно. Тогда у нас v принадлежит ядру, а значит матрица у нас вырожденная, то есть $det(A - \lambda E) = 0$.
Полученное уравнение мы будем называть **характеристическим многочленом**. Решаем его, получаем λ (возможно несколько), подставляем их в исходное, решаем уравнение на v. Теперь мы имеем и новый базис (матрица перехода по сути), и диагональный вид.

Замечание

Если существует ненулевой вектор v, такой что $(A-\lambda E)v = 0$, то это означает, что у матрицы $A - \lambda E)$ есть ненулевое решение однородной системы, то есть её ядро непусто, а раз у матрицы ненулевое ядро, то она не обратима, то есть вырождена.

Теорема

Теорема о ядре и образе: $dimkerA + dimImA = dimA$

Замечание

Матрица диагонализуема, если существует базис из собственных векторов.

5 Определители

Чтобы разобраться с темой определители, надо сначала познакомиться с перестановками, потом поймёте почему. Сейчас ВЫ НЕ ГОТОВЫ.

Определение

Перестановкой называется биективное отображение множества первых n натуральных чисел на себя. Формально:
Перестановкой степени n называется любая биекция:

$$\sigma\colon \{1, 2, \dots, n\} \to \{1, 2, \dots, n\}.$$

Множество всех перестановок степени n обозначается S_n и называется *симметрической группой* степени n.

Перестановки, на мой взгляд, - довольно непростая тема, потому что легко запутаться, поэтому давайте разберём примеры, поймём как это вообще всё работает: что и куда переходит.

Пример

Матричная запись перестановки и как мы можем сократить запись с помощью **циклических перестановок**

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = (1\ 3)(2\ 4)$$

Вот эта сокращённая запись означает **цикл**, (1 3) - это 1 перешла в 3, а 3 перешла в 1, цикл закончился, потом мы описываем второй цикл. Если элемент статичен (никуда не переходит), мы считаем, что он переходит в себя.

Пример

Умножение перестановок выполняется справа налево (как композиция функций).

$$\tau = (1\ 3\ 5)(2\ 4\ 6\ 7), \quad \sigma = (1\ 4\ 7)(2\ 3\ 5\ 6)$$

Вычисляем $\tau \circ \sigma$:

1. Для элемента 1: $\sigma(1) = 4$, $\tau(4) = 6 \Rightarrow 1 \to 6$
2. Для элемента 2: $\sigma(2) = 3$, $\tau(3) = 5 \Rightarrow 2 \to 5$
3. Для элемента 3: $\sigma(3) = 5$, $\tau(5) = 1 \Rightarrow 3 \to 1$
4. Для элемента 4: $\sigma(4) = 7$, $\tau(7) = 2 \Rightarrow 4 \to 2$
5. Для элемента 5: $\sigma(5) = 6$, $\tau(6) = 7 \Rightarrow 5 \to 7$
6. Для элемента 6: $\sigma(6) = 2$, $\tau(2) = 4 \Rightarrow 6 \to 4$
7. Для элемента 7: $\sigma(7) = 1$, $\tau(1) = 3 \Rightarrow 7 \to 3$

Результат:

$$\tau \circ \sigma = (1\ 6\ 4\ 2\ 5\ 7\ 3)$$

Определение

Транспозиция - это перестановка длины 2: $(i_1\ i_2)$

Теорема

Любую перестановку можно представить как произведение транспозиций.

Определение

Знак перестановки можно задать так:
$sgn(i_1\ i_2) = -1$
$sgn(\tau\ \sigma) = sgn(\tau) \cdot sgn(\sigma)$

Замечание

Знак **цикла** - это число транспозиций, в которое он раскладывается. Имеется в виду $(-1)^k$, где k - число транспозиций.

Вот теперь, маслята, вы готовы к детерминанту.

Определение

Определитель (или детерминант) - это:

$$det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & \dots & \dots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} = \sum_{\sigma \in S_n} sgn(\sigma) a_{1p_1} a_{2p_2} \dots a_{np_n}$$

На всякий случай уточню, что $\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ p_1 & p_2 & p_3 & \dots & p_n \end{pmatrix}$

Страшно? Видимо не были готовы, давайте разберём пример, чтобы стало хоть немного потятнее, и эта формула выглядела осмысленнее. Возьмём матрицу 2×2:

Пример

Заметьте, что первый индекс - это у нас строки, а второй индекс - это столбец, перестановки именно по столбцу! Запись a_{1p_1} означает, что 1 переходит в p_1. e - это тождественная перестановка, которая оставляет всё на своих местах.

$$det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = sgn(e) \cdot a_{11} a_{22} + sgn(1\ 2) \cdot a_{12} a_{21}$$

$\sigma \in S_2 = \{e, (1\ 2)\}$, потому что $n = 2$, S_1 нет смысла рассматривать.

Теорема

Свойства определителя:

1. Определитель линеен по строкам
2. $det(A) = det(A^T)$
3. Определитель кососимметричен
4. Если у нас есть линейно зависимые строки, то определитель равен нулю

Доказательство:

1. $$\begin{vmatrix} A_1 + B_1 \\ A_2 \\ \dots \\ A_n \end{vmatrix} = \begin{vmatrix} A_1 \\ A_2 \\ \dots \\ A_n \end{vmatrix} + \begin{vmatrix} B_1 \\ A_2 \\ \dots \\ A_n \end{vmatrix}$$

   Просто пользуемся дистрибутивностью: $\sum_{\sigma \in S_n} sgn(\sigma)(a_{1p_1} + b_{1p_1}) a_{2p_2} \dots a_{np_n} = \sum_{\sigma \in S_n} sgn(\sigma) a_{1p_1} a_{2p_2} \dots a_{np_n} + \sum_{\sigma \in S_n} sgn(\sigma) b_{1p_1} a_{2p_2} \dots a_{np_n}$

Аналогично доказывается для домножения на скаляр λ. ‘

2. Пользуемся всё тем же определением, при A^T у нас просто a_{1p_1} превращается в $a_{p_1 1}$, но мы всё также получаем суммы по всем перестановкам. Нетрудно убедиться на примере, что мы получили действительно одно и то же.

3. Иными словами перестановка двух строк определителя приводит к смене знака. Лучше смотреть док-во в лекции, но на самом деле всё по определению.

4. Очевидно из линейности + кососимметричности: $f(A, \lambda A) = \lambda f(A, A) = -\lambda f(A, A) \Rightarrow f(A, A) = 0$

Замечание

$$\begin{vmatrix} A & B \\ 0 & C \end{vmatrix} = detA \cdot detC$$

Потому что мы просто приводим A и C к треугольному виду как-то (мы точно знаем, что это возможно), а дальше просто берём произведение диагонали, а это и есть $detA \cdot detC$

$$\begin{vmatrix} A & B \\ & 0 \end{vmatrix} = (-1)^n \begin{vmatrix} B & A \\ 0 & C \end{vmatrix} = (-1)^n detB \cdot detC$$

По аналогии, $(-1)^n$ возникает, потому что мы переставляем столбцы, а у нас det кососимметричен

Теорема

$$\begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}^2 \leq (x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2)(x_3^2 + y_3^2 + z_3^2)$$

Доказательство:
Доказывается геометрически, помним, что определитель - это по сути объём фигуры, натянутый на три вектора (вектора - это наши столбцы). Дальше просто оцениваем объём и возводим обе части неравенства в квадрат.

Определение

Алгебраическое дополнение $A_{ij} = (-1)^{i+j} M_{ij}$, где M_{ij} - это минор матрицы, полученный вычёркиванием i-й строки и j-го столбца.

Теорема

Разложение по строке (столбцу): $detA = \sum_{i=1}^{n} a_{ij} A_{ij}$

Доказательство:

См. лекция, там довольно тяжело. Начало примерно 1:10:00.

Теорема

$A^{-1} = \frac{1}{detA}A^{ad}$, где A^{ad} - матрица алгебраических дополнений.