

MTC True Tech

Talents



M T
C

ЛИНЕЙНАЯ АЛГЕБРА 2



Аслан Байрамкулов

Руководитель группы моделирования
профиля MTC Big Data

MTC Тета



Меня хорошо видно и слышно?



Проверить, идет ли запись



Ставим «+», если все хорошо
«-», если есть проблемы

МТС Тета





Аслан Байрамкулов

Руководитель группы моделирования
профиля MTC Big Data

Довелось поработать в разных направлениях:
алготрейдинг, аналитика, A/B тестирование, classic ML.

В разных компаниях: WorldQuant, Yandex, X5 Retail Group.
Чуть более 2 лет руковожу группой моделирования профиля
клиента (ранее – экспериментальное направление) в MTC Big Data.

Сфера интересов: прикладная статистика и A/B тестирование,
механизмы принятия решений, теория игр, классический ML.

Образование: ВМК МГУ (исследование операций).

ПЛАН ЗАНЯТИЯ

- Определитель матрицы
- Ранг матрицы
- Обратная матрица
- СЛАУ
- Обусловленность СЛАУ
- Собственные значения
и собственные вектора матрицы
- Матричные разложения (SVD, Truncated SVD)

ОПРЕДЕЛИТЕЛЬ МАТРИЦЫ

М Т
С



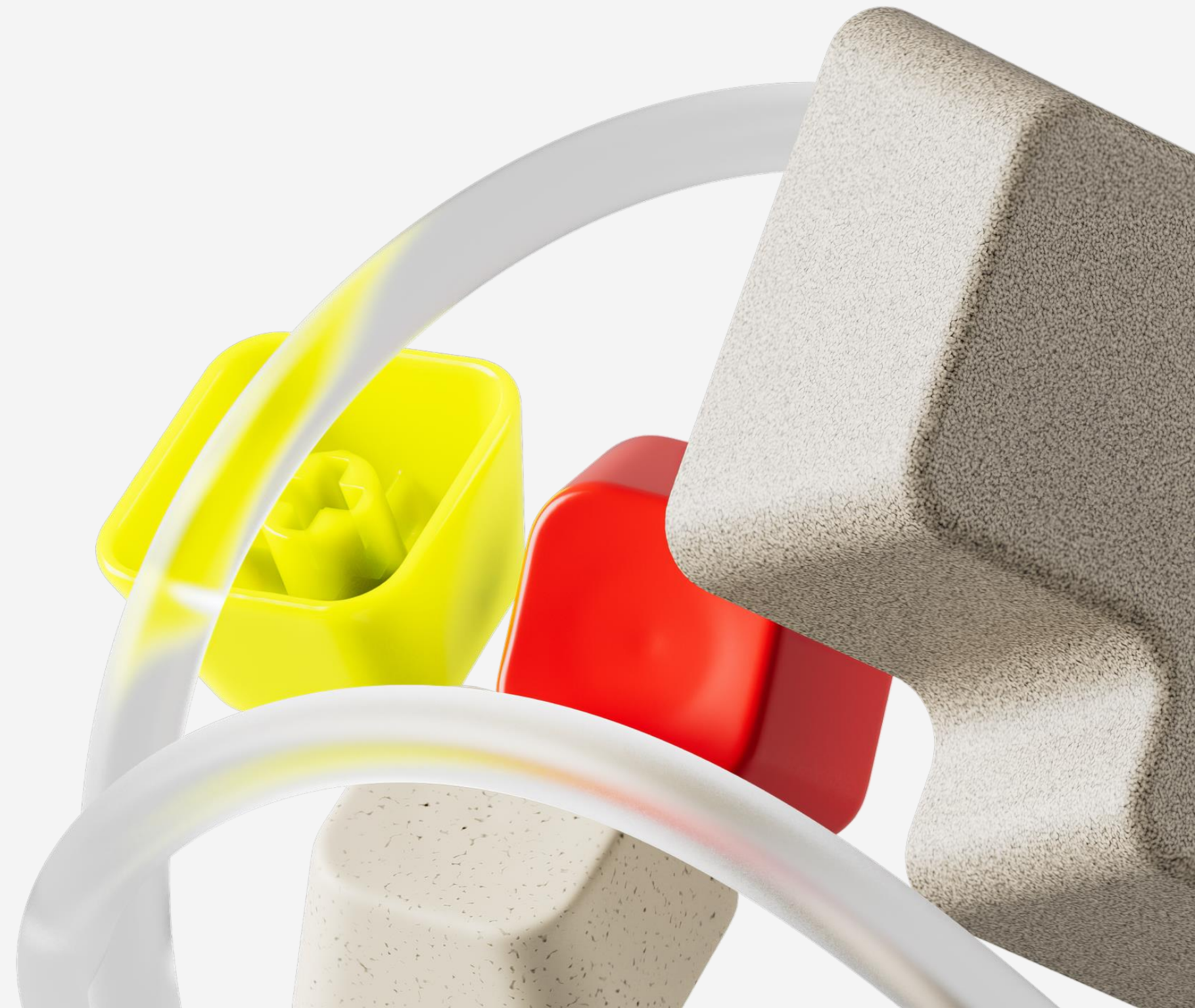
MTC Тета



Определитель 2x2: геометрический смысл

С помощью определителя можно вычислить площадь построенного на двух векторах параллелограмма на плоскости. Для этого каждый из векторов следует записать как отдельный столбец матрицы 2x2. Определитель с точностью до знака будет равен площади параллелограмма.

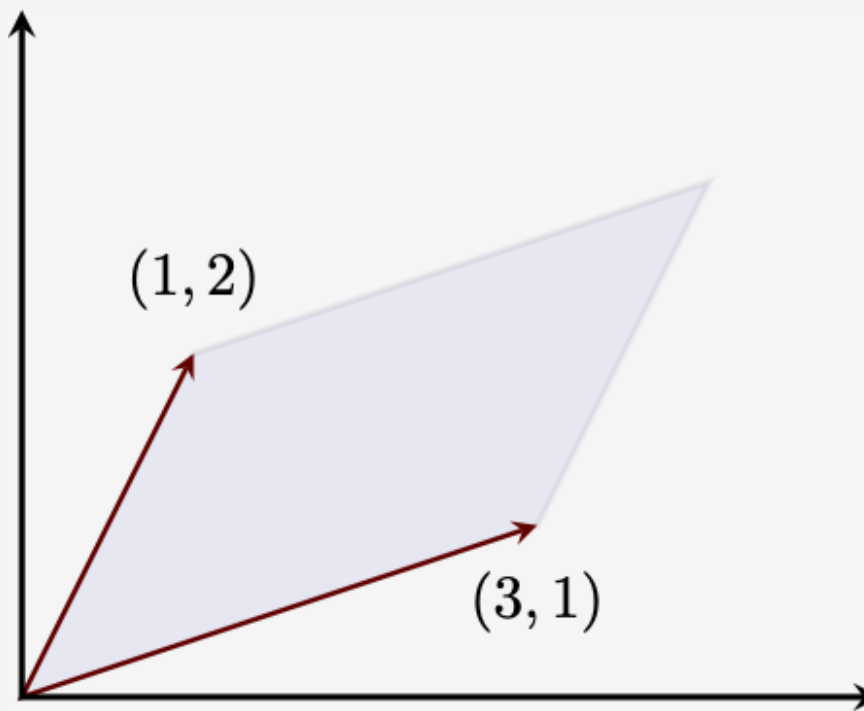
$$\det A = a_{11}a_{22} - a_{12}a_{21}, \quad S = |\det A|.$$



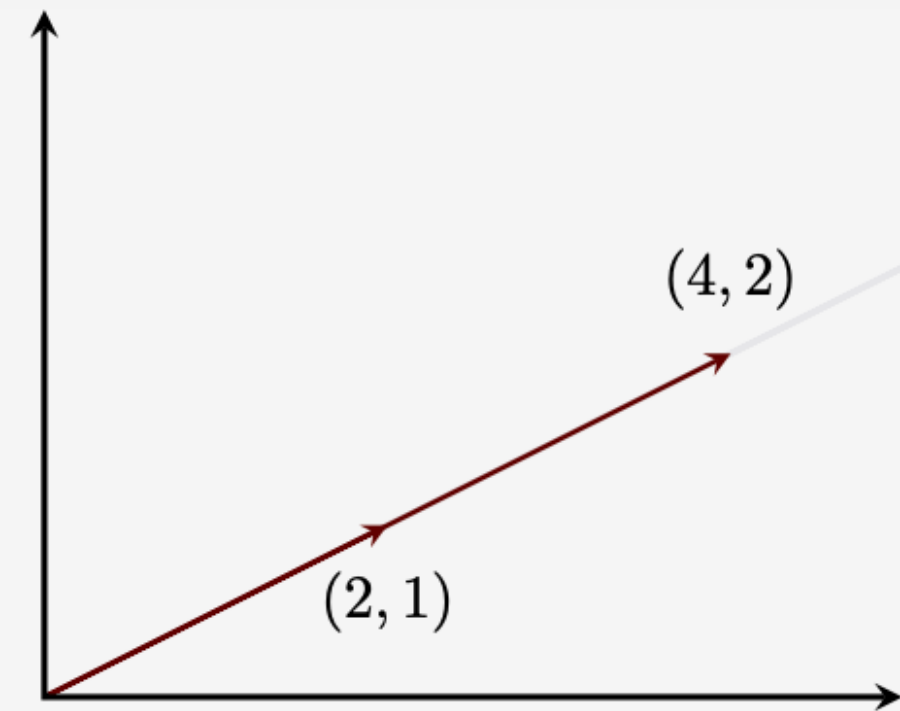
Определитель 2x2: геометрический смысл

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

$$S = |\det A| = |3 \cdot 2 - 1 \cdot 1| = 5.$$



Площадь параллелограмма.



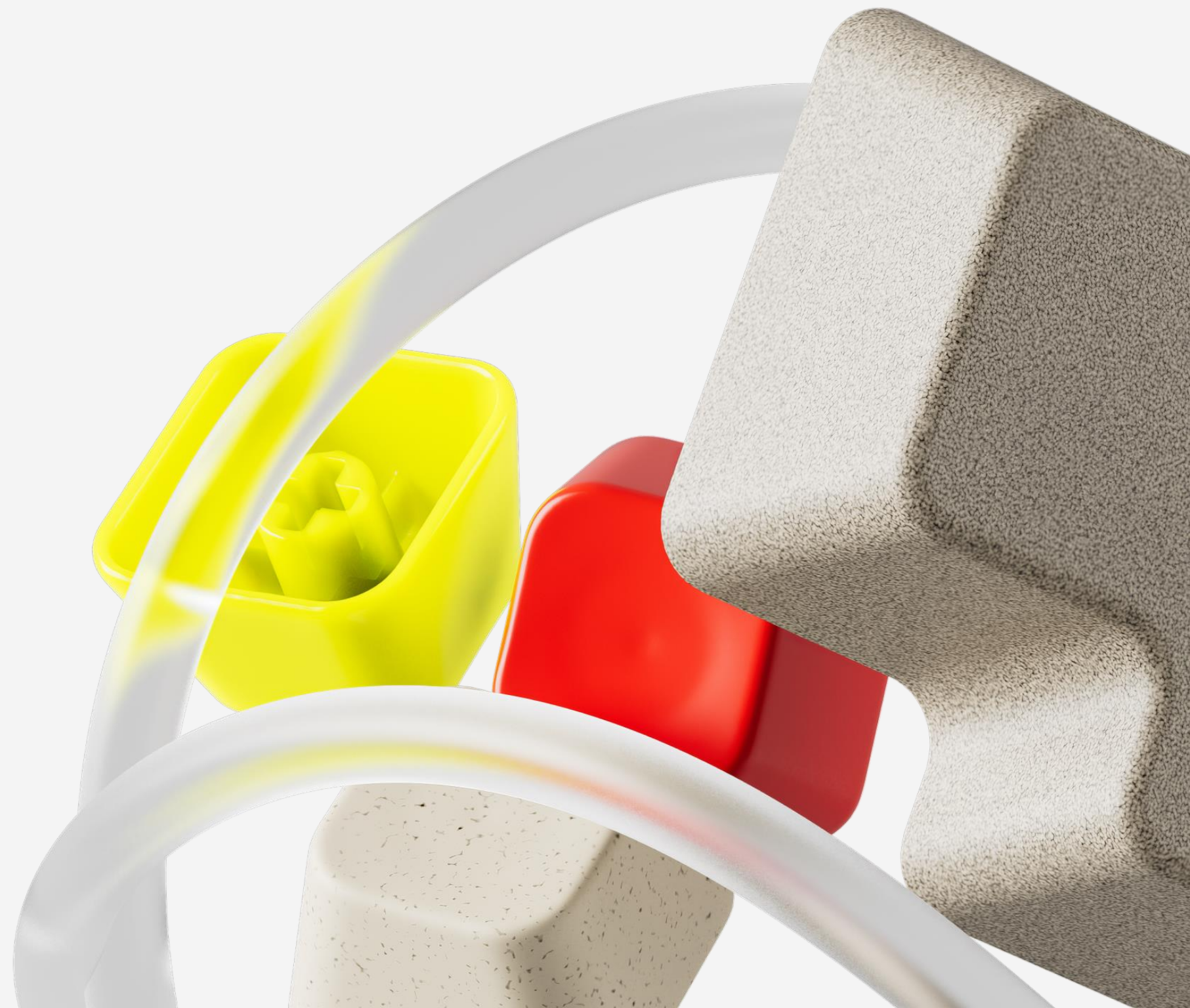
Вырожденный случай.

Определитель

- Геометрический смысл в общем случае — ориентированный объем многомерного параллелепипеда, построенного на векторах, которые являются столбцами этой матрицы
- Определение можно дать по-разному, например через рекурсивную формулу (разложение по строке):

$$\det A = \sum_{j=1}^n (-1)^{1+j} a_{1j} \bar{M}_j^1,$$

\bar{M}_j^1 — определитель матрицы, полученной из исходной вычеркиванием 1-строки и j-го столбца



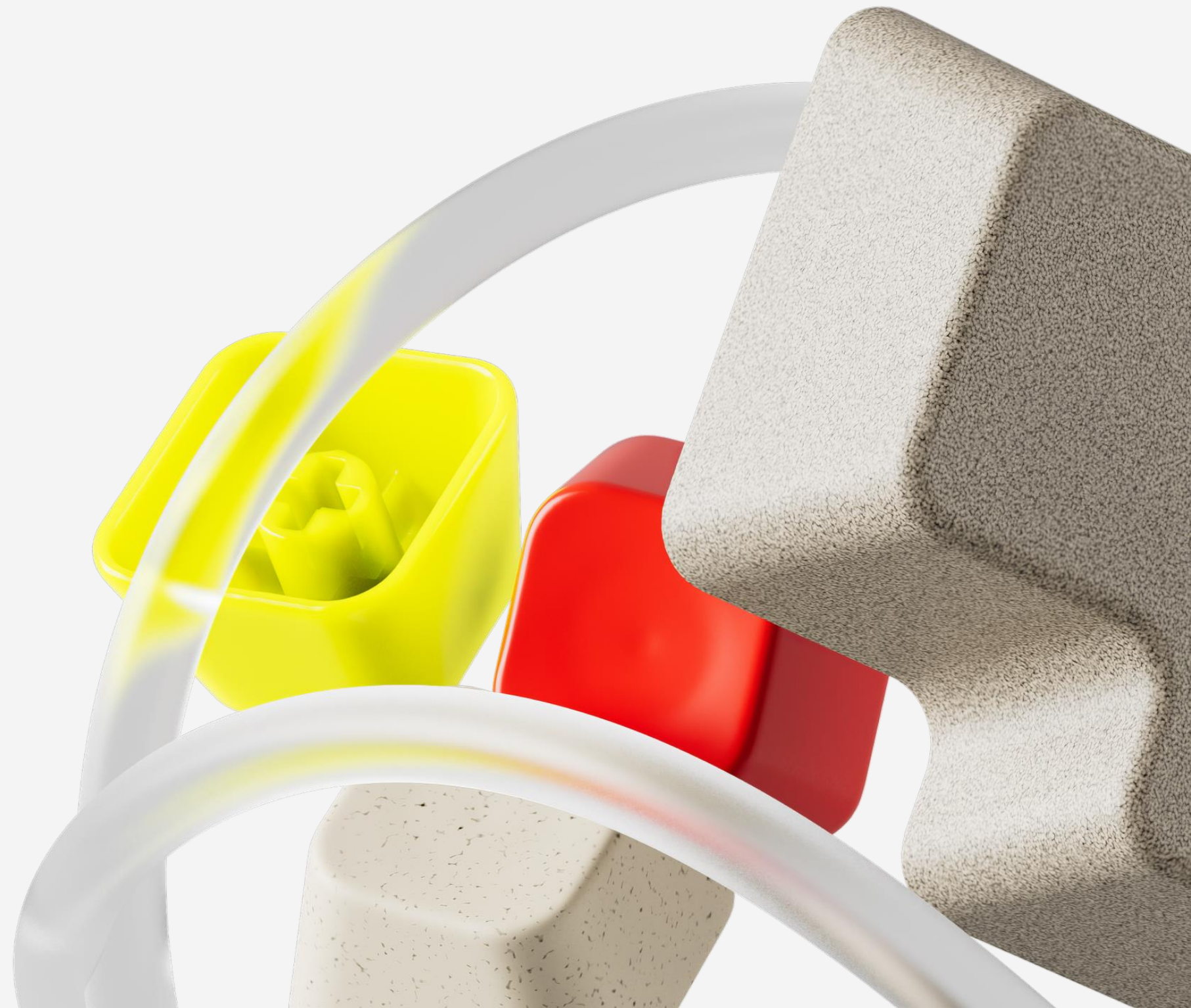
Определитель

Пример

$$\det \begin{pmatrix} 2 & 1 \\ 3 & 5 \end{pmatrix} = (-1)^0 2 * 5 + (-1)^1 1 * 3 = 7$$

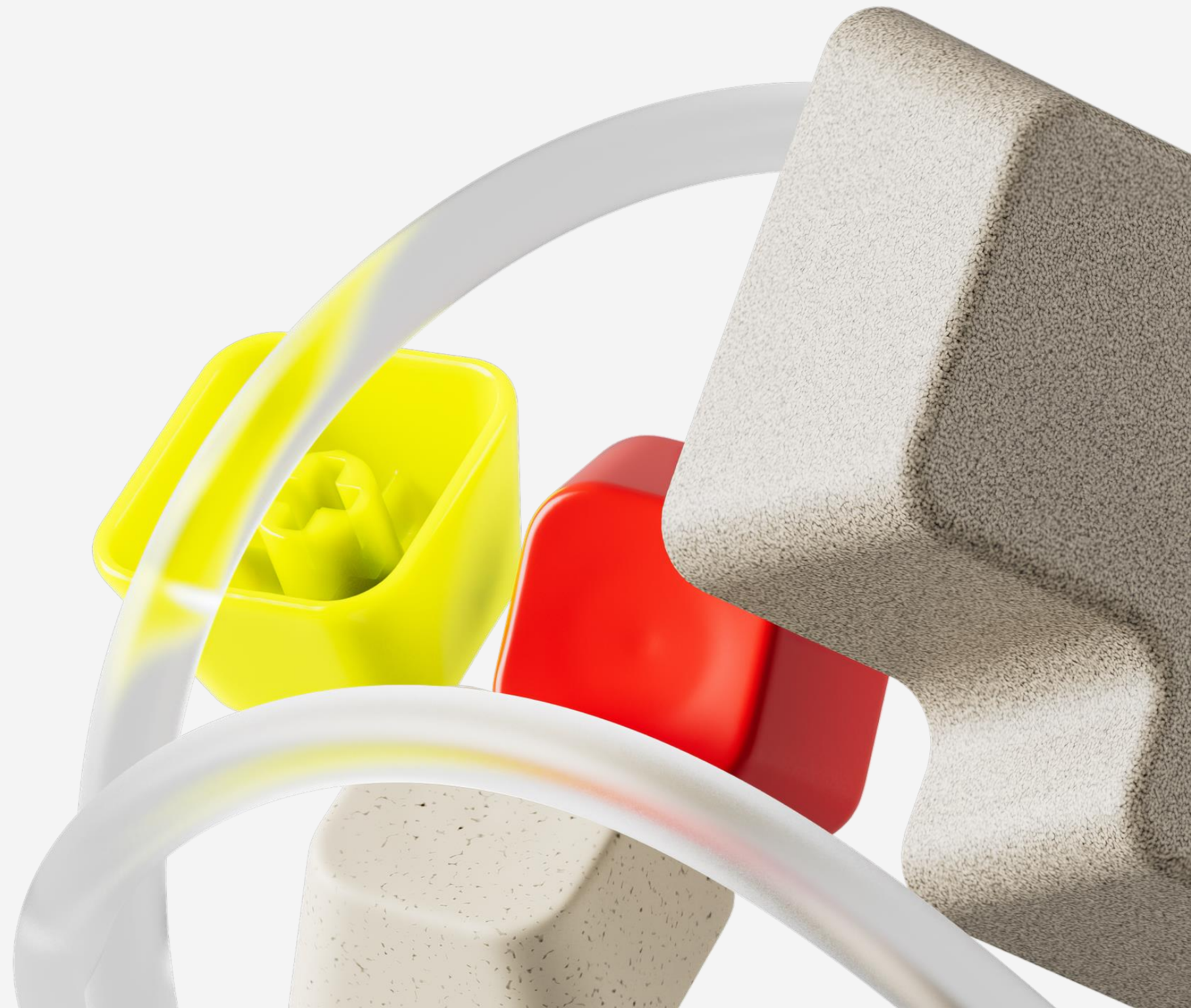
$$\det \begin{pmatrix} 2 & 1 & 7 \\ 3 & 5 & 0 \\ 1 & -4 & 4 \end{pmatrix} =$$

$$\begin{aligned} &= (-1)^0 2 * 5 * 4 + (-1)^1 2 * 0 * (-4) + (-1)^1 1 * 3 * 4 + \\ &+ (-1)^2 1 * 0 * 1 + (-1)^2 7 * 3 * (-4) + (-1)^3 7 * 5 * 1 = \\ &= 40 + 0 - 12 + 0 - 84 - 35 = -91 \end{aligned}$$



Определитель Свойства

- Определитель матрицы, содержащей линейно зависимые строки или столбцы, равен 0
- Определитель не меняется при транспонировании
- $A, B \in R^{n \times n} \Rightarrow \det(AB) = \det A \times \det B$



РАНГ МАТРИЦЫ

М Т
С



MTC Teta



Ранг матрицы

Строчный ранг —

максимальное число линейно независимых строк

Столбцовый ранг —

максимальное число линейно независимых столбцов

Строчный ранг равен столбцовому,

поэтому говорим просто о ранге матрицы



Ранг матрицы

- Рангом матрицы называют максимальный порядок её миноров, среди которых есть хотя бы один, не равный нулю
- Эквивалентные матрицы — матрицы, ранги которых равны между собой



Ранг матрицы Свойства

- При транспонировании матрицы ее ранг не изменяется
- Если вычеркнуть из матрицы нулевой ряд, то ранг матрицы не изменится
- При элементарных преобразованиях матрицы ее ранг не изменяется

Основные методы нахождения ранга

1. Метод окаймляющих векторов
2. Метод элементарных преобразований (алгоритм Гаусса)

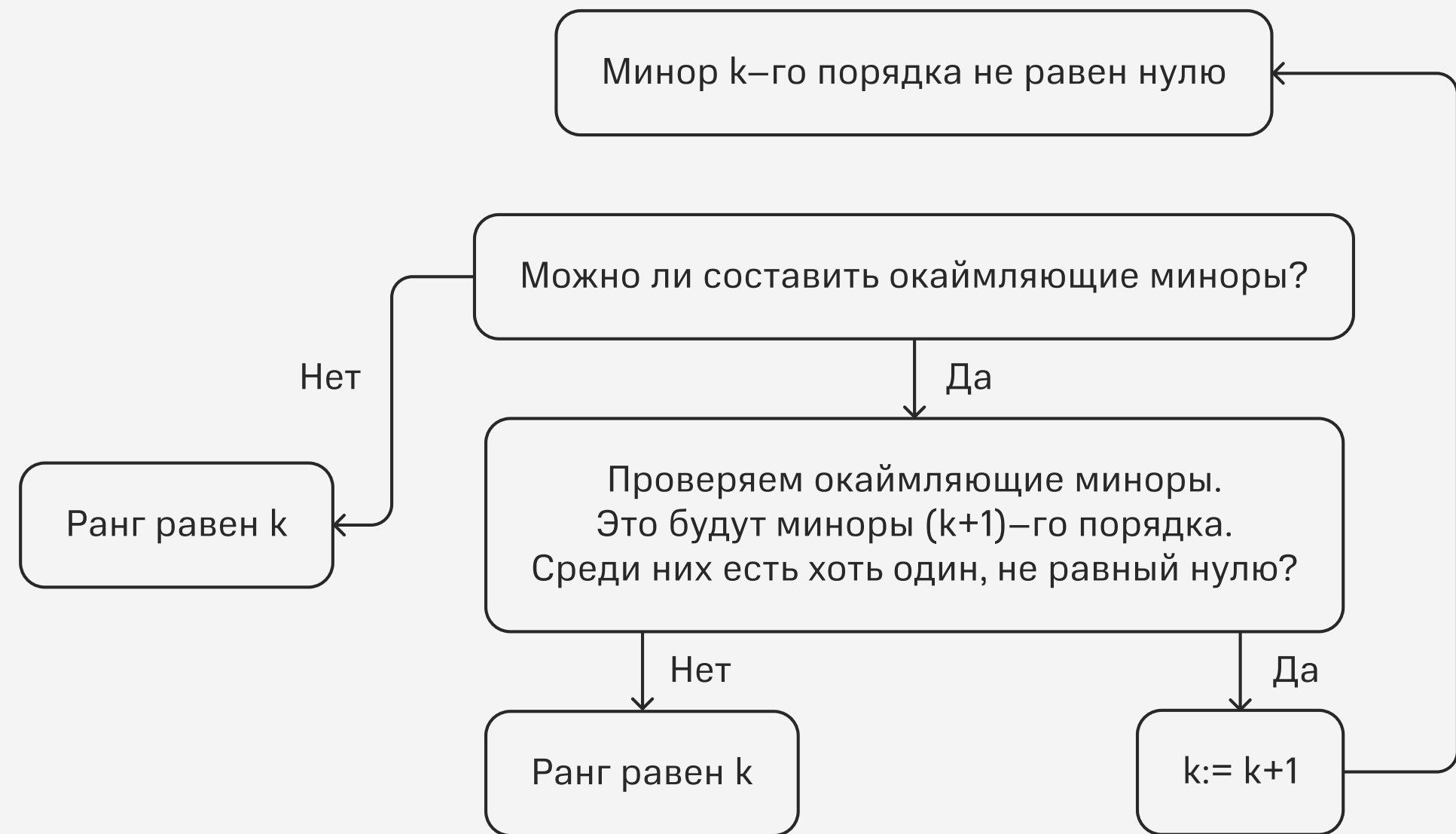


Ранг матрицы. Метод окаймляющих миноров

Пусть в матрице найден минор k -го порядка M , отличный от нуля.

Рассмотрим лишь те миноры $(k+1)$ -го порядка, которые содержат в себе минор M : если все они равны нулю, то ранг матрицы равен k .

В противном случае среди окаймляющих миноров найдется ненулевой минор $(k+1)$ -го порядка, и вся процедура повторяется.



Ранг матрицы. Метод окаймляющих миноров

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 & 4 \\ 3 & 4 & 2 & 6 & 8 \\ 1 & 2 & 1 & 3 & 4 \end{pmatrix}, \quad M_2 = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2 \neq 0.$$

$$M_3 = \begin{vmatrix} 1 & 2 & 1 \\ 3 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix} = 0, \quad M_3 = \begin{vmatrix} 1 & 2 & 3 \\ 3 & 4 & 6 \\ 1 & 2 & 3 \end{vmatrix} = 0, \quad M_3 = \begin{vmatrix} 1 & 2 & 4 \\ 3 & 4 & 8 \\ 1 & 2 & 4 \end{vmatrix} = 0.$$

Так как все окаймляющие миноры 3-го порядка равны нулю, то ранг матрицы равен 2.



Ранг матрицы. Метод элементарных преобразований

С помощью элементарных преобразований матрицу можно привести к такому виду, когда все ее элементы, кроме

$$a_{11}, a_{22}, \dots, a_{rr} \quad (r \leq \min(m, n))$$

равны нулю. Следовательно, ранг матрицы равен r .



Ранг матрицы. Метод элементарных преобразований

Элементарные преобразования

- Перемена мест двух строк (столбцов)
- Умножение всех элементов строки (столбца) на некоторое число $a \neq 0$
- Суммирование всех элементов одной строки (столбца) с соответствующими элементами иной строки (столбца), умноженными на некое вещественное число

Элементарные преобразования не изменяют ранг матрицы

Конечная цель преобразований матрицы — сделать её ступенчатой. Ранг ступенчатой матрицы равен количеству ненулевых строк.



ОБРАТНАЯ МАТРИЦА

М Т
С



МТС Тета



Обратная матрица

- Квадратная матрица A называется невырожденной (неособенной), если определитель матрицы A не равен нулю
- Если матрица A невырожденная, то существует и притом единственная матрица, такая, что:
$$A \cdot A^{-1} = A^{-1} \cdot A = E$$
- A^{-1} называется обратной матрицей

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



Обратная матрица. Свойства и основные методы нахождения

Метод присоединённой матрицы

- Присоединенная матрица — это матрица, составленная из алгебраических дополнений данной матрицы и транспонированная

Метод элементарных преобразований

- В методе используется расширенная матрица $(A|E)$

Итерационные методы

$$\det(A^{-1}) = \frac{1}{\det A}$$

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

$$(A^{-1})^T = (A^T)^{-1}$$

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



Обратная матрица

Пример

$$A = \begin{bmatrix} 2 & 7 \\ 3 & 9 \end{bmatrix}$$

$$\left[\begin{array}{cc|cc} 2 & 7 & 1 & 0 \\ 3 & 9 & 0 & 1 \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 2 & 7 & 1 & 0 \\ 1 & 3 & 0 & \frac{1}{3} \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 0 & 1 & 1 & -\frac{2}{3} \\ 1 & 3 & 0 & \frac{1}{3} \end{array} \right] \rightarrow$$

$$\rightarrow \left[\begin{array}{cc|cc} 0 & 1 & 1 & -\frac{2}{3} \\ 1 & 0 & -3 & \frac{7}{3} \end{array} \right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -3 & \frac{7}{3} \\ 0 & 1 & 1 & -\frac{2}{3} \end{array} \right]$$

$$A^{-1} = \begin{bmatrix} -3 & \frac{7}{3} \\ 1 & -\frac{2}{3} \end{bmatrix}$$



СЛАУ



MTC Тета



СЛАУ

- Систему из m линейных уравнений относительно n неизвестных x_1, \dots, x_n можно записать в общем виде:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases}$$

- Систему из m линейных уравнений относительно n неизвестных x_1, \dots, x_n можно записать в общем виде:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$



Или еще более кратко: $Ax = b$

где A — это матрица коэффициентов, b — столбец свободных членов, а x — вектор из неизвестных.

Причем возможны три случая:

- Система вовсе не имеет решения.
Такая система называется несовместной
- Система имеет единственное решение
- Система имеет бесконечно много решений



СЛАУ. Теорема Кронекера-Капелли

СЛАУ $Ax = b$ совместна тогда и только тогда, когда $\text{rg}(A) = \text{rg}(A|b)$, причём:

- система имеет единственное решение, если ранг равен числу неизвестных
- бесконечное множество решений, если ранг меньше числа неизвестных

Расширенная матрица получается из матрицы A приписыванием к ней справа столбца b .



СЛАУ. Решение через обратную матрицу

Решение уравнения $Ax = b$, если матрица A — невырожденная, единственно и запишется в виде:

$$x = A^{-1}b$$

Задача вычисления обратной матрицы и решения СЛАУ, строго говоря, сравнимы по сложности.



СЛАУ

Прикладная роль

- Матрица объекты-признаки (матрица, в строках которой содержатся описания исследуемых объектов, а в столбцах — значения их характеристик)
- Например, в первом столбце матрицы может содержаться
- Расстояние от квартиры до ближайшего метро, во втором — её площадь, в третьем — этаж, а в векторе-столбце b может стоять стоимость квартиры и т.д.
- Алгоритмы, основывающиеся на поиске коэффициентов линейной комбинации, называются линейными
- Основное преимущество — интерпретируемость (важно, например, в задаче кредитного скоринга)



СЛАУ. Случай несовместной системы

- Система несовместна, то есть мы не можем восстановить точную линейную зависимость целевой переменной b от столбцов матрицы A
- Ставится задача регрессии — задача приближённого восстановления зависимости целевой переменной от входных переменных, в которой целевая переменная принимает значения из R
- Решение таких задач сводится к минимизации функционала ошибки, определяющего качество нашего решения



СЛАУ. Случай несовместной системы

Пример — метод наименьших квадратов —
минимизация суммы квадратов ошибок регрессии
(нормы вектора ошибок)

$$||Ax - b||^2 \rightarrow \min_{x \in \mathbb{R}^n}$$

$$x = \operatorname{argmin}_{x \in \mathbb{R}^n} ||Ax - b||^2$$

$$x = (A^T A)^{-1} A^T b$$



ОБУСЛОВЛЕННОСТЬ СЛАУ

М Т
С



МТС Тета



Неустранимые погрешности

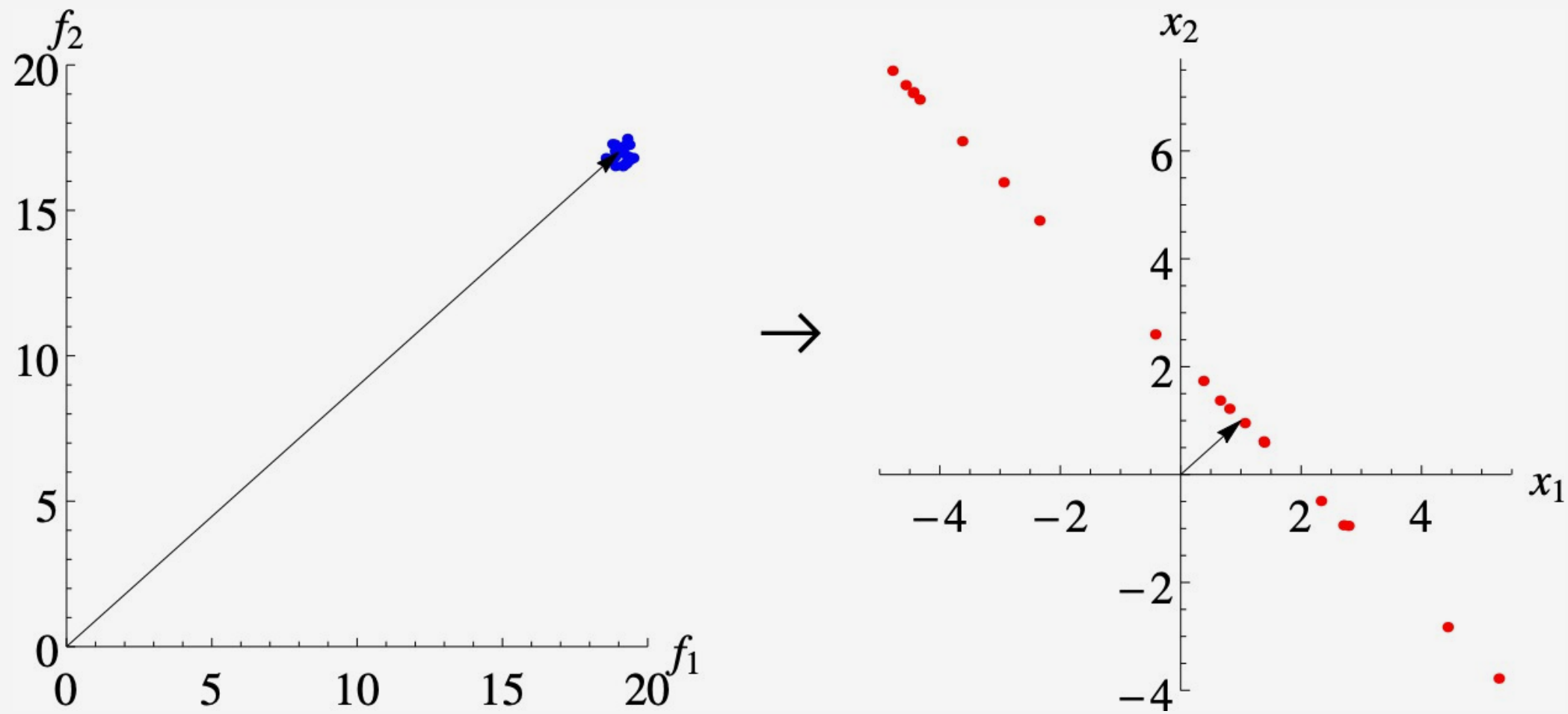
- Рассмотрим систему:

$$\begin{pmatrix} 10 & 9 \\ 9 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 19. \\ 17. \end{pmatrix}$$

- Причем матрица известна точно, а правая часть получена округлением до целого (погрешность не более 3%). Посмотрим, на какую точность можно рассчитывать при решении системы
- С точки зрения линейной алгебры, проблем при решении данной системы не должно быть, так как определитель не равен 0



Неустраняемые погрешности



Плохо обусловленные системы

Задача оказалась плохо обусловленной. Сравнительно небольшие возмущения системы уравнений привели к существенным отклонениям в решении.

Обусловленность задачи не связана с конкретным численным методом, это неустраняемая ошибка. Существуют способы снижения погрешности, вызванной плохой обусловленностью:

- Каким-то образом перейти к хорошо обусловленной эквивалентной системе
- Повысить точность определения коэффициентов СЛАУ и правой части

Плохо обусловленные системы являются обобщением понятия вырожденных систем. Системы «близкие» к вырожденным скорее всего будут плохо обусловлены.



Векторные нормы

В вычислительной математике широко распространены следующие нормы:

Максимальная или бесконечная норма (норма Чебышёва)

$$\|\mathbf{x}\|_{\infty} = \max_i |x_i|$$

Манхэттенская норма (L1)

$$\|\mathbf{x}\|_1 = \sum_i |x_i|$$

Евклидова норма (L2)

$$\|\mathbf{x}\|_e = \sqrt{(\mathbf{x}, \mathbf{x})} \equiv \sqrt{\sum_i x_i^2}$$



Матричные нормы

$$\|A\| > 0, \quad A \neq 0, \quad \|0\| = 0,$$

$$\|\alpha A\| = |\alpha| \|A\|,$$

$$\|A + B\| \leq \|A\| + \|B\|,$$

$$\|AC\| \leq \|A\| \|C\|$$

Если норма удовлетворяет только первым трем аксиомам, то такую норму называют аддитивной. Если удовлетворяются все четыре аксиомы, то мультипликативной. Всякую аддитивную норму умножением на достаточно большую положительную константу можно превратить в мультипликативную.



Матричные нормы

Первая норма (1-норма)

$$\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^n |a_{ij}|;$$

Спектральная норма (норма Гильберта)

$$\|A\|_2 = \max_{1 \leq i \leq m} \sigma_i = \left(\max_{1 \leq i \leq m} \lambda_i(A^T A) \right)^{1/2} = \left(\max_{1 \leq i \leq n} \lambda_i(AA^T) \right)^{1/2}$$



Матричные нормы

Бесконечная норма

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|;$$

Сферическая норма (норма Фробениуса)

$$\begin{aligned} \|A\|_F = \|A\|_E = N(A) &= \left(\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2 \right)^{1/2} = \\ &= \left(\sum_{j=1}^m \sigma_j^2 \right)^{1/2} = \left(\sum_{j=1}^m \lambda_j(A^T A) \right)^{1/2} = \sqrt{\text{tr}(A^T A)}. \end{aligned}$$



Матричные нормы



Матричная норма называется согласованной с векторной нормой, если:

$$\|A\mathbf{x}\|_V \leq \|A\|_M \|\mathbf{x}\|_V$$

для любой матрицы A и всех векторов \mathbf{x} .
Всякая норма матрицы согласована с какими-нибудь векторными нормами.



Матричные нормы

Пусть задана векторная норма. Тогда числовая функция:

$$\|A\| = \sup_{\|x\|_V \neq 0} \frac{\|Ax\|_V}{\|x\|_V} = \sup_{\|x\|_V = 1} \|Ax\|_V$$

является матричной нормой и называется нормой матрицы, подчиненной векторной норме.



Число обусловленности для СЛАУ

Рассмотрим СЛАУ и полученную из неё возмущением правой части. Оценим относительную погрешность решения, связав ее с относительной погрешностью правой части.

$$\mathbf{Ax} = \mathbf{b}$$

$$\Delta \mathbf{x} = \mathbf{z} - \mathbf{x}.$$

$$\mathbf{Az} = \mathbf{b} + \Delta \mathbf{b}.$$

$$\Delta \mathbf{x} = \mathbf{A}^{-1} \Delta \mathbf{b}$$



Число обусловленности для СЛАУ

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\|\mathbf{A}^{-1}\| \|\Delta \mathbf{b}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{A}^{-1}\| \|\mathbf{b}\|}{\|\mathbf{x}\|} \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}$$

Число обусловленности СЛАУ

$$\nu(\mathbf{A}, \mathbf{b}) = \frac{\|\mathbf{A}^{-1}\| \|\mathbf{b}\|}{\|\mathbf{x}\|} \equiv \frac{\|\mathbf{A}^{-1}\| \|\mathbf{b}\|}{\|\mathbf{A}^{-1} \mathbf{b}\|}$$

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \nu(\mathbf{A}, \mathbf{b}) \frac{\|\Delta \mathbf{b}\|}{\|\mathbf{b}\|}$$



Число обусловленности матрицы

Насколько плохо может быть обусловлена система с данной матрицей? Можно ли за счет выбора правой части сделать систему сколь угодно плохой? Оказывается, нет.

$$\nu(\mathbf{A}, \mathbf{b}) = \frac{\|\mathbf{A}^{-1}\| \|\mathbf{b}\|}{\|\mathbf{A}^{-1} \mathbf{b}\|} = \frac{\|\mathbf{A}^{-1}\| \|\mathbf{Ax}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\|.$$

Число обусловленности матрицы

$$\mu(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$$

Матрицу можно считать вырожденной, если ее число обусловленности превосходит $1/\epsilon$, где ϵ — относительная погрешность выполнения машинных операций.



Число обусловленности матрицы. Пример

$$\begin{pmatrix} 10 & 9 \\ 9 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 19. \\ 17. \end{pmatrix}, \quad \mathbf{A}^{-1} = \frac{1}{-1} \begin{pmatrix} 8 & -9 \\ -9 & 10 \end{pmatrix}$$

$$\|\mathbf{A}\|_{\infty} = \|\mathbf{A}\|_1 = 19, \quad \|\mathbf{A}^{-1}\|_{\infty} = \|\mathbf{A}^{-1}\|_1 = 19,$$

$$\mu_1(\mathbf{A}) = \mu_{\infty}(\mathbf{A}) = 19 \cdot 19 = 361.$$

$$\mu_e(\mathbf{A}) = \left| \frac{\lambda_{\max}(\mathbf{A})}{\lambda_{\min}(\mathbf{A})} \right| = \frac{\sqrt{82} + 9}{\sqrt{82} - 9} \approx 326.$$

Таким образом, погрешность в 3% в правой части решения приводит примерно к $\sim 1000\%$ ошибки в решении.



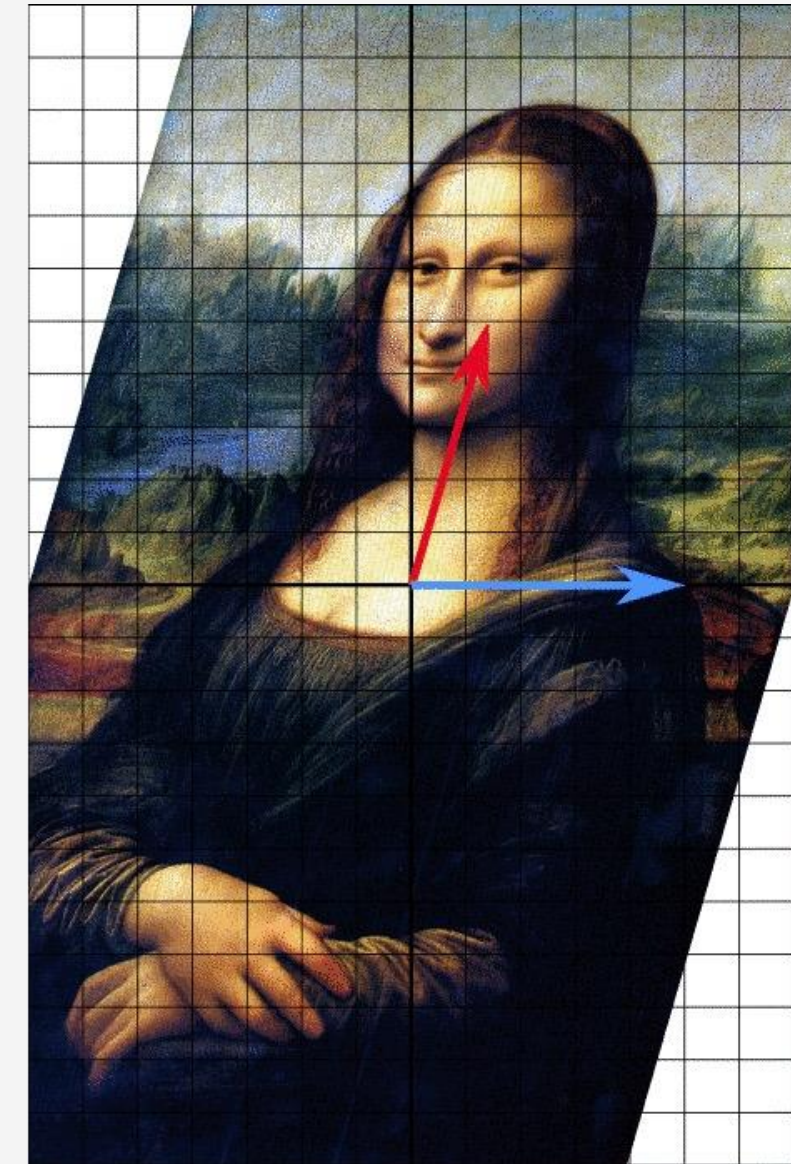
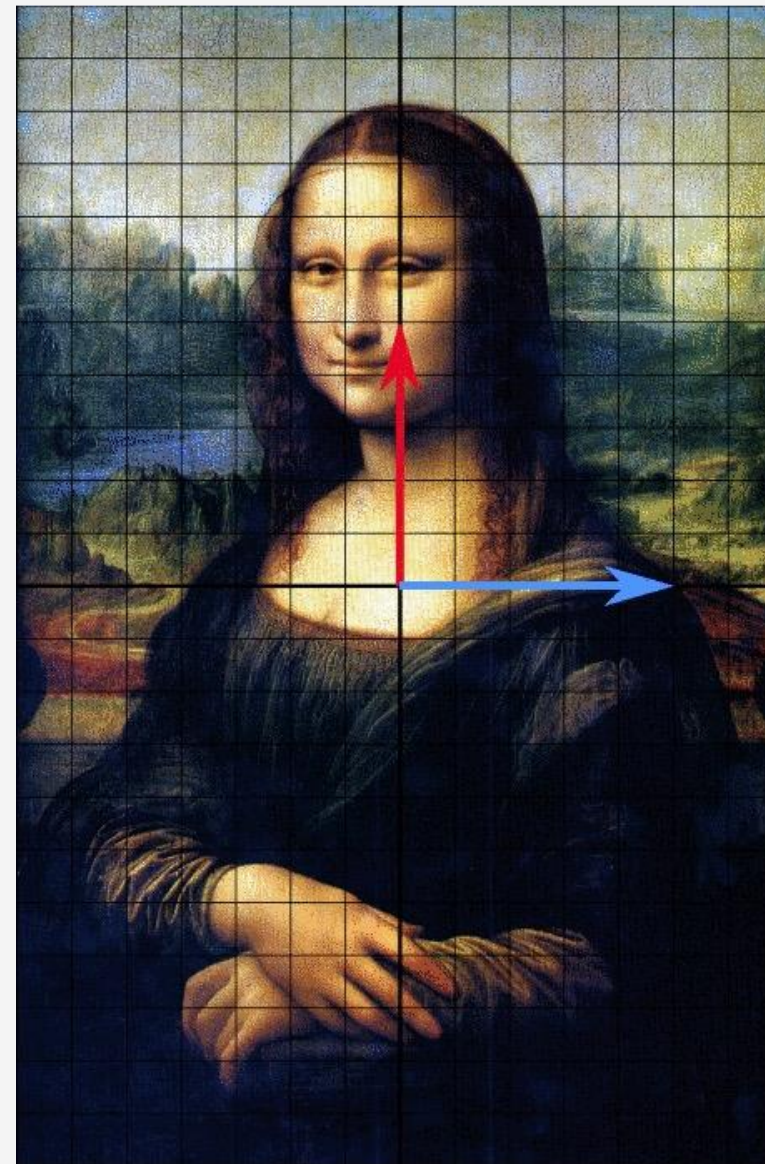
СОБСТВЕННЫЕ ЗНАЧЕНИЯ И ВЕКТОРЫ



Собственные векторы

Собственный вектор —

любой ненулевой вектор x , который отображается в коллинеарный ему вектор, оператором A , а соответствующий скаляр называется собственным значением оператора.



Собственные векторы

- Линейное преобразование может не иметь собственных векторов вообще, например поворот в двумерном пространстве (кроме нескольких исключительных случаев), или иметь n собственных векторов с различными собственными значениями
- Все корни характеристического многочлена и только они являются собственными значениями матрицы

$$\det(A - \lambda E) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Собственные векторы. Пример

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

$$\begin{aligned} \det(A - \lambda I) &= \left| \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} 2 - \lambda & 1 \\ 1 & 2 - \lambda \end{vmatrix} \\ &= 3 - 4\lambda + \lambda^2 \\ &= (\lambda - 3)(\lambda - 1). \end{aligned}$$

$$(A - I)\mathbf{v}_{\lambda=1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$1v_1 + 1v_2 = 0$$

$$\mathbf{v}_{\lambda=1} = \begin{bmatrix} v_1 \\ -v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$(A - 3I)\mathbf{v}_{\lambda=3} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-1v_1 + 1v_2 = 0;$$

$$1v_1 - 1v_2 = 0$$

$$\mathbf{v}_{\lambda=3} = \begin{bmatrix} v_1 \\ v_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$



МАТРИЧНЫЕ РАЗЛОЖЕНИЯ

MT C



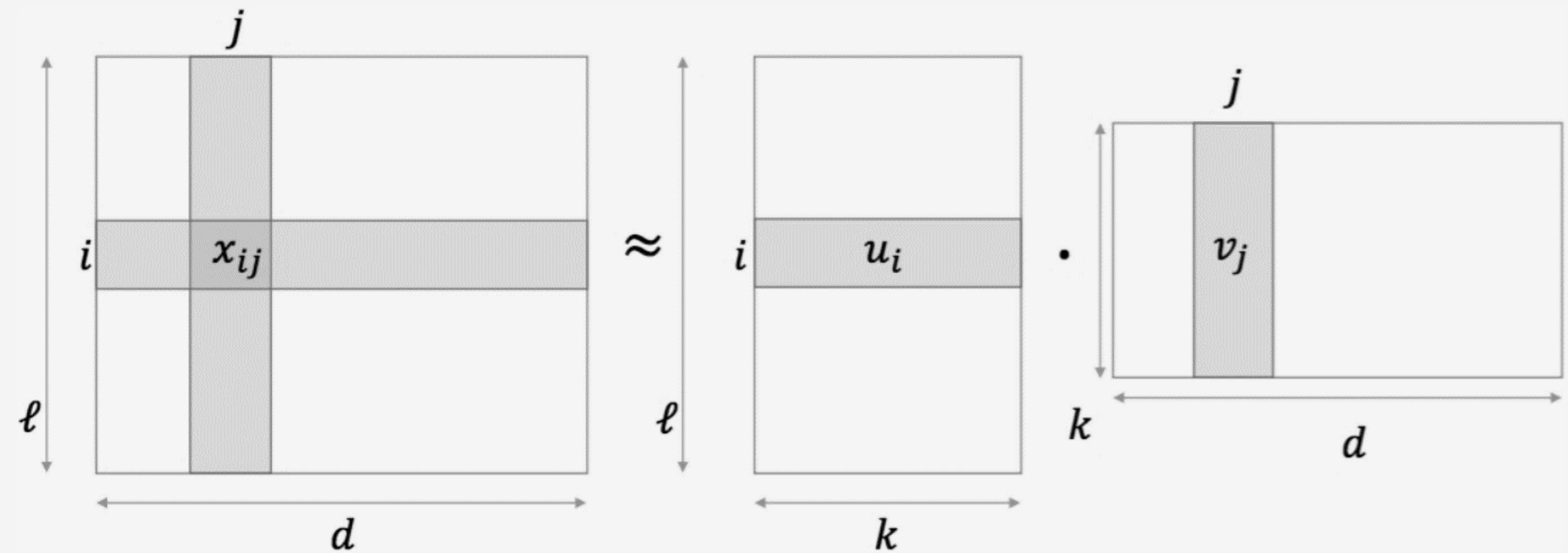
MTC Тета



Постановка задачи

В задаче матричного разложения требуется приблизить матрицу произведением двух других, где матрицы U и V задают новое признаковое описание объектов и изначальных признаков.

$$X_{l,n} \approx U_{l,k} \cdot V_{k,n}^T$$



Постановка задачи

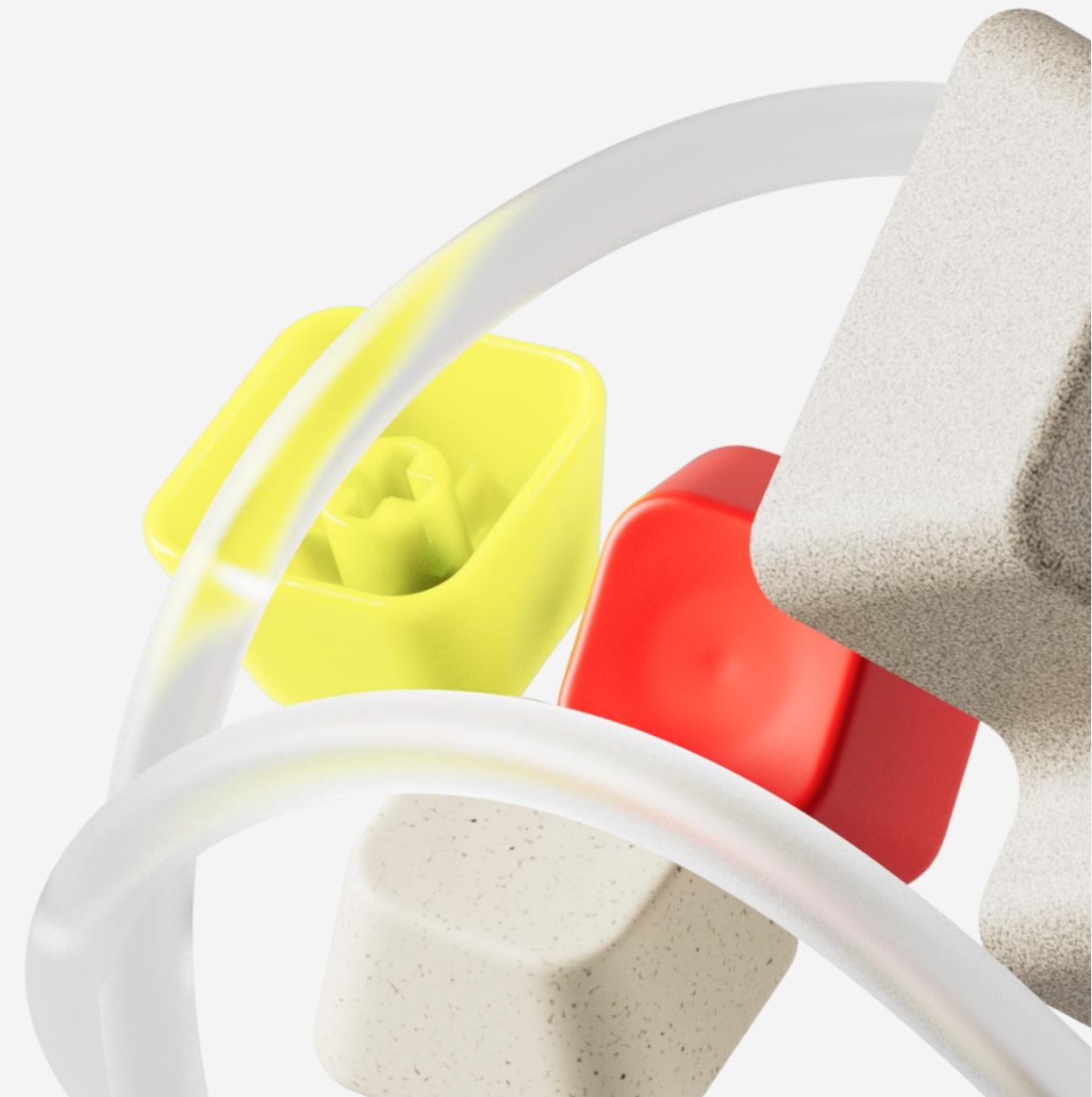
- Формально задачу можно свести к поиску минимума:

$$\|X - U \cdot V^T\| \rightarrow \min$$

- Если использовать норму Фробениуса, то задача записывается следующим образом:

$$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2} \quad \sum_{i,j} (x_{ij} - \langle u_i, v_j \rangle)^2 \rightarrow \min$$

где u_i — новое описание объекта i , а v_j — новое описание признака j .

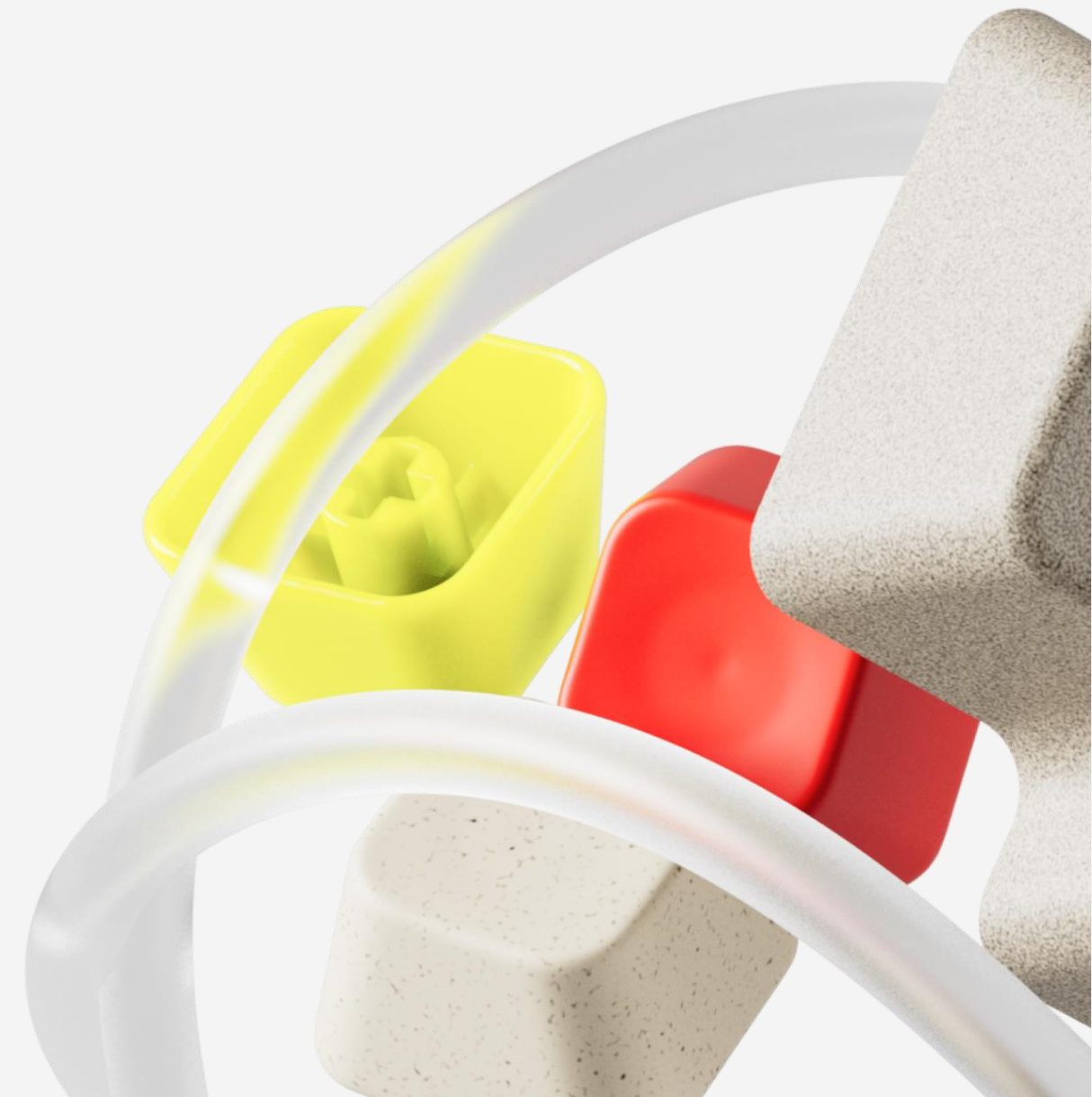


SVD разложение

Сингулярное разложение — это способ представить некоторую исходную матрицу в виде произведения трех других

$$X = U\Sigma V^T$$

где U — ортогональная матрица, Σ — диагональная матрица, V — ортогональная матрица

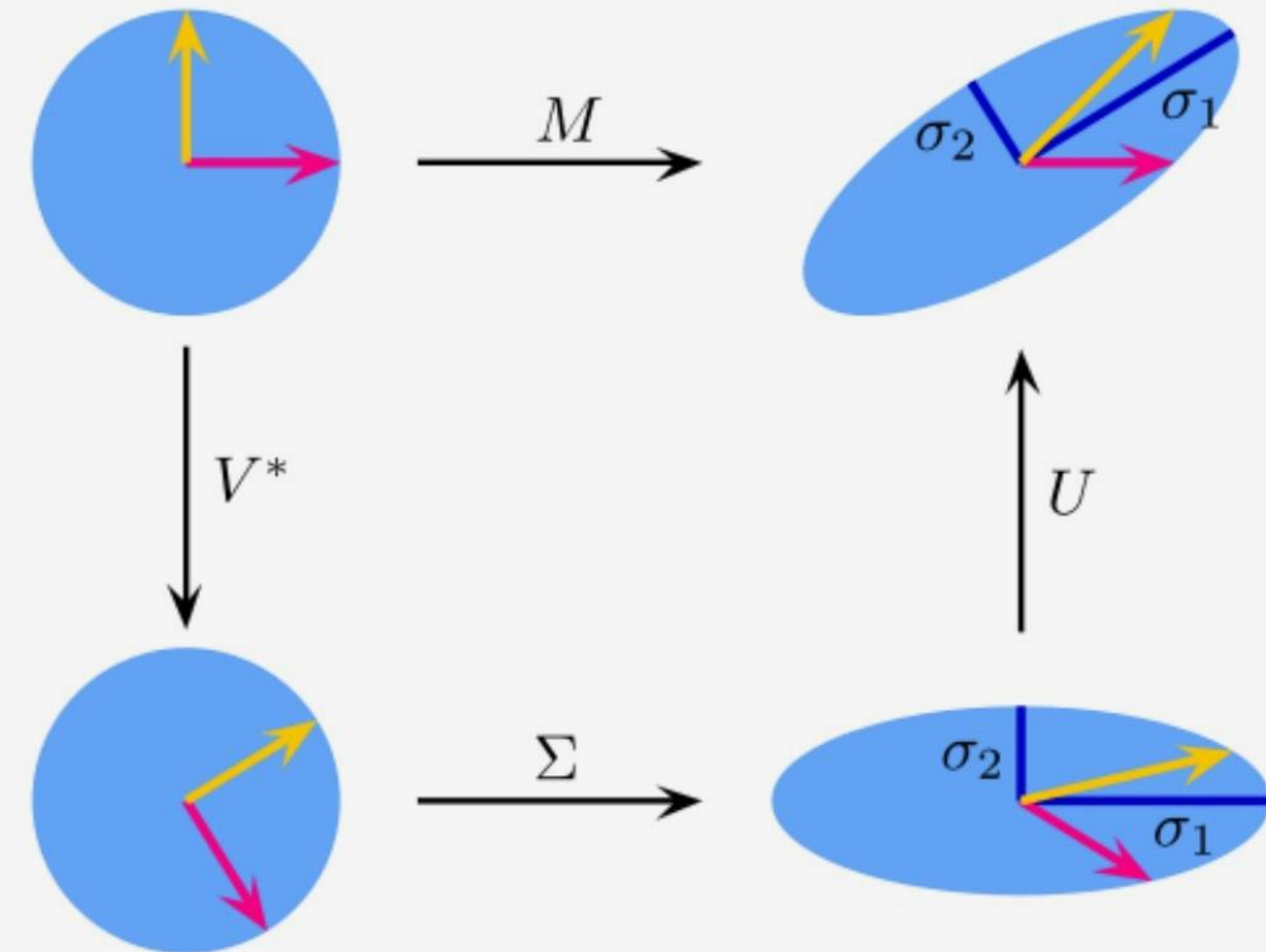


SVD разложение

Геометрическая интерпретация

Если рассматривать матрицу M как задающую некоторое отображение, то оно раскладывается на составляющие:

1. Сначала пространство поворачивают
2. Потом растягивают вдоль осей координат
3. Затем снова поворачивают
(говоря более точно, здесь могут быть не только повороты, но и вращения некоторых осей, но в упрощённом виде это выглядит так)



$$M = U \cdot \Sigma \cdot V^*$$

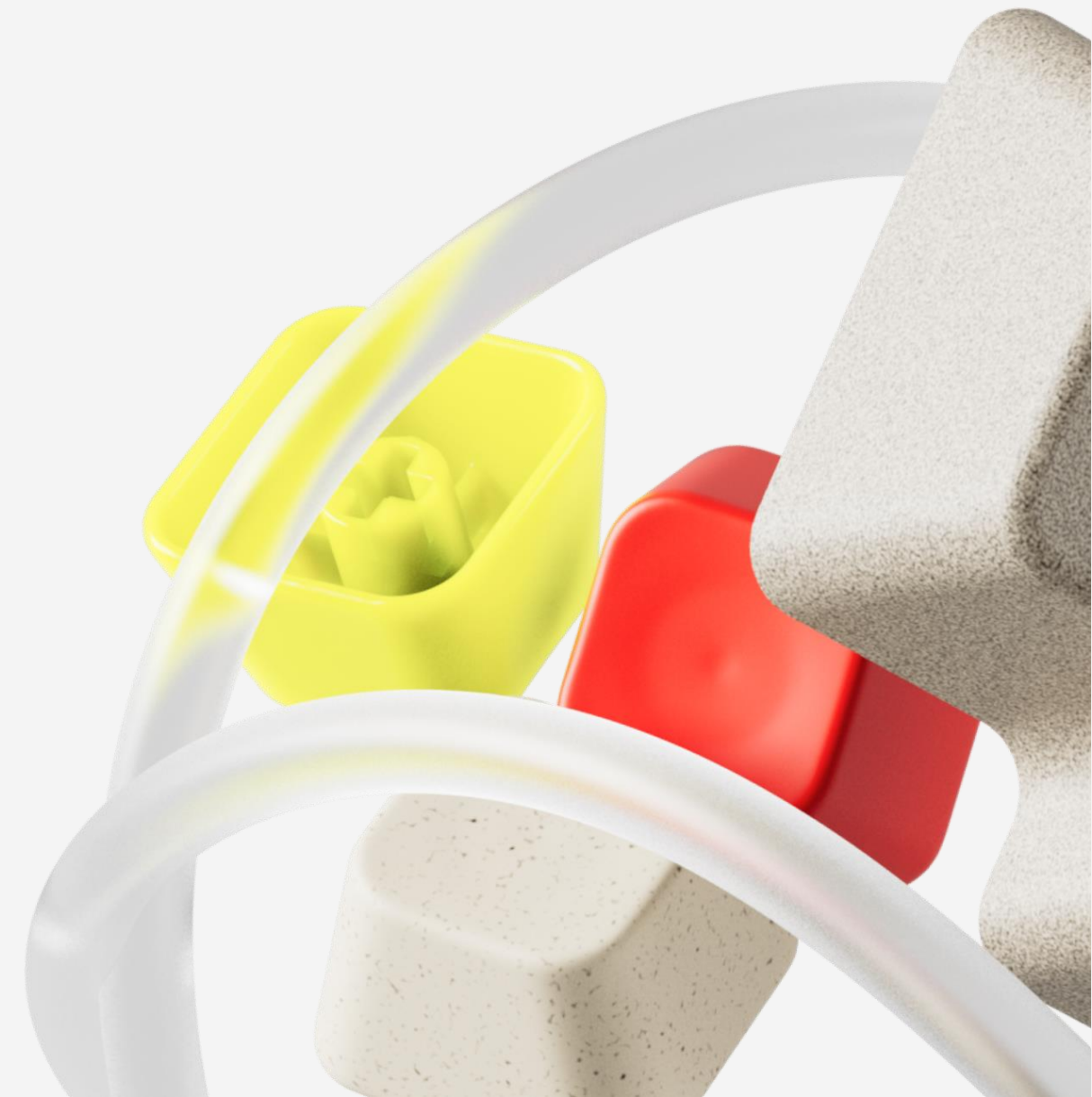
SVD разложение

Сингулярное разложение матриц может быть полезно для рассматриваемой задачи. Можно рассмотреть сингулярное разложение матрицы X :

$$X = \tilde{U} \Sigma \tilde{V}^T$$

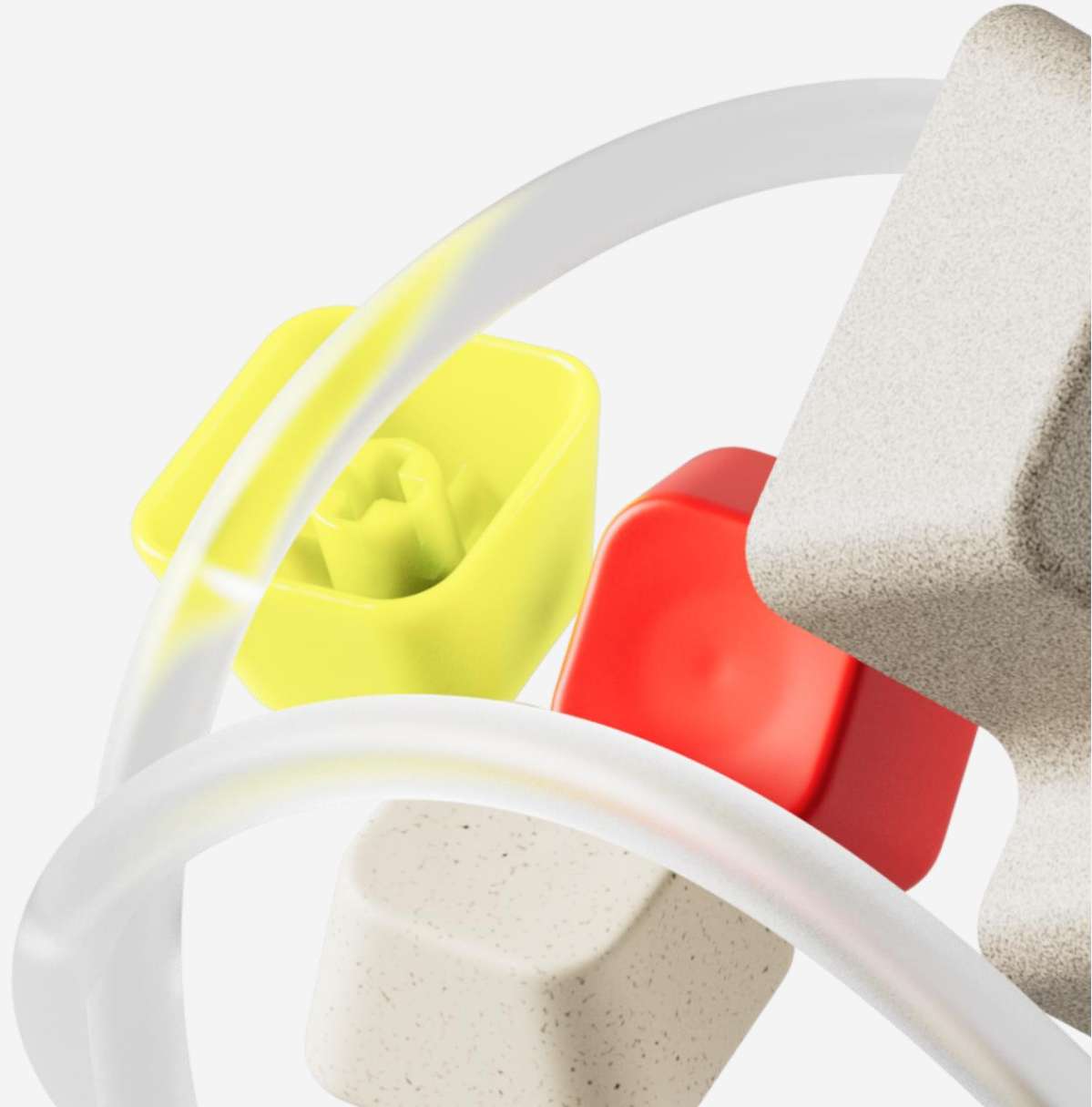
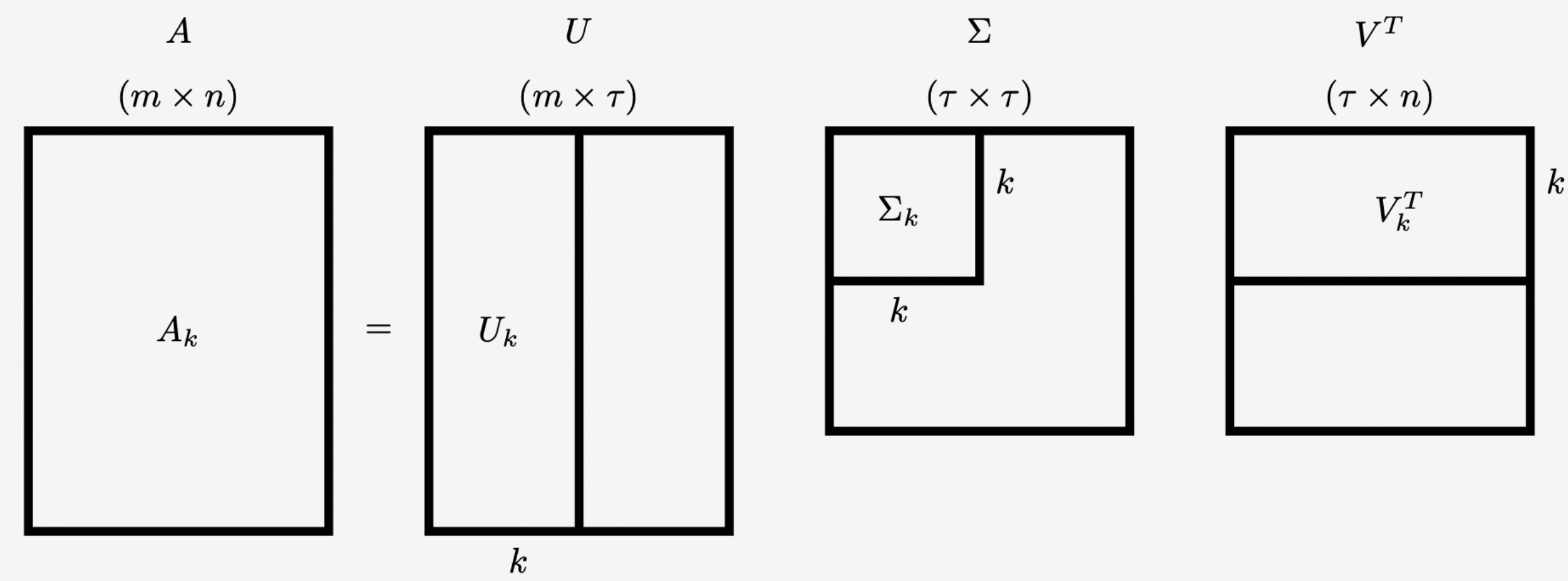
Теперь у каждой матрицы можно взять какую-то часть: от матрицы U первые k столбцов, от матрицы Σ — квадрат размера $k \times k$, от матрицы V — первые k строк.

Таким образом, определяются усеченные матрицы сингулярного разложения.



SVD разложение

Усечение



SVD разложение

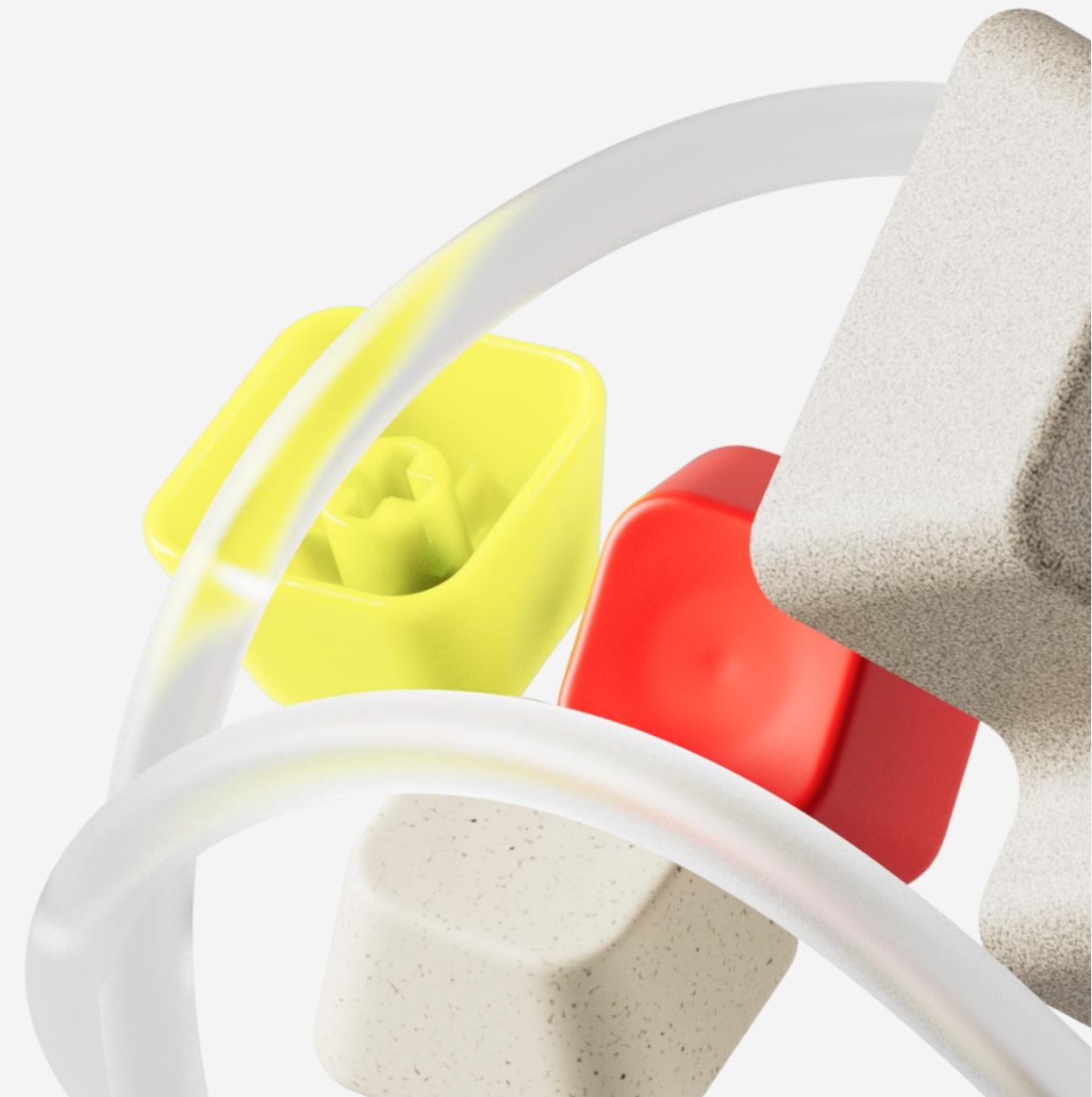
Если в задаче матричного разложения
взять в качестве искомых матриц

$$U = \tilde{U}_k \Sigma_k, \quad V = \tilde{V}_k$$

то это будет решением оптимизационной задачи,
которое даст наилучшее приближение исходной
матрицы по норме Фробениуса. При этом в качестве
искомых матриц можно использовать другие.
Это указывает на неоднозначность решения
задачи матричного разложения.

$$U = \tilde{U}_k, \quad V = \tilde{V}_k \Sigma_k$$

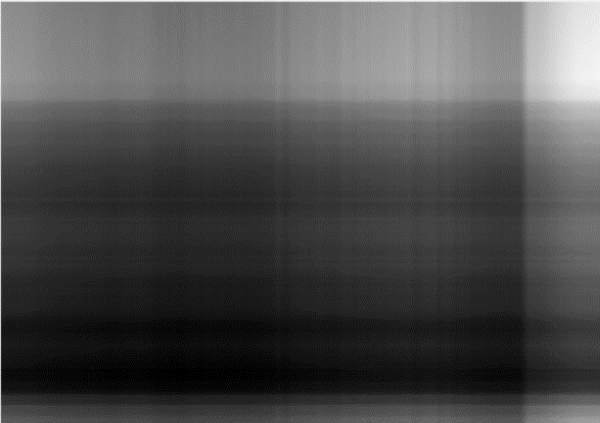
$$U = \tilde{U}_k \sqrt{\Sigma_k}, \quad V = \tilde{V}_k \sqrt{\Sigma_k}.$$



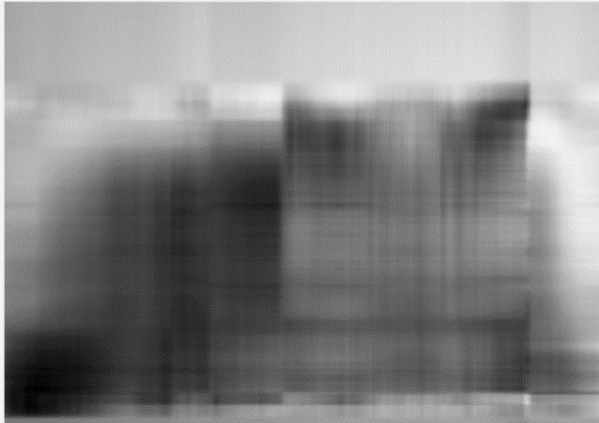
Сжатие изображений на базе SVD

Compression at different orders

Order = 1 | RMSE = 38.80



Order = 5 | RMSE = 18.84



Order = 20 | RMSE = 12.87



Order = 50 | RMSE = 10.45



Order = 200 | RMSE = 5.41



Order = 400 | RMSE = 2.20



Order = 800 | RMSE = 0.09

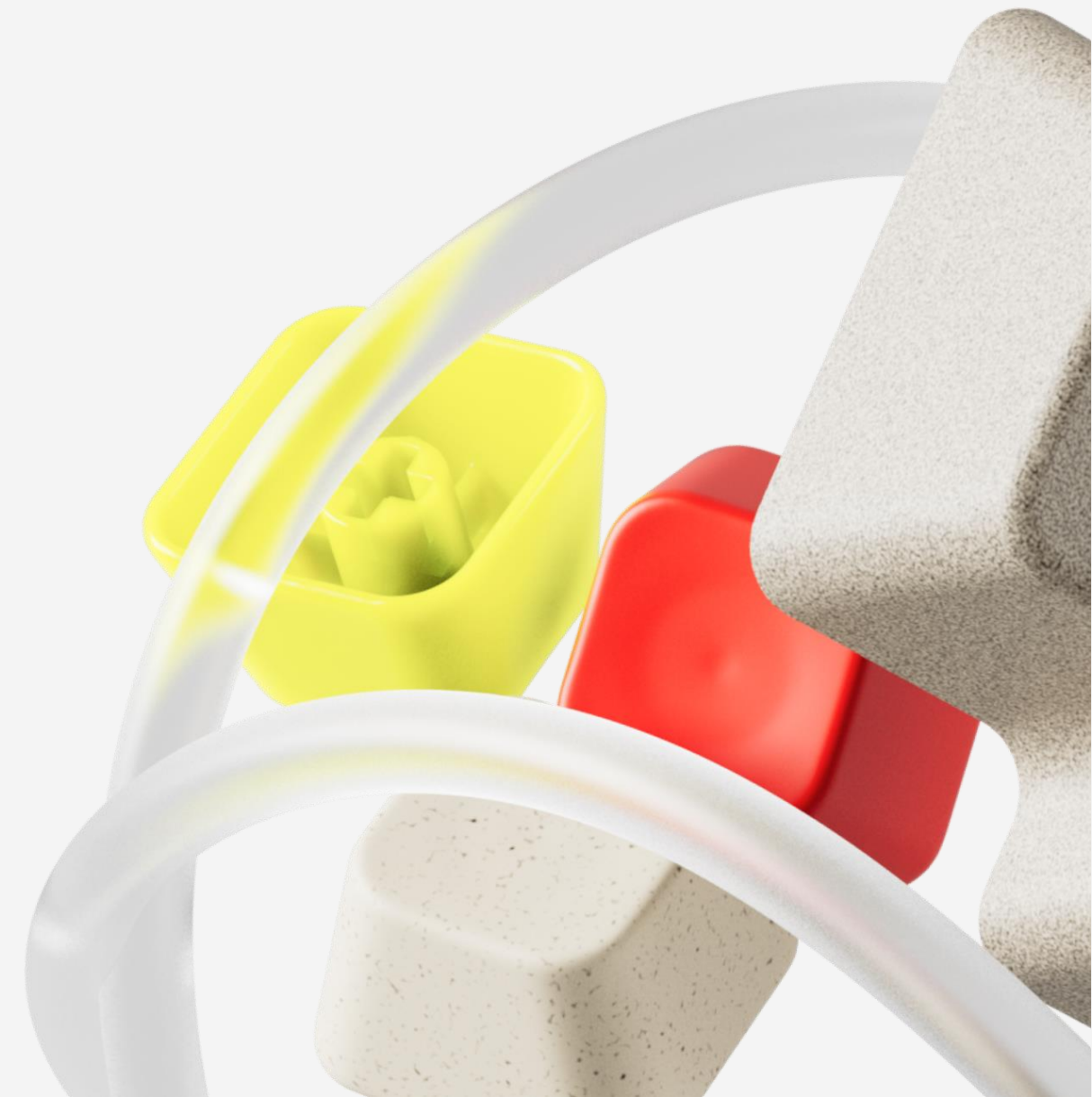


Init image



Применение матричных разложений

- Понижение размерности
- Заполнение пропусков в матрицах
- Рекомендательные системы
- Профилирование
- Анализ текстов
- Широко в науке...
- И другие...



Что мы узнали

- Разобрали матричную терминологию: определитель, ранг, обратная матрица
- Рассмотрели работу со СЛАУ
- Обсудили вопросы обусловленности СЛАУ и матриц
- Определили понятия собственных векторов и значений
- Детально рассмотрели механизм разложения матриц и рассмотрели наиболее актуальное в анализе данных — SVD разложение



Полезные ССЫЛКИ

- <https://www.cis.upenn.edu/~jean/math-deep.pdf>
- <https://mml-book.github.io/>
- https://youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab&si=PCfYtyqnveHtF2My

