Семинар 3

Конкретные векторные пространства

Основной объект изучения – пространство столбцов $\mathbb{R}^n$. Любая матрица $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ определяет отображение $\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$ заданное $x \mapsto Ax$. В каком-то смысле это единственный пример, а потому – самый важный.

Однако есть и другие, внешне не похожие, примеры:

1. Матрицы $\mathrm{M}_n(\mathbb{R})$. В качестве отображений нужно рассматривать $\mathrm{M}_n(\mathbb{R}) \to \mathrm{M}_m(\mathbb{R})$ заданное по правилу $X \mapsto \sum_i P_i X Q_i$, где $P_i, Q_i^t \in \mathrm{M}_{m\,n}(\mathbb{R})$.
2. Решения СЛУ $\{y \in \mathbb{R}^n \mid Ay = 0\}$ для некоторой $A \in \mathrm{M}_{m\,n}(\mathbb{R})$. Хороший вопрос: какие тут должны быть отображения между решениями?
3. Еще более интересный вопрос, а какие должны быть отображения между $\mathrm{M}_n(\mathbb{R})$ и $\{y \in \mathbb{R}^m \mid Ay = 0\}$?

Именно поэтому нужна общая теория, которая бы помогла нам одним языком описать все эти случаи и придумать правильные определения, когда они не очевидны.

Абстрактные векторные пространства

В определении векторного пространства надо зафиксировать откуда берутся коэффициенты. Вариантов несколько: вещественные числа $\mathbb{R}$, комплексные числа $\mathbb{C}$, рациональные числа $\mathbb{Q}$.[1] Для простоты, все определения будем формулировать с вещественными числами.[2]

Определение. Векторное пространство над $\mathbb{R}$ это следующие данные:

1. множество V.[3]
2. операция сложения векторов, т.е. отображение $+\colon V \times V \to V$ вида $(v_1, v_2) \mapsto v_1 + v_2$.
3. операция умножения векторов на число, т.е. отображение $\cdot\colon \mathbb{R} \times V \to V$ вида $(r, v) \mapsto rv$.

И эти данные удовлетворяют следующим аксиомам:

1. Для любых $v, u, w \in V$ верно $(v + u) + w = v + (u + w)$.
2. Существует вектор $0 \in V$ такой, что для любого вектора $v \in V$ имеем $0 + v = v + 0 = v$.
3. Для любого вектора $v \in V$ существует вектор $-v$ такой, что $v + (-v) = (-v) + v = 0$.
4. Для любых векторов $v, u \in V$ верно $v + u = u + v$.
5. Для любых $r \in \mathbb{R}$ и $v, u \in V$ верно $r(v + u) = rv + ru$.
6. Для любых $r_1, r_2 \in \mathbb{R}$ и $v \in V$ верно $(r_1 + r_2)v = r_1 v + r_2 v$.
7. Для любых $r_1, r_2 \in \mathbb{R}$ и $v \in V$ верно $(r_1 r_2)v = r_1(r_2 v)$.
8. Для любого $v \in V$ верно $1v = v$.

Обычно элементы V называют *векторами*, а элементы $\mathbb{R}$ называют *скалярами*. Даже в абстрактном случае, полезно думать геометрически, представляя себе в первую очередь $\mathbb{R}^n$ как главные пример.

[1] На самом деле годится все что угодно, если в этом «чем угодно» можно складывать, вычитать, умножать и делить. Такие объекты называются *полями*.

[2] Определение векторного пространства может показаться жутко сложным и формальным. Не стоит бросаться учить наизусть все, что в нем находится. Главное понимать как с ним работать. Действительно, никто из нас не знает строгого определения $\mathbb{R}$, но это не мешает нам с ним работать!

[3] Это будет как раз множество векторов.

Замечания Полезно понимать, что указанный выше набор аксиом – это минимальный набор. Но из него можно вывести простыми но нудными манипуляциями еще полезные свойства, которые надо иметь перед глазами. В дальнейшем окажется, что все разумные свойства, к которым вы привыкли в $\mathbb{R}^n$ выполняются в произвольном векторном пространстве. Потому интуитивно надо все время думать про $\mathbb{R}^n$, эта интуиция вас не подведет. А вот пример полезных свойств:

1. Нулевой вектор всегда единственный.
2. Для любого вектора $v \in V$ противоположный вектор $-v$ всегда единственный и совпадает с $-1 \cdot v$.
3. Если умножить число 0 на любой вектор, то получится нулевой вектор, то есть $0 \cdot v = 0$ (тут слева ноль – это число, а справа – вектор).
4. Аналогично, если умножить нулевой вектор на что угодно, то получится нулевой вектор, то есть $\lambda \cdot 0 = 0$ (тут с обеих сторон ноль означает нулевой вектор).

Определение. Пусть V – векторное пространство над $\mathbb{R}$, тогда подмножество $U \subseteq V$ называется подпространством, если

1. $0 \in U$
2. Если $u, v \in U$, то и $v + u \in U$.
3. Если $r \in \mathbb{R}$ и $v \in U$, то и $rv \in U$.

Замечания

- Отметим, что всякое подпространство само является векторным пространством.
- Обратите внимание, что первое свойство равносильно тому, что подпространство U обязательно не пусто. Действительно, если есть 0 вектор, то подпространство не пусто. Если же U не пусто, то там есть какой-то вектор v, тогда мы v можем умножить на число 0 и получим нулевой вектор внутри U по третьему свойству.

Примеры Пространства:

1. Пусть $V = \mathbb{R}^n$ с обычными операциями сложения и умножения на скаляр.
2. Пусть V $\mathrm{M}_{m\,n}(\mathbb{R})$ с обычными операциями сложения и умножения на скаляр.
3. Пусть $V = \{0\} = \mathbb{R}^0$.
4. Пусть $\mathbb{R}[x] = \{a_0 + a_1 x + \ldots + a_n x^n \mid a_i \in \mathbb{R},\, n \in \mathbb{N}\}$ – множество многочленов от переменной x с вещественными коэффициентами с обычными операциями сложения и умножения на число.
5. Пусть $C[0,1] = \{f\colon [0,1] \to \mathbb{R} \mid f \text{ непрерывна } \}$ – множество непрерывных функций на отрезке $[0,1]$, тогда оно является векторным пространством над $\mathbb{R}$.

Подпространства (а значит тоже пространства):

1. $\{y \in \mathbb{R}^n \mid Ay = 0\}$ для некоторой $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ с обычными операциями сложения и умножения на скаляр.
2. $\{X \in \mathrm{M}_n(\mathbb{R}) \mid AXB = 0\}$ для некоторых $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ и $B \in \mathrm{M}_{n\,k}(\mathbb{R})$ с обычными операциями сложения и умножения на скаляр.

Линейные комбинации

Пусть V – векторное пространство над $\mathbb{R}$. Главные вопрос: как надо про него думать? Ответ прост: это куча из элементов и эти элементы можно складывать и умножать на числа. То есть, мы всегда можем вытащить элементы[4] $v_1, \ldots, v_n \in V$ и начать их складывать с коэффициентами, получив некий новый вектор $r_1 v_1 + \ldots + r_n v_n \in V$, где $r_i \in \mathbb{R}$. Поэтому все, что можно сказать про векторное пространство, обязательно формулируется в терминах таких выражений. Поэтому предлагается изучать подобные выражения.

Пусть $v_i \in V$ и $r_i \in \mathbb{R}$ как выше, тогда выражение $r_1 v_1 + \ldots + r_n v_n$ называется *линейной комбинацией* векторов $v_1, \ldots, v_n$. Линейная комбинация называется *тривиальной*, если все $r_i = 0$, в противном случае *нетривиальной*. Вектора $v_1, \ldots, v_n$ называются *линейно зависимыми*, если существует нетривиальная линейная комбинация этих векторов равная нулю (нулевому вектору), т.е. $v_1, \ldots, v_n$ – линейно зависимы, если существуют $r_1, \ldots, r_n \in \mathbb{R}$ так, что хотя бы один из r_i не равен нулю и $r_1 v_1 + \ldots + r_n v_n = 0$. В противном случае вектора называются *линейно независимыми*.

Пусть $v_i \in V$ – произвольный набор векторов, тогда набор $v_1, \ldots, v_n$ называется порождающим V (или просто порождающим), если любой вектор из V представляется в виде их линейной комбинации, то есть для любого $v \in V$ найдутся числа $r_1, \ldots, r_n \in \mathbb{R}$ (не обязательно единственный) такой, что $v = r_1 v_1 + \ldots + r_n v_n$. Обратите внимание, что если набор $v_1, \ldots, v_n$ является порождающим для V, то V – минимальное подпространство внутри V, содержащее все эти векторы. Любое другое подпространство хотя бы один из них не содержит.

Примеры

1. Рассмотрим $V = \mathbb{R}^3$ и пусть

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

   Тогда вектора v_1, v_2, v_3 линейно независимы. Вектора v_1, v_2, v_4 тоже линейно независимы. Но вот вектора v_1, v_2, v_3, v_4 уже зависимы, так как $v_1 + v_2 + v_3 - v_4 = 0$. Также зависимы вектора v_1, v_2 и v_5, ибо $v_1 + v_2 - v_5 = 0$.

2. Один вектор $v \in V$ линейно зависим тогда и только тогда, когда он равен нулю.

3. Два вектора $v, u \in V$ линейно зависимы тогда и только тогда, когда они пропорциональны, т.е. либо $v = \lambda u$ либо $u = \lambda v$.

Базис

Утверждение. *Пусть V – векторное пространство над $\mathbb{R}$ и пусть $v_1, \ldots, v_n \in V$. Следующие утверждения эквивалентны:*

1. *Вектора $v_1, \ldots, v_n$ линейно независимы и набор $v_1, \ldots, v_n$ является порождающим.*

2. *Вектора $v_1, \ldots, v_n$ – максимальный линейно независимый набор, то есть для любого $u \in V$ вектора $v_1, \ldots, v_n, u$ уже линейно зависимы.*

3. *Вектора $v_1, \ldots, v_n$ – минимальный порождающий набор, то есть для любого i набор $\{v_1, \ldots, v_n\} \setminus \{v_i\}$ уже не порождающий.*

Доказательство. (1)⇒(2). Возьмем любой $u \in V$, тогда $u = \lambda_1 v_1 + \ldots + \lambda_n v_n$, тогда $\lambda_1 v_1 + \ldots + \lambda_n v_n - u = 0$ – линейная зависимость между векторами.

(2)⇒(1). Пусть $u \in V$, тогда существует какая-то линейная зависимость $\lambda_1 v_1 + \ldots + \lambda_n v_n + \mu u = 0$. Если $\mu = 0$, то вектора v_i линейно зависимы, но это не так. Значит $\mu \neq 0$. Тогда на него можно поделить и получим $u = -\frac{\lambda_1}{\mu} v_1 - \ldots - \frac{\lambda_n}{\mu} v_n$.

(1)⇒(3). Так как вектора $v_1, \ldots, v_n$ линейно независимы, то ни один из v_i не выражается через оставшиеся (иначе бы мы нашли нетривиальную линейную комбинацию). А значит, выкинув любой из v_i мы обязательно получим систему, линейные комбинации которой не содержат v_i, то есть не порождающую систему.

[4] называемые векторами

(3)$\Rightarrow$(1). Надо доказать линейную независимость векторов. Если бы это было не так, то нашлась бы нетривиальная линейная комбинация этих векторов и как в доказательстве (2)$\Rightarrow$(1) мы бы могли выразить какой-то v_i через оставшиеся. Давайте предположим для простоты, что это v_n. Но тогда $v_1, \ldots, v_{n-1}$ была бы меньшая система порождающих. Действительно, любой вектор v расписывается как $r_1v_1+\ldots+r_{n-1}v_{n-1}+r_nv_n$. Но теперь мы можем выразить v_n через $v_1, \ldots, v_{n-1}$. Подставляя это выражение, получаем разложение для v через $v_1, \ldots, v_{n-1}$. Получено противоречие с минимальностью системы векторов $v_1, \ldots, v_n$. Значит она была линейно независима. $\square$

Если в векторном пространстве V существует система векторов $v_1, \ldots, v_n$ обладающая одним из свойств выше, то мы будем называть такую систему векторов *базисом* пространства V.

Описание всех векторных пространств с базисами

Пусть V – векторное пространство над $\mathbb{R}$ и пусть у нас есть базис $v_1, \ldots, v_n \in V$. Тогда любой вектор $u \in V$ единственным образом представляется в виде $u = a_1v_1 + \ldots + a_nv_n$. Надо лишь объяснить единственность. Если $u = b_1v_1 + \ldots + b_nv_n$, то $(a_1 - b_1)v_1 + \ldots + (a_n - b_n)v_n = 0$ – линейная комбинация равна нулю. Но так как v_i линейно независимы, это может быть лишь тривиальная линейная комбинация, т.е. все $a_i - b_i = 0$. Таким образом получаем биекцию между множеством векторов V и пространством столбцов $\mathbb{R}^n$, а именно каждому вектору v сопоставим столбец его координат. Тогда обратное отображение действует по правилу $u = \sum_{i=1}^n a_iv_i \mapsto \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. Кроме того обратите, что это сопоставление переводит сумму векторов в сумму столбцов, а также вектор умноженный на число соответствует столбцу умноженному на число. То есть эта биекция согласована со структурой векторного пространства. По сути это означает, что эти два пространства устроены одинаково с точки зрения теории векторных пространств. Любое свойство V, которое можно вытащить через «интерфейс» векторного пространства будет выполнено тогда и только тогда, когда то же самое свойство будет выполнено в $\mathbb{R}^n$.

Теперь важное замечание. Базис существует всегда! Только не всегда он состоит из конечного числа векторов. Я не хочу обсуждать бесконечные базисы. Однако, важно понимать следующее.

Утверждение. *Пусть V – векторное пространство. Любые два базиса V имеют одинаковое число элементов.*[5]

Если V – векторное пространство, то число элементов в базисе называется *размерностью* векторного пространства V и обозначается $\dim V$. Если конечного базиса нет, будем писать $\dim V = \infty$.[6]

Примеры Для простых пространств размерность в точности равна числу коэффициентов, которые необходимы для задания векторов: $\dim \mathbb{R}^n = n$ или $\dim \mathrm{M}_{m\,n}(\mathbb{R}) = mn$. Чуть позже мы увидим, что для $V = \{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in \mathrm{M}_{m\,n}(\mathbb{R})$, $\dim V$ равна числу свободных переменных СЛУ $Ay = 0$. То есть это число характеризует на сколько много есть решений у системы.

Утверждение. *Пусть V – векторное пространство и пусть $U \subseteq V$ – подпространство.*

1. $\dim U \leqslant \dim V$. *В частности, если V обладает конечным базисом, то и U обязательно обладает конечным базисом.*

2. $\dim U = \dim V$ *тогда и только тогда, когда $U = V$.*

Размерность – это величина, показывающая на сколько векторное пространство большое и характеризует «количество степеней свободы» в пространстве. Кроме того, это понятие согласовано с нашей интуицией: прямая $\mathbb{R}^1$ имеет размерность 1, плоскость $\mathbb{R}^2$ – размерность 2, а пространство $\mathbb{R}^3$ – размерность 3.

Смысл базиса

Напомню, что «по-простому» векторное пространство – это все что угодно, где элементы можно складывать и умножать на числа. Формально надо проверить еще какие-то аксиомы, но все, что возникает в реальной жизни, обязательно будет удовлетворять им. Главные примеры – $\mathbb{R}^n$ и $\mathrm{M}_{m\,n}(\mathbb{R})$.

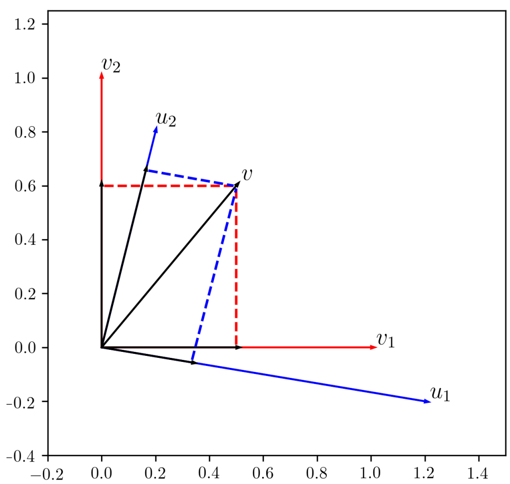
[5]Это утверждение верно и для конечных и для бесконечных базисов, но для работы с бесконечными базисами надо знать как сравнивать бесконечные множества.

[6]На самом деле, теория множеств позволяет различать какие-то из бесконечных множеств, но мы этого делать не будем.

Пусть V – векторное пространство. Напомню, что базис – это набор векторов $v_1, \ldots, v_n$ который линейно независим и через них все выражается. То есть уравнение $x_1 v_1 + \ldots + x_n v_n = 0$ имеет только нулевое решение $x_i = 0$ и любой вектор $v \in V$ представляется (по безысходности единственным образом) в виде $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$. Смысл базиса вот в чем: если вы его выбрали, то вы можете отождествить V с $\mathbb{R}^n$ следующим образом: если $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$, то ему соответствует единственный столбец $\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$. Таким образом, про любое векторное пространство, когда это необходимо, можно думать как просто про $\mathbb{R}^n$. Но надо помнить, что (1) часто необязательно выбирать базис и без выбора базиса задачи могут проще решаться и (2) базис можно выбрать не единственным образом, и если его выбрать по-другому вычисления могут стать либо проще либо сложнее (как повезет).

Смена базиса

На рисунке ниже изображена плоскость с двумя базисами: красный v_1, v_2 и синий u_1, u_2. При этом вектор v можно разложить как по одному, так и по другому базису. В зависимости от этого у него будут разные координаты.



Давайте теперь поговорим, как это устроено в общем случае. Пусть у нас есть векторное пространство V. Можно считать, что $V = \mathbb{R}^n$ для удобства. Пусть у вас в V есть два базиса: $e_1, \ldots, e_n$ и $f_1, \ldots, f_n$.[7] И пусть у нас есть вектор $v \in V$. Тогда он раскладывается по обоим базисам:

$$v = x_1 e_1 + \ldots + x_n e_n = \begin{pmatrix} e_1 & \ldots & e_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad v = y_1 f_1 + \ldots + y_n f_n = \begin{pmatrix} f_1 & \ldots & f_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Вопрос: а как связаны координаты x_i и координаты y_i? Вот на этот вопрос мы и попытаемся ответить. Для начала надо знать, как связаны базисы e_i и f_i. По определению базиса для e_i каждый вектор f_i представляется в виде:

$$f_i = c_{1i} e_1 + \ldots + c_{ni} e_n = \begin{pmatrix} e_1 & \ldots & e_n \end{pmatrix} \begin{pmatrix} c_{1i} \\ \vdots \\ c_{ni} \end{pmatrix}$$

Если положить $C = (c_{ij}) \in \mathrm{M}_n(\mathbb{R})$, то предыдущий набор равенств можно записать кратко в виде:

$$\begin{pmatrix} f_1 & \ldots & f_n \end{pmatrix} = \begin{pmatrix} e_1 & \ldots & e_n \end{pmatrix} \begin{pmatrix} c_{11} & \ldots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \ldots & c_{nn} \end{pmatrix} = \begin{pmatrix} e_1 & \ldots & e_n \end{pmatrix} C$$

[7] Напомню, что базисы обязательно имеют одинаковый размер.

Такая матрица называется *матрицей перехода* от e_i к f_i. Теперь запишем наш вектор v так

$$v = \begin{pmatrix} f_1 & \dots & f_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} C \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{и} \quad v = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Так как разложение по базису e_i однозначно (по определению базиса), то получаем связь на координаты

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = C^{-1} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Запоминать это правило надо так: если от базиса e к базису f мы перешли с помощью умножения справа на матрицу C, то на координатах у нас отображение в обратную сторону с помощью умножения на матрицу C слева (то есть тоже с другой стороны). Еще полезно держать перед глазами вот эту таблицу.

базис	новый $\xleftarrow{\cdot C}$ старый
координаты	новые $\xrightarrow{C\cdot}$ старые

Смена координат в $\mathbb{R}^n$

В случае, когда мы работаем в $\mathbb{R}^n$ вот как можно думать про равенство

$$\begin{pmatrix} f_1 & \dots & f_n \end{pmatrix} = \begin{pmatrix} e_1 & \dots & e_n \end{pmatrix} C$$

В этом случае каждый вектор f_i – это вектор столбец высоты n. Потому левая часть равенства (когда там записаны n столбцов) представляет из себя матрицу n на n. То есть $(f_1 \dots f_n) \in \mathrm{M}_n(\mathbb{R})$. Аналогично можно думать, что $(e_1 \dots e_n) \in \mathrm{M}_n(\mathbb{R})$. Таким образом можно найти матрицу перехода: $C = (e_1 \dots e_n)^{-1}(f_1 \dots f_n)$. Тут есть важный частный случай, предположим, что e_i – стандартный базис, т.е. e_i имеет 1 на i-ом месте и нули в остальных местах. Тогда матрица перехода $C = (f_1 \dots f_n)$. То есть C составлена из координат f_i в стандартном базисе.

Линейные оболочки

Пусть V – векторное пространство. Для простоты можно думать, что это $\mathbb{R}^n$. И пусть у нас задан произвольный набор векторов $v_1, \dots, v_k \in V$.[8] Понятно, что конечный набор не образует подпространство, но можно рассмотреть наименьшее подпространство, содержащее данные вектора. Это подпространство обозначается $\langle v_1, \dots, v_k \rangle$ и состоит из всех линейных комбинаций v_i, т.е.

$$\langle v_1, \dots, v_k \rangle = \{\lambda_1 v_1 + \ldots + \lambda_k v_k \in V \mid \lambda_i \in \mathbb{R}\}$$

Заметим, что $\dim\langle v_1, \dots, v_k \rangle \leqslant k$ и равенство достигается тогда и только тогда, когда v_i линейно независимы. Кроме того, по определению $v_1, \dots, v_k$ являются порождающими для линейной оболочки $\langle v_1, \dots, v_k \rangle$.

Пусть вектора $v_1, \dots, v_k$ линейно зависимы. Выделим среди них наибольшее линейно независимое подмножество. После перенумерации векторов, можно считать, что это $v_1, \dots, v_m$, где $m \leqslant k$. Тогда каждый вектор v_j при $j > m$ будет линейно выражаться через первые m векторов. Последнее означает, что $\langle v_1, \dots, v_k \rangle = \langle v_1, \dots, v_m \rangle$, но теперь вектора справа линейно независимы.

Выделение базиса из системы векторов

Дано Пусть $v_1, \dots, v_m \in \mathbb{R}^n$ – вектора и $V = \langle v_1, \dots, v_m \rangle$ – их линейная оболочка.

Задача Среди векторов $v_1, \dots, v_m$ найти базис пространства V и разложить оставшиеся вектора по этому базису.

[8] Вектора могут быть любые, могут быть линейно зависимыми или независимыми, могут быть хоть все нулевыми или просто одинаковыми.

Алгоритм

1. Запишем вектора $v_1, \ldots, v_m$ по столбцам в матрицу $A \in \mathrm{M}_{n\,m}(\mathbb{R})$. Например, при $n = 3$, $m = 5$

$$A = \begin{pmatrix} v_{11} & v_{21} & v_{31} & v_{41} & v_{51} \\ v_{12} & v_{22} & v_{32} & v_{42} & v_{52} \\ v_{13} & v_{23} & v_{33} & v_{43} & v_{53} \end{pmatrix}$$

2. Приведем матрицу A элементарными преобразованиями строк к улучшенному ступенчатому виду. Например

$$A' = \begin{pmatrix} 1 & 0 & a_{31} & 0 & a_{51} \\ 0 & 1 & a_{32} & 0 & a_{52} \\ 0 & 0 & 0 & 1 & a_{53} \end{pmatrix}$$

3. Пусть $k_1, \ldots, k_r$ – номера главных позиций в матрице A'. Тогда вектора $v_{k_1}, \ldots, v_{k_r}$ образуют базис V. Например, в примере выше это вектора v_1, v_2 и v_4.

4. Пусть v_i – вектор соответствует неглавной позиции в A'. Тогда в i-ом столбце A' записаны координаты разложения v_i через найденный базис выше. Например, в примере выше $v_3 = a_{31}v_1 + a_{32}v_2$ и $v_5 = a_{51}v_1 + a_{52}v_2 + a_{53}v_4$.

Пример Пусть

$$v_1 = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 5 \\ 12 \\ 7 \end{pmatrix}, v_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_5 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$$

Тогда

$$\begin{pmatrix} 1 & 1 & 5 & 1 & -1 \\ 3 & 2 & 12 & 1 & 1 \\ 2 & 1 & 7 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 5 & 1 & -1 \\ 0 & 0 & 0 & -1 & 2 \\ 2 & 1 & 7 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 1 & 5 & 1 & -1 \\ 0 & 0 & 0 & -1 & 2 \\ 1 & 0 & 2 & 0 & 1 \end{pmatrix} \mapsto$$
$$\mapsto \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 1 & 1 & 5 & 1 & -1 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 3 & 1 & -2 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 2 & 0 & 1 \\ 0 & 1 & 3 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Тогда v_1, v_2 и v_4 – базис линейной оболочки. $v_3 = 2v_1 + 3v_2$ и $v_5 = v_1 - 2v_4$.

Подпространства в $\mathbb{R}^n$

Мы хотим понять как устроены все возможные подпространства в $\mathbb{R}^n$. Для начала надо понять, а как вообще задавать подпространства в $\mathbb{R}^n$. Существует два способа:

1. С помощью образующих векторов: $U \subseteq \mathbb{R}^n$ задано в виде $U = \langle v_1, \ldots, v_k \rangle$, где $v_1, \ldots, v_k \in \mathbb{R}^n$ – некоторые вектора. В этом случае часто бывает полезно, чтобы вектора v_i были линейно независимыми.

2. С помощью СЛУ: $U \subseteq \mathbb{R}^n$ задано в виде $U = \{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ – некоторая матрица. В этом случае часто бывает полезно, чтобы строки матрицы A были линейно независимыми.

Любое пространство можно задать любым из этих двух способов, а значит, если пространство задано одним из этих способов, его можно задать и другим.

Фундаментальная система решений (ФСР)

Так как $\mathbb{R}^n$ обладает конечным базисом, то и $\{y \in \mathbb{R}^n \mid Ay = 0\}$, где $A \in \mathrm{M}_{m\,n}(\mathbb{R})$, тоже обладает конечным базисом, причем количество базисных элементов не превосходит n. Любой базис такого пространства называется *фундаментальной системой решений*. Наша задача научиться находить его.

Дано Матрица $A \in \mathrm{M}_{m\,n}(\mathbb{R})$.

Задача Найти базис пространства $U = \{y \in \mathbb{R}^n \mid Ay = 0\}$.

Алгоритм

1. Приведем матрицу A к улучшенному ступенчатому виду. Пусть например она имеет вид

$$\begin{pmatrix}1 & a_{12} & 0 & a_{14} & 0 & a_{16}\\ 0 & 0 & 1 & a_{24} & 0 & a_{26}\\ 0 & 0 & 0 & 0 & 1 & a_{36}\end{pmatrix}$$

2. Теперь $\dim U$ равна количеству свободных переменных. ФСР строится так: для каждой свободной переменной будет свой базисный вектор. Такую свободную переменную полагаем 1, а остальные свободные переменные 0. После чего рассчитываем значения главных переменных. В примере выше, свободные переменные x_2, x_4 и x_6. Тогда ФСР

$$v_2 = \begin{pmatrix}-a_{12}\\ \underline{1}\\ 0\\ \underline{0}\\ 0\\ \underline{0}\end{pmatrix},\quad v_4 = \begin{pmatrix}-a_{14}\\ \underline{0}\\ -a_{24}\\ \underline{1}\\ 0\\ \underline{0}\end{pmatrix},\quad v_6 = \begin{pmatrix}-a_{16}\\ \underline{0}\\ -a_{26}\\ \underline{0}\\ -a_{36}\\ \underline{1}\end{pmatrix},$$

В векторах выше подчеркнуты позиции свободных переменных, которые мы задаем сами.

Ранг системы векторов

Пусть V – некоторое векторное пространство. Системой векторов называется последовательность $(v_1, \ldots, v_k)$ из векторов V, в которой векторы v_i могут повторяться.[9]

По определению рангом системы $(v_1, \ldots, v_k)$ называется максимальное количество линейно независимых векторов в этой системе. Ранг такой системы будет обозначаться $\operatorname{rk}(v_1, \ldots, v_k)$.

Утверждение. *Если $(v_1, \ldots, v_k)$ – некоторая система векторов в векторном пространстве V, то* $\operatorname{rk}(v_1, \ldots, v_k) = \dim\langle v_1, \ldots, v_k\rangle$.

Матричный ранг

Пусть $A \in \operatorname{M}_{m\,n}(\mathbb{R})$ – некоторая матрица. Сейчас я определю пять разных определений ранга матрицы. Все эти ранги между собой совпадают и полученная величина будет просто называться рангом матрицы A и обозначаться $\operatorname{rk} A$.

Определение. Пусть $A_1, \ldots, A_n \in \mathbb{R}^m$ – столбцы матрицы A, то есть $A = (A_1|\ldots|A_n)$. Тогда столбцовым рангом матрицы A называется ранг системы $(A_1, \ldots, A_n)$, то есть $\operatorname{rk}_{\text{столб}} A = \operatorname{rk}(A_1, \ldots, A_n)$.

Определение. Пусть $A_1, \ldots, A_m \in \mathbb{R}^n$ – строки матрицы A, то есть $A^t = (A_1|\ldots|A_m)$. Тогда строковым рангом матрицы A называется ранг системы $(A_1, \ldots, A_m)$, то есть $\operatorname{rk}_{\text{стр}} A = \operatorname{rk}(A_1, \ldots, A_m)$.

Определение. Факториальным рангом матрицы A называется следующее число

$$\min\{k \mid A = BC,\ \text{где } B \in \operatorname{M}_{m\,k}(\mathbb{R}),\, C \in \operatorname{M}_{k\,n}(\mathbb{R})\}$$

то есть это минимальное число k такое, что матрица A представима в виде произведения матриц BC, где общая размерность для B и C, по которой они перемножаются, есть k.

Определение. Тензорным рангом матрицы A называется следующее число

$$\min\{k \mid A = x_1y_1^t + \ldots + x_ky_k^t,\ \text{где } x_i \in \mathbb{R}^m,\, y_i \in \mathbb{R}^n\}$$

то есть это минимальное число k такое, что матрица A представима в виде суммы k «тощих» матриц вида xy^t, где $x \in \mathbb{R}^m$ и $y \in \mathbb{R}^n$.

Если я в матрице A выделю какой-нибудь набор из k строк и одновременно набор из k столбцов, а потом возьму матрицу составленную из элементов на пересечении этих строк и столбцов, то я получу квадратную матрицу размера k. Такие матрицы мы будем называть квадратными подматрицами матрицы A.

[9] В подобной ситуации повторяющиеся векторы различаются по индексу – «ключу».

Определение. Минорным рангом матрицы A называется размер наибольшей невырожденной квадратной подматрицы.[10]

Главное для нас следующее утверждение.

Утверждение. *Для любой матрицы $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ все пять видов ранга совпадают и не превосходят $\min(m, n)$.*

Примеры

1. В начале заметим, что матрица имеет ранг 0 тогда и только тогда, когда $A = 0$.
2. Ранг матрицы A равен единице тогда и только тогда, когда она не нулевая и все столбцы пропорциональны одному общему столбцу (или что эквивалентно, все строки пропорциональны одной общей строке). Если воспользоваться определением факториального ранга, то мы видим, что тогда матрица A имеет вид $A = xy^t$, где $x \in \mathbb{R}^m$ и $y \in \mathbb{R}^n$ – ненулевые вектора.

Свойства ранга

Прежде всего надо запомнить как ранг связан с матричными операциями.

Утверждение. *Пусть $A, B \in \mathrm{M}_{m\,n}(\mathbb{R})$, тогда*

$$|\operatorname{rk} A - \operatorname{rk} B| \leqslant \operatorname{rk}(A + B) \leqslant \operatorname{rk} A + \operatorname{rk} B$$

Надо понимать, что, во-первых, все эти эффекты можно увидеть на диагональных матрицах; во-вторых, все границы неравенств достигаются. Смысл этого утверждения вот в чем: если вы шевелите матрицу A с помощью матрицы B, то ранг A может измениться не более чем на ранг B в любую сторону. Теперь посмотрим на матрицы: $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, $B = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ и $C = -A$. Тогда $\operatorname{rk}(A + B) = \operatorname{rk} A + \operatorname{rk} B$ и $\operatorname{rk}(A + C) = \operatorname{rk} A - \operatorname{rk} C$.

Утверждение. *Пусть $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ и $B \in \mathrm{M}_{n\,k}(\mathbb{R})$, тогда*

$$\operatorname{rk} A + \operatorname{rk} B - n \leqslant \operatorname{rk}(AB) \leqslant \min(\operatorname{rk} A, \operatorname{rk} B)$$

Как и в предыдущем случае, все обе границы неравенства достигаются и все можно пронаблюдать на диагональных матрицах. Пусть $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ и $B = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Тогда $\operatorname{rk}(AA) = \operatorname{rk} A$ и $\operatorname{rk}(AB) = \operatorname{rk} A + \operatorname{rk} B - 2$.

Утверждение. *Пусть $A \in \mathrm{M}_n(\mathbb{R})$ – квадратная матрица. Тогда $\operatorname{rk} A = n$ тогда и только тогда, когда A невырождена, т.е. $\det A \neq 0$.*

Таким образом на ранг можно смотреть как на степень невырожденности матрицы A. Самый высокий ранг у невырожденных матриц, самый маленький у нулевой, но есть еще и промежуточные состояния.

Утверждение. *Если матрица $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ находится в ступенчатом виде и имеет k ступенек, то ее ранг равен k.*

Это утверждение вместе со следующим дают эффективный способ считать ранг.

Утверждение. *Для любой матрицы $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ и любых невырожденных матриц $C \in \mathrm{M}_m(\mathbb{R})$ и $D \in \mathrm{M}_n(\mathbb{R})$ верно: $\operatorname{rk} A = \operatorname{rk}(CA) = \operatorname{rk}(AD)$.*[11]

В частности ранг матрицы не меняется при элементарных преобразованиях столбцов и строк. Обычно этим пользуются для нахождения ранга. Более того, если $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ имеет ранг r, то элементарными преобразованиями строк и столбцов она приводится к виду

$$A \mapsto \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix}, \text{ где } E \in \mathrm{M}_r(\mathbb{R}) - \text{единичная матрица}$$

Следствием данного замечания является следующее.

Утверждение. *Для любых матриц $A \in \mathrm{M}_{m\,n}(\mathbb{R})$ и $B \in \mathrm{M}_{s\,t}(\mathbb{R})$ имеем*

$$\operatorname{rk} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \operatorname{rk} A + \operatorname{rk} B$$

[10] На самом деле можно дать более сильное определение, а именно, минорный ранг – это размер любой максимальной невырожденной подматрицы. То есть мы берем какую-то квадратную подматрицу, которая невырождена, а любая большая подматрица уже вырождена. Оказывается, что все максимальные невырожденные подматрицы имеют одинаковый размер и он называется минорным рангом.

[11] В частности ранг не меняется при элементарных преобразованиях строк и столбцов.