Математика для Data Science. Линейная алгебра.
Решения задач

Содержание

Замечание. Таким цветом отмечены ссылки на сайт Stepik, а вот этим цветом — ссылки на страницы внутри этого файла.

Нейронные сети

Задача 1

На предыдущем шаге мы немного слукавили: в нашем описании нейронной сети с одним скрытым слоем не хвататет одного важного элемента. Какого — узнаете на следующем шаге. А в этой задаче вам предлагается убедиться, что без дополнительных элементов добавление линейного слоя ничего не дает с точки зрения вычислительной способности нейросети.

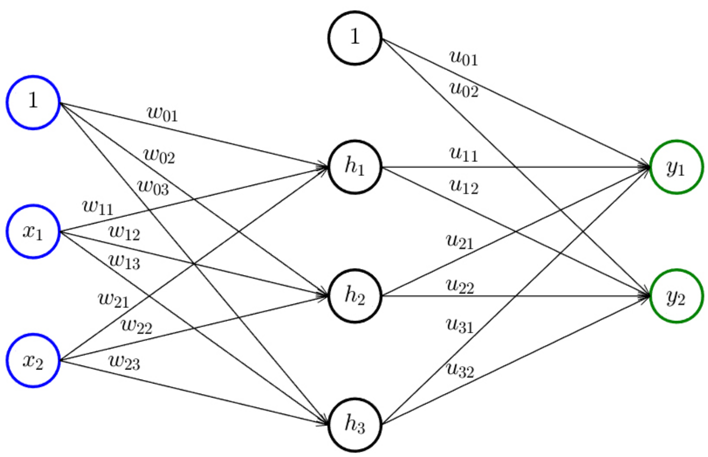
Вернемся к нейронной сети с предыдущего шага. Пусть:

- $m+1$ — число нейронов во входном слое с учетом нейрона сдвига
- $k+1$ — число нейронов в скрытом слое с учетом нейрона сдвига
- n — размер выхода

Докажите, что какими бы ни были значения весов в такой нейросети, для нее найдется эквивалентная однослойная линейная нейронная сеть с

- $m+1$ нейроном во входном слое с учетом нейрона сдвига
- n выходными нейронами

Нейросети называются *эквивалентными* в случае, если для любого входного вектора их выход одинаков.



Переформулируем эту задачу.

- Пусть даны матрицы W и U соответствующих размеров. При помощи этих матриц каждому вектору $\vec{x}$ ставится в соответствие вектор $\vec{y}$ (зависящий от $\vec{x}$), как это было описано на предыдущем шаге.
- Докажите, что найдётся матрица A, такая что для любого $\vec{x}$ соответствующий $\vec{y}$ можно вычислить так: $\vec{y} \ = A^T\vec{x}'$. Эта матрица A и будет задавать однослойную нейронную сеть, которая эквивалентна нашей двухслойной сети.

Подсказка. Композиция линейных функций — тоже линейная функция.

Решение. Обозначим $x_0 := 1$ и $h_0 := 1$. Тогда $h_t = \sum\limits_{i=1}^{m+1} w_{it}x_{i-1}$ и $y_l = \sum\limits_{j=1}^{k+1} u_{jl}h_{j-1} = \sum\limits_{j=1}^{k+1} u_{jl} \sum\limits_{i=1}^{m+1} w_{i,j-1}x_{i-1} = \sum\limits_{j=1}^{k+1}\sum\limits_{i=1}^{m+1}(u_{jl}w_{i,j-1})x_{i-1} = \sum\limits_{i=1}^{m+1}\left(\sum\limits_{j=1}^{k+1} u_{jl}w_{i,j-1}\right)x_{i-1}$, то есть выход y_l — сумма входных нейронов $x_1, x_2, \ldots, x_m$ с коэффициентами, указанными в больших скобках. Требумая в задаче эквивалентная однослойная линейная нейронная сеть задаётся как раз такими весами.

Задача 2

Рассмотрим нейронную сеть, у которой на нулевом слое n_0 нейронов, на первом слое n_1 нейронов, ..., на k-ом слое n_k нейронов, если не учитывать нейроны сдвига.

Сколько будет параметров у такой модели?

Подсказка. Вспомните, сколько параметров у однослойной нейронной сети, то есть какой размер имеет матрица весов.

Решение. Поскольку на нулевом слое n_0 нейронов, на первом слое n_1 нейронов, то матрица весов первого линейного слоя будет иметь размерность $n_1 \times (n_0 + 1)$ с учетом нейрона сдвига. А значит, число параметров будет $n_1(n_0 + 1)$. Аналогично, рассматривая второй слой, получим $n_2(n_1 + 1)$ параметров. И так далее до выходного слоя.

Суммируя полученные выражения, получаем, что всего параметров у модели будет $n_1(n_0 + 1) + n_2(n_1 + 1) + \cdots + n_k(n_{k-1} + 1)$

Подпространства и линейные комбинации

Задача 1

Пусть V – подпространство в $\mathbb{R}^n$, порождённое векторами $\vec{a}_1, \ldots, \vec{a}_k$.

Докажите, что

1. Если $\vec{x} \in V$, $\vec{y} \in V$, то и $\vec{x} + \vec{y} \in V$. Другими словами – если $\vec{x}$ и $\vec{y}$ являются линейными комбинациями порождающих векторов V, то и $\vec{x} + \vec{y}$ тоже является линейной комбинацией порождающих векторов V.

2. Если $\vec{x} \in V$ и $c \in \mathbb{R}$, то $c\vec{x} \in V$.

В частности, из второго свойства следует, что

- $0\vec{x} = \vec{0} \in V$, то есть V содержит ноль
- если $\vec{x} \in V$, то и $-\vec{x} = (-1)\vec{x} \in V$, то есть V содержит противоположные векторы для всех своих векторов

Подсказка. Найдите, с какими коэффициентами $\vec{x} + \vec{y}$ и $c\vec{x}$ выражаются через векторы $\vec{a}_1, \ldots, \vec{a}_k$.

Решение.

1. Пусть $\vec{x}$ и $\vec{y}$ являются линейными комбинациями векторов $\vec{a}_1, \ldots, \vec{a}_k$, докажем, что $\vec{x} + \vec{y}$ тоже является линейной комбинацией этих векторов. Действительно, пусть $\vec{x} = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_k\vec{a}_k$, где $\alpha_i \in \mathbb{R}$ — коэффициенты в линейной комбинации. Аналогично, пусть $\vec{y} = \beta_1\vec{a}_1 + \beta_2\vec{a}_2 + \cdots + \beta_k\vec{a}_k$, где $\beta_i \in \mathbb{R}$. Тогда $\vec{x} + \vec{y} = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_k\vec{a}_k + \beta_1\vec{a}_1 + \beta_2\vec{a}_2 + \cdots + \beta_k\vec{a}_k = (\alpha_1 + \beta_1)\vec{a}_1 + (\alpha_2 + \beta_2)\vec{a}_2 + \cdots + (\alpha_k + \beta_k)\vec{a}_k$, то есть $\vec{x} + \vec{y}$ тоже является линейной комбинацией векторов $\vec{a}_1, \ldots, \vec{a}_k$ с коэффициентами $(\alpha_i + \beta_i)$. А значит, $\vec{x} + \vec{y} \in V$.

2. Как и в первом пункте, пусть $\vec{x} = \alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_k\vec{a}_k$, где $\alpha_i \in \mathbb{R}$. Тогда $c\vec{x} = c(\alpha_1\vec{a}_1 + \alpha_2\vec{a}_2 + \cdots + \alpha_k\vec{a}_k) = (c\alpha_1)\vec{a}_1 + (c\alpha_2)\vec{a}_2 + \cdots + (c\alpha_k)\vec{a}_k$. Итак, $c\vec{x}$ является линейной комбинацией векторов $\vec{a}_1, \ldots, \vec{a}_k$ с коэффициентами α_i. Другими словами, $c\vec{x} \in V$.

Задача 2

Набор векторов $\vec{a}_1, \ldots, \vec{a}_k$ порождает пространство V. Докажите, что следующие утверждения эквивалентны:

1. этот набор линейно независим

2. при удалении любого вектора из набора оставшиеся векторы не порождают V

То есть нужно доказать, что $1 \Rightarrow 2$ и $2 \Rightarrow 1$.

Заметьте, что утверждение этой задачи можно переформулировать так. Следующие утверждения эквивалентны:

1. набор $\vec{a}_1, \ldots, \vec{a}_k$ линейно зависим
2. найдётся такой вектор $\vec{a}_i$, что при удалении его из набора оставшиеся векторы всё равно порождают V

Подсказка. Воспользуйтесь результатами предыдущего шага.

Решение. Докажем, что из первого утверждения следует второе, а затем наоборот.

$\mathbf{1} \Rightarrow \mathbf{2}$. Итак, пусть дано, что $\vec{a}_1, \ldots, \vec{a}_k$ линейно независимы. Мы хотим доказать, что при удалении любого вектора из набора оставшиеся векторы не порождают V. Для простоты записи будем считать, что удаляем мы $\vec{a}_1$ (аналогичные рассуждения будут верны для удаления любого $\vec{a}_i$).

Докажем от противного: то есть, пусть после удаления $\vec{a}_1$ оставшиеся векторы всё ещё порождают V. Тогда, поскольку $a_1 \in V$, он должен выражаться через $\vec{a}_2, \ldots, \vec{a}_k$, но тогда по утверждению с предыдущего шага векторы $\vec{a}_1, \ldots, \vec{a}_k$ должны быть линейно зависимы, что неверно. Мы получили противоречие, значит, наше изначальное предположение было неверно. То есть мы доказали, что при удалении любого вектора из набора оставшиеся векторы не порождают V.

$\mathbf{2} \Rightarrow \mathbf{1}$. Докажем в обратную сторону. То есть, пусть дано, что при удалении любого вектора из набора оставшиеся векторы не порождают V. А мы хотим доказать, что $\vec{a}_1, \ldots, \vec{a}_k$ линейно независимы. Опять же будем доказывать от противного. Пусть векторы $\vec{a}_1, \ldots, \vec{a}_k$ линейно зависимы. Тогда по утверждению с предыдущего шага хотя бы один из этих векторов представляется в виде линейной комбинации оставшихся векторов. И вновь для простоты записи будем считать, что это вектор $\vec{a}_1$. Итак, $\vec{a}_1 = \lambda_2 a_2 + \lambda_3 a_3 + \cdots + \lambda_k a_k$, но тогда, если мы вместо $\vec{a}_1$ будем писать получившуюся линейную комбинацию, то фактически все векторы V мы можем выразить, не используя $\vec{a}_1$. А это и значит, что, если мы удалим $\vec{a}_1$, то $\vec{a}_2, \ldots, \vec{a}_k$ всё равно будут порождать V. Мы пришли к противоречию. Итак, векторы $\vec{a}_1, \ldots, \vec{a}_k$ действительно будут линейно независимы

Задача 3

Обозначение. Утверждение "V порождено векторами $\vec{a}_1, \ldots, \vec{a}_k$"записывают так: $V = \operatorname{Span}\{\vec{a}_1, \ldots, \vec{a}_k\}$.

Тем самым, на прошлом шаге мы доказали, что $\operatorname{Span}\{\begin{pmatrix}1\\1\end{pmatrix}\} = \operatorname{Span}\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}3\\3\end{pmatrix}\}$.

В каждом из следующих пунктов докажите, что два набора векторов порождают одно и то же подпространство. В пунктах $1, 2, 4, 6$ считайте, что векторы лежат в $\mathbb{R}^n$.

1. $\operatorname{Span}\{\vec{a}_1\} = \operatorname{Span}\{\vec{a}_1, 100\vec{a}_1\}$
2. $\operatorname{Span}\{\vec{a}_1, \vec{a}_2\} = \operatorname{Span}\{\vec{a}_1, \vec{a}_1, \vec{a}_2\}$ (мы нигде не говорили, что векторы в наборе не должны повторяться)
3. $\operatorname{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} = \operatorname{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}4\\7\\0\end{pmatrix}\}$
4. $\operatorname{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\} = \operatorname{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k, c_1\vec{a}_1 + c_2\vec{a}_2 + \cdots + c_k\vec{a}_k\}$, где $c_1, \ldots, c_k$ – любые числа.
5. $\operatorname{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} = \operatorname{Span}\{\begin{pmatrix}1\\100\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\}$
6. $\operatorname{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\} = \operatorname{Span}\{\vec{a}_1 + c\vec{a}_2, \vec{a}_2, \ldots, \vec{a}_k\}$, где c – любое число

Подсказка. Докаазательства аналогичны доказательству с прошлого шага.

Решение. Доказательство будем строить так: сначала докажем, что любой элемент первого множества является элементом второго, а затем — наоборот (что любой элемент второго множества является элементом первого).

1. $\mathrm{Span}\{\vec{a}_1\} = \mathrm{Span}\{\vec{a}_1, 100\vec{a}_1\}$

   - Любой элемент $\mathrm{Span}\{\vec{a}_1\}$ имеет вид $\lambda\vec{a}_1$. Но $\lambda\vec{a}_1 = \lambda\vec{a}_1 + 0 \cdot (100\vec{a}_1)) \in \mathrm{Span}\{\vec{a}_1, 100\vec{a}_1\}$
   - Любой элемент $\mathrm{Span}\{\vec{a}_1, 100\vec{a}_1\}$ имеет вид $\lambda_1\vec{a}_1 + \lambda_2 100\vec{a}_1 = (\lambda_1 + 100\lambda_2)\vec{a}_1 \in \mathrm{Span}\{\vec{a}_1\}$

2. $\mathrm{Span}\{\vec{a}_1, \vec{a}_2\} = \mathrm{Span}\{\vec{a}_1, \vec{a}_1, \vec{a}_2\}$ (мы нигде не говорили, что векторы в наборе не должны повторяться)

   - $\vec{x} \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2\} \Rightarrow \vec{x} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + 0\vec{a}_2 \in \mathrm{Span}\{\vec{a}_1, \vec{a}_1, \vec{a}_2\}$
   - $\vec{x} \in \mathrm{Span}\{\vec{a}_1, \vec{a}_1, \vec{a}_2\} \Rightarrow \vec{x} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_1 + \lambda_3\vec{a}_2 = (\lambda_1 + \lambda_2)\vec{a}_1 + \lambda_3\vec{a}_2 \in \mathrm{Span}\{\vec{a}_1, \vec{a}_1\}$

3. $\mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} = \mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}4\\7\\0\end{pmatrix}\}$

   - $\vec{x} \in \mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} \Rightarrow \vec{x} = \lambda_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\0\end{pmatrix} = \lambda_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + 0 \begin{pmatrix}4\\7\\0\end{pmatrix} \in \mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}4\\7\\0\end{pmatrix}\}$
   - $\vec{x} \in \mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}4\\7\\0\end{pmatrix}\} \Rightarrow \vec{x} = \lambda_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\0\end{pmatrix} + \lambda_3 \begin{pmatrix}4\\7\\0\end{pmatrix} = \begin{pmatrix}\lambda_1 + 4\lambda_3\\ \lambda_2 + 7\lambda_3\\ 0\end{pmatrix} = (\lambda_1 + 4\lambda_3) \begin{pmatrix}1\\0\\0\end{pmatrix} + (\lambda_2 + 7\lambda_3) \begin{pmatrix}0\\1\\0\end{pmatrix} \in \mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}4\\7\\0\end{pmatrix}\}$

4. $\mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\} = \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k, c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_k\vec{a}_k\}$, где $c_1, \dots, c_k$ – любые числа.

   - $\vec{x} \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\} \Rightarrow \vec{x} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \dots + \lambda_k\vec{a}_k = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \dots + \lambda_k\vec{a}_k + 0(\vec{a}_k, c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_k\vec{a}_k) \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k, c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_k\vec{a}_k\}$
   - $\vec{x} \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k, c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_k\vec{a}_k\} \Rightarrow \vec{x} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \dots + \lambda_k\vec{a}_k + \lambda_{k+1}(c_1\vec{a}_1 + c_2\vec{a}_2 + \dots + c_k\vec{a}_k) = (\lambda_1 + c_1\lambda_{k+1})\vec{a}_1 + (\lambda_2 + c_2\lambda_{k+1})\vec{a}_2 + \dots + (\lambda_k + c_k\lambda_{k+1})\vec{a}_k \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\}$

   Замечание. Этот пункт — обобщение предыдущего (в нём $\vec{a}_1 = \vec{e}_1$, $\vec{a}_2 = \vec{e}_2$ и $c_1 = 4$, $c_2 = 7$)

5. $\mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} = \mathrm{Span}\{\begin{pmatrix}1\\100\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\}$

   Эта задача — частный случай следующего пункта, в его обозначениях $\vec{a}_1 = \vec{e}_1$, $\vec{a}_2 = \vec{e}_2$ и $c = 100$.

6. $\mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\} = \mathrm{Span}\{\vec{a}_1 + c\vec{a}_2, \vec{a}_2, \dots, \vec{a}_k\}$, где c – любое число.

   - $\vec{x} \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\} \Rightarrow \vec{x} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \dots + \lambda_k\vec{a}_k = \lambda_1(\vec{a}_1 + c\vec{a}_2) + (\lambda_2 - \lambda_1 c)\vec{a}_2 + \dots + \lambda_k\vec{a}_k \in \mathrm{Span}\{\vec{a}_1 + c\vec{a}_2, \vec{a}_2, \dots, \vec{a}_k\}$
   - $\vec{x} \in \mathrm{Span}\{\vec{a}_1 + c\vec{a}_2, \vec{a}_2, \dots, \vec{a}_k\} \Rightarrow \vec{x} = \lambda_1(\vec{a}_1 + c\vec{a}_2) + \lambda_2\vec{a}_2 + \dots + \lambda_k\vec{a}_k = \lambda_1\vec{a}_1 + (\lambda_1 c + \lambda_2)\vec{a}_2 + \dots + \lambda_k\vec{a}_k \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k\}$

Базис, размерность, ранг

Задача 1

Вот две несложных задачи на понимание определения базиса.

1. Докажите, что любой базис пространства является порождающим набором этого пространства.
2. Докажите, что набор векторов $\vec{e}_1, \dots, \vec{e}_n$ является базисом пространства $\mathbb{R}^n$.

Подсказка. Утверждения доказываются по определению.

Решение.

1. Пусть $\vec{a}_1, \vec{a}_, \ldots, \vec{a}_n$ — базис пространства V. Докажем, что $V = \mathrm{Span}\{\vec{a}_1, \vec{a}_, \ldots, \vec{a}_n\}$. Действительно, если $\vec{x} \in V$, то по определению базиса $\vec{x}$ выражается через базис: $\vec{x} = \lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \cdots + \lambda_n\vec{a}_n$, то есть принадлежит $\mathrm{Span}\{\vec{a}_1, \vec{a}_, \ldots, \vec{a}_n\}$

2. $\vec{e}_1, \ldots, \vec{e}_n$ является базисом пространства $\mathbb{R}^n$, так как для любого $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ выполнено $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$. Единственность разложения докажем от противного: пусть $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n = y_1\vec{e}_1 + y_2\vec{e}_2 + \cdots + y_n\vec{e}_n$, где хотя бы для одного i выполнено $x_i \neq y_i$. Но тогда на i-ом месте вектора $\vec{x}$ с одной стороны должно стоять x_i, а с другой — y_i, что невозможно.

Задача 2

Оказывается, что базисы это в точности линейно независимые порождающие наборы. В этой задаче мы просим вас это доказать. Ниже строгая формулировка.

Дано пространство V и набор векторов $\vec{a}_1, \ldots, \vec{a}_k$ из V. Докажите, что следующие утверждения эквивалентны:

1. этот набор является базисом V

2. этот набор линейно независим и порождает V

То есть нужно доказать, что $1 \Rightarrow 2$ и $2 \Rightarrow 1$.

Комбинируя эту задачу с задачей из прошлого урока, мы доказали эквивалентность трёх утверждений:

1. набор является базисом V

2. набор линейно независим и порождает V

3. набор порождает V и при удалении любого элемента из набора оставшиеся векторы не порождают V

Подсказка для $2 \Rightarrow 1$. Пусть набор не является базисом, то есть какой-то вектор можно представить в виде двух разных линейных комбинаций векторов $\vec{a}_1, \ldots, \vec{a}_k$. Используя эти две комбинации, постройте нетривиальную линейную комбинацию, равную $\vec{0}$.

Решение. Докажем, что из первого утверждения следует второе, а затем наоборот.

$1 \Rightarrow 2$. Пусть $\vec{a}_1, \ldots, \vec{a}_k$ — базис V. Докажем, что этот набор линейно независим и порождает V. То, что набор порождает V, уже было доказано в задаче 1 этого урока. Остаётся доказать линейную независимость. Докажем её от противного: пусть $\vec{a}_1, \ldots, \vec{a}_k$ линейно зависимы. Но тогда по утверждению с шага прошлого урока хотя бы один из его векторов представляется в виде линейной комбинации оставшихся векторов. Как и раньше, для простоты записи пусть $\vec{a}_1 = \lambda_2 a_2 + \lambda_3 a_3 + \cdots + \lambda_k a_k$. Тогда $\vec{a}_1$ и $\lambda_2 a_2 + \lambda_3 a_3 + \cdots + \lambda_k a_k$ это два разных представления вектора $\vec{a}_1$ в нашем базисе, что невозможно по определению базиса. Итак, мы получили противоречие, и наше предположение было неверным. То есть $\vec{a}_1, \ldots, \vec{a}_k$ на самом деле линейно независимы.

$2 \Rightarrow 1$. Пусть $\vec{a}_1, \ldots, \vec{a}_k$ линейно независимы и порождают V. Докажем, что тогда этот набор — базис V. Вновь докажем от противного. Пусть набор не является базисом, тогда некоторый вектор $\vec{x} \in V$ можно представить в виде двух разных линейных комбинаций векторов $\vec{a}_1, \ldots, \vec{a}_k$. То есть, $\vec{x} = \alpha_1\vec{a}_1 + \cdots + \alpha_k\vec{a}_k = \beta_1\vec{a}_1 + \cdots + \beta_k\vec{a}_k$, где $\alpha_i \neq \beta_i$ для хотя бы одного i. Но тогда $(\alpha_1 - \beta_1)\vec{a}_1 + \cdots + (\alpha_k - \beta_k)\vec{a}_k = \vec{0}$ — нетривиальная линейная комбинация (ведь $\alpha_i - \beta_i \neq 0$). То есть векторы $\vec{a}_1, \ldots, \vec{a}_k$ линейно зависимы. Противоречие.

Задача 3

Напоминание-обозначение. Пусть дана функция $f : X \to Y$ и элемент $x \in X$. Тогда элемент $f(x) \in Y$ называется значением функции f на элементе x. Множество всех значений, которые принимает функция f, называется областью значений функции f. Область значений также называют образом функции f.

Пусть дана матрица A размера m на n. Тем самым она задаёт линейное отображение $A : \mathbb{R}^n \to \mathbb{R}^m$. Докажите, что образ A совпадает с линейным подпространством $\mathbb{R}^m$, порождённым столбцами матрицы A.

Подсказка. Вспомните, как можно интерпретировать столбцы матрицы.

Решение. В задаче прошлой недели мы доказали, что $A(\vec{e}_j)$ совпадает с j-ым столбцом матрицы A. Таким образом, если $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$, то, так как отображение линейно, $A(x) = x_1 A(\vec{e}_1) + \cdots + x_n A(\vec{e}_n)$, то есть в правой части записана линейная комбинация столбцов матрицы A.

Задача 4

Докажите, что ранг матрицы A равен размерности образа A.
Используя предыдущую задачу, можно переформулировать это утверждение так:
Ранг матрицы A равен размерности пространства, порождённого столбцами A.

Подсказка. Рассмотрите векторы $A(\vec{e}_1), \ldots, A(\vec{e}_n)$.

Решение. Для любого $\vec{x}$ выполнено $A(x) = x_1 A(\vec{e}_1) + \cdots + x_n A(\vec{e}_n)$. Тем самым в образе A лежат все линейные комбинации векторов $A(\vec{e}_1), \ldots, A(\vec{e}_n)$ и только они. Значит, образ A это $\mathrm{Span}\{A(\vec{e}_1), A(\vec{e}_2), \ldots, A(\vec{a}_n)\}$. Рассмотрим максимальный набор линейно независимых векторов среди столбцов A. Докажем, что этот набор является базисом $\mathrm{Span}\{A(\vec{e}_1), \ldots, A(\vec{a}_n)\}$. Этот набор линейно независим по построению. Если какой-то из оставшихся столбцов не выражается через векторы набора, то этот столбец можно добавить в набор. При этом набор останется линейно независимым. Значит, изначальный набор не был максимальным. Противоречие. Получается, все оставшиеся столбцы выражаются через вектора набора. Значит, все векторы из $\mathrm{Span}\{A(\vec{e}_1), \ldots, A(\vec{a}_n)\}$ выражаются через векторы набора. Следовательно, набор является базисом $\mathrm{Span}\{A(\vec{e}_1), \ldots, A(\vec{a}_n)\}$. С одной стороны, размерность образа A равна числу элементов набора, с другой стороны ранг A равен числу элементов набора.

Метод Гаусса

Задача 1

Примените метод Гаусса к матрице $\begin{pmatrix} 0 & 6 & 8 & 17 \\ 3 & 13 & 18 & 8 \\ 2 & 8 & 10 & 4 \\ 1 & 3 & 2 & 0 \end{pmatrix}$.

Если что-то в методе Гаусса осталось не ясным – обязательно обсудите это с преподавателем.

Комментарий. Для вычислений удобно переставлять столбцы так, чтобы коэффициенты элементарных преобразований получались симпатичными (целыми, а не дробными). Например, посмотрим на матрицу из задачи. На первом шаге нужно выбрать столбец, у которого на последнем месте стоит не ноль. Таких столбцов три – первый, второй и третий, мы можем выбрать любой. Чтобы получить "симпатичные"коэффициенты, в нашем случае лучше выбрать первый столбец (потому что 3 и 2 делятся на 1 нацело). Можете попробовать выбрать второй или третий столбец, и увидите, что возникают дроби. Конечно, во многих случаях дробей избежать не удастся, но в нашем случае их можно избежать (мы специально так подобрали коэффициенты матрицы).

Естественно, этот комментарий актуален, только если мы считаем руками на бумаге, а не на компьютере.

Подсказка. Перечитайте алгоритм метода Гаусса на Степике или в шпаргалке.

Решение. Пока что на последнем месте последней строки стоит ноль, так что исправим это, переставив местами первый и последний столбцы, получим матрицу $\begin{pmatrix} 17 & 6 & 8 & 0 \\ 8 & 13 & 18 & 3 \\ 4 & 8 & 10 & 2 \\ 0 & 3 & 2 & 1 \end{pmatrix}$.

Далее вычтем из второго столбца последний, домноженный на 3, а из третьего — последний, домноженный на 2. Получим $\begin{pmatrix} 17 & 6 & 8 & 0 \\ 8 & 4 & 12 & 3 \\ 4 & 2 & 6 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Для удобства поменяем местами второй и третий столбцы: $\begin{pmatrix} 17 & 8 & 6 & 0 \\ 8 & 12 & 4 & 3 \\ 4 & 6 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Теперь вычтем из первого столбца третий, умноженный на 2, а из второго столбца — третий, умноженный на 3. Получим $\begin{pmatrix} 5 & -10 & 6 & 0 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

И наконец вычтем из первого столбца второй, умноженный на (-0.5), или, другими словами, прибавим к первому столбцу второй, умноженный на 0.5: $\begin{pmatrix} 0 & -10 & 6 & 0 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Задача 2

Докажите, что ранг матрицы, полученной методом Гаусса, равен количеству ненулевых столбцов.

Другими словами, нужно доказать, что набор из всех ненулевых столбцов полученной матрицы линейно независим.

Подсказка. Вспомните, как описывались свойства столбцов у полученной матрицы.

Решение. В результате метода Гаусса получается матрица, в которой каждый следующий ненулевой столбец "длиннее" предыдущего. То есть номер последней ненулевой координаты у него больше, чем номер последней ненулевой координаты у всех предыдущих столбцов. Пусть мы нашли нетривиальную линейную комбинацию этих столбцов, которая равняется нулю. Посмотрим на самый длинный из столбцов, входящих в комбинацию с ненулевым коэффициентом. Обозначим его $\vec{a} = (a_1, \ldots, a_k, 0, \ldots, 0)^T$, при этом $a_k \neq 0$. Сложим вместе все оставшиеся столбцы с ненулевыми коэффициентами, взятыми из линейной комбинации. Мы получим $\vec{b} = (b_1, \ldots, b_m, 0, \ldots\ldots, 0)^T$, где $k > m$, так как все сложенные столбцы были "короче" $\vec{a}$.

Напоминаем, что здесь T — это значок транспонирования, то есть мы просто для экономии места написали столбцы как транспонированную строчку.

Нам осталось доказать, что $\vec{a} + \vec{b} \neq 0$. Это очевидно, так как k-ая координата вектора $\vec{a} + \vec{b}$ равна $a_k + 0 = a_k \neq 0$.

Задача 3

Дан набор векторов. Как проверить, является ли он линейно независимым?

Комментарий. В этой и всех последующих задачах можно пользоваться умением вычислять ранг любой матрицы. Для этого не нужно каждый раз описывать метод Гаусса.

Подсказка. Посмотрите на матрицу, чьи столбцы равны векторам этого набора.

Решение. Составим из данных векторов матрицу и применим к ней метод Гаусса. Как мы помним, ранг матрицы — это максимальное число линейно независимых столбцов этой матрицы. То есть, если в нашей матрице все столбцы линейно независимы, то её ранг будет равен количеству столбцов. Другими словами, после метода Гаусса у матрицы не должно быть ни одного нулевого столбца.

Задача 4

Подпространство порождено набором векторов $\vec{a}_1, \ldots, \vec{a}_k$. Дан вектор $\vec{x}$. Как понять, лежит $\vec{x}$ в этом подпространстве или нет?

Подсказка. Если вектор лежит в подпространстве, порождённом набором векторов, то его можно выразить через векторы этого набора.

Решение. Если вектор $\vec{x}$ лежит в $\mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\}$, то он представим в виде линейной комбинации: $\vec{x} = \lambda_1 \vec{a}_1 + \dots + \lambda_k \vec{a}_k$. А тогда из пункта 4 этой задачи следует, что ранги у

- матрицы со столбцами $\vec{a}_1, \dots, \vec{a}_k$ и
- матрицы со столбцами $\vec{a}_1, \dots, \vec{a}_k, \vec{x}$

должны совпадать.

Если же вектор не лежит в $\mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\}$, то его невозможно представить в виде линейной комбинации, то его добавление к $\vec{a}_1, \dots, \vec{a}_k$ увеличит ранг на 1.

Задача 5

Дан набор векторов $\vec{a}_1, \dots, \vec{a}_k$ в $\mathbb{R}^n$. Как понять, является этот набор базисом $\mathbb{R}^n$ или нет?

Подсказка. Задача — частный случай следующей.

Решение. Сначала проверим, является ли набор линейно независимым – это мы умеем делать при помощи метода Гаусса. Затем нужно понять, порождает ли этот набор все векторы $\mathbb{R}^n$. Для этого достаточно для каждого i проверить, лежит ли $\vec{e}_i$ в $\mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\}$. Это можно сделать по предыдущей задаче. Если какой-то $\vec{e}_i$ не лежит в $\mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\}$, то набор не являтеся базисом, так как не порождает один из векторов $\mathbb{R}^n$, а именно $\vec{e}_i$. Если все $\vec{e}_i$ порождаются набором, то и все вектора $\mathbb{R}^n$ порождаются набором.

Задача 6

Подпространство $V \subset \mathbb{R}^n$ порождено векторами $\vec{a}_1, \dots, \vec{a}_k$. Дан набор векторов $\vec{b}_1, \dots, \vec{b}_l$. Как понять, является этот набор базисом V или нет?

Подсказка. Нужно проверить линейную независимость и то, что первый набор векторов и второй набор векторов порождают одно и то же пространство.

Решение. Для начала нужно как в задаче 3 проверить, что $\vec{b}_1, \dots, \vec{b}_l$ линейно независимы. Далее необходимо проверить, что $\mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\} = \mathrm{Span}\{\vec{b}_1, \dots, \vec{b}_l\}$. Для этого нужно

- проверить, что $\vec{a}_i$ выражается через $\vec{b}_1, \dots, \vec{b}_l$ для любого i. Другими словами, $\vec{a}_i \in \mathrm{Span}\{\vec{b}_1, \dots, \vec{b}_l\}$. Ведь тогда и $\mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\} \subset \mathrm{Span}\{\vec{b}_1, \dots, \vec{b}_l\}$
- Аналогочино, нужно проверить, что $\vec{b}_j$ выражается через $\vec{a}_1, \dots, \vec{a}_k$ для любого j. То есть что $\vec{b}_j \in \mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\}$. А тогда $\mathrm{Span}\{\vec{b}_1, \dots, \vec{b}_l\} \subset \mathrm{Span}\{\vec{a}_1, \dots, \vec{a}_k\}$

А проверять, выражается ли вектор через набор векторов, мы научились в задаче 4.

Задача 7

Подпространство $V \subset \mathbb{R}^n$ порождено векторами $\vec{a}_1, \dots, \vec{a}_k$. Как найти какой-нибудь базис этого подпространства? Как найти размерность этого подпространства?

Подсказка. Элементарные преобразования не меняют линейную оболочку (Span).

Решение. Составим матрицу со столбцами $\vec{a}_1, \dots, \vec{a}_k$ и примениим метод Гаусса. Получившиеся в результате ненулевые столбцы можно взять в качестве базиса подпространства. Действительно, ведь элементарные преобразования не изменяли линейные оболочки столбцов. А размерностью подпространства будет количество полученных ненулевых столбцов, то есть как раз ранг (к слову, это согласуется с задачей из прошлого урока про ранг матрицы).