Математика для Data Science. Линейная алгебра.
Условия 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.

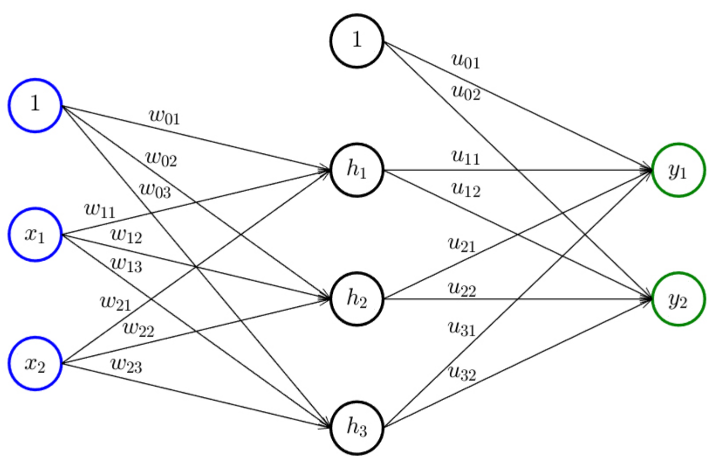
Вернемся к нейронной сети с предыдущего шага. Пусть:

- $m + 1$ — 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.

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?

Пример 1. Пусть у однослойной нейронной сети 2 нейрона в нулевом и 4 нейрона в первом слое (без учета нейронов сдвига). Тогда число параметров модели — $(2 + 1) \cdot 4 = 12$.

Пример 2. Пусть у двухслойной нейронной сети 3 нейрона в нулевом, 5 нейрона в первом и 7 нейронов во втором слое (без учета нейронов сдвига). Тогда число параметров модели равно $(3 + 1) \cdot 5 + (5 + 1) \cdot 7 = 20 + 42 = 62$.

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

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 с теорией.

Задача 3

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

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

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. $\mathrm{Span}\{\vec{a}_1\} = \mathrm{Span}\{\vec{a}_1, 100\vec{a}_1\}$
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)
3. $\mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} = \mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}4\\7\\0\end{pmatrix}\}$
4. $\mathrm{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\} = \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k, c_1\vec{a}_1 + c_2\vec{a}_2 + \cdots + c_k\vec{a}_k\}$, где $c_1, \ldots, c_k$ – любые числа.
5. $\mathrm{Span}\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\} = \mathrm{Span}\{\begin{pmatrix}1\\100\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\}$
6. $\mathrm{Span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\} = \mathrm{Span}\{\vec{a}_1 + c\vec{a}_2, \vec{a}_2, \ldots, \vec{a}_k\}$, где c – любое число

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 $\mathbb{R}^n$.

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 $\vec{a}_1, \dots, \vec{a}_k$ 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 $\vec{a}_1, \dots, \vec{a}_k$. Используя эти две комбинации, постройте нетривиальную линейную комбинацию, равную $\vec{0}$.

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.

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.

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.

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.

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.

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?

Задача 5

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

Задача 6

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

Задача 7

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