

第二章 向量和矩阵基础

第 3 讲 向量空间

黄定江

DaSE @ ECNU

djhuang@dase.ecnu.edu.cn

1 3.1 向量空间的基本概念

2 3.2 向量子空间

3 3.3 线性无关性

4 3.4 生成集、基底和坐标

5 3.5 秩

6 3.6 仿射子空间

1 3.1 向量空间的基本概念

2 3.2 向量子空间

3 3.3 线性无关性

4 3.4 生成集、基底和坐标

5 3.5 秩

6 3.6 仿射子空间

向量空间引例：鸢尾花数据集降维

sepal_length	sepal_width	petal_length	petal_width	species
5.1	3.5	1.4	0.2	Iris-setosa
4.9	3	1.4	0.2	Iris-setosa
4.7	3.2	1.3	0.2	Iris-setosa
4.6	3.1	1.5	0.2	Iris-setosa
5	3.6	1.4	0.2	Iris-setosa
5.4	3.9	1.7	0.4	Iris-setosa
7	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.9	3.1	4.9	1.5	Iris-versicolor
5.5	2.3	4	1.3	Iris-versicolor
6.5	2.8	4.6	1.5	Iris-versicolor
5.7	2.8	4.5	1.3	Iris-versicolor
6.3	3.3	6	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica
7.1	3	5.9	2.1	Iris-virginica
6.3	2.9	5.6	1.8	Iris-virginica
6.5	3	5.8	2.2	Iris-virginica
7.6	3	6.6	2.1	Iris-virginica
4.9	2.5	4.5	1.7	Iris-virginica
7.3	2.9	6.3	1.8	Iris-virginica

PCA

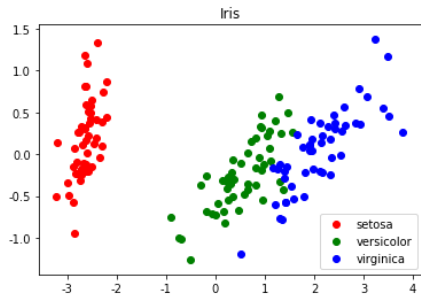


图 1: Iris 数据集 (左), PCA 降维可视化 (右)

向量空间：数据处理空间的出发点

定义 1

设 \mathbb{V} 是由 n 维向量组成的非空集合， \mathbb{K} 是一个数域。在 \mathbb{V} 上定义了加法，在 \mathbb{K} 与集合 \mathbb{V} 上定义了数乘，并且 $\forall \mathbf{a}, \mathbf{b} \in \mathbb{V}$ 及任意数 $k \in \mathbb{K}$ ，有 $\mathbf{a} + \mathbf{b}, k\mathbf{a} \in \mathbb{V}$ ，则称 \mathbb{V} 对于向量的加法和数乘两种运算封闭， \mathbb{V} 为数域 \mathbb{K} 上的 n 维向量空间或者线性空间。

例 1

数域 \mathbb{K} 上的 n 维向量, 按照如下定义的加法和数乘运算, 构成数域 \mathbb{K} 上的向量空间。

考虑向量空间 $\mathbb{V} = \mathbb{K}^n$, 任意两个向量 $\mathbf{a}, \mathbf{b} \in \mathbb{V}$, $\lambda \in \mathbb{K}$ 满足:

1. 加法

$$\mathbf{a} + \mathbf{b} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix} \in \mathbb{V}$$

2. 数乘

$$\lambda \mathbf{a} = \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix} \in \mathbb{V}$$

例 2

数域 \mathbb{K} 上的 $m \times n$ 矩阵, 按照如下定义的加法和数乘运算, 构成数域 \mathbb{K} 上的向量空间。考虑矩阵空间 $\mathbb{V} = \mathbb{K}^{m \times n}$, 任意两个矩阵 $A, B \in \mathbb{V}$, $\lambda \in \mathbb{K}$ 满足:

1. 加法

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix} \in \mathbb{V}$$

2. 数乘

$$\lambda A = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix} \in \mathbb{V}$$

例 3

假设元素属于数域 \mathbb{C} :

- 令 λ 所在的数域 $\mathbb{K} = \mathbb{C}$, 定义加法为复数加法、数乘为复数乘法, 根据复数的加法和乘法, 我们可以知道复数域 \mathbb{C} 是自身上的向量空间。
- 令 λ 所在的数域 $\mathbb{K} = \mathbb{R}$, 定义加法为实部与实部相加, 虚部与虚部相加, 而数乘则是将实数分别乘至实部和虚部。容易知道, 复数域 \mathbb{C} 是实数域 \mathbb{R} 上的向量空间。

例 4

数域 \mathbb{R} 上的次数小于 n 的一元多项式, 即

$$\mathbb{P}_n = \{p: p(x) = a_{n-1}x^{n-1} + \cdots + a_1x + a_0, \text{ 其中 } a_0, a_1, \cdots, a_{n-1} \in \mathbb{R}\}$$

构成 \mathbb{R} 上的向量空间。这是因为对于 $\forall p_1, p_2 \in \mathbb{P}_n$ 及任意数 $k \in \mathbb{K}$, 有 $p_1 + p_2, kp_1 \in \mathbb{P}_n$ 。

设有

$$p_1 = a_{1,n-1}x^{n-1} + \cdots + a_{1,1}x + a_{1,0}, \quad p_2 = a_{2,n-1}x^{n-1} + \cdots + a_{2,1}x + a_{2,0}$$

$\lambda \in \mathbb{R}$, 则有

$$p_1 + p_2 = (a_{1,n-1} + a_{2,n-1})x^{n-1} + \cdots + (a_{1,1} + a_{2,1})x + (a_{1,0} + a_{2,0})$$

和

$$\lambda p_1 = \lambda a_{1,n-1}x^{n-1} + \cdots + \lambda a_{1,1}x + \lambda a_{1,0}$$

- 1 3.1 向量空间的基本概念
- 2 3.2 向量子空间
- 3 3.3 线性无关性
- 4 3.4 生成集、基底和坐标
- 5 3.5 秩
- 6 3.6 仿射子空间

3.2.1 子空间

定义 2

设 \mathbb{X} 是 \mathbb{K} 上的 n 维线性空间, \mathbb{Y} 是 \mathbb{X} 的子集且满足: 若 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Y}$, 则 $\boldsymbol{x} + \boldsymbol{y} \in \mathbb{Y}$; 若 $a \in \mathbb{K}, \boldsymbol{x} \in \mathbb{Y}$, 则 $a\boldsymbol{x} \in \mathbb{Y}$, 则称 \mathbb{Y} 是 \mathbb{X} 的线性子空间, 简称子空间。

例 5

非空的线性空间一定会有的子空间: 自身和 $\{\mathbf{0}\}$ 。我们把只含零向量的子集称为零子空间。零子空间和线性空间本身统称为平凡子空间, 其它子空间叫做非平凡子空间。

例 6

图2中只有 D 是 \mathbb{R}^2 的子空间。在 A 和 C 中封闭性被违反。 B 则不包括 0 。

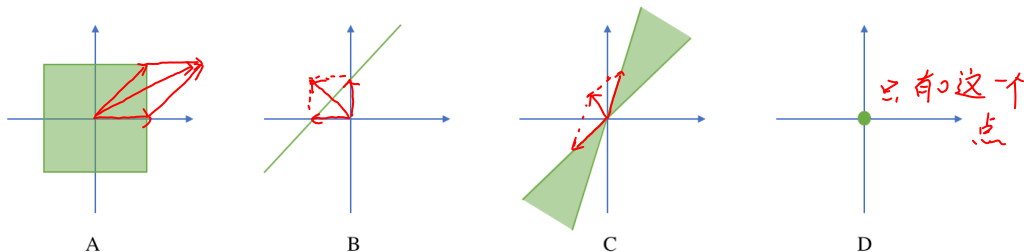


图 2: \mathbb{R}^2 中的一些子集

例 7

1. 线性方程组 $Ax = 0$ 的解空间是 \mathbb{R}^n 中常见的子空间
2. 线性方程组 $Ax = b$ 的解空间, 当 $b \neq 0$ 时, 不是子空间
3. 线性方程组 $Ax = 0$ 的解空间和 $Bx = 0$ 的解空间的交集也是 \mathbb{R}^n 中的子空间

并集不是

3.2 子空间的交、和、直和

定理 1

设 \mathbb{Y}_1 与 \mathbb{Y}_2 都是数域 \mathbb{K} 上的线性空间 \mathbb{X} 的子空间。若用 $\mathbb{Y}_1 \cap \mathbb{Y}_2$ 表示 \mathbb{Y}_1 与 \mathbb{Y}_2 中的公共元素集合, 则 $\mathbb{Y}_1 \cap \mathbb{Y}_2$, 也是 \mathbb{X} 的子空间, 且称 $\mathbb{Y}_1 \cap \mathbb{Y}_2$ 为 \mathbb{Y}_1 与 \mathbb{Y}_2 的交。

定理 2

给定 \mathbb{X} 的两个子空间 $\mathbb{Y}_1, \mathbb{Y}_2$, 若用 $\mathbb{Y}_1 + \mathbb{Y}_2$ 表示全体形如 $\mathbf{y}_1 + \mathbf{y}_2 (\mathbf{y}_1 \in \mathbb{Y}_1, \mathbf{y}_2 \in \mathbb{Y}_2)$ 的向量组成的集合, 则 $\mathbb{Y}_1 + \mathbb{Y}_2$ 也是 \mathbb{X} 的子空间, 且称 $\mathbb{Y}_1 + \mathbb{Y}_2$ 为 \mathbb{Y}_1 与 \mathbb{Y}_2 的和。

定义 3

如果 \mathbb{Y} 中的每个向量 \boldsymbol{x} 可唯一地表成 $\boldsymbol{x} = \boldsymbol{y}_1 + \boldsymbol{y}_2$ ($\boldsymbol{y}_1 \in \mathbb{Y}_1, \boldsymbol{y}_2 \in \mathbb{Y}_2$) 的形式, 则称 \mathbb{Y} 为 \mathbb{Y}_1 与 \mathbb{Y}_2 的直和。记作 $\mathbb{Y} = \mathbb{Y}_1 + \mathbb{Y}_2$ 或 $\mathbb{Y}_1 \oplus \mathbb{Y}_2$

定理 3

$\mathbb{Y}_1 + \mathbb{Y}_2$ 为直和的充分必要条件是: 由 $\boldsymbol{y}_1 + \boldsymbol{y}_2 = \mathbf{0}$ ($\boldsymbol{y}_1 \in \mathbb{Y}_1, \boldsymbol{y}_2 \in \mathbb{Y}_2$) 可推出 $\boldsymbol{y}_1 = \boldsymbol{y}_2 = \mathbf{0}$ 。

推论 1

$\mathbb{Y}_1 + \mathbb{Y}_2$ 为直和的充分必要条件是: 由 $\mathbb{Y}_1 \cap \mathbb{Y}_2 = \{\mathbf{0}\}$ 。

- 1 3.1 向量空间的基本概念
- 2 3.2 向量子空间
- 3 3.3 线性无关性**
- 4 3.4 生成集、基底和坐标
- 5 3.5 秩
- 6 3.6 仿射子空间

3.3.1 线性表出定义

- 一个向量能否用其它向量表示？
- 一个线性空间里的所有向量最少可以用几个向量表示出来？
- 这些向量之间又是什么关系？
- 空间里的坐标是如何确定的？
- 为什么向量空间要用数乘和向量加法定义？
- 为什么封闭性是向量空间非常重要的性质？

这些问题就是我们接下来要讨论的问题。

3.3.1 线性表出定义

定义 4

设向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_s$ 是数域 \mathbb{K} 上的 n 维向量组, k_1, k_2, \dots, k_s 是数域 \mathbb{K} 上的一组数, 那么表达式

$$k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 + \dots + k_s \mathbf{a}_s.$$

称为向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_s$ 的一个线性组合, 而 k_1, k_2, \dots, k_s 称为组合系数。

定义 5

若向量 \mathbf{b} 是向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_s$ 的一个线性组合, 即

$$\mathbf{b} = k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 + \dots + k_s \mathbf{a}_s,$$

则称 \mathbf{b} 可以由向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_s$ 线性表出。

例 8

零向量 $\mathbf{0}$ 总可以写成其它一些向量的线性组合。

例 9

例如, 设向量组

$$\mathbf{a}_1 = (2, -1, 3, 1)$$

$$\mathbf{a}_2 = (4, -2, 5, 4)$$

$$\mathbf{a}_3 = (2, -1, 4, -1)$$

则有 $\mathbf{a}_3 = 3\mathbf{a}_1 - \mathbf{a}_2$, 这表示 \mathbf{a}_3 可以由 $\mathbf{a}_1, \mathbf{a}_2$ 线性表出。

3.3.2 线性相关/无关

定义 6

设 $\mathbf{a}_i \in \mathbb{K}^n (i = 1, 2, \dots, r)$. 若在 \mathbb{K} 中存在 r 个不全为零的数 $\lambda_i (i = 1, 2, \dots, r)$, 使 $\sum_{i=1}^r \lambda_i \mathbf{a}_i = 0$, 则称向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ 线性相关. 反之, 如果向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ 不线性相关, 即只有 $\lambda_1, \lambda_2, \dots, \lambda_r$ 全为零时, 才能使得 $\sum_{i=1}^r \lambda_i \mathbf{a}_i = 0$, 则称向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ 线性无关.

定义 7

向量组的一部分组称为一个极大线性无关组, 如果这个部分组本身线性无关, 但从原向量组的其余向量中任取一个添加进去后, 所得的部分组都线性相关.

例 10

- 一个地理例子可能有助于阐明线性独立性的概念。
- 在上海的一个人在描述宣城的位置时可能会说：“您可以先向西北行驶 180 公里到常州，再向西南行驶 244.8 公里，才能到宣城。”
- 这是描述宣城位置的充分信息，因为地理坐标系可能被视为二维矢量空间（忽略高度和地球表面）。
- 这个人可能会加上“它在这里以西约 282.9 公里处。”
- 尽管这个说法是正确的，但鉴于先前的信息就可以找到宣城。



图 3: 一个线性相关和线性无关的例子

例 11

在向量组

$$\mathbf{a}_1 = (2, -1, 3, 1) \quad \mathbf{a}_2 = (4, -2, 5, 4) \quad \mathbf{a}_3 = (2, -1, 2, 3)$$

中, 由 $\mathbf{a}_1, \mathbf{a}_2$ 组成的部分组就是一个极大线性无关组。首先, $\mathbf{a}_1, \mathbf{a}_2$ 线性无关, 因为由

$$\begin{aligned} k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 &= k_1(2, -1, 3, 1) + k_2(4, -2, 5, 4) \\ &= (2k_1 + 4k_2, -k_1 - 2k_2, 3k_1 + 5k_2, k_1 + 4k_2) = (0, 0, 0, 0) \end{aligned}$$

就有 $k_1 = k_2 = 0$, 同时, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 线性相关 ($\mathbf{a}_2 = \mathbf{a}_1 + \mathbf{a}_3$)。不难看出, $\mathbf{a}_2, \mathbf{a}_3$ 也是一个极大线性无关组。

3.3.3 等价

定义 8

设 a_1, a_2, \dots, a_s 和 b_1, b_2, \dots, b_t 是数域 \mathbb{K} 上的两个向量组, 如果向量组 a_1, a_2, \dots, a_s 中每一个向量 $a_i (i = 1, 2, \dots, s)$ 都可以用向量组 b_1, b_2, \dots, b_t 线性表出, 那么称向量组 a_1, a_2, \dots, a_s 可以用向量组 b_1, b_2, \dots, b_t 线性表出。如果两个向量组互相可以线性表出, 则称为它们等价。

例 12

例如, 设

$$\mathbf{a}_1 = (1, 0), \mathbf{a}_2 = (0, 1);$$

$$\mathbf{b}_1 = (1, 1), \mathbf{b}_2 = (-1, 1),$$

则向量组 $\mathbf{a}_1, \mathbf{a}_2$ 与向量组 $\mathbf{b}_1, \mathbf{b}_2$ 是等价的。

$$\mathbf{a}_1 = \frac{1}{2}\mathbf{b}_1 - \frac{1}{2}\mathbf{b}_2, \quad \mathbf{a}_2 = \frac{1}{2}\mathbf{b}_1 + \frac{1}{2}\mathbf{b}_2$$

$$\mathbf{b}_1 = \mathbf{a}_1 + \mathbf{a}_2, \quad \mathbf{b}_2 = -\mathbf{a}_1 + \mathbf{a}_2$$

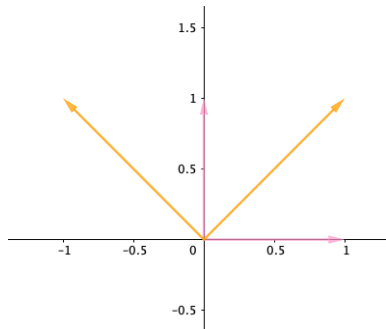


图 4: 向量组等价

- 1 3.1 向量空间的基本概念
- 2 3.2 向量子空间
- 3 3.3 线性无关性
- 4 3.4 生成集、基底和坐标**
- 5 3.5 秩
- 6 3.6 仿射子空间

3.4.1 生成集

定义 9

设 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ 是 \mathbb{V} 的一组向量, 则这组向量所有可能的线性组合 $\sum_{k=1}^r \lambda_k \mathbf{a}_k$ 所成的集合是 \mathbb{V} 的一个子空间, 称为由 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ 张成的子空间, 记作 $L(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r)$ 或 $\text{span}(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r)$ 。
 $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$ 叫做 \mathbb{V} 的一个生成集。

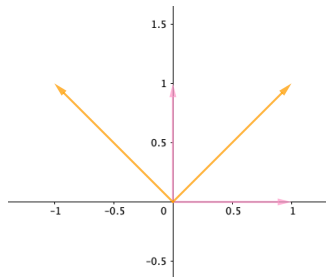


图 5: 张成子空间

定理 4

两个向量组张成相同的子空间的充分必要条件是: 这两个向量组等价。

3.4.2 基底与维数

定义 10

如果在向量空间 \mathbb{V} 中有 n 个线性无关的向量 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$, 且 \mathbb{V} 中任一向量都可以用它们线性表出, 则称 \mathbb{V} 为 \mathbb{K} 上的 n 维线性空间, n 称为 \mathbb{V} 的维数, 记作 $\dim(\mathbb{V}) = n$ 。而 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ 就是 \mathbb{V} 的一组基。

例 13

复数域 \mathbb{C} 在 \mathbb{C} 上和 \mathbb{R} 上是两个不同的向量空间。

- 因为在 \mathbb{C} 上它是一维的，数 1 就是一组基；
- 而在 \mathbb{R} 上它是二维的，数 1 与 i 就是一组基。

这个例子告诉我们，维数是和所考虑的数域有关的。

定理 5

令 V 是一向量空间, $B \subseteq V, B \neq \emptyset$ 下列命题等价:

- B 是 V 的一个基
- B 是最小生成集
- B 是 V 中的极大线性无关组
- V 中每一个向量能被 B 线性表出

标准基

定义 11

如果一组基中的每一个向量长度均为 1，我们称其为**标准基**。

在后面的课程中，我们将会严格说明**向量的长度**。

例 14

- 在右图中，粉色的两个向量组成的基是一组标准基。

- 在 \mathbb{R}^3 中，常用基 $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$ 就是一组标准基。

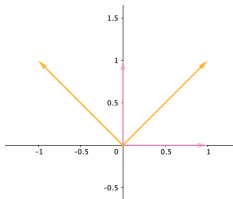


图 6: 标准基和基

确定一组基

例 15

对于一个由向量 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 张成的向量空间 $\mathbb{U} \subseteq \mathbb{R}^4$

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 1 \end{pmatrix},$$

我们关心 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 是否是 \mathbb{U} 的一组基。为此，我们需要确认 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ 是否线性无关。因

此，我们需要解 $\sum_{i=1}^3 \lambda_i \mathbf{x}_i = \mathbf{0}$

这是一个关于下面这个矩阵的一个线性方程组, 并且我们对这个矩阵作行初等变换可将其化成阶梯型

$$(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

从而我们可以发现 $\mathbf{x}_1, \mathbf{x}_2$ 是线性无关的, $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 = \mathbf{0}$ 只有 $\lambda_1 = \lambda_2 = 0$ 时成立。因此 $\{\mathbf{x}_1, \mathbf{x}_2\}$ 是 \mathbb{U} 的一组基。

这个例子说明, \mathbb{U} 是 \mathbb{R}^4 中的 2 维向量空间。如果我们添加向量

$\mathbf{e}_3 = (0, 0, 1, 0)^T, \mathbf{e}_4 = (0, 0, 0, 1)^T$, 那么因为 $\mathbf{e}_1 = \mathbf{x}_1 - \mathbf{e}_3, \mathbf{e}_2 = \mathbf{x}_2 - \mathbf{e}_3 - \mathbf{e}_4$, 则

$\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4$ 可以由 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_3, \mathbf{e}_4$ 线性表出, 也就是 $\mathbf{x}_1, \mathbf{x}_2, \mathbf{e}_3, \mathbf{e}_4$ 生成的子空间为 \mathbb{R}^4 。

3.4.2 子空间的扩张

定理 6

设 $\mathbb{Y} = L(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m)$ 是 n 维空间 \mathbb{X} 的一个 m 维子空间, 则向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ 可扩张为 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_n$ 使得 $\mathbb{X} = L(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_n)$ 。

注意: 其中 $L(\mathbf{a}_{m+1}, \dots, \mathbf{a}_n)$ 也是 \mathbb{X} 的一个子空间。

$$L(\mathbf{a}_{m+1}, \dots, \mathbf{a}_n) \oplus L(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m) = L(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m, \mathbf{a}_{m+1}, \dots, \mathbf{a}_n).$$

定理 7

维数公式: $\dim(\mathbb{Y}_1 + \mathbb{Y}_2) = \dim \mathbb{Y}_1 + \dim \mathbb{Y}_2 - \dim(\mathbb{Y}_1 \cap \mathbb{Y}_2)$

- 对于直和: $\dim(\mathbb{Y}_1 \oplus \mathbb{Y}_2) = \dim \mathbb{Y}_1 + \dim \mathbb{Y}_2$

定义 12

如果一个向量空间 V 中任一向量都能被 n 个线性无关的向量线性表出时, V 称为有限维线性空间, 否则, 称为无限维线性空间。

有限维线性空间

1. n 维向量空间
2. $n \times m$ 维矩阵空间
3. 最高次为 n 次的多项式空间
4. 复数域

无限维线性空间

1. 所有的多项式构成的空间
2. 一阶可导函数空间
3. 傅里叶变换后的频域空间

3.4.3 坐标

定义 13

在 n 维向量空间 \mathbb{V} 中, n 个线性无关的向量 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 称为 \mathbb{V} 的一组基。设 \mathbf{a} 是 \mathbb{V} 中任一向量, 于是 $\epsilon_1, \epsilon_2, \dots, \epsilon_n, \mathbf{a}$ 线性相关, 因此 \mathbf{a} 可以被基 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 线性表出:

$$\mathbf{a} = a_1\epsilon_1 + a_2\epsilon_2 + \dots + a_n\epsilon_n,$$

其中系数 a_1, a_2, \dots, a_n 是被向量 \mathbf{a} 和基 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 唯一确定的, 这组数就称为 \mathbf{a} 在基 $\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 下的坐标, 记为 (a_1, a_2, \dots, a_n) 。

例 16

在向量空间 P_n 中,

$$1, x, x^2, \dots, x^{n-1}$$

是 n 个线性无关的向量, 而且每一个次数小于 n 的数域 \mathbb{K} 上的多项式都可以被它们线性表出, 所以 P_n 是 n 维的, 而 $1, x, x^2, \dots, x^{n-1}$ 就是它的一组基。

在这组基下, 多项式 $f(x) = a_0 + a_1x + \dots + a_{n-1}x^{n-1}$ 的坐标就是它的系数 $(a_0, a_1, \dots, a_{n-1})$ 。

如果在 \mathbb{V} 中取另外一组基

$$\epsilon_1' = 1, \epsilon_2' = (x - a), \dots, \epsilon_n' = (x - a)^{n-1}.$$

那么按泰勒展开公式

$$f(x) = f(a) + f'(a)(x - a) + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x - a)^{n-1}.$$

因此, $f(x)$ 在基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 下的坐标是

$$\left(f(a), f'(a), \cdots, \frac{f^{(n-1)}(a)}{(n-1)!} \right).$$

例 17

在 n 维向量空间 V 中, 显然

$$\begin{cases} \epsilon_1 = (1, 0, \cdots, 0), \\ \epsilon_2 = (0, 1, \cdots, 0), \\ \cdots \cdots \cdots \\ \epsilon_n = (0, 0, \cdots, 1) \end{cases}$$

是一组基。任意向量 $a = (a_1, a_2, \cdots, a_n) = a_1\epsilon_1 + a_2\epsilon_2 + \cdots + a_n\epsilon_n$ 。

所以 (a_1, a_2, \cdots, a_n) 就是向量 a 在这组基下的坐标。

不难证明,

$$\begin{cases} \varepsilon_1' = (1, 1, \cdots, 1), \\ \varepsilon_2' = (0, 1, \cdots, 1), \\ \cdots \cdots \cdots \\ \varepsilon_n' = (0, 0, \cdots, 1) \end{cases}$$

是 \mathbb{V} 中 n 个线性无关的向量。

在基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 下, 对向量 $\mathbf{a} = (a_1, a_2, \cdots, a_n)$, 有

$$\mathbf{a} = a_1 \varepsilon_1' + (a_2 - a_1) \varepsilon_2' + \cdots + (a_n - a_{n-1}) \varepsilon_n'.$$

因此, \mathbf{a} 在基 $\varepsilon_1', \varepsilon_2', \cdots, \varepsilon_n'$ 下的坐标为

$$(a_1, a_2 - a_1, \cdots, a_n - a_{n-1}).$$

- 1 3.1 向量空间的基本概念
- 2 3.2 向量子空间
- 3 3.3 线性无关性
- 4 3.4 生成集、基底和坐标
- 5 3.5 秩**
- 6 3.6 仿射子空间

3.5.1 秩、矩阵的秩

定义 14

向量组 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r$ 的极大线性无关组中所含向量的个数称为这个向量组的秩, 记作 $\text{rank}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$.

定义 15

矩阵 \mathbf{A} 的行 (列) 向量组的秩称为 \mathbf{A} 的行秩 (列秩), 其中矩阵的行秩和列秩相等, 它们都称为矩阵 \mathbf{A} 的秩, 记作 $\text{rank}(\mathbf{A})$.

定理 8

$$\dim L(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r) = \text{rank}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_r\}$$

例 18

设矩阵 A

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

矩阵 A 的行向量组为

$$\mathbf{a}_1 = (1, 1, 3, 1) \quad \mathbf{a}_2 = (0, 2, -1, 4)$$

$$\mathbf{a}_3 = (0, 0, 0, 5) \quad \mathbf{a}_4 = (0, 0, 0, 0).$$

可以证明, $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ 是向量组 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ 的一个极大线性无关组。因此, 向量组 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ 的秩为 3, 换句话说, 矩阵 A 的行秩为 3。

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

A 的列向量组为

$$\mathbf{b}_1 = (1, 0, 0, 0)^T$$

$$\mathbf{b}_2 = (1, 2, 0, 0)^T$$

$$\mathbf{b}_3 = (3, -1, 0, 0)^T$$

$$\mathbf{b}_4 = (1, 4, 5, 0)^T.$$

用同样的方法可以证明, $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ 线性无关, 且 $\mathbf{b}_3 = \frac{7}{2}\mathbf{b}_1 - \frac{1}{2}\mathbf{b}_2$, 所以 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ 是向量组 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ 的一个极大线性无关组, 于是向量组 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ 的秩为 3, 换句话说, 矩阵 A 的列秩为 3。

例 19

一般在推荐系统中，数据往往使用“用户——物品”矩阵来表示的。用户对其接触过的物品进行评分，评分表示了用户对于物品的喜爱程度，分数越高，表示用户越喜欢这个物品。而这个矩阵往往是稀疏的，空白项是用户还未接触到的物品，推荐系统的任务则是选择其中的部分物品推荐给用户。这就需要对矩阵中的空白项进行补全。

	物品1	物品2	物品3	物品4	物品5	物品6	物品7	物品8	物品9	物品10
用户1	3					5			2	
用户2			3		5			2		
用户3		1		2			5			
用户4			3					3		5
用户5	5				2					

	物品1	物品2	物品3	物品4	物品5	物品6	物品7	物品8	物品9	物品10
用户1	3					5			2	
用户2			3		5			2		
用户3		1		2			5			
用户4			3					3		5
用户5	5				2					

设 \mathbb{E} 为可以被观察到评分的 (用户, 物品) 指标集, \mathbf{M} 为观察评分矩阵, \mathbf{M}_{ij} 为观测到的用户 i 对物品 j 的评分, \mathbf{X} 为预测评分矩阵, \mathbf{X}_{ij} 为预测的用户 i 对物品 j 的评分。矩阵补全问题可以转化为寻找与观测到数据集 \mathbb{E} 中所有项匹配的低秩矩阵 \mathbf{X} 。形式化如下

$$\min_{\mathbf{X}} \quad \text{rank}(\mathbf{X})$$

$$s.t. \quad \mathbf{X}_{ij} = \mathbf{M}_{ij} \quad \forall i, j \in \mathbb{E}$$

或者转化为限定在秩为 r 的条件下, 求矩阵使得观测到的评分与预测的评分最接近:

$$\min_{\mathbf{X}} \quad \sum_{ij} (\mathbf{X}_{ij} - \mathbf{M}_{ij})^2 \quad \forall i, j \in \mathbb{E}$$

$$s.t. \quad \text{rank}(\mathbf{X}) = r$$

- 1 3.1 向量空间的基本概念
- 2 3.2 向量子空间
- 3 3.3 线性无关性
- 4 3.4 生成集、基底和坐标
- 5 3.5 秩
- 6 3.6 仿射子空间**

3.6.1 仿射子空间

定义 16

令 \mathbb{V} 是一线性空间, $\mathbf{x}_0 \in \mathbb{V}$ 且 $\mathbb{U} \subseteq \mathbb{V}$ 是一线性子空间, 则子集

$$\mathbb{L} = \mathbf{x}_0 + \mathbb{U} := \{\mathbf{x}_0 + \mathbf{u} | \mathbf{u} \in \mathbb{U}\} \subseteq \mathbb{V}$$

是一仿射子空间。我们定义线性子空间的维数为仿射子空间的维数。

- 注意, 如果 $\mathbf{x}_0 \notin \mathbb{U}$, 则仿射子空间 \mathbb{L} 不是一个线性子空间。
- 若 \mathbb{U} 有一基底 $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$, 则 \mathbb{L} 中的每一个元素 \mathbf{x} 均可写成 $\mathbf{x}_0 + k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 + \dots + k_m \mathbf{a}_m$ 。
这一结论通过定义是容易知道的。

超平面

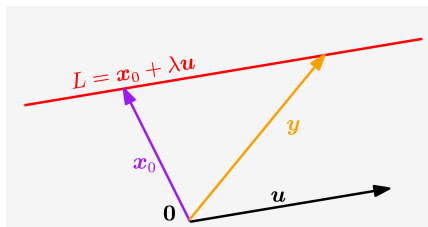


图 7: 仿射子空间

例 20

 \mathbb{R}^3 中常见的仿射子空间

1. 零维仿射子空间：单点集 $\{x_0\}$
2. 一维仿射子空间：直线 $\{x_0 + ku\}$
3. 二维仿射子空间：平面 $\{x_0 + k_1u_1 + k_2u_2\}$
4. \mathbb{R}^3 本身
5. \mathbb{R}^n 中的 $n-1$ 维仿射子空间称为超平面。在二维空间中一条直线是一个超平面；在三维空间中一个平面就是它的超平面；在四维空间中，一个超平面是三维空间。

例 21

我们已经知道线性方程组 $Ax = b, b \neq 0$ 的解空间不是一个线性空间，但是它的解空间是一个仿射空间。

设 $Ax = 0$ 的解空间为 V ，它是一个子空间，且 x_0 是 $Ax = b$ 的一个特解，则可以证明仿射空间 $x_0 + V$ 是方程组 $Ax = b, b \neq 0$ 的解空间。

事实上， $\forall x \in x_0 + V$ ， x 必可以写成 $x = x_0 + x_1$ ，其中 $x_1 \in V$ 。显然：

$$Ax = A(x_0 + x_1) = Ax_0 + Ax_1 = 0 + b = b.$$

说明 $x_0 + V \subseteq \{x | Ax = b\}$.

反之， $\forall x$ 满足 $Ax = b$ ，则 $Ax - Ax_0 = A(x - x_0) = 0$ ，则 $x - x_0 \in V$ ， $x \in x_0 + V$ 。

说明 $\{x | Ax = b\} \subseteq x_0 + V$.

综上，线性方程组 $Ax = b, b \neq 0$ 的解空间为 $x_0 + V$ ，这是一个仿射空间。

向量空间小结

向量空间上的运算封闭性

- 加法
- 数乘
- ...

“非严格”向量空间

- 输入空间
- 输出空间
- 特征空间

如何把数据向量从高维向量空间“拉回”到低维子空间？机器学习中输入和输出空间以及特征空间可以考虑为向量空间，要实现运算任务，还需要附加额外的数学结构，那可以附加哪些数学结构呢？