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

3.1.1 向量空间

先前已经介绍了什么是向量，什么是矩阵。接下来介绍线性空间（向量空间）的概念，向量空间是向量所在的空间，在这个空间上，向量以及向量的运算共同构造出了这样一个空间。尽管名称是向量空间，我们可以把数或者矩阵的空间也叫做向量空间，因为它也满足向量空间的定义。

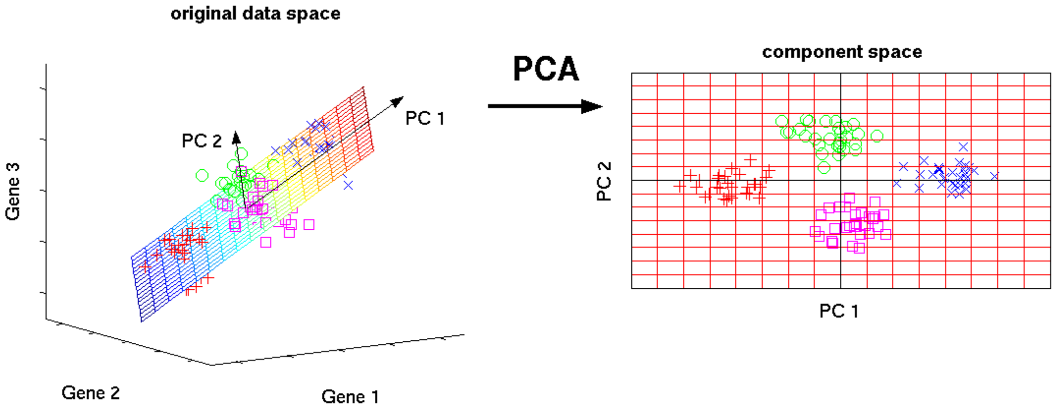
original data space
PCA
component space
PC 1
PC 2
Gene 3
Gene 2
Gene 1
PC 2
PC 1

图 1: 降维

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 $\mathbb{V}$ 上定义了的加法，在 $\mathbb{K}$ 与集合 $\mathbb{V}$ 上定义了数乘，并且 $\forall \boldsymbol{a}, \boldsymbol{b} \in \mathbb{V}$ 及任意数 $k \in \mathbb{K}$, 有 $\boldsymbol{a}+\boldsymbol{b}, k\boldsymbol{a} \in \mathbb{V}$，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 $\boldsymbol{a},\boldsymbol{b}\in\mathbb{V}$，$\lambda\in\mathbb{R}$ 满足:

1. 加法

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

2. 数乘

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

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 $\mathbb{V}=\mathbb{K}^{m\times n}$，任意的两个矩阵 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{V}$，$\lambda\in\mathbb{R}$ 满足:

1. 加法

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

2. 数乘

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

例 3

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

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

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

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

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 $\{0\}$。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 $\mathbf{0}$。

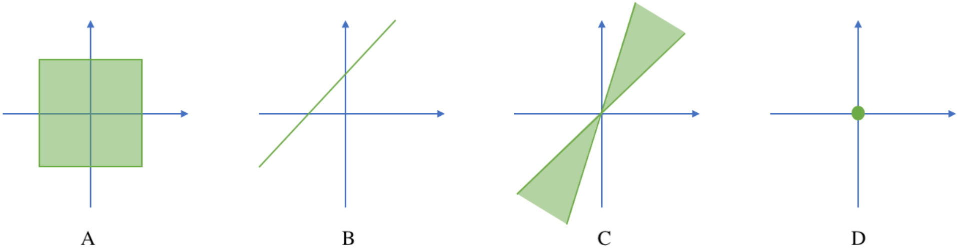

图 2: $\mathbb{R}^2$ 中的一些子集

例 7

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

3.2 子空间的交、和、直和

定理 1

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

若用 $\mathbb{Y}_1 + \mathbb{Y}_2$ 表示全体形如 $\boldsymbol{y}_1 + \boldsymbol{y}_2(\boldsymbol{y}_1 \in \mathbb{Y}_1, \boldsymbol{y}_2 \in \mathbb{Y}_2)$ 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 $\boldsymbol{x}=\boldsymbol{y}_1+\boldsymbol{y}_2(y_1\in\mathbb{Y}_1,\boldsymbol{y}_2\in\mathbb{Y}_2)$ 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

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

定理 4

$\mathbb{Y}_1+\mathbb{Y}_2$ 为直和的必要充分条件是：由 $\mathbb{Y}_1\cap\mathbb{Y}_2=\{0\}$。

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。

定义 4

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

$$k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \cdots + k_s\boldsymbol{a}_s.$$

称为向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_s$ 的**一个线性组合**，而 $k_1, k_2, \cdots, k_s$ 称为**组合系数**。

定义 5

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

$$\boldsymbol{b} = k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \cdots + k_s\boldsymbol{a}_s,$$

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

例 8

例如，设向量组

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

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

- 零向量 $\mathbf{0}$ 总可以写成其它一些向量的线性组合。

3.3.2 线性相关/无关

定义 6

设 $\boldsymbol{a}_i \in \mathbb{K}^n(i=1,2,\cdots,r)$. 若在 $\mathbb{K}$ 中存在 r 个不全为零的数 $\lambda_i(i=1,2,\cdots,r)$, 使 $\sum_{i=1}^{r}\lambda_i\boldsymbol{a}_i=0$, 则称向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ **线性相关**. 反之，如果向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ 不线性相关, 即只有 $\lambda_1,\lambda_2,\cdots,\lambda_r$ 全为零时，才能使得 $\sum_{i=1}^{r}\lambda_i\boldsymbol{a}_i=0$, 则称向量组 $\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_r$ **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 9

- 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 10

在向量组

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

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

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

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

3.3.3 等价

定义 8

设 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_s$ 和 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_t$ 是数域 $\mathbb{K}$ 上的两个向量组，如果向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_s$ 中每一个向量 $\boldsymbol{a}_i (i = 1, 2, \cdots, s)$ 都可以用向量组 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_t$ 线性表出，那么称向量组 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_s$ 可以用向量组 $\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_t$ 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 11

例如，设

$$\boldsymbol{a}_1 = (1,0), \boldsymbol{a}_2 = (0,1);$$
$$\boldsymbol{b}_1 = (1,1), \boldsymbol{b}_2 = (-1,1),$$

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

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

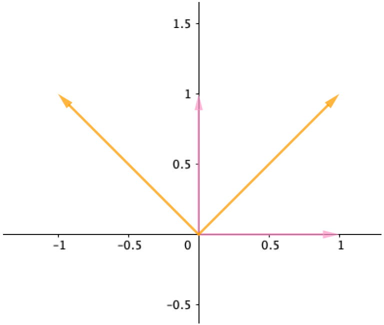

图 4: 向量组等价

3.4.1 生成集

定义 9

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

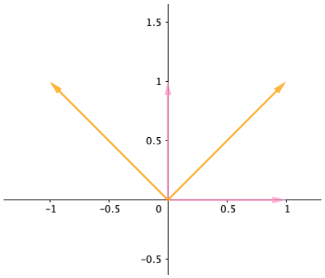

图 5: 张成子空间

定理 5

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 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$，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 $\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n$ 就是 $\mathbb{V}$ 的一组**基**。

例 12

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 6

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

- $\mathbb{B}$ 是 $\mathbb{V}$ 的一个基
- $\mathbb{B}$ 是最小生成集
- $\mathbb{B}$ 是 $\mathbb{V}$ 中的极大线性无关组
- $\mathbb{V}$ 中每一个向量能被 $\mathbb{B}$ 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 13

- 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。

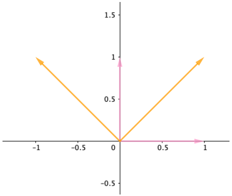

图 6: 标准基和基

确定一组基

例 14

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

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

我们关心 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ 是否是 $\mathbb{U}$ 的一组基。为此，我们需要确认 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ 是否线性无关。因此，我们需要解 $\sum_{i=1}^{3} \lambda_i \boldsymbol{x}_i = \mathbf{0}$

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

$$(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

从而我们可以发现 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 是线性无关的,$\lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 = \mathbf{0}$ 只有 $\lambda_1 = \lambda_2 = 0$ 时成立。因此 $\{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ 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 $\boldsymbol{e}_3 = (0, 0, 1, 0)^{\mathrm{T}}, \boldsymbol{e}_4 = (0, 0, 0, 1)^{\mathrm{T}}$，那么因为 $\boldsymbol{e}_1 = \boldsymbol{x}_1 - \boldsymbol{e}_3, \boldsymbol{e}_2 = \boldsymbol{x}_2 - \boldsymbol{e}_3 - \boldsymbol{e}_4$，则 $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 可以由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 线性表出，也就是 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{e}_3, \boldsymbol{e}_4$ 生成的子空间为 $\mathbb{R}^4$。

3.4.2 子空间的扩张

定理 7

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

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

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

定理 8

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 $\mathbb{V}$ 中任一向量都能被 n 个线性无关的向量线性表出时，$\mathbb{V}$ 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 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 称为 $\mathbb{V}$ 的一组基。设 $\boldsymbol{a}$ 是 $\mathbb{V}$ 中任一向量，于是 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n, \boldsymbol{a}$ 线性相关，因此 $\boldsymbol{a}$ 可以被基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 线性表出：

$$\boldsymbol{a} = a_1\varepsilon_1 + a_2\varepsilon_2 + \cdots + a_n\varepsilon_n,$$

其中系数 $a_1, a_2, \cdots, a_n$ 是被向量 $\boldsymbol{a}$ 和基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 唯一确定的，这组数就称为 $\boldsymbol{a}$ 在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ **下的坐标**，记为 $(a_1, a_2, \cdots, a_n)$。

例 15

在向量空间 P_n 中，

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

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 $1, x, x^2, \cdots, x^{n-1}$ 就是它的一组基。
在这组基下，多项式 $f(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1}$ 的坐标就是它的系数 $(a_0, a_1, \cdots, a_{n-1})$。
如果在 $\mathbb{V}$ 中取另外一组基

$$\varepsilon_1{}' = 1, \varepsilon_2{}' = (x-a), \cdots, \varepsilon_n{}' = (x-a)^{n-1}.$$

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 16

在 n 维向量空间 $\mathbb{V}$ 中，显然

$$\begin{cases} \boldsymbol{\varepsilon}_1 = (1, 0, \cdots, 0), \\ \boldsymbol{\varepsilon}_2 = (0, 1, \cdots, 0), \\ \qquad \cdots\cdots\cdots \\ \boldsymbol{\varepsilon}_n = (0, 0, \cdots, 1) \end{cases}$$

是一组基。任意向量 $\boldsymbol{a} = (a_1, a_2, \cdots, a_n) = a_1\boldsymbol{\varepsilon}_1 + a_2\boldsymbol{\varepsilon}_2 + \cdots + a_n\boldsymbol{\varepsilon}_n$。
所以 $(a_1, a_2, \cdots, a_n)$ 就是向量 $\boldsymbol{a}$ 在这组基下的坐标。

不难证明，

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

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 $\boldsymbol{a} = (a_1, a_2, \cdots, a_n)$，有

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

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

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

3.5.1 秩、矩阵的秩

定义 14

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

定义 15

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

定理 9

$$\dim L(\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r) = \mathrm{rank}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_r\}$$

例 17

设矩阵 $\boldsymbol{A}$

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

矩阵 $\boldsymbol{A}$ 的行向量组为

$$\boldsymbol{a}_1=(1,1,3,1)\quad \boldsymbol{a}_2=(0,2,-1,4)$$
$$\boldsymbol{a}_3=(0,0,0,5)\quad \boldsymbol{a}_4=(0,0,0,0).$$

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

$$\begin{aligned} b_1 &= (1,0,0,0)^{\mathrm{T}} \\ b_2 &= (1,2,0,0)^{\mathrm{T}} \\ b_3 &= (3,-1,0,0)^{\mathrm{T}} \\ b_4 &= (1,4,5,0)^{\mathrm{T}}. \end{aligned}$$

用同样的方法可以证明，b_1, b_2, b_4 线性无关，且 $b_3 = \frac{7}{2} b_1 - \frac{1}{2} b_2$，所以 b_1, b_2, b_4 是向量组 b_1, b_2, b_3, b_4 的一个极大线性无关组，于是向量组 b_1, b_2, b_3, b_4 的秩为 3，换句话说，矩阵 A 的列秩为 3。

例 18

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， $\boldsymbol{M}$ 为观察评分矩阵， $\boldsymbol{M}_{ij}$ 为观测到的用户 i 对物品 j 的评分， $\boldsymbol{X}$ 为预测评分矩阵， $\boldsymbol{X}_{ij}$ 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合 $\mathbb{E}$ 中所有项匹配的最低秩矩阵 $\boldsymbol{X}$。形式化如下

$$\begin{aligned} \min_{\boldsymbol{X}} \quad & \operatorname{rank}(\boldsymbol{X}) \\ s.t. \quad & \boldsymbol{X}_{ij} = \boldsymbol{M}_{ij} \quad \forall i,j \in \mathbb{E} \end{aligned}$$

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：

$$\begin{aligned} \min_{\boldsymbol{X}} \quad & \sum_{ij} (\boldsymbol{X}_{ij} - \boldsymbol{M}_{ij})^2 \quad \forall i,j \in \mathbb{E} \\ s.t. \quad & \operatorname{rank}(\boldsymbol{X}) = r \end{aligned}$$

3.6.1 仿射子空间

定义 16

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

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

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 $x_0 \notin \mathbb{U}$，则仿射子空间 $\mathbb{L}$ 不是一个线性子空间。
- 若 $\mathbb{U}$ 有一基底 $\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_m$，则 $\mathbb{L}$ 中的每一个元素 $\boldsymbol{x}$ 均可写成 $\boldsymbol{x}_0 + k_1\boldsymbol{a}_1 + k_2\boldsymbol{a}_2 + \dots + k_m\boldsymbol{a}_m$。
  这一结论通过定义是容易知道的。

超平面

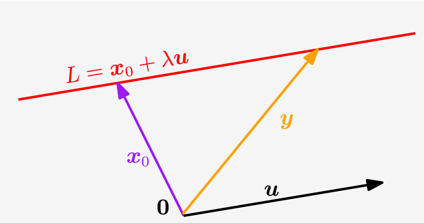

图 7: 仿射子空间

例 19

$\mathbb{R}^3$ 中常见的仿射子空间

1. 零维仿射子空间：单点集 $\{\boldsymbol{x}_0\}$
2. 一维仿射子空间：直线 $\{\boldsymbol{x}_0 + k\boldsymbol{u}\}$
3. 二维仿射子空间：平面 $\{\boldsymbol{x}_0 + k_1\boldsymbol{u}_1 + k_2\boldsymbol{u}_2\}$
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 20

我们已经知道线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \boldsymbol{b}\neq\boldsymbol{0}$ 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 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{0}$ 的解空间为 $\mathbb{V}$，它是一个子空间；设 $\boldsymbol{x}_0$ 是 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的一个特解。
$\forall \boldsymbol{x}\in\boldsymbol{x}_0+\mathbb{V}$，$\boldsymbol{x}$ 必可以写成 $\boldsymbol{x}=\boldsymbol{x}_0+\boldsymbol{x}_1$，其中 $\boldsymbol{x}_1\in\mathbb{V}$。显然：

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{A}(\boldsymbol{x}_0+\boldsymbol{x}_1)=\boldsymbol{A}\boldsymbol{x}_0+\boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{0}+\boldsymbol{b}=\boldsymbol{b}.$$

说明 $\boldsymbol{x}_0+\mathbb{V}\subseteq\{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\}$
反之,$\forall\boldsymbol{x}$ 满足 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, 则 $\boldsymbol{A}\boldsymbol{x}-\boldsymbol{A}\boldsymbol{x}_0=\boldsymbol{A}(\boldsymbol{x}-\boldsymbol{x}_0)=\boldsymbol{0}$, 则 $\boldsymbol{x}-\boldsymbol{x}_0\in\mathbb{V}$，$\boldsymbol{x}\in\boldsymbol{x}_0+\mathbb{V}$。
说明 $\{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}\}\subseteq\boldsymbol{x}_0+\mathbb{V}$
综上，线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \boldsymbol{b}\neq\boldsymbol{0}$ 的解空间为 $\boldsymbol{x}_0+\mathbb{V}$，这是一个仿射空间。

- 向量空间
- 向量子空间
- 向量组的线性相/无关
- 向量组的等价
- 生成集与向量组张成子空间
- 向量空间的维数、基和坐标
- 仿射子空间