《线性代数》

1-向量介绍 (Introduction to Vectors)

杨启哲

上海师范大学信机学院计算机系

2024 年 2 月 28 日

主要内容

- 向量加法和数乘
- 向量长度和点积
- 矩阵

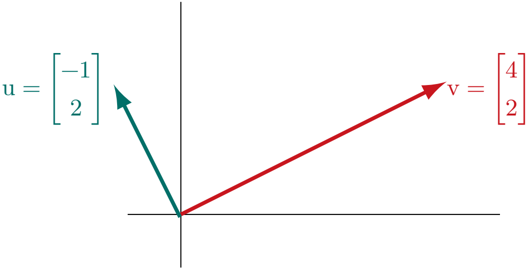
教材参考章节

- 第 2 章 2.1, 2.2, 第 5 章 5.1

向量加法和数乘

首先我们来考察一下二维空间中的向量，用**列向量 (colomn vectors)** 来表示：

- $u = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, $v = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$



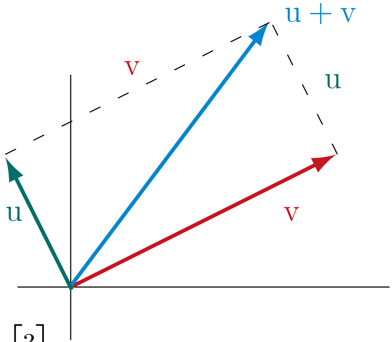
符号说明

在本课程的课件中，我们使用 $\mathrm{x}, \mathrm{y}, \mathrm{z} \ldots$ 等符号来表示向量，使用 $x, y, z, \ldots$ 等符号来表示标量的值；特别的对于一个向量 u 来说，我们经常使用 u_i 表示其第 i 个分量的值。

向量加法 (Vector Addition)

向量加法:

$$\underset{\mathbf{u}}{\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}} + \underset{\mathbf{v}}{\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}} = \underset{\mathbf{u}+\mathbf{v}}{\begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \end{bmatrix}}$$

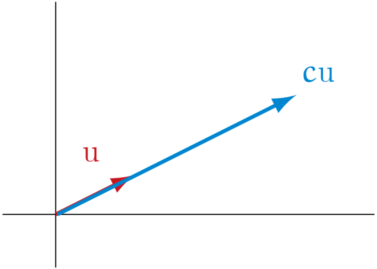


- 一个例子: $\begin{bmatrix} -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

向量数乘

$$c \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}$$

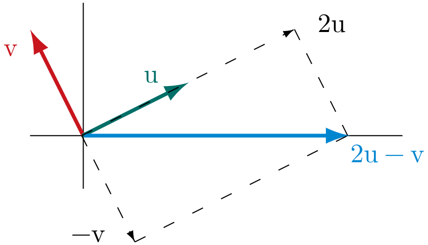
其中 c 是一个标量，也称为 scalar.



- 一个例子：$3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$

考虑如下的例子:

$$2\mathrm{u} - \mathrm{v} = 2\begin{bmatrix}2\\1\end{bmatrix} + (-1)\begin{bmatrix}-1\\2\end{bmatrix} = \begin{bmatrix}4\\2\end{bmatrix} - \begin{bmatrix}-1\\2\end{bmatrix} = \begin{bmatrix}5\\0\end{bmatrix}$$



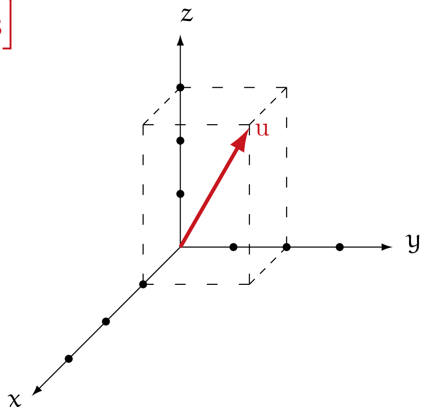
向量的线性组合

$$\underset{c\mathrm{u}}{c\begin{bmatrix}u_1\\u_2\end{bmatrix}} + \underset{d\mathrm{v}}{d\begin{bmatrix}v_1\\v_2\end{bmatrix}} = \underset{c\mathrm{u}+d\mathrm{v}}{\begin{bmatrix}cu_1 + dv_1\\cu_2 + dv_2\end{bmatrix}}$$

向量在三维的情况

- 二维向量 $\begin{bmatrix} x \\ y \end{bmatrix}$ 可以视作在 2 维 xOy 平面上从 $(0,0)$ 指向 (x,y) 的一个有向线段。
- 三维向量也是类似的，只不过是在 3 维空间中从 $(0,0,0)$ 指向 (x,y,z) 的一个有向线段，

  比如考察向量 $u = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$：



给定 3 维的向量:

$$
\mathrm{u}=\begin{bmatrix}1\\0\\3\end{bmatrix},\ \mathrm{v}=\begin{bmatrix}1\\2\\1\end{bmatrix},\ \mathrm{w}=\begin{bmatrix}2\\3\\-1\end{bmatrix}
$$

其线性组合也是类似的:

$$
\begin{aligned}
\mathrm{u}=1\mathrm{u} &= \begin{bmatrix}1\\0\\3\end{bmatrix}\\
\mathrm{u}+2\mathrm{v} &= \begin{bmatrix}1\\0\\3\end{bmatrix}+2\begin{bmatrix}1\\2\\1\end{bmatrix}=\begin{bmatrix}3\\4\\5\end{bmatrix}\\
\mathrm{u}+4\mathrm{v}-2\mathrm{w} &= \begin{bmatrix}1\\0\\3\end{bmatrix}+4\begin{bmatrix}1\\2\\1\end{bmatrix}-2\begin{bmatrix}2\\3\\-1\end{bmatrix}=\begin{bmatrix}1\\2\\9\end{bmatrix}
\end{aligned}
$$

三个问题

假设 u, v, w 是三维空间的三个向量:

- 所有形如 cu 的线性组合对应的几何直观是什么?
- 所有形如 $cu + dv$ 的线性组合对应的几何直观是什么?
- 所有形如 $cu + dv + ew$ 的线性组合对应的几何直观是什么?

答案当然依赖于 u, v, w 的具体取值，但是我们可以通过一些例子来感受一下。

三个问题 (II)

上海师范大学 Shanghai Normal University

我们考虑之前的三个向量:

$$u = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \ v = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \ w = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$$

1. 形如 cu 的线性组合填满了一条通过 $(0, 0, 0)$ 的线。
2. 形如 $cu + dv$ 的线性组合填满了一个通过 $(0, 0, 0)$ 的平面。
   - $(2, 3, -1)$ 便不在此平面上，因此 w 无法表示成 $cu + dv$ 的线性组合。
3. 形如 $cu + dv + ew$ 的线性组合填满了整个三维空间。
   - 这意味着对于任何的 $a, b, c \in \mathbb{R}$

$$\begin{aligned} x + y + 2z &= a \\ 2y + 3z &= b \\ 3x + y - z &= c \end{aligned}$$

   存在一个解。

问题 1.

描述一下由向量 $u = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $v = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ 的线性组合在三维空间中所组成的平面。

解 2.

形如 $cu + dv$ 的线性组合填满了一个通过 $(0, 0, 0)$ 的平面，我们有:

$$cu + dv = \begin{bmatrix} c \\ c + d \\ d \end{bmatrix}$$

c, d 是任意的，所以该平面包含了所有第二维是第一维和第三维的和的向量。

阶段总结

1. 二维向量是一个具有两个分量的有序对，可以视作二维平面原点出发的一个有向线段。
2. 向量的加法和数乘是针对对应分量进行操作。
3. 向量之间的线性组合是指形如 $cu + dv + ew$ 的向量。
4. 在三维空间中，向量的线性组合可以填满一条线，一个平面，或者整个三维空间。

向量长度和点积

向量的一些几何性质

符号说明

为了节省空间，列向量 $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ 有时使用 (u_1, u_2, u_3) 表示。

我们来考察两维中的向量，给定两个向量 $u = (u_1, u_2)$，$v = (v_1, v_2)$：

1. 向量 u 的长度 $\|u\|$ 是多少？

$$\|u\| = \sqrt{u_1^2 + u_2^2}$$

2. 向量 u 与 v 的夹角 θ 是多少？

$$\cos\theta = \frac{u \cdot v}{\|u\|\|v\|}$$

这些问题我们都可以使用点积 (dot product) 的概念来解决。

Shanghai Normal University

定义 3 **[点积].**

向量 $u = (u_1, u_2)$，$v = (v_1, v_2)$ 的点积 $u \cdot v$ 定义为：

$$u \cdot v = u_1 v_1 + u_2 v_2$$

一般的，对于向量 $u = (u_1, u_2, \ldots, u_n)$，$v = (v_1, v_2, \ldots, v_n)$，其点积定义为：

$$u \cdot v = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \sum_{i=1}^{n} u_i v_i$$

点积的一些性质

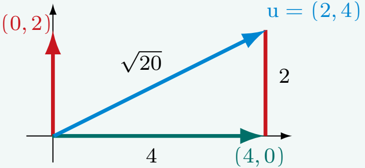
1. u 和 v 是垂直的 (perpendicular) 当且仅当 $u \cdot v = 0$。
2. 点积是可交换的，即 $u \cdot v = v \cdot u$。

定义 4 **[向量的长度].**

向量 $\mathrm{u} = (u_1, u_2, \ldots, u_n)$ 的长度 $\|\mathrm{u}\|$ 定义为：

$$\|\mathrm{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2} = \sqrt{\mathrm{u} \cdot \mathrm{u}}$$

勾股定理 (Pythagorean Law)



一般来说，对于垂直的 u 和 v：

$$\|\mathrm{u}\|^2 + \|\mathrm{v}\|^2 = \|\mathrm{u} + \mathrm{v}\|^2 = \|\mathrm{u} - \mathrm{v}\|^2$$

定义 5.

长度为 1 的向量 u 被称作为单位向量，即 $\|\mathrm{u}\| = 1$.

例 6.

如下的向量都是单位向量:

$$\mathrm{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \mathrm{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \mathrm{u} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$

单位化 (Normalization)

给定一个向量 $u = (2, 2, 1)$，其长度为:

$$\|u\| = \sqrt{2^2 + 2^2 + 1^2} = \sqrt{9} = 3$$

我们可以将其缩小成一个单位向量:

$$(\frac{2}{3}, \frac{2}{3}, \frac{1}{3})$$

引理 7.

给定一个非零向量 $u = (u_1, \ldots, u_n)$，则:

$$\frac{u}{\|u\|} = (\frac{u_1}{\|u\|}, \ldots, \frac{u_n}{\|u\|})$$

是一个与 u 同方向的单位向量。

互相垂直的向量

现在我们来说明点积值和向量之间的夹角的关系:

引理 8.

令 $\mathrm{u}, \mathrm{v} \in \mathbb{R}^2$，如果 u 和 v 是垂直的，则：$\mathrm{u} \cdot \mathrm{v} = 0$。

证明. 由勾股定理，我们有:

$$\|\mathrm{u}\|^2 + \|\mathrm{v}\|^2 = \|\mathrm{u} - \mathrm{v}\|^2$$

将其展开有:

$$\begin{aligned} u_1^2 + u_2^2 + v_1^2 + v_2^2 &= (u_1 - v_1)^2 + (u_2 - v_2)^2 \\ u_1 v_1 + u_2 v_2 &= 0 \end{aligned}$$

$\square$

说明

1. $\mathrm{u} \cdot \mathrm{v} = 0 \iff$ u 和 v 是垂直的。
2. $\mathbf{0} \cdot \mathrm{u} = 0$，零向量 $\mathbf{0}$ 与任何向量都是垂直的。

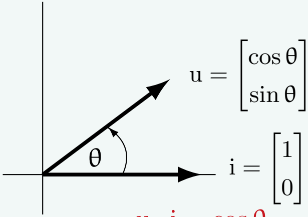
定理 9.

令 $u, v \in \mathbb{R}^2$ 是两个单位向量，θ 是 u 和 v 之间的夹角，则：

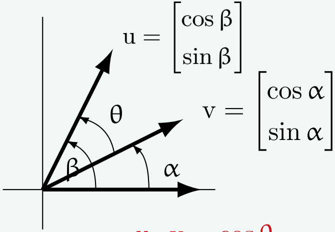
$$\cos\theta = u \cdot v$$

几何视角

1. $v = i = (1, 0)$，则有 $\cos\theta = u \cdot v = u_1$。
2. $v \neq (1, 0)$，则可以视其旋转了 α 角度，其中 $v = (\cos\alpha, \sin\alpha)$。



$u \cdot i = \cos\theta$



$u \cdot v = \cos\theta$

定理 9的证明.

- 若 $v = (1,0)$，则易得 $u \cdot v = u_1 = \cos\theta$.
- 若 $v \neq (1,0)$，则可以视 v 是由 $(1,0)$ 旋转了 α 角度得到的单位向量，即 $v = (\cos\alpha, \sin\alpha)$; 同理令 $u = (\cos\beta, \sin\beta)$，则其夹角为 $\theta = \beta - \alpha$，并且:

$$u \cdot v = \cos\beta\cos\alpha + \sin\beta\sin\alpha = \cos(\beta - \alpha) = \cos\theta$$

$\square$

如果 u, v 不是单位向量，怎么求其夹角？

- 将其单位化。

定理 10.

令 $\mathrm{u}, \mathrm{v} \in \mathbb{R}^2$ 是两个非零向量，θ 是 u 和 v 之间的夹角，则：

$$\cos\theta = \frac{\mathrm{u}\cdot\mathrm{v}}{\|\mathrm{u}\|\|\mathrm{v}\|}$$

现在我们来看一些例子：

定理 11 **[Cauchy-Schwarz-Buniakowsky].**

$$|u \cdot v| \leqslant \|u\| \|v\|$$

证明. 我们用一个不用余弦定理的方法来证明。

注意到：

$$\begin{aligned}(u_1^2+u_2^2)(v_1^2+v_2^2)-(u_1v_1+u_2v_2)^2 &= u_1^2v_2^2+u_2^2v_1^2-2u_1u_2v_1v_2\\ &= (u_1v_2-u_2v_1)^2 \geqslant 0\end{aligned}$$

$\square$

定理 12 **[三角不等式].**

$$\|\mathrm{u}+\mathrm{v}\| \leqslant \|\mathrm{u}\| + \|\mathrm{v}\|$$

证明. 我们同样使用一个不用余弦定理的方法来证明。

注意到:

$$\begin{aligned} \|\mathrm{u}+\mathrm{v}\|^2 &= \|\mathrm{u}\|^2 + \|\mathrm{v}\|^2 + 2\mathrm{u}\cdot\mathrm{v} \\ \mathrm{u}\cdot\mathrm{v} &\leqslant \|\mathrm{u}\|\|\mathrm{v}\| \end{aligned}$$

从而我们有:

$$\|\mathrm{u}+\mathrm{v}\|^2 \leqslant (\|\mathrm{u}\| + \|\mathrm{v}\|)^2$$

□

问题

如何证明 $\|\mathrm{u}+\mathrm{v}\|^2 = \|\mathrm{u}\|^2 + \|\mathrm{v}\|^2 + 2\mathrm{u}\cdot\mathrm{v}$?

阶段总结

1. 向量的点积是相应部分的乘积的和，即 $\mathrm{u} \cdot \mathrm{v} = u_1 v_1 + \ldots u_n v_n$。
2. 向量的长度是点积的平方根，其对应的单位向量为：$\frac{\mathrm{u}}{\|\mathrm{u}\|}$，长度为 1。
3. $\mathrm{u} \cdot \mathrm{v} = 0$ 意味着两个向量 u, v 是垂直的。
4. 向量夹角的余弦值等于点积除以长度的乘积，即:

$$\cos\theta = \frac{\mathrm{u} \cdot \mathrm{v}}{\|\mathrm{u}\|\|\mathrm{v}\|}$$

5. 柯西-施瓦茨不等式和三角不等式。

$$|\mathrm{u} \cdot \mathrm{v}| \leqslant \|\mathrm{u}\|\|\mathrm{v}\|, \qquad \|\mathrm{u} + \mathrm{v}\| \leqslant \|\mathrm{u}\| + \|\mathrm{v}\|$$

矩阵

上海师范大学 Shanghai Normal University

给定三个向量:

$$u=\begin{bmatrix}1\\-1\\0\end{bmatrix},\ v=\begin{bmatrix}0\\1\\-1\end{bmatrix},\ w=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

其线性组合可以表示为 $x_1u+x_2v+x_3w$，也就是:

$$x_1\begin{bmatrix}1\\-1\\0\end{bmatrix}+x_2\begin{bmatrix}0\\1\\-1\end{bmatrix}+x_3\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}x_1\\x_2-x_1\\x_3-x_2\end{bmatrix}$$

矩阵表示

我们可以利用矩阵来表示上述行为:

$$Ax=\begin{bmatrix}1&0&0\\-1&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}x_1\\x_2-x_1\\x_3-x_2\end{bmatrix}$$

我们可以理解成，矩阵 A 作用在一个列向量 x 上，其结果是矩阵 A 中的列向量的线性组合。

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b$$

例 13.

上述的矩阵 A 被称作差分矩阵(difference matrix)，因为其得到的向量是原向量的差分。

$$A \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} = \begin{bmatrix} 1-0 \\ 4-1 \\ 9-4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

另一个视角来看待矩阵的作用

我们现在用行的角度来看待矩阵作用在向量上的结果:

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (1,0,0)\cdot x \\ (-1,1,0)\cdot x \\ (0,-1,1)\cdot x \end{bmatrix} = b$$

补充说明

这是大多数中文教材中定义的方式。

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = b$$

- 之前我们讨论的是线性组合，即给定三个向量 u, v, w 和三个数 x_1, x_2, x_3，求其线性组合 $x_1 u + x_2 v + x_3 w$; 将矩阵 A 看成 $\begin{bmatrix} u & v & w \end{bmatrix}$ 的话，即知道了 A 和 x，求 b。
- 现在我们来考虑另一个问题：给定矩阵 A 和 b，求 x。

一个大家更为熟知的形式：线性方程组 (Linear Equations):

$$
\begin{aligned}
x_1 + 0x_2 + 0x_3 &= b_1 \\
-x_1 + x_2 + 0x_3 &= b_2 \\
0x_1 - x_2 + x_3 &= b_3
\end{aligned}
$$

不难验证，其解可以表示为：

$$
\begin{aligned}
x_1 &= b_1 \\
x_2 &= b_1 + b_2 \\
x_3 &= b_1 + b_2 + b_3
\end{aligned}
$$

上述过程也可以写成另一个矩阵:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = {\color{red}A^{-1}} b$$

上述矩阵称为差分矩阵 A 的逆矩阵 A^{-1}。

我们再来进行一下对比:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix},\ A^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

$$Ax = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix},\ A^{-1}b = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

另一个例子 (I)

给定三个向量：

$$u = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \ v = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \ w = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

其线性组合可以表示为 $x_1 u + x_2 v + x_3 w$，也就是：

$$x_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

矩阵表示

我们将上述矩阵记为 C，也被称为循环差分矩阵 (cyclic difference matrix)：

$$Cx = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 - x_3 \\ x_2 - x_1 \\ x_3 - x_2 \end{bmatrix}$$

另一个例子 (II)

与 A 的不同的是，$C\mathbf{x}=\mathbf{b}$ 不一定有解，例如：

$$C\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}1&0&-1\\-1&1&0\\0&-1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}x_1-x_3\\x_2-x_1\\x_3-x_2\end{bmatrix}=\begin{bmatrix}1\\3\\5\end{bmatrix}$$

没有解。

$x_1\mathbf{u}+x_2\mathbf{v}+x_3\mathbf{w}$ 的几何直观

换个角度讲，$x_1\mathbf{u}+x_2\mathbf{v}+x_3\mathbf{w}$ 的所有线性组合并没有充满了整个 3 维空间。事实上，其仅仅覆盖了如下的一个平面：

$$b_1+b_2+b_3=0$$

更多的讨论

无关性和相关性 (Independence and Dependence)

- 在上述 $Ax = b$ 的例子中，$xu + yv + zw = \mathbf{0}$ 当且仅当 $x = y = z = 0$。
  后面的课程中我们会看到：
  1. u, v, w 是线性无关的。
  2. $Ax = \mathbf{0}$ 只有一个解，称 A 是可逆矩阵 (invertible matrix)。
- 在上述 $Cx = b$ 的例子中，存在任意多个 x, y, z 满足 $xu + yv + zw = \mathbf{0}$。
  后面的课程中我们会看到：
  1. u, v, w 是线性相关的。
  2. $Cx = \mathbf{0}$ 有无穷多个解，称 C 是一个奇异矩阵 (singular matrix)。

阶段总结

1. 矩阵作用在向量 Ax = 矩阵 A 的列向量的线性组合。
2. $\mathrm{Ax} = \mathrm{b}$ 的解为 $\mathrm{x} = \mathrm{A}^{-1}\mathrm{b}$。
3. $\mathrm{Cx} = \mathbf{0}$ 存在无穷多个解，C 没有逆矩阵。

注意

我们并没有给出严格的相关定义，但我们已经描述了这些关键的想法。