教材、参考书

• 教材:

• Gilbert Strang, Introduction to Linear Algebra, fifth Edition, Wellesley-Cambridge Press, 2016. (影印版,清华大学出版社)

• 参考书:

- Gilbert Strang, Linear Algebra and Learning from Data,
 Wellesley-Cambridge Press, 2019.
- R. A. Horn and C. R. Johnson, Matrix Analysis. New York: Cambridge University Press, 1989.
- R. A. Horn and C. R. Johnson, Topics in Matrix Analysis. New York: Cambridge University Press, 1991.

考试与作业

- 日常练习(占期末总评10%,随堂布置,5-8题/次,独立完成,提交拍照电子版,期末统一讲评);
- 大作业(占期末总评20%,以一至三人组队完成);
- 期末考试(占期末总评70%,时间暂定第17周)。

关于大作业

- 挑选教材 Chapters 10-12 任意一节自学;
- 使用中文叙述学习体会(翻译+学习心得);
- 完成课后习题至少5题;
- 使用 Matlab/Python/C 编程;
- 使用 Latex 打印书写作业;
- 2人或3人合作;需在作业中注明每个人的分工,以及相应完成部分(按照每人的实际贡献给分)。

注:需自学Latex排版软件。

第一讲:向量空间与子空间

1.1 向量与线性组合

- ① $3\mathbf{v} + 5\mathbf{w}$ 是向量 $\mathbf{v} = \mathbf{v} = \mathbf{v} + \mathbf{v} = \mathbf{v} + \mathbf{v} = \mathbf{v} = \mathbf{v} = \mathbf{v} + \mathbf{v} = \mathbf{v}$
- ② 给定向量 $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 与 $\mathbf{w} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ 则它们的线性组合是 $\mathbf{z} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5 \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3+10 \\ 3+15 \end{bmatrix} = \begin{bmatrix} 13 \\ 18 \end{bmatrix}$.
- ③ 线性组合 $c \begin{bmatrix} 1 \\ 1 \end{bmatrix} + d \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ 撑满整个 xy 平面. 他们生成所有向量 $\begin{bmatrix} x \\ y \end{bmatrix}$.
- 4 线性组合 $c \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + d \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ 撑满 xyz 空间中的一个平面.
- $\begin{cases} c+2d=1\\ c+3d=0\text{ 无解,因为右端向量} \\ c+4d=0 \end{cases} \begin{bmatrix} 1\\0\\0 \end{bmatrix}$ 不在这个平面内.

1.2 长度与点积

① 向量
$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
 和 $\mathbf{w} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$ 的 "点积" 是 $\mathbf{v} \cdot \mathbf{w} = (1)(4) + (2)(5) = 4 + 10 = \mathbf{14}$.

② 向量
$$\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$
 与 $\mathbf{w} = \begin{bmatrix} 4 \\ -4 \\ 4 \end{bmatrix}$ 是垂直的,因为它们的点积等于零:
$$(1)(4) + (3)(-4) + (2)(4) = \mathbf{0}.$$

③
$$\mathbf{v} = \begin{vmatrix} 1 \\ 3 \\ 2 \end{vmatrix}$$
 的长度平方是 $\mathbf{v} \cdot \mathbf{v} = 1 + 9 + 4 = 14$. 它的长度是 $\|\mathbf{v}\| = \sqrt{14}$.

①
$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{v}}{\sqrt{14}} = \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$
 有长度 $\|\mathbf{u}\| = 1$. 验证 $\frac{1}{14} + \frac{9}{14} + \frac{4}{14} = 1$.

- ② 向量 v 和 w 的夹角 θ 满足 $\cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\| \|\mathbf{w}\|}$.
- ③ 例如:向量 $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ 和 $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ 的夹角 θ 满足 $\cos \theta = \frac{1}{(1)(\sqrt{2})}$.那么这个夹角等于 $\theta = 45^\circ$.
- **4 柯西施瓦茨不等式:** 所有的夹角都满足 $|\cos \theta| \le 1$. 故对所有的向量都有 $|\mathbf{v} \cdot \mathbf{w}| \le ||\mathbf{v}|| ||\mathbf{w}||$.

1.3 矩阵

①
$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$
 是 3 乘 2 的矩阵: $m = 3$ 行和 $n = 2$ 列.

②
$$A\mathbf{x} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
 可视为列的线性组合 $A\mathbf{x} = x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$.

③ 线性方程组的矩阵形式
$$A\mathbf{x} = \mathbf{b}: \begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$
 取代

$$2x_1 + 5x_2 = b_1$$
$$3x_1 + 7x_2 = b_2$$

④ Ax = b 的解可表示为 $x = A^{-1}b$,当且仅当 A^{-1} 存在时。

1.4 矩阵乘法法则

- ① 具有n个列的矩阵A与具有n个行的矩阵可相乘: $A_{m \times n}B_{n \times p} = C_{m \times p}$.
- ② AB = C 中的每个元素实际是一个点积: $C_{ij} = (A \text{ 的第 } i \text{ 行}) \cdot (B \text{ 的第 } j$ 列).

- ③ 结合率: AB 乘以 C 等于 A 乘以 BC. 特别地, (AB)x = A(Bx).
- ④ 交換律通常不能满足:即难以得到 AB = BA.并且更多情况下,A 与 B 不可乘.
- ⑤ 矩阵乘法的分块形式: $A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}$ 乘以 $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ 等于 $A_1B_1 + A_2B_2$.

1.5 逆矩阵

- 1 若有一个方阵 A 具有一个逆 A^{-1} ,则 $A^{-1}A = I$ 与 $AA^{-1} = I$.
- 2 验证可逆的计算算法是消元法: A 必须有 n 个(非零)主元。
- 3 验证可逆的代数法则是行列式: A 的行列式 det A 不等于零.
- 4 验证可逆的方程方法是: 齐次方程 Ax = 0: x = 0 只有零解.
- 5 当 A 和 B 维数相同且都可逆时,AB: $(AB)^{-1} = B^{-1}A^{-1}$.
- 6 $AA^{-1} = I$ 表示 n 个方程以计算 A^{-1} 的 n 个列。高斯消去法: $[A\ I]$ 至 $[I\ A^{-1}]$.

1.6 转置

- 1 Ax, AB, A^{-1} 的转置依次是 x^TA^T , B^TA^T , $(A^T)^{-1}$.
- 2 点积(内积)可表示为 $x \cdot y = x^T y$, 它是 $(1 \times n)(n \times 1) = (1 \times 1)$.
- 3 关于 A^T 的一个常见运算: $Ax \cdot y$ 会等于 $x \cdot A^Ty$, 因为 $(Ax)^Ty = x^TA^Ty = x^T(A^Ty)$.

1.7 向量空间

向量空间及其子空间是理解方程 Ax = b 解的重要手段。

由最重要的向量空间开始:它们可记为 \mathbb{R}^1 , \mathbb{R}^2 , \mathbb{R}^3 , \mathbb{R}^4 , 每个空间 \mathbb{R}^n 包含 所有 n 维向量的集合。例如, \mathbb{R}^5 包含所有具有5个分量的向量的集合,因此称之为: 5 维空间。

定义: 向量空间 R^n 由所有具有 n 个分量的列向量 v 组成。

向量 v 的分量都是实数,这就是为什么我们使用字母 R. 当一个向量的 n 个分量出现复数时,记为向量空间 \mathbf{C}^n .

向量空间 \mathbb{R}^2 表示常见的 x-y 平面。 \mathbb{R}^2 中的每个向量具有两个元素。"空间" 这个词要求我们考虑整个平面的所有向量。每个向量给出 x 和 y 平面坐标系 里的一个点: v=(x,y).

类似地, \mathbb{R}^3 中的向量对应所有的点 (x,y,z),是三维空间; \mathbb{R}^1 是一条直线,是一维空间。

向量空间:满足加法、数乘运算封闭,并且对应运算满足如下八条性质:

向量加法法则:

- x + (y + z) = (x + y) + z
- 存在唯一的"零向量"使得 x + 0 = x
- 对每个 x, 存在唯一的"负向量"-x满足 x + (-x) = 0

向量的(实)数乘法则

- 1 乘以 x 等于 x
- $c(\mathbf{x} + \mathbf{y}) = c\mathbf{x} + c\mathbf{y}$

除列向量外还有向量,除 R^n 外还有其它向量空间,所有向量空间都必须遵守 八条合理的规则。

实向量空间是一组"向量"以及向量加法和实数乘法规则。

加法和乘法必须产生空间中的向量,并且必须满足八个条件。

以下列举 R^n 外的常见三个向量空间:

M 所有 2 乘 2 实矩阵构成的向量空间

F 所有实函数 f(x) 构成的向量空间

Z 仅包含零向量的向量空间

F 是无限维的; Z 是零维的; 问题: M 是几维的?

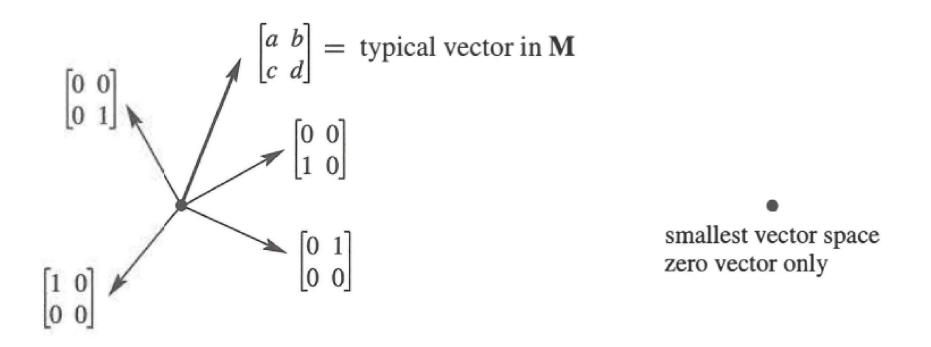


Figure 3.1: "Four-dimensional" matrix space M. The "zero-dimensional" space Z.

1.8 子空间

从通常的三维空间 \mathbf{R}^3 开始。选择穿过原点 (0,0,0) 的平面. 这个平面本身就是一个向量空间。如果我们在平面中添加两个向量,他们的和在平面上。如果我们将平面内向量乘以 2 或 -5,它仍然在平面内。三维空间中的平面不是 \mathbf{R}^2 (即使它看起来像 \mathbf{R}^2)。向量具有三个分量,它们属于 \mathbf{R}^3 。平面是一个向量空间在 \mathbf{R}^3 的内部。

这说明了线性代数中最基本的思想之一。穿过 (0,0,0) 的平面是向量空间 \mathbf{R}^3 的子空间。

定义 向量空间的一个子空间,它包含满足以下关系的所有向量(包括0)的集

合: 若 v 与 w 是该集合中的任意两个向量,则

(i) $\mathbf{v} + \mathbf{w}$ 属于它 (ii) $c\mathbf{v}$ 也属于它

上述条件也可表述为: 若 v 与 w 的所有线性组合属于这个集合。

例: \mathbb{R}^3 所有可能的子空间:

(L) 任意穿过 (0, 0, 0)的直线 (P) 任意穿过 (0, 0, 0)的平面

 (\mathbf{R}^3) 全空间 (\mathbf{Z}) 仅包含原点 (0, 0, 0) 集合

如果我们只保留平面或直线的一部分,子空间的要求就不成立。看看以下 \mathbb{R}^2 中的这些例子——它们不是子空间。

- 例 1 只保留分量为正或零的向量 (x,y) (这是四分之一平面)。包含向量 (2,3),但不包含向量 (-2,-3)。因此,当我们试图乘以 c=-1 时,违反了规则 (ii)。四分之一平面不是子空间。
- 例 2 即使包括了其分量均为负值的向量。现在我们有两架四分之一平面。满足要 (ii);我们可以用任何 c 乘法。但规则 (i) 依然失败。 $\mathbf{v}=(2,3)$ 和 $\mathbf{w}=(-3,-2)$ 的和是 (-1,1),它位于两个四分之一平面之外。**两个四分之一平面不构成子空间**。

例 3 在所有2乘2矩阵的向量空间M内,有两个子空间:

(U) 所有上三角矩阵
$$\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$
 (D) 所有对角矩阵 $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$

在 U 中添加任意两个矩阵,总和在 U 中。对角矩阵的和还是对角的。在这种情况下,D 也是 U 的子空间! 当然,当 a, b, 和 d 都等于零时,零矩阵就在这些子空间中。

小结

- 1 标准 n 维向量空间 \mathbf{R}^n 包含所有具有 n 个元素的实列向量.
- 2 若 v 和 w 在向量空间 S 内,则它们的任意线性组合也在 S 内.
- 3 **S** 中的"向量"可以是 **x** 的矩阵或函数。1-点空间 **Z** 由 x = 0 组成。
- 4 \mathbf{R}^n 的子空间是 \mathbf{R}^n 内的向量空间。示例:直线 $\mathbf{y}=3\mathbf{x}$ 在 \mathbf{R}^2 内。

习题

Problem Set 3.1: 3, 10, 15, 17, 18, 30

第二讲:线性方程组的几何解释,矩阵的列空间

2.1 向量与线性方程

线性代数的中心问题是求解方程组。这些方程是线性的,这意味着未知量只会乘以数字——我们永远看不到 x 乘以 y (非线性项)。

以下线性方程组规模很小,但可以有不同视角来看它:

两个方程
$$x - 2y = 1$$
 π $3x + 2y = 11$ (1)

行视角.

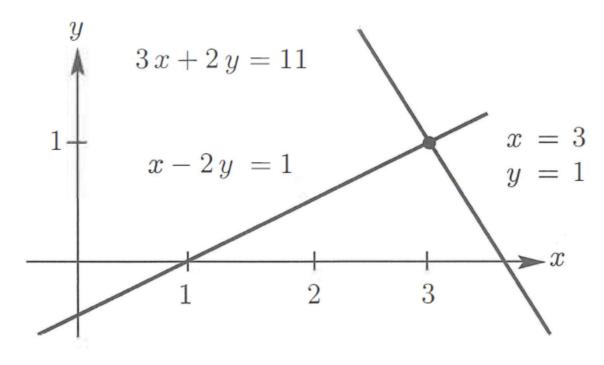


Figure 2.1: Row picture: The point (3, 1) where the lines meet solves both equations.

行: 行视角展示为两条直线相交于单一点(方程的解)。

列视角.

把同一个线性方程看作"向量方程"。我们需要看到的不是数字而是向量。如果将原始方程拆分为列而不是行,则会得到一个向量方程:

列向量的线性组合等于 b
$$x \begin{bmatrix} 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 11 \end{bmatrix} = \mathbf{b}.$$
 (2)

左侧有两个列向量。问题是找到那些等于右边向量的线性组合。我们将第一列乘以 x,第二列乘以 y,然后相加。如果正确选择 x=3 和 y=1(与前面的数字相同),则会生成 3(column 1) +1(column 2) = b。

列: 列视角展示方程左端列向量的线性组合以生成右端向量 b。

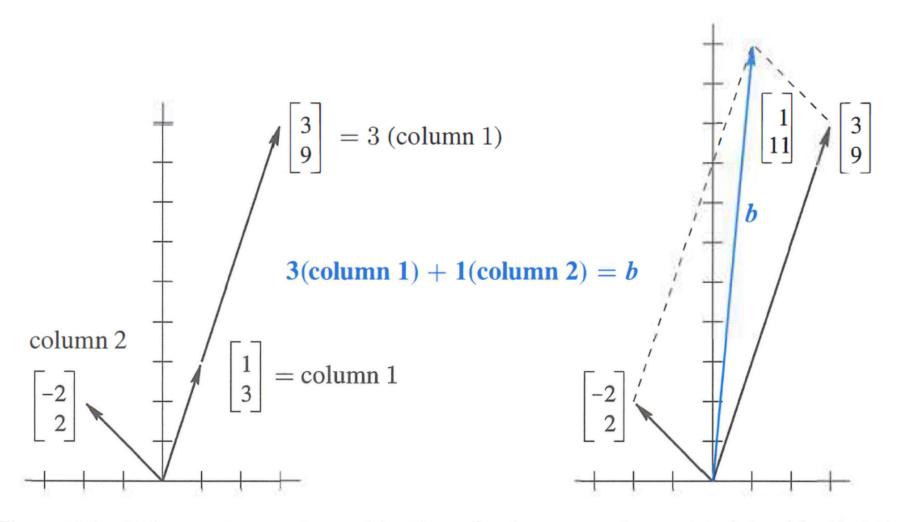


Figure 2.2: Column picture: A combination of columns produces the right side (1, 11).

强调: 向量方程的左侧是列的线性组合。问题是找到正确的系数 x = 3 和 y = 1。我们将标量乘法和向量加法合并为一个步骤。这一步至关重要,因为它包含两个基本操作: 分别乘以 3 和 1, 然后相加.

线性组合
$$3\begin{bmatrix}1\\3\end{bmatrix}+\begin{bmatrix}-2\\2\end{bmatrix}=\begin{bmatrix}1\\11\end{bmatrix}$$
.

2.2 三个方程三个未知数

设有三个未知数 x,y,z 满足三个线性方程:

$$x + 2y + 3z = 6$$

$$Ax = b 2x + 5y + 2z = 4 (3)$$

$$6x + 3y + z = 2$$

我们寻找能够同时求解所有三个方程的数 x,y,z。这些所需的数可能存在,也可能不存在。对于这个系统,它们确实存在。当未知数的数目与有效的方程的数目相匹配时,通常只有一个解。

两种视角:

行 行视角展示为三个平面相交于单一点.

列 列视角展示方程左端列向量的线性组合以生成右端向量 b = (6,4,2).

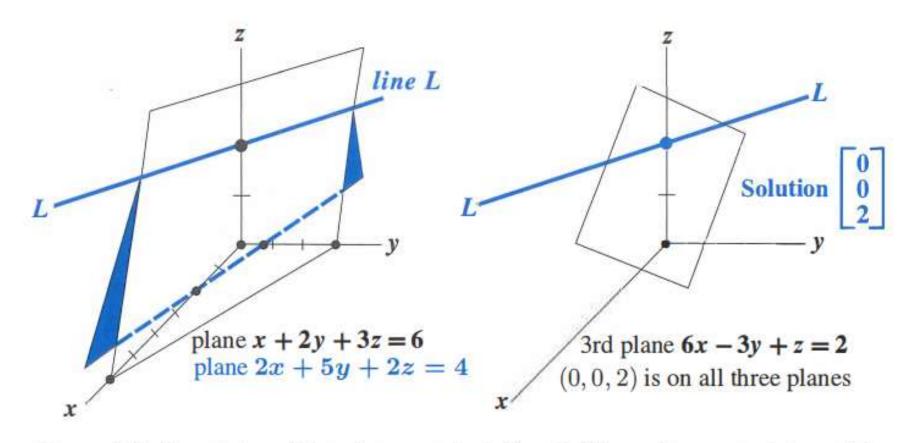


Figure 2.3: Row picture: Two planes meet at a line L. Three planes meet at a point.

列视角以向量形式展示方程组 Ax = b:

组合列
$$x \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} + y \begin{bmatrix} 2 \\ 5 \\ -3 \end{bmatrix} + z \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix} = b$$
 (4)

未知项是系数 x, y, z。 我们想用正确的数 x, y, z 乘以三个列向量,得到 $\mathbf{b} = (6, 4, 2)$ 。

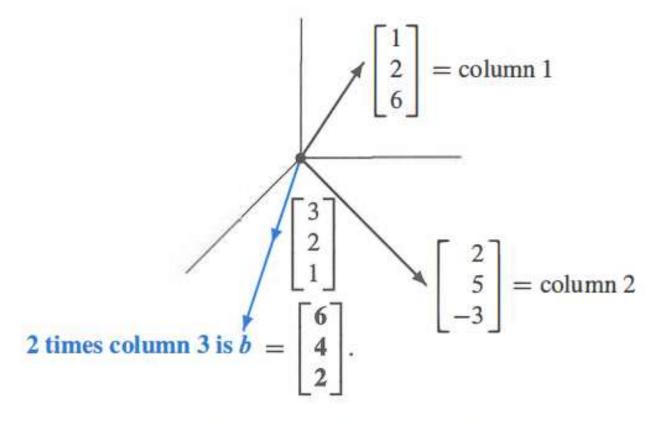


Figure 2.4: Column picture: Combine the columns with weights (x, y, z) = (0, 0, 2).

2.3 方程的矩阵形式

矩阵方程
$$Ax = b$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 2 \\ 6 & -3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}.$$
(5)

基本问题: A 乘以x 的意义?

行相乘 Ax 可由点积获得,记每一行与x 做点积:

$$A\boldsymbol{x} = \begin{bmatrix} (\ \mathbf{1} \ \mathbf{1}) \cdot \boldsymbol{x} \\ (\ \mathbf{1} \ \mathbf{2}) \cdot \boldsymbol{x} \\ (\ \mathbf{1} \ \mathbf{3}) \cdot \boldsymbol{x} \end{bmatrix}$$
(6)

列相乘 Ax 可由一组列向量做线性组合:

$$Ax = x(5 1) + y(5 2) + z(5 3).$$
 (7)

小结-1

- ① $A\mathbf{x} = \mathbf{b}$ 的列视角: A 的 n 个向量的某一个线性组合能生成 \mathbf{b} .
- ② $Ax = x_1a_1 + \cdots + x_na_n = b : a_1, a_2, \ldots, a_n$ 对应 A 的列.
- ③ 当 $\mathbf{b} = \mathbf{0}$, 一个显然的组合 $\mathbf{x} = (0, ..., 0)$ 生成零向量。
- **4** $A\mathbf{x} = \mathbf{b}$ 的行视角: m 个行定义的 m 个平面相较于 x.
- **⑤** 可使用点积表示这些平面: $(\mathbf{row} \ \mathbf{1}) \cdot \mathbf{x} = b_1, \dots, (\mathbf{row} \ \mathbf{m}) \cdot \mathbf{x} = b_m$.
- **⑤** 当 b = 0, 所有的平面将都会经过 x = (0, 0, ..., 0).

2.4 列空间

列空间是最重要的子空间,它直接关联到矩阵 A。我们尝试求解 A**x**= **b**。如果 A 不是可逆的,系统对于某些 **b** 是可解的,对于其他 **b** 是不可解的。那些可解的方程所对应的 **b** B 形成了 A 的"列空间"。

请记住,Ax 是 A 列的线性组合。为了获得每个可能的 b,我们使用每个可能的 x。从 A 列开始,取它们的所有线性组合。这会产生 A 的列空间它是由列向量组成的向量空间。

定义 列空间 C(A) 由列的所有线性组合组成。这些组合都是可能的向量Ax。它们填充列空间 C(A).

注意: C(A) 不仅仅包含 A 的列, 还包含其所有的线性组合.

要解 Ax = b,也就是要将 b 展开成 A 的列的线性组合。故右侧向量 b 必须是在列空间中,由左侧的 A 生成,否则没有解。

系统 Ax = b 可解当且仅当 b 在A 的列空间中

当 b 位于列空间中时,它是某一线性组合。该组合中的系数为我们提供了一个解。

假设 $A \in m$ 乘 n 矩阵。它的列有 m 个分量(不是 n)。所以列属于 \mathbf{R}^m 。 A 的列空间是 \mathbf{R}^m 的一个子空间,而不是 \mathbf{R}^n 的。所有列组合 $A\mathbf{x}$ 满足子空间的规则 (i) 和 (ii) 。当我们添加线性组合或乘以标量,仍然会生成列的组合。

例 4

$$A$$
x 等于 $\begin{bmatrix} 1 & 0 \\ 4 & 3 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 可表示为 $x_1 \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}$.

两列的所有组合的列空间填满 \mathbf{R}^3 中的一个平面。我们绘制了一个特定的 \mathbf{b} (列的组合)。这个 $\mathbf{b} = A\mathbf{x}$ 位于平面上。

该平面的厚度为零,因此 \mathbb{R}^3 中的大多数右侧 \mathbb{B} 都不在在列空间中。即对于大多数 \mathbb{B} 在 2 个未知数中没有 3 个方程的解。

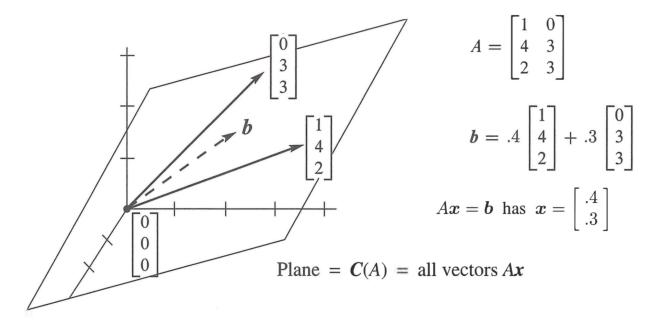


Figure 3.2: The column space C(A) is a plane containing the two columns. Ax = b is solvable when b is on that plane. Then b is a combination of the columns.

为代替 R^m 中的列,我们可以从向量空间 V 中的任何一组 S 向量开始。为得 到 V 的子空间 SS,我们取该集合中向量的所有组合:

S = V 中向量的集合 (无需是子空间)

SS = S中向量的所有线性组合 (必然是子空间)

 $SS = all c_1 \mathbf{v}_1 + \cdots + c_N \mathbf{v}_N =$ 子空间 SS 由 S "张成(span)"。

当 S 是列集合时,SS 是列空间。当只有一个非零向量 v 在 S 中,子空间 SS 是穿过 v 的线。始终 SS 是包含 S 的最小的子空间。这是创建子空间的基本方法。

强调: 列s"张成" 列空间.

子空间 SS 由 S 张成, 它包含 S 中所有向量的线性组合.

例 5 描述以下矩阵的列空间 (它们都是 \mathbb{R}^2 的子空间)

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
 and $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}$

解

单位矩阵 I 的列空间是全空间 R^2 。每个向量都是矩阵 I 的列的线性组合。在向量空间语言中,C(I) 是 R^2 。

A 的列空间仅为一条线。第二列 (2,4) 是第一列 (1,2) 的倍数。这些向量是不同的,我们关注的是向量张成的空间。即 A 的列空间包含 (1,2),(2,4) 以及沿着该行的所有其他向量 (c,2c)。方程式 $A\mathbf{x} = \mathbf{b}$ 只有当 \mathbf{b} 在这条线上时,才可解。

B 有三列,列空间 C(B) 全部为 R^2 。每个 b 都是可以实现的。例如,向量 b = (5, 4) 可以表示成第二列 + 第三列,可得 x 是 (0, 1, 1)。它又可以表示成 2 乘以第一列 + 第三列,可得另一个 x 是 (2, 0, 1)。 B 与 I 有相同的列空间,即所有 b 都是可实现的。但后者 x 需要有个额外的分量,并且有更多的解决方案和更多的组合来提供 b。

小结-2

- 1. A 的列空间包含 A 列的所有组合: \mathbf{R}^m 的子空间。
- 2. 列空间包含所有向量Ax。因此,当b位于C(A)中时,Ax = b是可解的。
- 3. Ax = b 有解仅当 b 落在 A 的列空间。

习题

自学 Worked Example 3.1B

Problem Set 2.1: 19, 29

Problem Set 3.1: 19, 23, 26, 32