

MIT 18.06 线性代数

第一讲：线性方程组的几何意义

整理自 MIT 18.06 课程及相关公开笔记

引言

线性代数的基本问题之一是求解 n 个未知数、 n 个方程组成的线性方程组。本讲的核心在于建立**方程组**、**矩阵**与**向量空间**之间的联系，我们将从行图像 (Row Picture)、列图像 (Column Picture) 和矩阵图像 (Matrix Picture) 三个维度来理解方程 $A\mathbf{x} = \mathbf{b}$ 。

1 线性方程组与矩阵的一般形式

由 n 个方程以及 n 个未知数构成的线性方程组的一般形式为：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

对应的矩阵形式（记为 $A\mathbf{x} = \mathbf{b}$ ）为：

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

其中 A 称为**系数矩阵**， $\mathbf{x}$ 为**未知数向量**， $\mathbf{b}$ 为**常数向量**。

2 二维情形的几何解释

我们通过具体方程组来建立直观理解：

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases} \implies \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

2.1 行图像 (Row Picture)

行图像是指：一次只考虑系数矩阵 A 中的一行（即一个方程），将该方程在二维坐标系中绘制成一条直线。所有直线的公共交点即为方程组的解。

对于本例，第一条直线过原点且斜率为 2，第二条直线斜率为 0.5 且截距为 1.5。两直线交点为 $(1, 2)$ ，代入原方程验证成立。

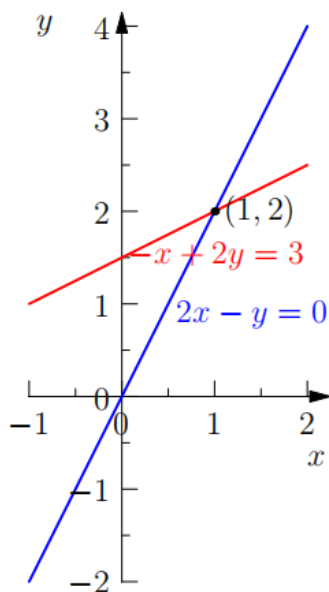


图 1: 二维行图像: 直线 $2x - y = 0$ 与 $-x + 2y = 3$ 相交于点 $(1, 2)$

2.2 列图像 (Column Picture)

列图像将矩阵 A 按列拆分为列向量, 将矩阵乘法理解为列向量的**线性组合**。原方程改写为:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

即寻找合适的系数 x, y , 使得两个列向量的组合恰好等于目标向量 $\mathbf{b}$ 。

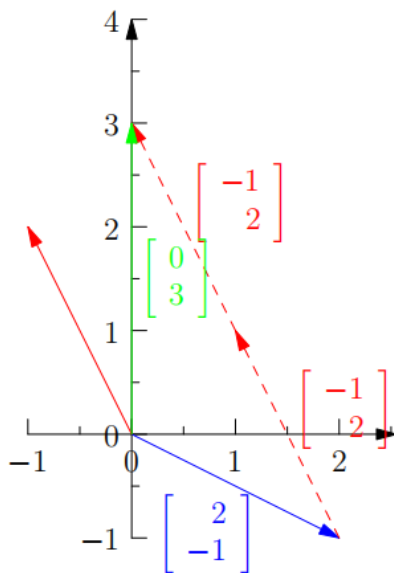


图 2: 二维列图像: 蓝色向量 $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ 与红色向量 $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ 的线性组合得到绿色向量 $\mathbf{b}$

从几何上看, 1 倍的蓝色向量加上 2 倍的红色向量 (首尾相接) 正好得到绿色向量 $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$, 因此解为 $x = 1, y = 2$ 。

2.3 矩阵乘法的两种核心计算视角

以矩阵 $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$ 与向量 $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 的乘法为例：

方法 1： **列视角（线性组合）**：将 $A\mathbf{x}$ 视为 A 的各列分别乘以 $\mathbf{x}$ 的对应分量后求和。

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 10 \\ 6 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$$

方法 2： **行视角（点积）**：将 $A\mathbf{x}$ 的结果向量中的每个分量视为 A 的对应行向量与 $\mathbf{x}$ 的点积。

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \cdot 2 + 2 \cdot 5 \\ 1 \cdot 1 + 2 \cdot 3 \end{bmatrix} = \begin{bmatrix} 12 \\ 7 \end{bmatrix}$$

这两种视角贯穿整个线性代数课程，列视角更强调“组合”的思想，而行视角更便于数值计算。

3 高维情形的推广

3.1 三维行图像

对于三维方程组，每一个方程在三维空间中确定一个平面。三个平面的交点（可能是点、直线或平面）即为方程组的解；若无公共交点，则方程组无解。

考虑如下三维示例：

$$\begin{cases} 2x - y = 0 \\ -x + 2y - z = -1 \\ -3y + 4z = 4 \end{cases} \implies \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

在三维中，直接观察三个平面的交点已经比较困难。通常的思路是先联立其中两个平面得到一条交线，再求这条直线与第三个平面的交点。对于四维及更高维度，行图像完全无法直观绘制，因此行图像在实用性和可推广性上受到极大限制。

3.2 三维列图像

将上述三维方程的系数矩阵按列拆分：

$$x \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} + z \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

列图像将问题转化为：在三维空间中给定三个列向量，寻找一组系数 x, y, z ，使它们的线性组合等于右侧向量 $\mathbf{b}$ 。

对于这个特定的例子，观察可知 $\mathbf{b}$ 恰好等于第三个列向量 $\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$ ，因此解为 $x = 0, y = 0, z = 1$ 。

这在行图像中并不直观，但通过列图像的组合视角一眼便能看出。

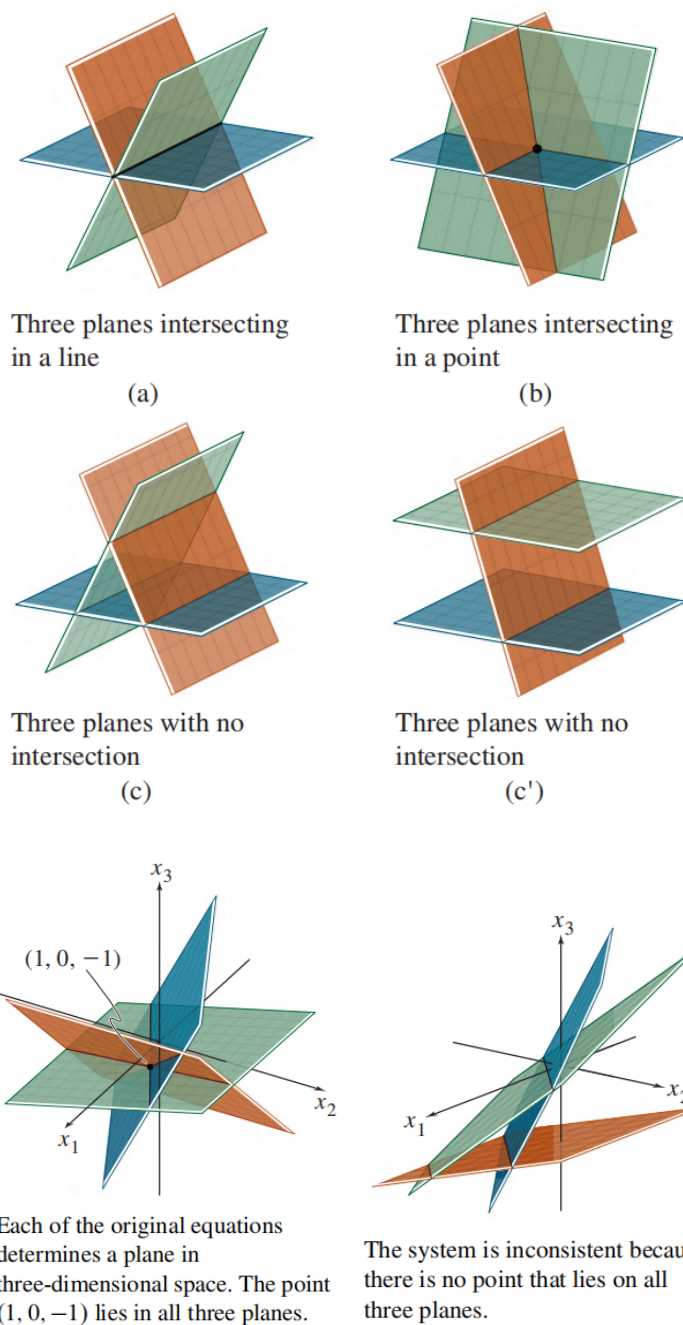


图 3: 三维行图像：三个平面相交于一点

3.3 n 维情形的抽象理解

- **行图像**：每个方程在 n 维空间中定义一个 $n-1$ 维的**超平面**。方程组的解就是这些超平面之间的交集（可能是点、线、面等）或无解（例如 n 个“平行”但不相交的超平面）。
- **列图像**：将 A 的 n 个列向量视为 n 维空间中的向量，寻找一组系数使得它们的线性组合等于 b 。

对于高维情况，列图像的优势极为突出：它避开了高维几何绘图的困难，将问题转化为纯粹的向量组合问题，便于用线性代数的工具（如线性无关、秩、列空间等）进行系统分析。

4 线性组合的覆盖范围与解的存在性

一个核心问题是：对于任意给定的 $\mathbf{b}$ ，方程 $A\mathbf{x} = \mathbf{b}$ 是否一定有解？

从列图像的角度来看，这个问题等价于：矩阵 A 的所有列向量的线性组合能否覆盖整个 n 维空间？

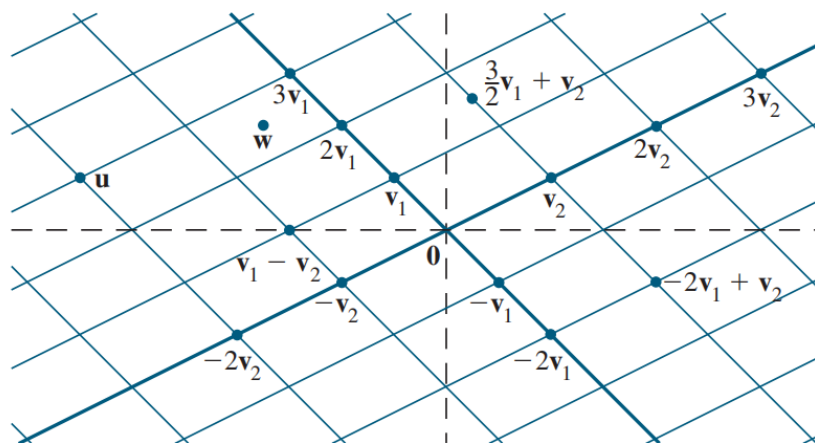


FIGURE 8 Linear combinations of $\mathbf{v}_1$ and $\mathbf{v}_2$.

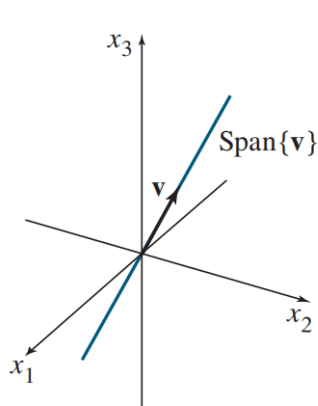


FIGURE 10 $\text{Span}\{\mathbf{v}\}$ as a line through the origin.

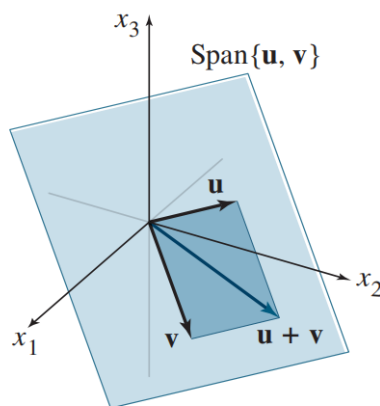


FIGURE 11 $\text{Span}\{\mathbf{u}, \mathbf{v}\}$ as a plane through the origin.

图 4: 向量张成的空间

4.1 非奇异矩阵（可逆矩阵）

如果矩阵 A 的列向量是**线性无关**的，即没有任何一个列向量可以由其他列向量线性表示，那么这些列向量的所有线性组合可以铺满整个 n 维空间。此时对于**任意**的 $\mathbf{b}$ ，方程组都有唯一解。这样的矩阵称为**非奇异矩阵**（可逆矩阵）。

4.2 奇异矩阵（不可逆矩阵）

如果矩阵 A 的列向量是**线性相关**的，例如在三维空间中三个列向量恰好共面（即其中一个向量是另外两个向量的线性组合），那么这些列向量的所有线性组合只能落在这个平面上，无法覆盖整个三维空间。此时，若 $\mathbf{b}$ 恰好落在这个平面内，则方程组有无穷多解；若 $\mathbf{b}$ 在这个平面外，则方程组无解。这样的矩阵称为**奇异矩阵**（不可逆矩阵）。

4.3 改变 $\mathbf{b}$ 对两种图像的影响

这是体现列图像优越性的一个关键点：

- 若采用**行图像**，当我们改变 $\mathbf{b}$ 时，方程组中每个方程的常数项都发生了变化，这意味着所有直线（或平面、超平面）的位置都发生了移动，必须重新绘制全部图形才能找到新的交点。
- 若采用**列图像**，改变 $\mathbf{b}$ 仅仅意味着目标向量的终点发生了变化，而矩阵 A 的列向量（即参与组合的基向量）**完全没有改变**，我们只需要为新的 $\mathbf{b}$ 寻找一组新的组合系数即可。

总结

- 线性方程组 $A\mathbf{x} = \mathbf{b}$ 具有行、列、矩阵三种等价的几何解释。
- 行图像直观但受限于低维空间，在高维时难以操作且改变 $\mathbf{b}$ 时需要完全重绘。
- 列图像将问题统一为“寻找列向量的线性组合”，引入了**线性组合**、**列空间**和**线性无关**等核心概念，这些概念是贯穿整个线性代数课程的主线。
- 矩阵是否可逆（奇异与否）直接决定了方程组是否对任意 $\mathbf{b}$ 都有解，这本质上是列向量能否张成整个空间的问题。