第一章　线性方程组

2021 年 11 月 14 日

前言

- 教材

  D.C. Lay, 《线性代数及其应用 (第 4/5 版)》，机械工业出版社.

- 参考阅读

  1. Sheldon Axler, 《线性代数应该这样学》, 人民邮电出版社;
  2. 柯朗 & 罗宾, 《什么是数学–对思想和方法的基本研究》，复旦大学出版社;
  3. E.T. 贝尔, 《数学精英》，商务印书馆;
  4. 伽莫夫, 《从一到无穷大: 科学中的事实与臆测》，科学出版社.

- 教学团队

Barbara Dietz	dietz@lzu.edu.cn
徐洪亚	xuhongya@lzu.edu.cn
赵继泽	zhaojz@lzu.edu.cn
陆汉涛	luht@lzu.edu.cn

- " 成功的学习是批判性地在个体头脑中重建知识的过程"

章节及大班计划学时

- 第一章　线性方程组 (6 学时)
- 第二章　矩阵代数 (6 学时)
- 第三章　行列式 (4 学时)
- 第四章　向量空间初步 (4 学时)
- 第五章　特征值与特征向量 (4 学时)
- 第六章　内积空间 (4 学时)
- 第七章　对称矩阵与二次型 (2 学时)

内容

- 什么是求解线性方程组

  把待解方程组用一系列更容易解的等价方程组代替，直到得出最后结果.

- 解的情况分类

$$
\begin{array}{ccc}
\text{无解} & \longrightarrow & \text{不相容} \\
\left.\begin{array}{c}\text{有唯一解} \\ \text{无穷多解}\end{array}\right\} & \longrightarrow & \text{相容}
\end{array}
$$

- 方程组的矩阵记号

  系数矩阵与增广矩阵

- 等价方程组的得到: 初等行变换

  包括倍加, 对换, 和倍乘

- 求解过程
  1. 消元 (顺序), 把增广矩阵化为阶梯形矩阵;
  2. 判断解的情况;
  3. 回代 (逆序, 从下往上), 把阶梯形矩阵变为简化阶梯形矩阵;
  4. 写出答案.

内容

- 阶梯形矩阵的定义

  对先导元素位置的约束

- 主元, 主元列

给定矩阵化为任何一个阶梯形矩阵时, 其主元的位置是固定的.

$$\Downarrow$$

定理 1: 简化阶梯形矩阵的唯一性

- 基本变量的个数 $=$ 主元的数目
- 方程组中的基本变量与自由变量的选择

  选主元列对应的变量为基本变量, 其它变量作为自由变量

  (注: *基本变量与自由变量的选择不是绝对的.*)
- 解集的参数表示

  把基本变量用自由变量表示出来

解的存在性与唯一性定理

把增广矩阵化简成阶梯形矩阵之后, 观察:

- 有解 $\Longleftrightarrow$ 增广矩阵的最右列不是主元列

  增广矩阵的阶梯形不会出现形如 $[0 \cdots 0\ b]$, $b \neq 0$ 的行

- 在有解的情况下:
  - 唯一解 $\Longleftrightarrow$ 没有自由变量

    基本变量数目 (主元数目) $=$ 变量个数
  - 无穷多解 $\Longleftrightarrow$ 有自由变量

    基本变量数目 (主元数目) $<$ 变量个数

内容

内容

- 向量

  由 n 个数组成的有序数组；只有一列的矩阵

  记号:

$$\mathbf{u} = \vec{u} = \begin{bmatrix} a \\ b \end{bmatrix} = (a, b), \quad \vec{u} \in \mathbb{R}^2$$

  矩阵 $A_{m\times n}$ 可以视为由 n 个 m 维列向量排列而成, 即: $A = [\vec{a}_1 \ \vec{a}_2 \ \cdots \ \vec{a}_n]$.

- 向量的运算: 加法和数乘 (线性运算)

- $\mathbb{R}^n$ 中向量的代数性质 (即线性空间的代数性质, 或定义)

$\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, c, d 是数

- $\vec{u} + \vec{v} = \vec{v} + \vec{u}$,
- $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$,
- $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$,
- $\vec{u} + (-\vec{u}) = -\vec{u} + \vec{u} = \vec{0}$,
- $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$,
- $(c + d)\vec{u} = c\vec{u} + d\vec{u}$,
- $c(d\vec{u}) = (cd)\vec{u}$,
- $1\vec{u} = \vec{u}$.

凡是在其上引进了具有上述性质的加法和数乘两种运算的集合, 并且集合关于这两种运算封闭, 则该集合就称之为一个线性空间. 线性空间中的元素称为向量.

- 线性组合

$$\vec{y} = c_1\vec{v}_1 + \cdots + c_p\vec{v}_p$$

- 生成子集 (子空间): $\mathrm{Span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$

  由 $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p$ 的所有线性组合所形成的集合

- 向量方程与线性方程组

$$\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n & \vec{b}\end{bmatrix} = [A\ \vec{b}]$$

$$\Updownarrow$$

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}$$

方程组有解的充要条件的另一种陈述

$$\text{方程组 } [A\ \vec{b}] \text{ 有解}$$

$$\Updownarrow$$

$$\boxed{\text{向量 } \vec{b} \text{ 可以表示成系数矩阵 } A \text{ 的列向量的线性组合}}$$

或者说 $\vec{b}$ 可被这些向量线性表出, 即

$$\vec{b} \in \mathrm{Span}\{\vec{a}_1, \vec{a}_2, \cdots, \vec{a}_n\}$$

内容

- 矩阵-向量积的定义

$$A\vec{x} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} := x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}.$$

或者写成分量形式

$$\left[A\vec{x}\right]_i = a_{ij}x_j$$

- 矩阵-向量积 $A\vec{x}$ 的性质 (线性性质)

$$A\left(\vec{u}+\vec{v}\right) = A\vec{u} + A\vec{v}, \qquad A\left(c\vec{u}\right) = c\left(A\vec{u}\right)$$

- 因此可以把矩阵 $A_{m\times n}$ 看作定义了一个从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的一个线性映射, 即 $A : \mathbb{R}^n \to \mathbb{R}^m$.

关于矩阵方程的解读

- 求解 $A\vec{x} = \vec{b} \quad (A \in \mathbb{R}^{m\times n})$ 即为寻找像 $\vec{b} \in \mathbb{R}^m$ 在 $\mathbb{R}^n$ 中的原像.
- $A\vec{x} = \vec{b}$ 有解等价于 $\vec{b}$ 关于映射 A 有原像.

关于线性方程组的三种等价陈述

- 线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2, \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m. \end{cases}$$

- 向量方程

$$\vec{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad j = 1, \dots, n; \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}; \quad x_1\vec{a}_1 + x_2\vec{a}_2 + \cdots + x_n\vec{a}_n = \vec{b}.$$

- 矩阵方程

  利用矩阵-向量积的定义：$A\vec{x} = \vec{b}$, A 是系数矩阵.

内容

- 齐次线性方程组 $A\vec{x}=\vec{0}$

  解集一定是某个子空间 $\mathrm{Span}\{\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_p\}$. 记为 $\mathrm{Nul}\,A$ (或 $\mathrm{Ker}\,A$).

  所需要的独立向量 $\vec{v}_i$ 的数目 $=$ 自由变量的数目.

- 非齐次线性方程组 $A\vec{x}=\vec{b}$, $\vec{b}\neq\vec{0}$

  解集 $\{\vec{w}\}$, 通式 $\vec{w}=\vec{p}+\vec{v}$; 或者写成 $\vec{x}=\mathrm{Nul}\,A+\vec{p}$.

  - $\vec{v}$ 是相应的齐次方程组 $A\vec{x}=\vec{0}$ 的解.
  - $\vec{p}$ 称为特解, 只需满足 $A\vec{p}=\vec{b}$ 即可;
  - $\vec{p}$ 的取法不是唯一的, 但是一定有 $\vec{p}_1-\vec{p}_2\in\mathrm{Nul}\,A$.

内容

向量组的线性相关与线性无关

定义

向量组 $S=\{\vec{v}_1,\vec{v}_2,\cdots,\vec{v}_p\}$ 如果能把零向量 $\vec{0}$ 非平凡表出,

$$\boxed{\text{存在不全为零的系数 } \{c_i\},\text{ 使得 } c_1\vec{v}_1+c_2\vec{v}_2+\cdots+c_p\vec{v}_p=\vec{0}}$$

则我们称该向量组是线性相关的；反之则是线性无关的.

推论

向量组 S 线性相关 $\Longleftrightarrow S$ 中至少有一个向量可被其它向量线性表出.

例

- 包含 $\vec{0}$ 的向量组必定线性相关
- 只有一个向量的集合 $\{\vec{v}\}$
  $\vec{v} \neq 0 \Longleftrightarrow$ 线性无关
- 包含两个向量的集合 $\{\vec{v}_1, \vec{v}_2\}$
  $\vec{v}_1$ 与 $\vec{v}_2$ 成比例 $\Longleftrightarrow$ 线性相关
- 包含三个向量的集合 $\{\vec{u}, \vec{v}, \vec{w}\}$
  如果 $\vec{u}$, $\vec{v}$ 线性无关, 而 $\{\vec{u}, \vec{v}, \vec{w}\}$ 线性相关 $\Longleftrightarrow$ $\vec{w} \in \mathrm{Span}\{\vec{u}, \vec{v}\}$

例

- 包含 $\vec{0}$ 的向量组必定线性相关
- 只有一个向量的集合 $\{\vec{v}\}$
  $\vec{v} \neq 0 \iff$ 线性无关
- 包含两个向量的集合 $\{\vec{v}_1, \vec{v}_2\}$
  $\vec{v}_1$ 与 $\vec{v}_2$ 成比例 $\iff$ 线性相关
- 包含三个向量的集合 $\{\vec{u}, \vec{v}, \vec{w}\}$
  如果 $\vec{u}$, $\vec{v}$ 线性无关, 而 $\{\vec{u}, \vec{v}, \vec{w}\}$ 线性相关 $\iff$ $\vec{w} \in \mathrm{Span}\{\vec{u}, \vec{v}\}$

定理 (定理 7)

若向量组 S 线性相关, 则 S 中 $\exists\ \vec{v}_j\,(j \geq 1)$, *s.t.* $\vec{v}_j$ 能被前面的 $j-1$ 个向量线性表出.

矩阵各列的线性无关

矩阵 A 各列向量组成的向量组线性无关 $\Longleftrightarrow A\vec{x}=\vec{0}$ 只有零解.

定理 (定理 8)

$\mathbb{R}^n$ 中任意向量组 $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_p\}$, 当 $p > n$ 时必线性相关.

内容

变换

- 变换 (函数, 映射): 定义了从一个集合到另一个集合的对应关系
- 记号和术语:

$$\begin{array}{rccc} T: & A & \to & B \\ & x & \mapsto & T(x), \qquad x \in A,\ T(x) \in B \end{array}$$

A 称为定义域, B 是取值空间; $T(x)$ 称为 x 在 T 下的像. 像的集合 (可以记为 $T(A)$) 称为值域, $T(A) \subset B$.

线性变换

满足以下两个条件的变换称为线性变换:

1. $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$
2. $T(c\vec{u}) = cT(\vec{u})$

推论

若 $T: A \to B$ 是线性变换, A 和 B 是线性空间, 则

- $$T\left(\vec{0}_A\right) = \vec{0}_B, \qquad T(a\vec{u}+b\vec{v}) = aT(\vec{u}) + bT(\vec{v}).$$
  即线性变换保持向量的加法与数乘.
- 值域 $T(A)$ 是一个线性空间.
- $T^{-1}(\vec{0}_B)$, *i.e.*, B 中零向量在 T 下的原像 (记为 $\mathrm{Ker}\ T$) 也是一个线性空间.
- 两个线性变换的复合依然是线性变换.

矩阵变换

定义

矩阵 $A_{m\times n}$ 给出了从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的映射 (变换): $\vec{v}=A\vec{u}$, $\vec{u}\in\mathbb{R}^n$, $\vec{v}\in\mathbb{R}^m$.

推论

- 矩阵变换属于线性变换.
- 上面对应的线性变换 $T:\mathbb{R}^n\to\mathbb{R}^m$ 可以定义为 $T(\vec{u}):=A\vec{u}$.

事实上, 任意的从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的一个线性变换都可以用一个 $m\times n$ 的矩阵表示, 即有限维线性空间之间的线性变换都有其对应的矩阵表示.

线性变换的矩阵

问题: 对于一个给定的线性变换 $T: \mathbb{R}^n \to \mathbb{R}^m$, 如何找到一个矩阵 A, 使得对于 $\forall\, \vec{x} \in \mathbb{R}^n$, 有 $T(\vec{x}) = A\vec{x}$?

思路:

- 对于空间 $\mathbb{R}^n$ 引进一组单位向量 $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$, 其中

$$\vec{e}_1 := \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \vec{e}_2 := \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad \text{etc.}$$

则 $\forall \vec{x} \in \mathbb{R}^n$ 都可以写成

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n = [\vec{e}_1 \ \ \vec{e}_2 \ \ \cdots \ \ \vec{e}_n]\,\vec{x}$$

矩阵 $[\vec{e}_1 \ \ \vec{e}_2 \ \ \cdots \ \ \vec{e}_n]$ 构成了一个 $n \times n$ 的单位矩阵，记为 I_n. x_i 是向量 $\vec{x}$ 相对于基 $\{\vec{e}_1, \cdots, \vec{e}_n\}$ 的第 i 个坐标.

- 求出 A

$$\begin{aligned} T(\vec{x}) &= T[x_1\vec{e}_1 + \cdots + x_n\vec{e}_n] = x_1T(\vec{e}_1) + \cdots + x_nT(\vec{e}_n) \\ &= [T(\vec{e}_1) \ \ \cdots \ \ T(\vec{e}_n)] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\vec{x} \end{aligned}$$

$A_{m\times n}$ 称为线性变换 T 在基 $\{\vec{e}_1, \cdots, \vec{e}_n\}$ 下的矩阵.

存在性问题 (解的存在性)

考虑线性变换 $T: \mathbb{R}^n \to \mathbb{R}^m$:

- 存在性问题:

  $\forall\, \vec{b} \in \mathbb{R}^m$, ? $\exists\, \vec{x} \in \mathbb{R}^n$, s.t. $T(\vec{x}) = \vec{b}$. 即 T 是不是把 $\mathbb{R}^n$ 映上 (onto)到 $\mathbb{R}^m$ (满射).

  等价陈述: 对应的 $A\vec{x} = \vec{b}$ 是不是总是有解的, 其中 $A = [T(\vec{e}_1) \ \cdots \ T(\vec{e}_n)]$.

@ 回答: 其充要条件是矩阵 A 的各列能够生成 $\mathbb{R}^m$, 即 $\mathrm{Span}\{T(\vec{e}_1), \cdots, T(\vec{e}_n)\} = \mathbb{R}^m$; 等价于要求 A 的每一行都有主元.

例

对于 $T: \mathbb{R}^n \to \mathbb{R}^m$, 若 $n < m$, 则 T 肯定不是满射.

考虑线性变换 $T\colon \mathbb{R}^n \to \mathbb{R}^m$:

- 唯一性问题

  $\forall\, \vec{b} \in \mathbb{R}^m$, $T(\vec{x}) = \vec{b}$ 至多有唯一解 (单射).

  等价陈述: 对应的 $A\vec{x} = \vec{b}$ 是不是至多有唯一解.

@ 回答: 其充要条件是 $A\vec{x} = \vec{0}$ 只有零解, 即 $\mathrm{Nul}\, A = \{\vec{0}\}$ (或者是 $\mathrm{Ker}\, T = \{\vec{0}\}$). 要求 A 的各列是线性无关的; 等价于要求 A 的每一列都有主元.

例

对于 $T\colon \mathbb{R}^n \to \mathbb{R}^m$, 若 $n > m$, 则 T 肯定不是单射.

可逆变换

定义

我们称一个变换 $T: A \to B$ 是可逆的, 若当存在变换 $G: B \to A$, 能够使得 $GT = \mathbf{1}_A$ 以及 $TG = \mathbf{1}_B$.

注: $\mathbf{1}_A$ 和 $\mathbf{1}_B$ 分别是集合 A 和 B 上的恒等变换.

- 变换 T 可逆 $\iff$ T 是双射, 即 T 既是满射又是单射.
- T 是有限维空间上的线性变换, $T(\vec{x}) = A\vec{x}$, 则 T 可逆当且仅当矩阵 A 等价于单位矩阵.

例

有 $A_{n\times n}$, A 可逆 (双射) $\iff$ $\mathrm{Nul}\, A = \{\vec{0}\}$.

维数公式

定理

$A_{m\times n}$ 是从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的一个线性映射，有

$$\dim \operatorname{Nul} A + \dim \operatorname{Range} A = n.$$

本章作业

- 1.1 节: 1, 3, 13, 19, 25
- 1.2 节: 1, 9, 13, 17, 23
- 1.3 节: 5, 11, 15, 21, 29
- 1.4 节: 1, 3, 7, 9, 17, 21, 33, 39
- 1.5 节: 11, 17, 27, 29, 37
- 1.6 节: 5, 13
- 1.7 节: 7, 11, 27, 29, 33, 37, 39
- 1.8 节: 1, 3, 7, 9, 19, 23, 27, 31, 39
- 1.9 节: 1, 5, 15, 17, 33, 39
- 1.10 节: 1

课堂练习

- 1.1 节: 11, 17, 27
- 1.2 节: 3, 7, 19
- 1.3 节: 1, 13, 33
- 1.4 节: 5, 13, 29, 37
- 1.5 节: 5, 25, 33
- 1.6 节: 7
- 1.7 节: 17, 35
- 1.8 节: 5, 17, 25, 27, 30, 37
- 1.9 节: 7, 21, 37