Linear Algebra

Bowen Liu, Chenglong Yu

目录

I 线性方程组 **2**

1 引言 **2**

2 列向量空间与线性函数 **3**

2.1 列向量空间以及加法，数乘运算 3

3 高斯消元法和最简阶梯型 **8**

4 线性方程组解的结构 **12**

4.1 选讲：最简阶梯型唯一性和舒伯特胞腔 13

Part I

线性方程组

1 引言

PageRank 是 Google 最早的用来计算网页连接的算法，我们考虑一个简单的模型：例如我们有 $1,2,3,4$ 四处网址，相互的超链接情况如下：

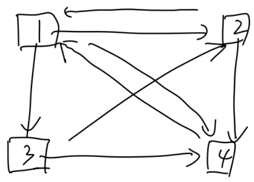

图 1: 简答的网页连接模型

其中 x_i 分别表示第 i 处网页的重要度。对于 1 号网页来说，它的重要度（流量）分成了三份，分别给了 $2,3,4$：

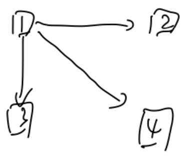

图 2: 1 号网页的流量流出

另一方面，有 $2,4$ 号网页的流量流入 1 号网页：

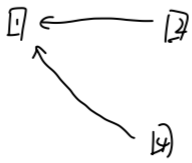

图 3: 1 号网页的流量流入

为了简单起见，我们认为这些流出的流量是被平均分配的，因此根据相互超链

接的关系得到如下的一些等式:

$$\begin{cases} x_1 = \frac{1}{2}x_1 + x_4 \\ x_2 = \frac{1}{3}x_1 + \frac{1}{2}x_3 \\ x_3 = \frac{1}{3}x_1 \\ x_4 = \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3. \end{cases}$$

通过消元法求解,我们可以得到所有的解形如

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = t \cdot \begin{pmatrix} \frac{2}{3} \\ \frac{4}{3} \\ \frac{4}{3} \\ 1 \end{pmatrix},$$

其中 t 为任意实数。

对于一般的网络连接形式,我们可以问以下问题:

1. 是否一定存在非零解?
2. 解是否都是形如如上的形式? 是否各个分量的比例相同,而且可以同为正?
3. 如何求解,以及哪种算法在实际中效率更高?

我们这门课的主要问题:求解线性方程组,了解解的结构。

2 列向量空间与线性函数

接下来,我们解释三个方面的问题

1. 在什么集合(空间)上求解?
2. 什么是线性方程?
3. 如何求解?

2.1 列向量空间以及加法,数乘运算

我们记 $\mathbb{R}$ 为全体实数的集合。在例子中我们有

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

其中 x_i 都是实数, 这样的 x 被称为一个长度为 4 的实数组. 如果我们考虑长度为 2 的实数组全体构成的集合

$$\{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid x_1, x_2 \in \mathbb{R}\},$$

这个集合记做 $\mathbb{R}^2$, 并且这个集合中的元素一一对应于平面上的点. 同时, $\mathbb{R}^2$ 中的点 P 也可以一一对应与从原点到这个点的向量 $\vec{OP}$, 如下图所示:

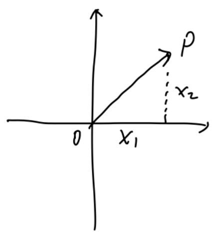

图 4: 点与向量的一一对应

类似的, 我们可以认为 $\mathbb{R}^3$ 一一对应于三维空间中的点.

定义 2.1.1 (列向量空间)**. 列向量空间** *(column vector space)*定义为如下的集合

$$\mathbb{R}^n = \{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \mid x_i \in \mathbb{R}\},$$

其中的元素称为**列向量** *(column vector)*

定义 2.1.2 (加法和数乘)**.** $\mathbb{R}^n$ 上可以定义如下两种运算*:*

(1) **加法** *(addition):*

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

(2) **数乘** *(scalar product):* 任取实数 $c \in \mathbb{R}$, 定义

$$c\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} cx_1 \\ cx_2 \\ cx_3 \\ \vdots \\ cx_n \end{pmatrix}.$$

注 2.1.1. 向量的加法和数乘有其对应的几何意义, 对于 $\mathbb{R}^2$ 来说, 其中向量的加法和数乘的几何意义可以通过如下具体图形象的展示:

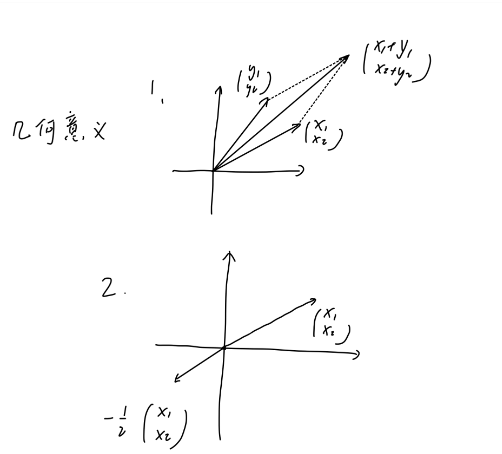

图 5: 点与向量的一一对应

接下来，我们考虑线形函数:

定义 2.1.3. 对于 $\mathbb{R}^n$ 上的函数 $F\colon \mathbb{R}^n \to \mathbb{R}$, 如果存在 $a_1, \ldots, a_n \in \mathbb{R}$ 是常数, 使得 F 有如下表达式

$$F(x) = a_1x_1 + \cdots + a_nx_n$$

那么称 F 是 $\mathbb{R}^n$ 上的**线性函数** *(linear function).*

例 2.1.1. 如下的 F 是 $\mathbb{R}^2$ 上的线性函数*:*

$$\begin{aligned} &F\colon \mathbb{R}^2 \to \mathbb{R} \\ &\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto (x_1+1)^2 - (x_1-1)^2 + (x_2-1)^2 - (x_1+1)^2 \end{aligned}$$

定理 2.1.1 (线性函数的运算性质等价描述)**.** 函数 $F\colon \mathbb{R}^n \to \mathbb{R}$ 是线性函数当且仅当 F 满足*:*

(1) 对任意 $x, y \in \mathbb{R}^n$, 有

$$F(x+y) = F(x) + F(y)$$

(2) 对任意 $c \in \mathbb{R}, x \in \mathbb{R}^n$, 有

$$F(cx) = cF(x)$$

证明. 如果 F 是线性函数, 可以直接验证 F 满足 (1), (2) 两条性质; 另一方面, 如果 F 满足 (1), (2), 我们记

$$a_1 = F(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}), \quad a_2 = F(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}), \quad \ldots$$

则

$$F(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) = a_1 x_1, \quad F(\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}) = a_2 x_2, \quad \ldots$$

那么任取 $v_1, \ldots, v_m \in \mathbb{R}^n$, 则

$$\begin{aligned} F(v_1 + v_2 + \cdots + v_m) &= F((v_1 + \cdots + v_{m-1}) + v_m) \\ &= F(v_1 + \cdots + v_{m-1}) + F(v_m) \\ &= F(v_1) + \cdots + F(v_m) \end{aligned}$$

因此,

$$F(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}) = F(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}) = F(\begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) + \cdots + F(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix}) = a_1 x_1 + \cdots + a_n x_n$$

□

命题 2.1.1. 如果 F 是线性函数, 那么 $F(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}) = 0$.

证明. 任意取 $0 \neq c \in \mathbb{R}$, 根据定义则有

$$F(c\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = cF(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}),$$

因此 $F(\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}) = 0$. □

例 2.1.2. *如下的* F_1, F_2 *不是* $\mathbb{R}^2$ *上的线性函数:*

$$F_1 \colon \mathbb{R}^2 \to \mathbb{R}$$
$$\begin{pmatrix}x_1\\x_2\end{pmatrix} \mapsto x_1 + x_4 + 4,$$

$$F_2 \colon \mathbb{R}^2 \to \mathbb{R}$$
$$\begin{pmatrix}x_1\\x_2\end{pmatrix} \mapsto x_1^2 + x_2^2$$

对于函数 F_1, 通过直接的计算得到 $F_1(\begin{pmatrix}0\\0\end{pmatrix}) \neq 0$, 从而利用命题*2.1.1*可知 F_1 不是线性函数. 对于函数 F_2, 我们可以发现

$$F_2(\begin{pmatrix}1\\0\end{pmatrix}) = 1, \quad F_2(2\begin{pmatrix}1\\0\end{pmatrix}) = 4 \neq 2,$$

从而根据线性函数的定义中第二条可知 F_2 不是线性函数.

定义 2.1.4. $\mathbb{R}^n$ 上 m 个线性函数 $F_1, F_2, \ldots, F_m$ 和 m 个实数 $b_1, b_2, \ldots, b_m$ 满足的方程组

$$\begin{cases} F_1(x) = b_1 \\ F_2(x) = b_2 \\ \vdots \\ F_m(x) = b_m \end{cases}$$

称为 n 个变元的**线性方程组** *(system of linear equations)*, 带入方程组使得其成立的 x 称为**线性方程组的解** *(solution of system of linear equations)*

注 2.1.2. 我们可以给线性方程组如下的一些几何解释:

(1) 在 $\mathbb{R}^2$ 中, 单个线性函数 $F_1(x) = a_1x_1 + a_2x_2$ 以及实数 b_1 给出的线性方程组 $a_1x_1 + a_2x_2 = b_1$ 的解是 $\mathbb{R}^2$ 中的一条直线.

(2) 在 $\mathbb{R}^2$ 中, 根据 *(1)* 的几何解释不难理解如下线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

的解是 $\mathbb{R}^2$ 中两条直线的交点. 注意, 在 $\mathbb{R}^2$ 中两条直线不一定相交, 即如上线性方程组不一定有有解, 但是如果有解一定只有唯一解.

(3) 在 $\mathbb{R}^3$ 中, 如下线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2 \end{cases}$$

的解可以看成是 $\mathbb{R}^3$ 中两个平面的交线. 注意: 在 $\mathbb{R}^3$ 中两个平面不一定相交, 即如上线性方程组不一定有有解, 并且如果相交, 也是交出一条线, 即此时解不唯一.

(4) 在更高维中也有同样的解释: 由一个线性函数给出的线性方程的解可以看成是一个低一维的超平面, 而多个线性函数给出的线性方程组的解则是这些超平面的交.

3 高斯消元法和最简阶梯型

根据注记2.1.2可知对于一个线性方程组其可能没有解, 并且即使有解也不一定只有唯一解, 那么该如何求解线性方程组呢? 在本节中我们将利用高斯消元法, 来求解一般的线性方程组. 我们先来看下面的一个简单的例子.

例 3.0.1.

$$\begin{cases} 4x_2 - x_3 = 7 & (r_1) \\ x_1 + 2x_2 = 5 & (r_2) \\ 2x_1 + x_3 = 3 & (r_3) \end{cases}$$

显然我们交换 r_1, r_2 不改变上述方程组的解, 因此我们得到:

$$\begin{cases} x_1 + 2x_2 = 5 & (r_1') \\ 4x_2 - x_3 = 7 & (r_2') \\ 2x_1 + x_3 = 3 & (r_3') \end{cases}$$

我们考虑如下操作: 保持 r_1', r_2' 不变, 用 r_3' 减去 $2r_1'$, 得到如下的方程组:

$$\begin{cases} x_1 + 2x_2 = 5 & (r_1'') \\ 4x_2 - x_3 = 7 & (r_2'') \\ -4x_2 + x_3 = -7 & (r_3'') \end{cases}$$

上述操作并不改变方程组的解, 因为可由 r_1'', r_2'', r_3'' 恢复出 r_1', r_2', r_3'. 类似的最后再保持 r_1'', r_2'' 不变, 用 r_3'' 加上 r_2'', 得到

$$\begin{cases} x_1 + 2x_2 = 5 \\ 4x_2 - x_3 = 7 \\ 0 = 0 \end{cases}$$

对于上述方程组我们可以用 x_3 来如下的表示 x_1, x_2, 其中 x_3 可以取任意的实数

$$x_1 = -\frac{1}{2}x_3 + \frac{3}{2}$$
$$x_2 = \frac{1}{4}x_3 + \frac{7}{4}$$

因此我们可以将方程组的解写作

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{4} \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{3}{2} \\ \frac{7}{4} \\ 0 \end{pmatrix}$$

注 3.0.1. 根据上述结果可以发现该线性方程组有无穷组解, 这对应于几何解释中 $\mathbb{R}^3$ 中三个平面相交出一条线.

回顾例3.0.1, 在解方程中我们主要用到了如下三种操作:

(E1) 交换方程组的某两行.

(E2) 将某一行乘以非零常数 c.

(E3) 将某一行的非零常数 c 倍加到另一行上.

我们称如上的三种操作为**基础行变换** (elementary row operations). 不难发现基础行变换均可逆, 并且其逆也是基础行变换.

定义 3.0.1. 有限个基础行变换的复合称为**行变换** *(row operations)*.

命题 3.0.1. *行变换均可逆, 并且其逆也为行变换.*

证明. 因为基础行变换可逆, 且其逆也为基础行变换, 并且操作 O_1O_2 的逆为 $O_2^{-1}O_1^{-1}$. □

推论 3.0.1. *行变换不改变线性方程组的解.*

由于作行变换只关注方程的系数以及右侧常数项, 因此对于如下的线性方程组

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

我们将其系数及常数项提出来记作

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

并将这个线性方程组记做 $Ax = b$, 这也引出了矩阵的概念.

定义 3.0.2. *由 $m \times n$ 个数 a_{ij} 排成的 m 行 n 列的 (实) 数表称为 m 行 n 列***矩阵** *(matrix), 记做 $(a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{R})$. 当 $m = n$ 时, $A \in M_{n\times n}(\mathbb{R})$ 被称为 n 阶***方阵** *(square matrix).*

例 3.0.2. $\mathrm{I}_n \in M_{n\times n}(\mathbb{R})$ *是只有 (i, i) 元为 1, 其余分量为零的矩阵, 称为***单位矩阵** *(identity matrix).*

对于线性方程组 $Ax = b$, A 称为**系数矩阵** (coefficient matrix), (A, b) 称为**增广矩阵** (augmented matrix), 并将上述方程记作 $Ax = b$. 现在我们即可以通过行变换来操作我们的增广矩阵, 使其最终的形式方便于我们求解, 那么究竟该操作到什么样子为止呢?

根据例3.0.1, 我们发现如果我们的增广矩阵有如下的形式, 线性方程组可以直接求解:

(1) 所有非零行在零行的上面.

(2) 对某一非零行, 称最左边的非零元为**主元** (pivot), 第 i 行的主元严格比第 $i + 1$ 行的主元靠左.

满足上述条件的矩阵称为**行阶梯型** (row echelon form), 并且如果主元所在列的其他元素均为零, 主元本身为 1, 则称此时为**最简行阶梯型** (reduced row echelon form).

定理 3.0.1. *矩阵 A 可通过行变换变成最简行阶梯型, 并且该最简行阶梯型不依赖于行变换的选取, 记作 $\operatorname{rref} A$.*

证明. 对 $m\times n$ 矩阵的列作归纳: 假设 $n=1$ 时, 对于 $m\times 1$ 矩阵

$$A=\begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}$$

如果 $a_{11}=\cdots=a_{m1}=0$, 则此时已经是最简行阶梯型. 若 $a_{11}=a_{21}=\cdots=a_{(k-1)1}=0, a_{k1}\neq 0$, 那么通过 (E1) 将 a_{k1} 换到第一行, 用 (E2) 将第一行乘以 $(a_{k1})^{-1}$ 使得主元变为 1, 再用 (E3) 将第一行以下变为零, 因此此时最简行阶梯型为

$$\operatorname{rref} A=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$$

并且不依赖于行列变换的选取. 假设对列数为 n 的时候成立, 对于 $m\times(n+1)$ 的矩阵 A, 将其写作 $A=(B,y)$, 其中 B 是 $m\times n$ 矩阵. 根据归纳假设 B 可由行变换得到最简行阶梯型, 记作 B', 将同样的变换作用在 A 上得到 $A'=(B',y')$. 如果 B' 没有非零行, 则次数 A' 已经是最简行阶梯型. 如果 B' 从 $k+1$ 行开始是零行, 则对

$$\begin{pmatrix}y'_{k+1}\\\vdots\\y'_m\end{pmatrix}$$

应用 $n=1$ 时的结论, 可做行变换得到最简行阶梯型, 同时也对 B' 作. 但由于行变换不改变零矩阵, 因此不改变 B', 得到的矩阵记作 A''. 考虑如下两种情况:

(1) 如果

$$\operatorname{rref}(\begin{pmatrix}y'_{k+1}\\\vdots\\y'_m\end{pmatrix})=\begin{pmatrix}0\\\vdots\\0\end{pmatrix}$$

则此时 A'' 已经是最简行阶梯型.

(2) 如果

$$\operatorname{rref}(\begin{pmatrix}y'_{k+1}\\\vdots\\y'_m\end{pmatrix})=\begin{pmatrix}1\\\vdots\\0\end{pmatrix}$$

则作 (E3) 将第 $k+1$ 行加到第 $1,2,\ldots,k$ 行, 将 $y'_1,\ldots,y'_k$ 变成零, 此时得到的矩阵也是最简行阶梯型. 由于上述操作只依赖于 B 和 y, 即 A 的最简行阶梯型只依赖于 A 本身.

$\square$

定义 3.0.3. 对于矩阵 A, $\mathrm{rref}\, A$ 的主元个数称为 A 的**秩** *(rank)*, 记作 $\mathrm{rank}\, A$.

注 3.0.2. 根据定义, 对于 $m \times n$ 矩阵, $\mathrm{rank}\, A \leq m$.

命题 3.0.2. 对于矩阵 A, B, 有 $\mathrm{rank}\, AB \leq \min\{\mathrm{rank}\, A, \mathrm{rank}\, B\}$

4 线性方程组解的结构

定义 4.0.1. 对于线性方程组的系数矩阵 A, $\mathrm{rref}\, A$ 中主元对应的未知元称为**主元** *(principal unknowns)*, 其余未知元称为**自由元** *(free unknowns)*.

例 4.0.1. 例如线性方程组 $Ax = b$, 其中

$$A = \begin{pmatrix} 1 & 0 & 3 & 5 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{pmatrix}$$

则 x_1, x_2, x_5 是主元, x_3, x_4, x_6 是自由元. 并且根据上述最简行阶梯型, 我们可以直接分析出方程组的解的情况:

1. 如果 b_4 或者 b_5 不是零, 则方程组 $Ax = b$ 无解.
2. 如果 $b_4 = b_5 = 0$, 则 x_3, x_4, x_6 取定任意实数后, 主元由方程组唯一确定:

$$\begin{aligned} x_1 &= b_1 - 3x_3 - 5x_4 - x_6 \\ x_2 &= b_2 - 2x_3 - x_4 - 2x_6 \\ x_5 &= b_3 \end{aligned}$$

定理 4.0.1 (线性方程组解的结构定理)**.** 对于方程 $Ax = b$, 用行变换将 (A, b) 化作最简行阶梯型 $(\overline{A}, \overline{b})$, 则

(1) 方程有解等价于 $\overline{A}$ 的零行对应的 $\overline{b}_i$ 也是零.

(2) 方程有解时自由元可以任意取值, 且自由元的每一组取值都唯一决定了一组解. 特别地, 方程有唯一解当且仅当没有自由元.

推论 4.0.1. 线性方程组 $Ax = b$

(1) 有解当且仅当 $\mathrm{rank}\, A = \mathrm{rank}(A, b)$.

(2) 有唯一解当且仅当 $\operatorname{rank} A$ 等于 A 的列数相同.

定义 4.0.2. 方程 $Ax = 0$ 称为**齐次线性方程组** *(system of homogeneous linear equations)*.

定理 4.0.2. 齐次线性方程组 $Ax = 0$ 的解在加法和数乘下封闭.

证明. 注意到

$$A(x+y) = Ax + Ay$$
$$A(cx) = cAx$$

□

定理 4.0.3. 对于线性方程组 $Ax = b$, 如果 $\widetilde{x}$ 是其某一解 (特解), 则 $Ax = b$ 的所有解均可唯一的表达为 $x = y + \widetilde{x}$, 其中 y 是 $Ax = 0$ 的解.

证明. 只需验证如下两点:

(1) 验证 $y + \widetilde{x}$ 是解.

(2) 验证当 x 是解时, $x = (x - \widetilde{x}) + \widetilde{x}$, 其中 $x - \widetilde{x}$ 满足 $Ax = 0$.

□

注 4.0.1. 从几何上来看, 齐次线性方程组 $Ax = 0$ 的解构成了 $\mathbb{R}^n$ 中的一个对加法数乘封闭的子集, 之后我们会用更抽象的观点去描述这种子集, 并称其为一个子空间. 而 $Ax = b$ 的解相当于是将这个子空间做了平移.

4.1 选讲：最简阶梯型唯一性和舒伯特胞腔