# 线性方程组

## 1.概念

包含未知数的一个**线性方程**是形如

的方程，其中与系数是实数或复数，通常是已知数。

而**线性方程组**是由一个或多个包含相同变量的线性方程组成的。例如

线性方程组的一组解是一组数，用这组数分别代替时所有方程的两边相等。例如代入方程组变成等式和.

方程组所有可能的解的集合称为线性方程组的**解集**。若是两个线性方程组有相同的解集，则它们等价。

求包含两个未知数的两个方程组成的方程组的解，等价于求两条直线的交点，例如

这两个方程的图形都是直线，两条直线可能交于一点、不相交(平行)、重合，分别对应于一个解、无解、无穷多解的情况。

## 2.矩阵

### 2.1矩阵引入

一个线性方程组包含的主要信息可以用**矩阵**表示，给出方程组

把每一个变量的系数写在对齐的一列中，矩阵

称之为该方程组的**系数矩阵**，而

称之为**增广矩阵**。

### 2.2线性方程组的求解

基本的思路是把方程组用一个更容易解的等价方程组代替。

粗略地说，我们用第一个方程组中第一个方程中含的项消去其他方程中的含的项，然后用第二个方程中含的项消去其他方程中含的项，以此类推，最后得到一个很简单的等价方程组(阶梯状的方程组)。

例如

可以将第一个方程\*4加上第3个方程，从而得到新的方程3.如下所示

就可以得到一个新的方程组和矩阵。以此类推最终得到阶梯状的方程组为

就这样，我们可以很容易地知道原方程组的解为。

下面我们引入行初等变换：

1. (倍加变换) 把某一行换成它本身与另一行的倍数的和。

2. (对数变换) 把两行对换。

3. (倍乘变换) 把某一行的所有元素乘以同一个非零数。

若其中一个矩阵经过若干行初等变换变换为另一个矩阵，我们称这两个矩阵为行等价的。

### 2.3主元

对于一个阶梯型矩阵，我们定义**主元元素**与**主元列**。下面是几个概念：

先导元素：(非零行中)该行最左边的非零元素。

主元位置：阶梯型矩阵中先导元素的位置

主元列：含有先导元素的列

我们举个例子，如下所示

可以知道，主元列是第1、2、4列，主元为1、2、-5.

通过先前的知识可以知道，矩阵可以对应于一个线性方程组，例如：

→

对应于主元列的变量和称为**基本变量**，其他变量()称为**自由变量**。

## 3.向量方程

### 3.1 向量

仅含一列的矩阵称为**列向量**，或简称**向量**。

若是正整数，表示所有个实数数列(或有序元组)的集合，通常写成列矩阵的形式，如

所有元素都是零的向量称为零向量，用表示。

下面举个例子：

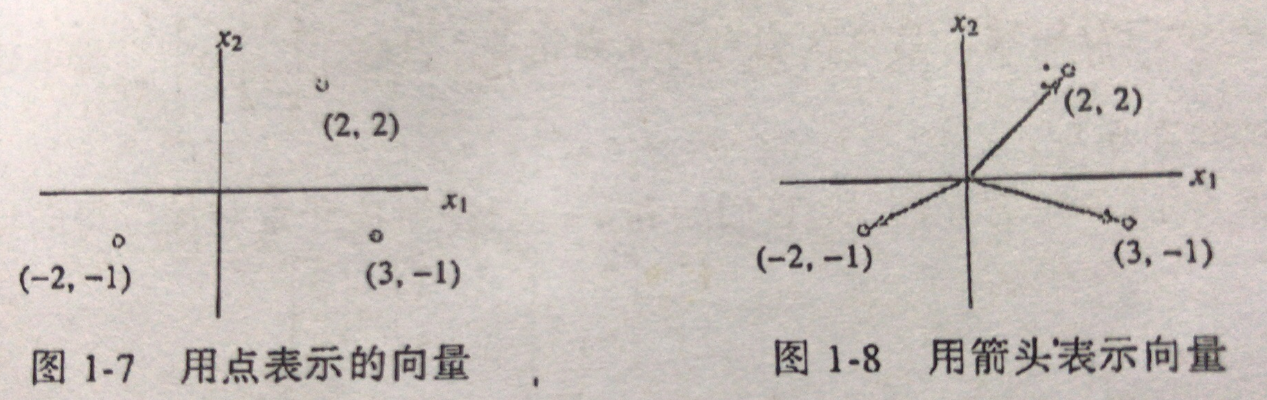
这里和是任意实数。所有的两个元素的向量的集记为，表示向量中的元素是实数，而指数2表示每个向量包含两个元素。

向量之间是可以做加减的，就是对应元素做加减得到新的向量；也支持和实数做标量乘法，就是实数与每个元素做乘法得到新的向量。

有时为了方便，我们会将列向量写成括号的形式，例如

→ (3,-1)

可以理解为是一个集合点对，看作是平面上所有点的集合，而向量的几何表示是一条由原点(0,0)指向点(3,-1)的有向线段，如图。



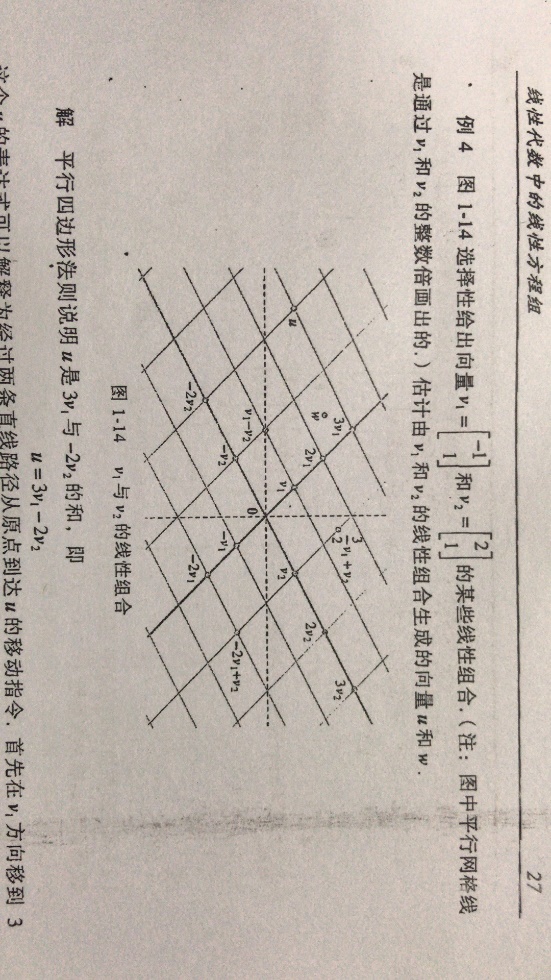
其他维度的也同理。

### 3.2线性组合

给定中向量和标量，向量

称为向量以为**权**的**线性组合**。

线性组合可以在几何上去理解，线性组合后的向量与其他向量一样，均是空间中的一个点。它的形成方式如下所示



也就是例如

可以当成是从原点开始沿着的方向走了三个单位，再沿着方向走了-2个单位。

线性代数的一个主要思想是研究可以表示为某一固定向量集合的线性组合的所有向量。若是中的向量，则的所有线性组合所称的集合用记号表示称为由所生成(或张成)的的子集，也就是说，是所有形如

的向量的集合，其中为标量。

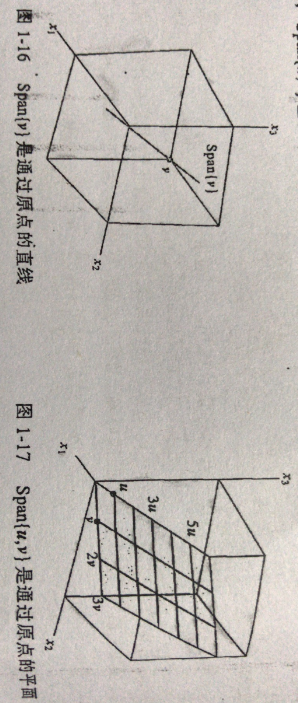
要判断向量是否属于，就是判断向量方程

是否有解，或等价地，判断增广矩阵的线性方程组是否有解。注意，一定包含零向量。

下面是与的几何解释。

设是中的向量，那么就是的所有数量倍数的集合，也就是通过和的直线上所有点的集合。

若和是的非零向量，不是的倍数，则是中通过、和的平面，可以结合我们之前的线性组合的几何理解上去理解。



## 4.矩阵方程

线性代数的一个基本思想是把向量的线性组合看作矩阵与向量的积。

若是矩阵，它的各列为。若是中向量，则与的积，记为，就是的各列**以中对应元素为权的线性组合**，即

注意维度要求，的列数必须要等于中的元素个数。例如

通过这个例子，相信能更好的理解矩阵相乘(即为什么要第一个矩阵的列数要等于第二个矩阵的行数)，还可以结合神经网络的前向传播来理解。

定理 设是矩阵，则下列命题是逻辑上等价的，也就是说，对某个，它们都成立或者都不成立。

1. 对中的每个，方程有解
2. 中的每个都是的列的一个线性组合
3. 的各列生成
4. 在每一列都有一个主元位置

其实要理解不难，a和d是等价的，每一列均有一个主元，即阶梯型的矩阵每一列均不

会出现 0=的情况，就可以从最后一行开始逐步往上求解，所以a也是成立的。

对于b，因为是一个可以到达最后一行的阶梯型的矩阵，所以自然是可以表示出任何

行向量。

### 4.1 齐次线性方程与线性相/无关

若它可以写成的形式，则称该线性方程组为**齐次线性方程组**。这样的方程组至少有一个解，即，这个解称为它的**平凡解**。我们往往更关注的是它的非平凡解。

我们将矩阵方程转换为向量方程，例如考虑方程

此方程当然有平凡解，即，不过我们更关心的是平凡解是否是唯一解。

我们定义，中的一组向量称为**线性无关**的，若向量方程

仅有平凡解，向量组(集)称为**线性相关**的，若存在不全为零的权，使

# 矩阵代数与行列式

## 概念与矩阵运算

若是矩阵，即有行列的矩阵，的第行第列的元素用表示，称为的

元素，如下图所示

设是相同维数的矩阵，与为数，则有

若乘积有定义，的第行第列的元素是的第行与的第列对应元素乘积之和。

若表示的元素，为矩阵，则

矩阵乘法的理解可以参照前面有关矩阵方程的介绍。

**矩阵的乘幂**表示为，其中是矩阵，表示个的乘积。为单位矩阵。

**矩阵的转置**表示为，其中是一个矩阵，则的转置是一个矩阵，它的列是由对应行构成的。

设与表示矩阵，其维数使下列和与积有定义，则

1. 对任意实数，

## 2.矩阵的逆

一个矩阵是可逆的，若存在一个矩阵使

且

这时称是的**逆阵**，可逆，且它的逆矩阵唯一。我们将记为，于是

且

不可逆矩阵也称为**奇异矩阵**，可逆矩阵也称为**非奇异矩阵**。

我们从这个定义可可以知道，可逆矩阵一定是方阵。后面还会引入左逆和右逆。下面是有关可逆矩阵的一些性质。

### 2.1 求的算法

将和排在一起构成增广矩阵，则对此矩阵进行行变换时，和受到同一变换。要么有一系列行变换把变成，同时变成；要么是不可逆的。例如



### 2.2 逆矩阵的应用-方程求解

可逆矩阵的一个重要性质就是用于求解方程组。

若可逆，则对每一个中的，方程有唯一解。

我们举一个例子来说明可逆矩阵的作用。比如我们要解如下方程

**解** 该方程组就是，可以看成

所以有

### 2.3 可逆矩阵的性质、判断是否可逆

下面是有关可逆矩阵的两个基本定理：

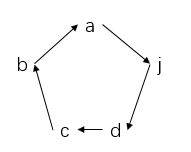
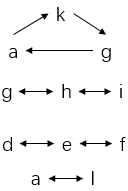
1. 若可逆，则有

接着是可逆矩阵定理。设为矩阵，则下列命题是等价的，即对某一特定的，它

们同时为真或同时为假。

1. 是可逆矩阵
2. 等价于单位矩阵
3. 有个主元位置
4. 方程仅有平凡解
5. 的各列线性无关
6. 线性变换是一对一的
7. 对中任意，方程至少有一个解
8. 的各列生成
9. 线性变换把映上上
10. 存在矩阵使
11. 存在矩阵使
12. 是可逆矩阵

它们的推导关系如下所示

c → b是因为在矩阵有个主元就说明它的简化阶梯形是。

关于d，可有j推出，即 推得 .

关于e，其实就是d，可参照4.1节。

利用这些定理，我们可以用于判断一个矩阵是否可逆。例如我们有一个矩阵，如下所示。

**解**

所以有3个主元位置，根据可逆矩阵定理c，是可逆的。

## 3.矩阵因式分解-LU分解

矩阵的因式分解是把表示为两个或更多矩阵的乘积，矩阵乘法是数据的综合（把两个或更多线性变换的作用结合成一个矩阵），矩阵的因式分解是数据的分解，把数据组成两个或者更多，这种结构可能更有用，或者更便于计算。

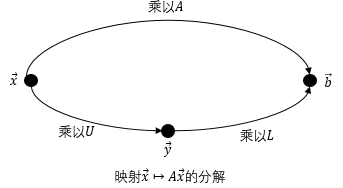
分解，是在工业与商业问题中常见的，解一系列具有相同系数矩阵中的线性方程

我们解决这个问题是可以通过计算来求解方程，但是若不可逆呢？

首先，我们假设是矩阵可以行化简为阶梯形而不必行对换，则可以写成形式，是下三角矩阵，主对角线元素全是1，是的一个等价的阶梯形矩阵，例如下图所示，这样一个分解称为分解，矩阵是可逆的，称为单位下三角矩阵。

求解可写成，因此可以由解下列方程来求解：

先解(1)再解(2)即可得到，每个方程都很容易解，因为和都是三角矩阵。



下面我们举一个例子来应用分解。

已知

应用的分解来解，其中.

**解**

所以.

### 分解算法

算法并不复杂，总的来说就是

1. 矩阵经过一系列行变换变成阶梯形矩阵，即为
2. 变换过程在每个主元列，把主元以下的元素除以主元即为在该列主元以下的元素。

以一个例子来讲解。求下列矩阵的分解

解

接着是在变换过程中求出矩阵。上式中标出的元素确定了化为的行变换，在每个主元列，把标出的元素除以主元后将结果放入。

容易证明，所求出的和满足.

### 算法性能分析

下列运算次数的计算适用于稠密矩阵（大部分元素非零），相当大，例如。

1. 计算的分解大约需要浮算，而求大约需要浮算。
2. 解和大约需要浮算，因为任意三角方程组可以用大约浮算接触。
3. 把乘以也需要浮算，但结果可能不如和得出的精确(由于计算及的舍入误差)
4. 若是稀疏矩阵（大部分元素为0），则和可能也是稀疏的，然而很可能是稠密的，显然用分解来解很可能比用快很多。

## 4.行列式

首先行列式是只对方阵才有定义的，即**只有方阵才能计算行列式**，可逆矩阵也类似。

当行列式不为0时，矩阵才可逆；当系数矩阵行列式不为0时，系数矩阵才线性无关，方程组才有解。

接下来我们会先介绍行列式是如何计算的。

### 4.1 行列式的计算

我们先给出矩阵和矩阵的行列式。

**余子式** 在阶行列式中，把元素所在的第行第列划去后，余下的阶行列式叫做元素的的余子式，记作，同时称为的代数余子式。

要注意的是，余子式与代数余子式都是行列式，行列式的阶越低越容易计算，它存在的意义就是为了更好的计算行列式。

下面给出行列式的计算公式。

这是按照第一行来计算行列式，更一般地，我们可以按照第行或者第列来计算。通过观察我们也可以知道，**三角形矩阵的行列式等于对角线上元素的乘积**。

让我们举个例子来加深理解。

计算，这里

**解** 按的第一列来进行计算，有

再对第一列进行展开，有

具体按照哪一列来展开要看情况。

### 4.2 行列式的性质

我们给出几个行列式的性质。

是一个方阵。

1. 若的某一行的倍数加到另一行得矩阵，则 =
2. 若的两行互换得矩阵，则 =
3. 若的某行乘以倍得到矩阵，则 =
4. ，这证明了行列式的列变换和行变换具有相同效果
5. 若与均为矩阵，则

这几个性质可以帮我们有效地计算行列式。我们举一个例子来说明。

计算，其中

**解** 思路是先将化简成阶梯形，再利用三角形矩阵的行列式等于对角线上元素的乘积来计算。

### 4.3 逆矩阵公式

下面给出一个利用行列式来计算矩阵的逆的公式。

**定理** 设是一个可逆的矩阵，则.

用一个例子来说明。求矩阵的逆。

**解** 九个代数余子式为

为代数余子式的矩阵的转置(例如到(1,2)的位置)，从而有

我们可以直接计算，但下列计算对上面的结果提供了一个检验并计算出的方法。

我们根据定理可知，，可知 .

所以，

### 4.4 用行列式表示面积或体积

若是一个矩阵，则由的列确定的平行四边形的面积为；若是一个矩阵，则由的列确定的平行六面体的体积也为。可引申到矩阵的情形。这也就是**行列式的几何意义**。

例如若为2阶对角矩阵，定理显然成立。

矩阵面积

所以可以得出，判断一个线性方程组是否有唯一解，即判断系数矩阵的列向量所构成的体积是否为0，即判断列向量是否两两位于不同平面上。