1 线性代数中的线性方程组

线性方程组

线性方程:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

其中, b 与系数 a_1,a_2,$\cdots$,a_n 是实数或复数, 通常为已知数. n 为任意正整数

线性方程组 – 由一个或几个包含相同变量的线性方程组成, 如:

$$\begin{cases} 2x_1 - x_2 + 1.5x_3 = 8 \\ x_1 \qquad - 4x_3 = -7 \end{cases}$$

若线性方程组的方程个数少于未知数个数, 称之为**欠定方程组**
若线性方程组的方程个数多余未知数个数, 称之为**超定方程组**
方程组所有可能的解的集合称为线性方程组的**解集**
若两个方程组有相同的解集, 则这两个方程组称为**等价**的

线性方程组解集情况:
1. 无解;
2. 有唯一解;

3. 有无穷多解.

当方程组无解时, 称线性方程组**不相容**
当方程组有唯一解或无穷多解时, 称线性方程组**相容**

线性方程组

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$

线性方程组的**系数矩阵**

$$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

线性方程组的**增广矩阵**

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

解线性方程组
思路: 将方程组用一个更容易求解的等价方程组替代

化简方程组的三种基本变换:
1. 倍加变换 - 将某方程替换为它与另一方程倍数的和;
2. 对换变换 - 交换两个方程的位置;
3. 倍乘变换 - 方程的所有系数乘以一个非 0 实数.

例. 简化如下方程组

$$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ 2x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

步骤 1 – ③+4①

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

步骤 2 – $\frac{1}{2}$②

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix}$$

步骤 3 – ③+3②

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & -3 & 13 & -9 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

步骤 4 – ②+4③

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 1 & 3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

步骤 5 – ①+(-③)

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

步骤 6 – ①+2②

$$\begin{bmatrix} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix} \Longrightarrow \begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{bmatrix}$$

若两个线性方程组的增广矩阵是行等价的, 则它们具有相同的解集

行化简与阶梯形矩阵

非零行 (列): 矩阵中至少包含一个非零元素的行 (列)
先导元素: 该行最左边的非零元素

定义　*一个矩阵称为阶梯形, 若它有以下三个性质:*
1. 所有非零行在零行之上;
2. 某一行先导元素的列位于上一行先导元素的右边;
3. 某一先导元素所在列下方元素都是零;
若还满足以下性质, 则称为简化阶梯形:
4. 每一非零行的先导元素是 1;
5. 每一先导元素是该元素所在列的唯一非零元素.

定理 1 (简化阶梯形矩阵的唯一性)
每个矩阵行等价于唯一的简化阶梯形矩阵.

定义　*矩阵 A 中的*主元位置*是 A 中对应于它的阶梯形中先导元素的位置.* 主元列*是 A 中含有主元位置的列.*

基本变量: 位于主元列的变量
自由变量: 位于非主元列的变量

定理 2 (存在与唯一性定理)
线性方程组相容的充要条件是增广矩阵的最右列不是主元列. 也就是说, 增广矩阵的阶梯形没有形如

$$[0 \ \cdots \ 0 \ b], b \neq 0$$

的行. 若线性方程组相容, 则它的解集可能有两种情形:
1）当没有自由变量时, 有唯一解;
2）若至少有一个自由变量, 则有无穷多解.

应用行化简算法解线性方程组:
1. 写出方程组的增广矩阵
2. 应用行化简算法把增广矩阵化为阶梯形, 确定方程组是否相容, 如果没有解则停止; 否则进行下一步
3. 继续行化简算法得到它的简化阶梯形
4. 写出简化阶梯形矩阵对应的方程组
5. 将每个非零方程改写为使用自由变量表示基本变量的形式

向量方程

列向量: 仅含一列的矩阵, 简称为向量. 如:

$$u = \begin{bmatrix} 3 \\ -1 \end{bmatrix} \quad \text{or} \quad u = (3, -1)$$

行向量: 仅含一行的矩阵. 如:

$$v = \begin{bmatrix} 2 & 5 \end{bmatrix}$$

所有两个元素的向量表示为 $\mathbb{R}^2$, $\mathbb{R}$ 表示元素为实数, 2 表示向量包含两个元素
向量加法:

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1+2 \\ -2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

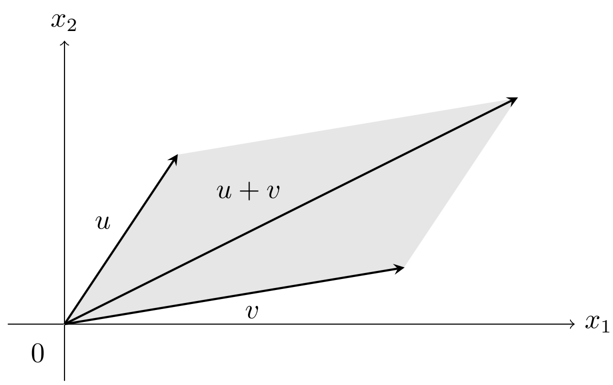
标量乘法:
若 $u = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, $c = 5$, 则:

$$cu = 5 \begin{bmatrix} 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 15 \\ -5 \end{bmatrix}$$

向量 $\begin{bmatrix} x \\ y \end{bmatrix}$ 的几何含义: 由原点 (0,0) 指向点 (x,y) 的有向线段

向量加法的平行四边形法则

若 $\mathbb{R}^2$ 中向量 $\mathbf{u}$ 和 $\mathbf{v}$ 用平面上的点表示, 则 $\mathbf{u}+\mathbf{v}$ 对应于以 $\mathbf{u}$,$\mathbf{0}$ 和 $\mathbf{v}$ 为三个顶点的平行四边形的第 *4* 个顶点, 如图.



所有元素都是零的向量称为**零向量**, 用 $\mathbf{0}$ 表示 ($\mathbf{0}$ 中元素的个数可由上下文确定)

$\mathbb{R}^n$ 中向量的代数性质

对 $\mathbb{R}^n$ 中一切向量 $\mathbf{u}$,$\mathbf{v}$,$\mathbf{w}$ 以及标量 c 和 d:

(i)	$\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$	(v)	$c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$
(ii)	$(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$	(vi)	$(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$
(iii)	$\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$	(vii)	$c(d\mathbf{u})=(cd)\mathbf{u}$
(iv)	$\mathbf{u}+(-\mathbf{u})=-\mathbf{u}+\mathbf{u}=\mathbf{0}$	$(viii)$	$1\mathbf{u}=\mathbf{u}$

给定 $\mathbb{R}^n$ 中向量 $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p$ 和标量 $c_1,c_2,\cdots,c_p$, 向量

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

称为向量 $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p$ 以 $c_1,c_2,\cdots,c_p$ 为**权**的**线性组合**.

向量方程

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

和增广矩阵为

$$[\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \quad \mathbf{b}] \tag{1.1}$$

的线性方程组有相同的解集. 特别地, $\mathbf{b}$ 可表示为 $\mathbf{a}_1,\mathbf{a}_2,\cdots,\mathbf{a}_n$ 的线性组合当且仅当对应于(1.1)式的线性方程组有解.

定义　若 $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p$ 是 $\mathbb{R}^n$ 中的向量, 则 $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p$ 的所有线性组合所成的组合用记号 $\mathrm{Span}\{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p\}$ 表示, 称为由 $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p$ 所**生成**（或**张成**）的 $\mathbb{R}^n$ 的**子集**. 也就是说, $\mathrm{Span}\{\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_p\}$ 是所有形如

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

的向量的集合, 其中 $c_1,c_2,\cdots,c_p$ 为标量.

矩阵方程 $A\mathbf{x} = \mathbf{b}$

定义　若 A 是 $m \times n$ 矩阵, 它的各列为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$. 若 $\boldsymbol{x}$ 是 $\mathbb{R}^n$ 中的向量, 则A 与 $\boldsymbol{x}$ 的积（记为 $A\boldsymbol{x}$）就是A 的各列以 $\boldsymbol{x}$ 中对应元素为权的线

性组合, 即

$$A\boldsymbol{x} = [\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_n]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$$

注意 $A\boldsymbol{x}$ 仅当 A 的列数等于 $\boldsymbol{x}$ 中的元素个数时才有意义.

定理 3　若 A 是 $m \times n$ 矩阵, 它的各列为 $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_n$, 而 $\boldsymbol{b}$ 属于 $\mathbb{R}^n$, 则矩阵方程

$$A\boldsymbol{x} = \boldsymbol{b}$$

与向量方程

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

有相同的解集. 它又与增广矩阵为

$$[\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_n\ \boldsymbol{b}]$$

的线性方程组有相同的解集.

定理 4　设 A 是 $m \times n$ 矩阵, 则下列命题是逻辑上等价的. 也就是说, 对某个 A, 它们都成立或者都不成立.

$a.$ 对 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$, 方程 $A\boldsymbol{x} = \boldsymbol{b}$ 有解.

$b.$ $\mathbb{R}^m$ 中的每个 $\boldsymbol{b}$ 都是 A 的列的一个线性组合.

$c.$ A 的各列生成 $\mathbb{R}^m$.

$d.$ A 在每一行都有一个主元位置.

计算 $\boldsymbol{A}\boldsymbol{x}$ 的行—向量规则

若乘积 $A\boldsymbol{x}$ 有定义, 则 $A\boldsymbol{x}$ 中的第 i 个元素是 A 的第 i 行元素与 $\boldsymbol{x}$ 的相应元素乘积之和.

矩阵的主对角线上元素为 1, 其他位置上元素为 0, 这个矩阵称为**单位矩阵**, 并记为 I.
如果矩阵为 $n \times n$ 单位矩阵, 记为 I_n.

定理 5　若 A 是 $m \times n$ 矩阵, $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中向量, c 是标量, 则
a. $A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$
b. $A(c\boldsymbol{u}) = c(A\boldsymbol{u})$

线性方程组的解集

若线性方程组可写成

$$A\boldsymbol{x} = \boldsymbol{0}$$

的形式, 则称为**齐次线性方程组**. 其中, A 是 $m \times n$ 矩阵, $\boldsymbol{0}$ 是 $\mathbb{R}^m$ 中的零向量.
齐次线性方程组至少有一个解, 即 $\boldsymbol{x} = \boldsymbol{0}$($\mathbb{R}^n$ 中的零向量), 这个解称为它的**平凡解**.
如果有一个非零向量 $\boldsymbol{x}$, 满足 $A\boldsymbol{x} = \boldsymbol{0}$, 这个解称为它的**非平凡解**.

齐次方程 $A\boldsymbol{x} = \boldsymbol{0}$ 有非平凡解当且仅当方程至少有一个自由变量.

$\boldsymbol{x} = s\boldsymbol{u} + t\boldsymbol{v}$ 为 $A\boldsymbol{x} = \boldsymbol{0}$ 的**参数向量形式**, 并称之为**参数向量方程**. 其中, s,t 为自由变量
$\boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{v}$ 为 $A\boldsymbol{x} = \boldsymbol{b}$ 的**参数向量形式**, 并称之为**参数向量方程**. 其中, t 为自由变量
例.
$x_1 = 0.3x_2 + 0.2x_3$

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 + 0.2x_3 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3x_2 \\ x_2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2x_3 \\ 0 \\ x_3 \end{bmatrix}$$
$$= x_2 \begin{bmatrix} 0.3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0.2 \\ 0 \\ 1 \end{bmatrix}$$

因此, $A\boldsymbol{x} = \boldsymbol{b}$ 的解集是一条通过 $\boldsymbol{p}$ 而平行于 $A\boldsymbol{x} = \boldsymbol{0}$ 的解集的直线. 也称为将 $\boldsymbol{v}$ 沿着 $\boldsymbol{p}$ 进行直线移动.

定理 6　设方程 $A\boldsymbol{x} = \boldsymbol{b}$ 对某个 $\boldsymbol{b}$ 是相容的, $\boldsymbol{p}$ 为一个特解, 则 $A\boldsymbol{x} = \boldsymbol{b}$ 的解集是所有形如 $\boldsymbol{w} = \boldsymbol{p} + \boldsymbol{v}_h$ 的向量的集, 其中 $\boldsymbol{v}_h$ 时齐次方程 $A\boldsymbol{x} = \boldsymbol{b}$ 的任意一个解.

把 (相容方程组的) 解集表示成参数向量形式

1. 把增广矩阵简化为简化阶梯形.

2. 把每个基本变量用自由变量表示.

3. 把一般解 $\boldsymbol{x}$ 表示成向量, 如果有自由变量, 其元素依赖于自由变量.

4. 把 $\boldsymbol{x}$ 分解为向量（元素为常数）的线性组合, 用自由变量作为参数.

线性方程组的应用

1. 经济学 - 部分的收支平衡
2. 化学式 - 等号两边原子守恒
3. 网络流 - 节点的进/出流量恒等

线性无关

定义　$\mathbb{R}^n$ 中一组向量 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 称为**线性无关**的, 若向量方程

$$x_1\boldsymbol{v}_1+x_2\boldsymbol{v}_2+\cdots+x_p\boldsymbol{v}_p=\boldsymbol{0}$$

仅有平凡解. 向量组（集）$\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 称为**线性相关**的, 若存在不全为零的权 $c_1,\cdots,c_p$, 使

$$c_1\boldsymbol{v}_1+c_2\boldsymbol{v}_2+\cdots+c_p\boldsymbol{v}_p=\boldsymbol{0}$$

矩阵 A 的各列线性无关, 当且仅当方程 $A\boldsymbol{x}=\boldsymbol{0}$ 仅有平凡解.

两个向量的集合 $\{\boldsymbol{v}_1,\boldsymbol{v}_2\}$ 线性相关, 当且仅当其中一个向量是另一个向量的倍数. 这个集合线性无关, 当且仅当其中任一个向量都不是另一个向量的倍数.

定理 7 (线性相关集的特征)
两个或更多个向量的集合 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 线性相关, 当且仅当 S 中至少有一个向量是其他向量的线性组合. 事实上, 若 S 线性相关, 且 $\boldsymbol{v}_1\neq\boldsymbol{0}$, 则某个 $\boldsymbol{v}_j(j>1)$ 是它前面向量 $\boldsymbol{v}_1,\cdots,\boldsymbol{v}_{j-1}$ 的线性组合.

定理 8　若一个向量组的向量个数超过每个向量的元素个数, 那么这个向量组线性相关. 就是说, $\mathbb{R}^n$ 中任意向量组 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 当 $p>n$ 时线性相关.

定理 9　若 $\mathbb{R}^n$ 中向量组 $S=\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_p\}$ 包含零向量, 则它线性相关.

线性变换介绍

由 $\boldsymbol{x}$ 到 $A\boldsymbol{x}$ 的对应是由一个向量集到另一个向量集的函数

定义　变换 (或映射)T 称为线性的, 若

(*i*) 对 T 的定义域中一切 $\boldsymbol{u}$,$\boldsymbol{v}$,$T(\boldsymbol{u}+\boldsymbol{v})=T(\boldsymbol{u})+T(\boldsymbol{v})$.

(*ii*) 对 T 的定义域中的一切 $\boldsymbol{u}$ 和数 c, $T(c\boldsymbol{u})=cT(\boldsymbol{u})$.

若 T 是线性变换, 则

$$T(\mathbf{0})=\mathbf{0}$$

且对 T 的定义域中一切向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 以及数 c 和 d, 有:

$$T(c\boldsymbol{u}+d\boldsymbol{v})=cT(\boldsymbol{u})+dT(\boldsymbol{v})$$

线性变换的矩阵

定理 10　设 T:$\mathbb{R}^n\to\mathbb{R}^m$ 为线性变换, 则存在唯一的矩阵 A, 使得对 $\mathbb{R}^n$ 中一切 $\boldsymbol{x}$, 有

$$T(\boldsymbol{x})=A\boldsymbol{x}$$

事实上, A 是 $m\times n$ 矩阵, 它的第 j 列是向量 $T(\boldsymbol{e}_j)$, 其中 $\boldsymbol{e}_j$ 是 $\mathbb{R}^n$ 中单位矩阵 $\boldsymbol{I}_n$ 的第 j 列:

$$A=[T(\boldsymbol{e}_1)\ \cdots\ T(\boldsymbol{e}_n)]$$

定义　映射 T:$\mathbb{R}^n\to\mathbb{R}^m$ 称为到 $\mathbb{R}^m$ 上的映射, 若 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$ 是 $\mathbb{R}^n$ 中至少一个 $\boldsymbol{x}$ 的像（也称为满射）.

定义　映射 $T:\mathbb{R}^n \to \mathbb{R}^m$ 称为**一对一映射**（或 *1:1*），若 $\mathbb{R}^m$ 中每个 $\boldsymbol{b}$ 是 $\mathbb{R}^n$ 中至多一个 $\boldsymbol{x}$ 的像（也称为单射）.

定理 11　设 $T:\mathbb{R}^n \to \mathbb{R}^m$ 为线性变换, 则 T 是一对一的当且仅当方程 $A\boldsymbol{x}=\mathbf{0}$ 仅有平凡解.

定理 12　设 $T:\mathbb{R}^n \to \mathbb{R}^m$ 是线性变换, 设 A 为 T 的标准矩阵, 则

a. T 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$, 当且仅当 A 的列生成 $\mathbb{R}^m$.

b. T 是一对一的, 当且仅当 A 的列线性无关.

商业、科学和工程中的线性模型

基尔霍夫电压定律

围绕一条回路同一方向的电压降 RI 的代数和等于围绕该回路的同一方向电动势的代数和.

如果有矩阵 A 使 $\boldsymbol{x}_1 = A\boldsymbol{x}_0$,$\boldsymbol{x}_2 = A\boldsymbol{x}_1$, 一般地,

$$\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k,\ k=0,1,2,\cdots$$

则称为**线性差分方程**（或**递归关系**）.

2 矩阵代数

矩阵运算

$m\times n$ 矩阵 $A=[a_{ij}]$ 的**对角线元素**是 $a_{11},a_{22},a_{33},\cdots$, 它们组成 A 的**主对角线**.

对角矩阵是一个方阵, 它的非对角线元素全是 0. 例如 $n\times n$ 单位矩阵 $\boldsymbol{I}_n$.

元素全是 0 的 $m\times n$ 矩阵称为**零矩阵**, 用 $\mathbf{0}$ 表示.

若两个矩阵有相同的维数 (即有相同的行数和列数), 而且对应元素相同, 则称该两个矩阵**相等**

若 r 是标量而 A 是矩阵, 则**标量乘法** rA 是一个矩阵, 它的每一列是 A 的对应列的 r 倍.

定理 1 设 A,B,C 是相同维数的矩阵, r 与 s 为数, 则有

a. $A+B=B+A$ b. $(A+B)+C=A+(B+C)$

c. $A+0=A$ d. $r(A+B)=rA+rB$

e. $(r+s)A=rA+sA$ f. $r(sA)=(rs)A$

定义 若 A 是 $m\times n$ 矩阵, B 是 $n\times p$ 矩阵, B 的列是 $\boldsymbol{b}_1,\cdots,\boldsymbol{b}_p$, 则乘积

AB 是 $m \times p$ 矩阵, 它的各列是 $A\boldsymbol{b}_1, \cdots, A\boldsymbol{b}_p$, 即

$$AB = A[\boldsymbol{b}_1\ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_p] = [A\boldsymbol{b}_1\ A\boldsymbol{b}_2\ \cdots\ A\boldsymbol{b}_p]$$

AB 的每一列都是 A 的各列的线性组合, 以 B 的对应列的元素为权.

计算 $\boldsymbol{AB}$ 的行列法则

若乘积 AB 有定义, 则 AB 的第 i 行第 j 列的元素是 A 的第 i 行与 B 的第 j 列对应元素乘积之和. 若 $(AB)_{ij}$ 表示 AB 的 (i,j) 元素, A 为 $m \times n$ 矩阵, 则

$$(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

定理 2　设 A 为 $m \times n$ 矩阵, B 和 C 的维数使下列各式的乘积有意义.

a. $A(BC) = (AB)C$　（乘法结合律）

b. $A(B+C) = AB + AC$　（乘法左分配律）

c. $(B+C)A = BA + CA$　（乘法右分配律）

d. $r(AB) = (rA)B = A(rB)$, r 为任意数

e. $I_m A = A = A I_m$　（矩阵乘法的恒等式）

乘积 AB 的因子关系为: A 被 B **右乘**, 或 B 被 A **左乘**

若 AB=BA, 我们称 A 和 B 彼此**可交换**

警告

1. 一般情况下, $AB \neq BA$.

2. 消去律对矩阵乘法不成立, 即若 $AB = AC$, 一般情况下, $B = C$ 并不成立.
3. 若乘积 AB 是零矩阵, 一般情况下, 不能断定 A=0 或 B=0.

给定 $m \times n$ 矩阵, 则 A 的**转置**是一个 $n \times m$ 矩阵, 用 A^T 表示, 它的列是由 A 的对应行构成的.

定理 3　设 A 与 B 表示矩阵, 其维数使下列和与积有定义, 则
a. $(A^T)^T = A$.
b. $(A+B)^T = A^T + B^T$.
c. 对任意数 r, $(rA)^T = rA^T$.
d. $(AB)^T = B^T A^T$.

若干个矩阵的乘积的转置等于它们的转置的乘积, 但相乘的顺序相反.

矩阵的逆

A 为 $n \times n$ 矩阵, 若存在一个 $n \times n$ 矩阵 C, 使得

$$CA = \boldsymbol{I}_n \qquad 且 AC = \boldsymbol{I}_n$$

则称 A **可逆**, 并且 C 是 A 的**逆**.

若 A 可逆, 它的逆是唯一的, 我们将它记为 A^{-1}, 则

$$A^{-1}A = I \qquad 且 AA^{-1} = I$$

不可逆矩阵也称为**奇异矩阵**.
可逆矩阵也称为**非奇异矩阵**.

定理 4　设 $A = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$. 若 $ad - bc \neq 0$, 则 A 可逆且

$$A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

若 $ad - bc = 0$, 则 A 不可逆.

数 $ad - bc$ 称为 A 的**行列式**, 记为

$$\det A = ad - bc$$

定理 5　若 A 是可逆 $n \times n$ 矩阵, 则对每一 $\mathbb{R}^n$ 中的 $\boldsymbol{b}$, 方程 $A\boldsymbol{x} = \boldsymbol{b}$ 有唯一解 $\boldsymbol{x} = A^{-1}\boldsymbol{b}$.

胡克定律
公式如下

$$\boldsymbol{y} = D\boldsymbol{f}$$

其中 D 为**弹性矩阵**, 它的逆称为**刚性矩阵**, $\boldsymbol{f}$ 表示它在各个点受的力, $\boldsymbol{y}$ 表示各个点的形变量.

定理 6
a. 若 A 是可逆矩阵, 则 A^{-1} 也可逆而且 $(A^{-1})^{-1} = A$.
b. 若 A 和 B 都是 $n \times n$ 可逆矩阵, 则 AB 也可逆, 且其逆是 A 和 B 的

逆矩阵按相反顺序的乘积, 即

$$(AB)^{-1} = B^{-1}A^{-1}$$

c. 若 A 可逆, 则 A^T 也可逆, 且其逆是 A^{-1} 的转置, 即 $(A^T)^{-1} = (A^{-1})^T$.

若干个 $n \times n$ 可逆矩阵的积也是可逆的, 其逆等于这些矩阵的逆按相反顺序的乘积.

把单位矩阵进行一次初等行变换, 就得到**初等矩阵**.

若对 $m \times n$ 矩阵 A 进行某种初等行变换, 所得矩阵可写成 EA, 其中 E 是 $m \times m$ 矩阵, 是由 I_m 进行同一行变换所得.

每个初等矩阵 E 是可逆的, E 的逆是一个同类型的初等矩阵, 它把 E 变回 I.

定理 7　$n \times n$ 矩阵 A 是可逆的, 当且仅当 A 行等价于 I_n, 这时, 把 A 化简为 I_m 的一系列初等行变化同时把 I_n 变成 A^{-1}.

求 $\boldsymbol{A}^{-1}$ 的算法

把增广矩阵 $[A\ I]$ 进行行化简. 若 A 行等价于 I, 则 $[A\ I]$ 行等价于 $[I\ A^{-1}]$, 否则 A 没有逆.

可逆矩阵的特征

定理 8 (可逆矩阵定理)

设 A 为 $n \times n$ 矩阵, 则下列命题是等价的, 即对某一特定的 A, 它们同时为真或同时为假.

a. A 是可逆矩阵.

b. A 行等价于 $n \times n$ 单位矩阵.

c. A 有 n 个主元位置.

d. 方程 $A\boldsymbol{x} = \boldsymbol{0}$ 仅有平凡解.

e. A 的各列线性无关.

f. 线性变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 是一对一的.

g. 对 $\mathbb{R}^n$ 中任意 $\boldsymbol{b}$, 方程 $A\boldsymbol{x} = \boldsymbol{b}$ 至少有一个解.

h. A 的各列生成 $\mathbb{R}^n$,

i. 线性变换 $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 把 $\mathbb{R}^n$ 映上到 $\mathbb{R}^n$.

j. 存在 $n \times n$ 矩阵 C 使 $CA = I$.

k. 存在 $n \times n$ 矩阵 D 使 $AD = I$.

l. A^T 是可逆矩阵.

设 A 和 B 为方阵, 若 $AB = I$, 则 A 和 B 都是可逆的, 且 $B = A^{-1}$, $A = B^{-1}$.

定理 9 　设 $T:\mathbb{R}^n \to \mathbb{R}^n$ 为线性变换, A 为 T 的标准矩阵. 则 T 可逆当且仅当 A 是可逆矩阵.

若一个 $m \times n$ 矩阵的主对角线以下元素全为 0, 则称之为**上三角矩阵**.

若一个 $m \times n$ 矩阵的主对角线以上元素全为 0, 则称之为**下三角矩阵**.

分块矩阵

形如

$$A = \left[\begin{array}{rrr|rr|r} 3 & 0 & -1 & 5 & 9 & -2 \\ -5 & 2 & 4 & 0 & -3 & 1 \\ \hline -8 & -6 & 3 & 1 & 7 & -4 \end{array}\right]$$

为矩阵 A 的 2×3 **分块矩阵**, 也可表示为

$$A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

设 A 为 $m \times n$ 矩阵, B 为 $n \times p$ 矩阵, 当 A 的列的分法与 B 的行的分法一致时, 可计算 AB. 如下:

$$A = \left[\begin{array}{rrr|rr} 2 & -3 & 1 & 0 & -4 \\ 1 & 5 & -2 & 3 & -1 \\ \hline 0 & -4 & -2 & 7 & -1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$$B = \left[\begin{array}{rr} 6 & 4 \\ -2 & 1 \\ -3 & 7 \\ \hline -1 & 3 \\ 5 & 2 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} A_{11}B_1 + A_{12}B_2 \\ A_{21}B_1 + A_{22}B_2 \end{bmatrix} = \left[\begin{array}{cc} -5 & 4 \\ -6 & 2 \\ \hline 2 & 1 \end{array}\right]$$

定理 10 (*AB* 的列行展开)
若 A 是 $m \times n$ 矩阵, B 是 $n \times p$ 矩阵, 则

$$\begin{aligned} AB &= [col_1(A)\ col_2(A)\ \cdots\ col_n(A)] \begin{bmatrix} row_1(B) \\ row_2(B) \\ \vdots \\ row_n(B) \end{bmatrix} \\ &= col_1(A)row_1(B) + \cdots + col_n(A)row_n(B) \end{aligned}$$

矩阵因式分解

设 A 是 $m \times n$ 矩阵, 它可以行化简为阶梯形 (化简步骤不包含对换变换), 则 A 可写成 $A = LU$. 其中, L 是 $m \times m$ 下三角矩阵, 主对角线元素全是 1; U 是 A 的一个 $m \times n$ 阶梯形矩阵.

LU 分解的算法
1. 如果可能的话, 用一系列的行倍加变换把 A 化为阶梯形 U（即 $L^{-1}A = U$）.
2. 填充 L 的元素使相同的行变换把 L 变为 I.

LU 分解图解:

$$
\begin{array}{ccccccccc}
A = \begin{bmatrix} 2 & 4 & 5 & -2 \\ -4 & -5 & -8 & 1 \\ 2 & -5 & 1 & 8 \\ -6 & 0 & -3 & 1 \end{bmatrix} & \Rightarrow & \begin{bmatrix} 2 & 4 & 5 & -2 \\ 0 & 3 & 2 & -3 \\ 0 & -9 & -4 & 10 \\ 0 & 12 & 12 & -5 \end{bmatrix} & \Rightarrow & \begin{bmatrix} 2 & 4 & 5 & -2 \\ 0 & 3 & 2 & -3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 4 & 7 \end{bmatrix} & \Rightarrow & \begin{bmatrix} 2 & 4 & 5 & -2 \\ 0 & 3 & 2 & -3 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix} = U \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & & 1 & 0 \\ -3 & & & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix} & & \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix} = L
\end{array}
$$

列昂惕夫投入产出模型

列昂惕夫投入产出模型或生产方程

$$\boldsymbol{x} = C\boldsymbol{x} + \boldsymbol{d}$$

定理 11　设 C 为某一经济体系的消耗矩阵, $\boldsymbol{d}$ 为最终需求. 若 C 和 $\boldsymbol{d}$ 的元素非负, C 的每一列的和小于 1, 则 $(I-C)^{-1}$ 存在, 产出向量

$$\boldsymbol{x} = (I-C)^{-1}\boldsymbol{d}$$

有非负元素, 且是下列方程的唯一解:

$$\boldsymbol{x} = C\boldsymbol{x} + d$$

计算机图形学中的应用

物体的平移并不直接对应于矩阵乘法, 因为平移并非线性变换, 所以引入**齐次坐标**

$\mathbb{R}^2$ 中每个点 (x, y) 对应于 $\mathbb{R}^3$ 中的点 $(x, y, 1)$, $(x, y, 1)$ 为 (x, y) 的**齐次坐标**

$(x, y, 1) \mapsto (x+h, y+k, 1)$ 的平移变换实现:

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x+h \\ y+k \\ 1 \end{bmatrix}$$

$\mathbb{R}^2$ 中任意线性变换可以通过齐次坐标乘以分块矩阵 $\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix}$ 实现, 其中 A 是 2×2 矩阵.

$(x, y, z, 1)$ 是 $\mathbb{R}^3$ 中点 (x, y, z) 的齐次坐标. 若 $H \neq 0$, 则 (X, Y, Z, H) 为 (x, y, z) 的齐次坐标, 且

$$x = \frac{X}{H}, y = \frac{Y}{H}, z = \frac{Z}{H}$$

点 (x, y, z) 在 xy 平面上的透视投影坐标为 $(\frac{x}{1 - z/d}, \frac{y}{1 - z/d}, 0)$. 其中, d 为 z 轴观测位置 $(0, 0, d)$

绕 $\mathbb{R}^2$ 中一点 p 的旋转是这样实现的: 首先把图形平移 $-p$, 然后绕原点旋转, 最后平移 p.

3

行列式

行列式介绍

有 3×3 矩阵 A

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

其中

$$\Delta = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

Δ 称为 3×3 矩阵 A 的**行列式**, 也可以写成

$$\begin{aligned} \Delta &= (a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}) - (a_{12}a_{21}a_{33} - a_{12}a_{23}a_{31}) + (a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}) \\ &= a_{11} \cdot \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \cdot \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \cdot \det \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \\ &= a_{11} \cdot \det A_{11} - a_{12} \cdot \det A_{12} + a_{13} \cdot \det A_{13} \end{aligned}$$

其中, A_{ij} 表示去除矩阵第 i 行和第 j 列元素后的内容.

例. A_{11} 表示如下:

$$\begin{bmatrix} \cancel{a_{11}} & \cancel{a_{12}} & \cancel{a_{13}} \\ \cancel{a_{21}} & a_{22} & a_{23} \\ \cancel{a_{31}} & a_{32} & a_{33} \end{bmatrix}$$

即

$$\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

行列式的两种表现形式:
原矩阵:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

第一种形式:

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

第二种形式:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

定义 当 $n \geqslant 2$, $n \times n$ 矩阵 $A = [a_{ij}]$ 的行列式是形如 $\pm a_{1j} \det A_{1j}$ 的 n 个项的和, 其中加号和减号交替出现, 这里元素 a_{11},a_{12},$\cdots$,a_{1n} 来自 A 的第一行, 即

$$\begin{aligned} \det A &= a_{11} \cdot \det A_{11} - a_{12} \cdot \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \cdot \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned}$$

给定 $A = [a_{ij}]$, A 的 (i, j) **余因子** C_{ij} 由下式给出

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

则

$$\det A = a_{11} \cdot C_{11} + a_{12} \cdot C_{12} + \cdots + a_{1n} \cdot C_{1n}$$

上述公式称为按 A 的**第一行的余因子展开式**

定理 1　$n \times n$ 矩阵 A 的行列式可按任意行或列的余因子展开式来计算. 按第 i 行的余因子展开式为:

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$$

按第 j 列的余因子展开式为:

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

定理 2　若 A 为三角阵, 则 $\det A$ 等于 A 的主对角线上元素的乘积

行列式的性质

定理 3 (行变换)
令 A 是一个方阵.
a. 若 A 的某一行的倍数加到另一行得矩阵 B, 则 $\det B = \det A$
b. 若 A 的两行互换得矩阵 B，则 $\det B = -\det A$
c. 若 A 的某行乘以 k 倍得到矩阵 B, 则 $\det B = k \det A$

定理 4　方阵 A 是可逆的当且仅当 $\det A \neq 0$

定理 5　若 A 为一个 $n \times n$ 矩阵, 则 $\det A^T = \det A$

定理 6 (乘法的性质)
若 A 和 B 均为 $n \times n$ 矩阵, 则 $\det AB = (\det A)(\det V)$

克拉默法则、体积和线性变换

对任意 $n \times n$ 矩阵 A 和任意的 $\mathbb{R}^n$ 中向量 b, 令 $A_i(b)$ 表示 A 中第 i 列由向量 **b** 替换得到的矩阵

$$A_i(b) = [\mathbf{a}_1 \ \cdots \ \mathbf{b} \ \cdots \ \mathbf{a}_n]$$

定理 7 (克拉默法则)
设 A 是一个可逆的 $n \times n$ 矩阵, 对 $\mathbb{R}^n$ 中任意向量 b, 方程 Ax=b 的唯一解可由下式给出

$$x_i = \frac{\det A_i(b)}{\det A}, i = 1, 2, \cdots, n$$

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

其中, 余因子组成的矩阵称为 A 的**伴随矩阵**, 记为 $adj\, A$

定理 8 (逆矩阵公式)
设 A 是一个可逆的 $n \times n$ 矩阵, 则 $A^{-1} = \frac{1}{\det A} adj\, A$

定理 9　若 A 是一个 2×2 矩阵, 则由 A 的列确定的平行四边形的面积为 $|\det A|$, 若 A 是一个 3×3 矩阵, 则由 A 的列确定的平行六面体的体积为 $|\det A|$

设 $\mathbf{a}_1$ 和 $\mathbf{a}_2$ 为非零向量, 则对任意数 c, 由 $\mathbf{a}_1$ 和 $\mathbf{a}_2$ 确定的平行四边形的面积等于由 $\mathbf{a}_1$ 和 $\mathbf{a}_2+c\mathbf{a}_1$ 确定的平行四边形的面积

定理 10　设 $T:\mathbb{R}^2\to\mathbb{R}^2$ 是由一个 2×2 矩阵 A 确定的线性变换, 若 S 是 $\mathbb{R}^2$ 中一个平行四边形, 则

$$\{T(S)\text{的面积}\} = |\det A|\cdot\{S\text{ 的面积}\}$$

若 T 是一个由 3×3 矩阵 A 确定的线性变换, 而 S 是 R^3 中的一个平行六面体, 则

$$\{T(S)\text{的体积}\} = |\det A|\cdot\{S\text{ 的体积}\}$$

4 向量空间

向量空间和子空间

定义　一个向量空间是由一些被称为向量的对象构成的非空集合 V, 在这个集合上定义两个运算, 称为加法和标量乘法（标量取实数）, 服从以下公理（或法则）, 这些公理必须对 V 中所有向量 $\mathbf{u}$,$\mathbf{v}$,$\mathbf{w}$ 及所有标量 c 和 d 均成立.

1. $\mathbf{u}$,$\mathbf{v}$ 之和表示为 $\mathbf{u}+\mathbf{v}$, 仍在 V 中
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$
4. V 中存在一个零向量 $\mathbf{0}$, 使得 $\mathbf{u}+\mathbf{0}=\mathbf{u}$
5. 对 V 中每个向量 $\mathbf{u}$, 存在 V 中向量 $-\mathbf{u}$, 使得 $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$
6. $\mathbf{u}$ 与标量 c 的标量乘法记为 $c\mathbf{u}$, 仍在 V 中
7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$
8. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$
9. $c(d\mathbf{u})=(cd)\mathbf{u}$
10. $1\mathbf{u}=\mathbf{u}$

对 V 中每个向量 $\mathbf{u}$ 和任意标量 c，有

$$0\mathbf{u} = \mathbf{0}$$

$$c\mathbf{0} = \mathbf{0}$$

$$-\mathbf{u} = (-1)\mathbf{u}$$

定义　向量空间 V 的一个**子空间**是 V 的一个满足以下三个性质的子集 H:

（i） V 中的零向量在 H 中

（ii） H 对向量加法封闭，即对 H 中任意向量 $\mathbf{u}$,$\mathbf{v}$，和 $\mathbf{u}+\mathbf{v}$ 仍在 H 中

（iii） H 对标量乘法封闭，即对 H 中任意向量 $\mathbf{u}$ 和任意标量 c，向量 $c\mathbf{u}$ 仍在 H 中

定理 1　若 v_1,v_2,$\cdots$,v_p 在向量空间 V 中，则 $\mathrm{Span}\{v_1,\cdots,v_p\}$ 是 V 的一个子空间

零空间、列空间和线性变换

考虑下列齐次方程组:

$$\begin{aligned} x_1-3x_2-2x_3&=0 \\ -5x_1+9x_2+\ x_3&=0 \end{aligned} \tag{4.1}$$

用矩阵的形式，此方程组可写成 $A\mathbf{x}=\mathbf{0}$，其中

$$A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$$

所有满足(4.1)的 $\mathbf{x}$ 的集合称为方程组(4.1)的**解集**
我们成满足 $A\mathbf{x}=\mathbf{0}$ 的所有 $\mathbf{x}$ 的集合为矩阵 A 的**零空间**

定义 矩阵 A 的零空间写成 $\operatorname{Nul} A$，是齐次方程 $A\mathbf{x}=\mathbf{0}$ 的全体解的集合. 用集合符号表示，即

$$\operatorname{Nul} A=\{\mathbf{x}:\mathbf{x}\in\mathbb{R}^n, A\mathbf{x}=\mathbf{0}\}$$

定理 2 $m\times n$ 矩阵 A 的零空间是 $\mathbb{R}^n$ 的一个子空间. 等价地，m 个方程，n 个未知数的齐次线性方程组 $A\mathbf{x}=\mathbf{0}$ 的全体解的集合是 $\mathbb{R}^n$ 的一个子空间

定义 $m\times n$ 矩阵 A 的**列空间**（记为 $\operatorname{Col} A$）是由 A 的列的所有线性组合组成的集合. 若 $A=[\mathbf{a}_1,\cdots,\mathbf{a}_n]$，则 $\operatorname{Col} A=\operatorname{Span}\{\mathbf{a}_1,\cdots,\mathbf{a}_n\}$

定理 3 $m\times n$ 矩阵 A 的列空间是 $\mathbb{R}^m$ 的一个子空间

$m\times n$ 矩阵 A 的列空间等于 $\mathbb{R}^m$ 当且仅当方程 $A\mathbf{x}=\mathbf{b}$ 对 $\mathbb{R}^m$ 中每个 $\mathbf{b}$ 有一个解

表 4.1: 对 $m\times n$ 矩阵 A，$\operatorname{Nul} A$ 与 $\operatorname{Col} A$ 之间的对比

$\operatorname{Nul} A$	$\operatorname{Col} A$
(i) $\operatorname{Nul} A$ 是 $\mathbb{R}^n$ 的一个子空间 (ii) $\operatorname{Nul} A$ 是隐式定义的，即仅给出了一个 $\operatorname{Nul} A$ 中向量必须满足的条件（$A\mathbf{x}=\mathbf{0}$） (iii) 求 $\operatorname{Nul} A$ 中的向量需要时间，需要对 $[A\quad \mathbf{0}]$ 作行变换 (iv) $\operatorname{Nul} A$ 与 A 的元素之间没有明显的关系 (v) $\operatorname{Nul} A$ 中的一个典型向量 $\mathbf{v}$ 具有 $A\mathbf{v}=\mathbf{0}$ 的性质 (vi) 给定一个特定的向量 $\mathbf{v}$，容易判断 $\mathbf{v}$ 是否在 $\operatorname{Nul} A$ 中. 仅需计算 $A\mathbf{v}$ (vii) $\operatorname{Nul} A=\{\mathbf{0}\}$ 当且仅当方程 $A\mathbf{x}=\mathbf{0}$ 仅有一个平凡解 (viii) $\operatorname{Nul} A=\{\mathbf{0}\}$ 当且仅当线性变换 $\mathbf{x}\mapsto A\mathbf{x}$ 是一对一的	(i) $\operatorname{Col} A$ 是 $\mathbb{R}^m$ 的一个子空间 (ii) $\operatorname{Col} A$ 是显式定义的，即明确指出如何构建 $\operatorname{Col} A$ 中的向量 (iii) 容易求出 $\operatorname{Col} A$ 中的向量. A 的列就是 $\operatorname{Col} A$ 中的向量，其余的可由 A 的列表示出来 (iv) $\operatorname{Col} A$ 与 A 的元素之间有明显的关系，因为 A 的列就在 $\operatorname{Col} A$ 中 (v) $\operatorname{Col} A$ 中一个典型向量 $\mathbf{v}$ 具有方程 $A\mathbf{x}=\mathbf{v}$ 是相容的性质 (vi) 给定一个特定的向量 $\mathbf{v}$，弄清 $\mathbf{v}$ 是否在 $\operatorname{Col} A$ 中需要时间，需要对 $[A\quad \mathbf{v}]$ 作行变换 (vii) $\operatorname{Col} A=\mathbb{R}^m$ 当且仅当方程 $A\mathbf{x}=\mathbf{b}$ 对每一个 $\mathbf{b}\in\mathbb{R}^m$ 有一个解 (viii) $\operatorname{Col} A=\mathbb{R}^m$ 当且仅当线性变换 $\mathbf{x}\mapsto A\mathbf{x}$ 将 $\mathbb{R}^n$ 映上到 $\mathbb{R}^m$

定义 由向量空间 V 映射到向量空间 W 内的**线性变换** T 是一个规划，它将 V 中每个向量 $\mathbf{x}$ 映射成 W 中唯一向量 $T(\mathbf{x})$，且满足：

（i） $T(\mathbf{u}+\mathbf{v})=T(\mathbf{u})+T(\mathbf{v})$，对 V 中所有 $\mathbf{u}$,$\mathbf{v}$ 均成立

（ii） $T(c\mathbf{u})=cT(\mathbf{u})$，对 V 中所有 $\mathbf{u}$ 及所有数 c 均成立

线性无关集和基

V 中向量的一个指标集 $\{v_1,\cdots,v_p\}$ 称为是**线性无关**的，如果向量方程

$$c_1\mathbf{v}_1+c_2\mathbf{v}_2+\cdots+c_p\mathbf{v}_p=\mathbf{0} \tag{4.2}$$

只有平凡解，即 $c_1=0,\cdots,c_p=0$.
集合 $\{\mathbf{v}_1,\cdots,\mathbf{v}_p\}$ 称为**线性相关**，如果(4.2)有一个非平凡的解，即存在某些权 $c_1,\cdots,c_p$ 不全为零，使得(4.2)式成立. 此时(4.2)式称为 $\mathbf{v}_1,\cdots,\mathbf{v}_p$ 之间的一个**线性相关关系**

定理 4 两个或多个向量组成的有编号的向量集合 $\{\mathbf{v}_1,\cdots,\mathbf{v}_p\}$（如果 $\mathbf{v}_1\neq\mathbf{0}$）是线性相关的，当且仅当某 $\mathbf{v}_j(j>1)$ 是其前面向量 $\mathbf{v}_1,\cdots,\mathbf{v}_{j-1}$ 的线性组合

定义 令 H 是向量空间 V 的一个子空间. V 中向量的指标集 $\mathcal{B}=\{\mathbf{b}_1,\cdots,\mathbf{b}_p\}$ 称为 H 的一个**基**，如果

（i） $\mathcal{B}$ 是一线性无关集

（ii） 由 $\mathcal{B}$ 生成的子空间与 H 相同，即 $H=\mathrm{Span}\{\mathbf{b}_1,\cdots,\mathbf{b}_p\}$

令 A 是一个可逆的 $n\times n$ 矩阵，比如 $A=\{\mathbf{a}_1,\cdots,\mathbf{a}_n\}$。则由可逆矩阵定理，A 的列组成 $\mathbb{R}^n$ 的一个基，这是因为它们是线性无关的且它们可以生成

$\mathbb{R}^n$

令 $\mathbf{e}_1,\cdots,\mathbf{e}_n$ 是 $n\times n$ 单位矩阵 I_n 的列，即

$$\mathbf{e}_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\mathbf{e}_2=\begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix},\cdots,\mathbf{e}_n=\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}$$

集合 $\{\mathbf{e}_1,\cdots,\mathbf{e}_n\}$ 称为 $\mathbb{R}^n$ 的**标准基**

定理 5 (生成集定理)
令 $S=\{\mathbf{v}_1,\cdots,\mathbf{v}_p\}$ 是 V 中的向量集，$H=\mathrm{Span}\{\mathbf{v}_1,\cdots,\mathbf{v}_p\}$

（i） 若 S 中某一个向量（比如说 $\mathbf{v}_k$）是 S 中其余向量的线性组合，则 S 中去掉 $\mathbf{v}_k$ 后形成的集合仍然可以生成 H

（ii） 若 $H\neq\{\mathbf{0}\}$，则 S 的某一子集是 H 的一个基

当 $\mathrm{Nul}\,A$ 包含非零向量时，我们的方法总可以产生一个线性无关集，从而由该方法可以得到 $\mathrm{Nul}\,A$ 的一个基

定理 6　矩阵 A 的主元列构成 $\mathrm{Col}\,A$ 的一个基

基是尽可能大的线性无关集

坐标系

定理 7 (唯一表示定理)
令 $\mathcal{B}=\{\mathbf{b}_1,\cdots,\mathbf{b}_n\}$ 是向量空间 V 的一个基，则对 V 中每个向量 $\boldsymbol{x}$，存在唯一的一组数 $c_1,\cdots,c_n$ 使得

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$$

定义　假设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是 V 的一个基，$\boldsymbol{x}$ 在 V 中，$\boldsymbol{x}$ 相对于基$\mathcal{B}$的坐标（或 $\boldsymbol{x}$ 的$\mathcal{B}$-坐标）是使得 $\boldsymbol{x}=c_1\boldsymbol{b}_1+\cdots+c_n\boldsymbol{b}_n$ 的权 $c_1,\cdots,c_n$
若 $c_1,\cdots,c_n$ 是 $\boldsymbol{x}$ 的 $\mathcal{B}$-坐标，则 $\mathbb{R}^n$ 中的向量

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

是 $\boldsymbol{x}$（相对于$\mathcal{B}$）的坐标向量或 $\boldsymbol{x}$ 的$\mathcal{B}$-坐标向量，映射 $x\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 称为（由$\mathcal{B}$ 确定的）坐标映射

令

$$P_{\mathcal{B}}=[\boldsymbol{b}_1\quad \boldsymbol{b}_2\quad \cdots\quad \boldsymbol{b}_n]$$

则向量方程

$$\boldsymbol{x}=c_1\boldsymbol{b}_1+c_2\boldsymbol{b}_2+\cdots+c_n\boldsymbol{b}_n$$

等价于

$$\boldsymbol{x}=P_{\mathcal{B}}[\boldsymbol{x}]_{\mathcal{B}}$$

我们称 $P_{\mathcal{B}}$ 为从 $\mathcal{B}$ 到 $\mathbb{R}^n$ 中标准基的**坐标变换矩阵**

定理 8　令 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 是向量空间 V 的一个基，则坐标映射 $\boldsymbol{x}\mapsto[\boldsymbol{x}]_{\mathcal{B}}$ 是一个由 V 映上到 $\mathbb{R}^n$ 的一对一的线性变换

向量空间的维数

定理 9　若向量空间 V 具有一组基 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$，则 V 中任意包含多余 n 个向量的集合一定线性相关

定理 10　若向量空间 V 有一组基含有 n 个向量，则 V 的每一组基一定恰好含有 n 个向量

定义　若 V 由一个有限集生成，则 V 称为**有限维的**，V 的维数写成 $\dim V$，是 V 的基中向量的个数. 零向量空间 $\{\mathbf{0}\}$ 的维数定义为零. 如果 V 不是由一有限集生成，则 V 称为**无穷维的**

定理 11　令 H 是有限维向量空间 V 的子空间，若有必要的话，H 中任一个线性无关集均可以扩充为 H 的一个基. H 也是有限维的并且

$$\dim H \leqslant \dim V$$

定理 12 (基定理)
令 V 是一个 p 维向量空间，$p\geqslant 1$，V 中任意含有 p 个元素的线性无关集必然是 V 的一个基. 任意含有 p 个元素且生成 V 的集合自然是 V 的一个基

$\operatorname{Nul} A$ 的维数是方程 $A\boldsymbol{x}=\mathbf{0}$ 中自由变量的个数，$\operatorname{Col} A$ 的维数是 A 中主元列的个数

秩

矩阵 A 中线性无关列的最大个数和 A^T 中线性无关列的最大个数 (即 A 中线性无关行的最大个数) 是相同的，这个公共值是矩阵 A 的**秩**

若 A 是一个 $m \times n$ 矩阵，A 的每一行具有 n 个元素，即可以视为 $\mathbb{R}^n$ 中一个向量. 其行向量的所有线性组合的集合称为 A 的**行空间**，记为 $\mathrm{Row}\, A$

定理 13　若两个矩阵 A 和 B 行等价，则它们的行空间相同. 若 B 是阶梯形矩阵，则 B 的非零行构成 A 的行空间的一个基同时也是 B 的行空间的一个基

定义　A 的秩即 A 的列空间的维数

定理 14 (秩定理)
$m \times n$ 矩阵 A 的列空间和行空间的维数相等，这个公共的维数（即 A 的秩）还等于 A 的主元位置的个数且满足方程

$$\mathrm{rank}\, A + \dim \mathrm{Nul}\, A = n$$

定理 15 (可逆矩阵定理（续）)
令 A 是一个 $n \times n$ 矩阵，则下列命题中的每一个均等价于 A 是可逆矩阵：

m. A 的列构成 $\mathbb{R}^n$ 的一个基
n. $\mathrm{Col}\, A = \mathbb{R}^n$
o. $\dim \mathrm{Col}\, A = n$
p. $\mathrm{rank}\, A = n$
q. $\mathrm{Nul}\, A = \{\mathbf{0}\}$
r. $\dim \mathrm{Nul}\, A = 0$

基的变换

定理 16　设 $\mathcal{B}=\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$ 和 $\mathcal{C}=\{\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n\}$ 是向量空间 V 的基，则存在一个 $n\times n$ 矩阵 $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathbf{P}}$ 使得

$$[x]_{\mathcal{C}}=\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathbf{P}}[\boldsymbol{x}]_{\mathcal{B}}$$

$\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathbf{P}}$ 的列是基 $\mathcal{B}$ 中向量的 $\mathcal{C}$-坐标向量，即

$$\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathbf{P}}=[[\boldsymbol{b}_1]_{\mathcal{C}}\quad[\boldsymbol{b}_2]_{\mathcal{C}}\quad\cdots\quad[\boldsymbol{b}_n]_{\mathcal{C}}]$$

差分方程中的应用

$$\begin{bmatrix}u_k & v_k & w_k\\ u_{k+1} & v_{k+1} & w_{k+1}\\ u_{k+2} & v_{k+2} & w_{k+2}\end{bmatrix}\begin{bmatrix}c_1\\ c_2\\ c_3\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\qquad \text{对所有}k\text{成立} \tag{4.3}$$

这个方程组的系矩阵称为信号的 **Casorati 矩阵**

如果对至少一个 k 值 Casorati 矩阵可逆，则(4.3)将蕴含 $c_1=c_2=c_3=0$，这就证明这三个信号是线性无关的

给定数 $a_0,\cdots,a_n,a_0$ 和 a_n 不为零，给定一个信号 $\{z_k\}$，方程

$$a_0y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=z_k\qquad\text{对所有}k\text{成立}$$

称为一个 n **阶线性差分方程**（或**线性递归关系**）
若 $\{z_k\}$ 是零序列，则方程是**齐次的**；否则，方程为**非齐次的**

定理 17　若 $a_n\neq0$ 且 $\{z_k\}$ 给定，只要 $y_0,\cdots,y_{n-1}$ 给定，方程

$$y_{k+n}+a_1y_{k+n-1}+\cdots+a_{n-1}y_{k+1}+a_ny_k=z_k$$

对所有k成立 有唯一解

定理 18 n 阶齐次线性差分方程

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_{n-1} y_{k+1} + a_n y_k = 0 \qquad \text{对所有}k\text{成立}$$

的解集 H 是一个 n 维向量空间

一个具有非负元素且各元素的数值相加等于 1 的向量称为**概率向量**

各向量均为概率向量的方阵为**随机矩阵**

马尔科夫链是一个概率向量序列 $\boldsymbol{x}_0$,$\boldsymbol{x}_1$,$\boldsymbol{x}_2$,$\cdots$ 和一个随机矩阵 P，满足

$$\boldsymbol{x}_1 = P\boldsymbol{x}_0, \boldsymbol{x}_2 = P\boldsymbol{x}_1, \boldsymbol{x}_3 = P\boldsymbol{x}_2, \cdots$$

用一阶差分方程描述:

$$bmx_{k+1} = P\boldsymbol{x}_k, k = 0, 1, 2, \cdots$$

若 P 是一个随机矩阵，则相对于 P 的**稳态向量**（或**平衡向量**）是一个满足

$$P\boldsymbol{q} = \boldsymbol{q}$$

的概率向量 $\boldsymbol{q}$
每一个随机矩阵有一个稳态向量

如果矩阵的某次幂 P^k 仅包含严格正的元素，则随机矩阵是正则的

一个向量序列 $\{\boldsymbol{x}_k : k = 1, 2, \cdots\}$ 当 $k \to \infty$ 时**收敛**到一个向量 $\boldsymbol{q}$，如果当 k 充分大时，$\boldsymbol{x}_k$ 中的元素无线接近 $\boldsymbol{q}$ 中对应的元素

定理 19　若 P 是一个 $n \times n$ 的正则随机矩阵, 则 P 具有唯一的稳态向量 $\boldsymbol{q}$. 进一步, 若 $\boldsymbol{x}_0$ 是任一个初始状态, 且 $\boldsymbol{x}_{k+1} = P\boldsymbol{x}_k, k = 0, 1, 2, \cdots$, 则当 $k \to \infty$ 时, 马尔科夫链 $\{\boldsymbol{x}_k\}$ 收敛到 $\boldsymbol{q}$

5 特征值与特征向量

特征向量与特征值

定义 A 为 $n \times n$ 矩阵，x 为非零向量，若存在数 λ 使 $A\boldsymbol{x} = \lambda\boldsymbol{x}$ 有非平凡解 $\boldsymbol{x}$，则称 λ 为 A 的特征值，$\boldsymbol{x}$ 称为对应于 λ 的特征向量

λ 是 A 的特征值当且仅当方程

$$(A - \lambda I)\boldsymbol{x} = \mathbf{0} \tag{5.1}$$

有非平凡解
方程(5.1)的所有解的集合就是矩阵 $A - \lambda I$ 的零空间
该集合是 $\mathbb{R}^n$ 的子空间，称为 A 的对应于 λ 的**特征空间**
特征空间由零向量和所有对应于 λ 的特征向量组成

定理 1 三角矩阵的主对角线的元素是其特征值

定理 2 $\lambda_1,\cdots,\lambda_r$ 是 $n\times n$ 矩阵 A 相异的特征值，$\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r$ 是与 $\lambda_1,\cdots,\lambda_r$ 对应的特征向量，那么向量集合 $\{\boldsymbol{v}_1,\cdots,\boldsymbol{v}_r\}$ 线性无关

特征方程

设 A 是 $n\times n$ 矩阵, U 是对 A 作行替换和行交换(不作行倍乘)所得到的任一阶梯形矩阵, r 是行交换的次数, 那么 A 的行列式 $\det A=(-1)^r u_{11}\cdots u_{nn}$. 如果 A 可逆, 那么 u_{11},$\cdots$,u_{nn} 都是主元（因为 $A\sim I_n$ 且 u_{ii} 没有归一化）. 否则, 至少有 u_{nn} 为零, 从而乘积 $u_{11}\cdots u_{nn}$ 为零. 因此

$$\det A=\begin{cases}(-1)^r u_{11}\cdots u_{nn} & A\\ 0 & A\end{cases}$$

定理 3 (可逆矩阵定理（续）)

设 A 是 $n\times n$ 矩阵，则 A 是可逆的当且仅当

s.0 不是 A 的特征值

t.A 的行列式不等于零

定理 4 (行列式的性质)

设 A 和 B 是 $n\times n$ 矩阵

a.A 可逆的充要条件是 $\det A\neq 0$

b.$\det AB=(\det A)(\det B)$

c.$\det A^T=\det A$

d. 若 A 是三角形矩阵，那么 $\det A$ 是 A 主对角线元素的乘积

e. 对 A 作行替换不改变其行列式值. 作一次行交换，行列式值符号改变一次. 数乘一行后，行列式值等于用此数乘原来的行列式值

数值方程 $\det(A-\lambda I)=0$ 称为 A 的**特征方程**

数 λ 是 $n\times n$ 矩阵 A 的特征值的充要条件是 λ 是特征方程 $\det(A-\lambda I)=0$ 的根

如果 A 是 $n\times n$ 矩阵, 那么 $\det(A-\lambda I)$ 是 n 次多项式, 称为 A 的**特征多项式**

把特征值 λ 作为特征方程根出现的次数称为 λ 的 **（代数）重数**

假如 A 和 B 是 $n \times n$ 矩阵，如果存在可逆矩阵 P，使得 $P_{-1}AP = B$，或等价地 $A = PBP_{-1}$，则称 **A 相似于 B**. 或简单说 A 和 B 是**相似的**. 把 A 变成 $P_{-1}AP$ 的变换称为**相似变换**

定理 5 　*若 $n \times n$ 矩阵 A 和 B 是相似的，那么它们有相同的特征多项式，从而有相同的特征值（和相同的重数）*

对角化

若

$$D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

则

$$对k \geqslant 1, D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}$$

若给定
$A = PDP^{-1}$
因此
$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PD^2P^{-1}$
同理
$A^3 = (PDP^{-1})A^2 = (PDP^{-1})(PD^2P^{-1}) = PD(P^{-1}P)D^2P^{-1} = PD^3P^{-1}$
一般对 $k \geqslant 1$，有

$$A^k = PD^K P^{-1}$$

如果方阵 A 相似于对角矩阵，即存在可逆矩阵 P 和对角矩阵 D，有 $A = PDP^{-1}$，则称 A **可对角化**

定理 6 (对角化定理)
$n \times n$ 矩阵 A 可对角化的充分必要条件是 A 有 n 个线性无关的特征向量
事实上,$A = PDP^{-1}$，D 为对角矩阵的充分必要条件是 P 的列向量是 A 的 n 个线性无关的特征向量. 此时，D 的主对角线上的元素分别是 A 的对应于 P 中特征向量的特征值

A 可对角化的充分必要条件是有足够的特征向量形成 $\mathbb{R}^n$ 的基，我们称这样的基为**特征向量基**

对角化步骤:
1. 求出 A 的特征值
2. 求 A 的三个线性无关的特征向量
3. 使用特征向量构造矩阵 P
4. 用与特征向量顺序对应的特征值构造矩阵 D

定理 7 有 n 个相异特征值的 $n \times n$ 矩阵可对角化

定理 8 设 A 是 $n \times n$ 矩阵，其相异的特征值是 $\lambda_1,\cdots,\lambda_p$
a. 对于 $a \leqslant k \leqslant p$，$\lambda_k$ 的特征空间的维数小于或等于 λ_k 的代数重数
b. 矩阵 A 可对角化的充分必要条件是所有不同特征空间的维数之和为 n. 即

（i）特征多项式可完全分解为线性因子

（ii）每个 λ_k 的特征空间的维数等于 λ_k 的代数重数

c. 若 A 可对角化，$\mathcal{B}_k$ 是对应于 λ_k 的特征空间的基，则集合 $\mathcal{B}_1,\cdots,\mathcal{B}_p$ 中所有向量的集合是 $\mathbb{R}^n$ 的特征向量基

特征向量与线性变换

设 V 是 n 维向量空间，W 是 m 维向量空间，T 是 V 到 W 的线性变换. V 的基 $\mathcal{B}$ 是 $\{\boldsymbol{b}_1,\cdots,\boldsymbol{b}_n\}$. 若 $\boldsymbol{x}=r_1\boldsymbol{b}_1+\cdots+r_n\boldsymbol{b}_n$，则

$$[\boldsymbol{x}]_{\mathcal{B}}=\begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$$

因为 T 是线性的，故

$$T(\boldsymbol{x})=T(r_1\boldsymbol{b}_1+\cdots+r_n\boldsymbol{b}_n)=r_1T(\boldsymbol{b}_1)+\cdots+r_nT(\boldsymbol{b}_n) \tag{5.2}$$

因为从 W 到 $\mathbb{R}^m$ 的坐标映射是线性的，故等式(5.2)可推出

$$[T(\boldsymbol{x})]_{\mathcal{C}}=r_1[T(\boldsymbol{b}_1)]_{\mathcal{C}}+\cdots+r_n[T(\boldsymbol{b}_n)]_{\mathcal{C}} \tag{5.3}$$

因为这些 $\mathcal{C}$-坐标向量都属于 $\mathbb{R}^m$，故向量等式(5.3)可以写为矩阵等式

$$[T(\boldsymbol{x})]_{\mathcal{C}}=M[\boldsymbol{x}]_{\mathcal{B}}$$

其中

$$M=[[T(\boldsymbol{b}_1)]_{\mathcal{C}}\quad [T(\boldsymbol{b}_2)]_{\mathcal{C}}\quad \cdots \quad [T(\boldsymbol{b}_n)]_{\mathcal{C}}] \tag{5.4}$$

矩阵 M 是 T 的矩阵表示，称为 **T 相对于基 $\mathcal{B}$ 和 $\mathcal{C}$ 的矩阵**

当 W=V，$\mathcal{C}=\mathcal{B}$ 时，(5.4)中的 M 称为 **T 相对于 $\mathcal{B}$ 的矩阵**，或简称为 **T 的 $\mathcal{B}$-矩阵**，记为 $[T]_{\mathcal{B}}$

定理 9（（对角矩阵表示））
设 $A=PDP^{-1}$，其中 D 为 $n\times n$ 对角矩阵，若 $\mathbb{R}^n$ 的基 $\mathcal{B}$ 由 P 的列向量组成，那么 D 是变换 $\boldsymbol{x}\mapsto A\boldsymbol{x}$ 的 $\mathcal{B}$-矩阵

复特征值

一个复数 λ 满足 $\det(A-\lambda I)=0$ 当且仅当在 $\mathbb{C}^n$ 中存在一个非零向量 $\boldsymbol{x}$, 使得 $A\boldsymbol{x}=\lambda\boldsymbol{x}$. 我们称这样的 λ 是(复)特征值, $\boldsymbol{x}$ 是对应于 λ 的(复)特征向量

$\mathbb{C}^n$ 中复向量 $\boldsymbol{x}$ 的共轭向量 $\bar{\boldsymbol{x}}$ 也是 $\mathbb{C}^n$ 中的向量，它的分量是 $\boldsymbol{x}$ 中对应分量的共轭复数，向量 $\operatorname{Re}\boldsymbol{x}$ 和 $\operatorname{Im}\boldsymbol{x}$ 称为复向量 $\boldsymbol{x}$ 的**实部**和**虚部**，分别由 $\boldsymbol{x}$ 的分量的实部和虚部组成

定理 10　设 A 是 2×2 实矩阵，有复特征值 $\lambda=a-bi(b\neq 0)$ 及对应的 $\mathbb{C}^2$ 中的复特征向量 $\boldsymbol{v}$，那么

$$A=PCP^{-1},\quad \text{其中} P=[\operatorname{Re}\boldsymbol{v}\quad \operatorname{Im}\boldsymbol{v}], C=\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

6

正交性和最小二乘法

内积、长度和正交性

如果 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是 $\mathbb{R}^n$ 中的向量, 则可以将 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 作为 $n\times 1$ 矩阵. 转置矩阵 $\boldsymbol{u}^T$ 是 $1\times n$ 矩阵, 且矩阵乘积 $\boldsymbol{u}^T\boldsymbol{v}$ 是一个 1×1 矩阵, 我们将其记为一个不加括号的实数(标量). 数 $\boldsymbol{u}^T\boldsymbol{v}$ 称为 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 的**内积**, 通常记作 $\boldsymbol{u}\cdot\boldsymbol{v}$, 也称为**点积**

定理 1 设 $\boldsymbol{v}$,$\boldsymbol{u}$ 和 $\boldsymbol{w}$ 是 $\mathbb{R}^n$ 中的向量,c 是一个数，那么

a. $\boldsymbol{u}\cdot\boldsymbol{v}=\boldsymbol{v}\cdot\boldsymbol{u}$

b. $(\boldsymbol{u}+\boldsymbol{v})\cdot\boldsymbol{w}=\boldsymbol{u}\cdot\boldsymbol{w}+\boldsymbol{v}\cdot\boldsymbol{w}$

c. $(c\boldsymbol{u})\cdot v=c(\boldsymbol{u}\cdot\boldsymbol{v})=\boldsymbol{u}\cdot(c\boldsymbol{v})$

d. $\boldsymbol{u}\cdot\boldsymbol{u}\geqslant 0$, 并且 $\boldsymbol{u}\cdot\boldsymbol{u}=0$ 成立的充分必要条件是 $\boldsymbol{u}=\boldsymbol{0}$

定义 向量 $\boldsymbol{v}$ 的**长度**（**或范数**）是非负数 $||\boldsymbol{v}||$, 定义为

$$||\boldsymbol{v}||=\sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}=\sqrt{v_1^2+v_2^2+\cdots+v_n^2} \text{ 且 } ||\boldsymbol{v}||^2=\boldsymbol{v}\cdot\boldsymbol{v}$$

长度为 1 的向量称为**单位向量**. 如果把一个非零向量除以其自身的长度, 即乘 $\dfrac{1}{||\boldsymbol{v}||}$, 就可以得到一个单位向量, 即 $\boldsymbol{u}=\dfrac{\boldsymbol{v}}{||\boldsymbol{v}||}$

把向量 $\boldsymbol{v}$ 化成单位向量 $\boldsymbol{u}$ 的过程，称为向量 $\boldsymbol{v}$ 的**单位化**

定义　*如果 $\boldsymbol{u}\cdot\boldsymbol{v}=0$，则 $\mathbb{R}^n$ 中的两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 是（相互）*正交的

定理 2 (毕达哥拉斯（勾股）定理)
两个向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 正交的充分必要条件是 $||\boldsymbol{u}+\boldsymbol{v}||^2=||\boldsymbol{u}||^2+||\boldsymbol{v}||^2$

如果向量 $\boldsymbol{z}$ 与 $\mathbb{R}^n$ 的子空间 W 中的任意向量都正交，则称 $\boldsymbol{z}$ **正交于 W**

与子空间 W 正交的向量 $\boldsymbol{z}$ 的全体组成的集合称为 W 的**正交补**，记作 $W^\perp$

1. 向量 $\boldsymbol{x}$ 属于 $W^\perp$ 的充分必要条件是向量 $\boldsymbol{x}$ 与生成空间 W 的任一向量都正交 2.$W^\perp$ 是 $\mathbb{R}^n$ 的一个子空间

定理 3　*假设 A 是 $m\times n$ 矩阵，那么 A 的行空间的正交补是 A 的零空间，且 A 的列空间的正交补是 A^T 的零空间:*

$$(\mathrm{Row}\,A)^\perp=\mathrm{Nul}\,A \text{ 且 } (\mathrm{Col}\,A)^\perp=\mathrm{Nul}\,A^T$$

正交集

$\mathbb{R}^n$ 中的向量集合 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 称为**正交集**，如果集合中的任意两个不同向量都正交，即当 $i\neq j$ 时，$\boldsymbol{u_i}\cdot\boldsymbol{u_j}=0$

定理 4 如果 $S=\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是由 $\mathbb{R}^n$ 中非零向量构成的正交集，那么 S 是线性无关集，因此构成 S 所生成的子空间的一组基

定义 $\mathbb{R}^n$ 中子空间 W 的一个**正交基**是 W 的一个基，也是正交集

定理 5 假设 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 $\mathbb{R}^n$ 中子空间 W 的正交基，对 W 中的每个向量 $\boldsymbol{y}$，线性组合 $\boldsymbol{y}=c_1\boldsymbol{u}_1+\cdots+c_p\boldsymbol{u}_p$ 中的权可以由 $c_j=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u_j}}{\boldsymbol{u_j}\cdot\boldsymbol{u_j}}$（$j=1,\cdots,p$）计算

对 $\mathbb{R}^n$ 中给出的非零向量 $\boldsymbol{u}$，考虑 $\mathbb{R}^n$ 中一个向量 $\boldsymbol{y}$ 分解为两个向量之和的问题，一个向量是向量 $\boldsymbol{u}$ 的倍数，另一个向量与 $\boldsymbol{u}$ 正交. 我们期望写成

$$\boldsymbol{y}=\hat{\boldsymbol{y}}+\boldsymbol{z} \tag{6.1}$$

其中 $\hat{\boldsymbol{y}}=\alpha\boldsymbol{u}$，$\alpha$ 是一个数，$\boldsymbol{z}$ 是一个垂直于 $\boldsymbol{u}$ 的向量. 对给定数 α，记 $\boldsymbol{z}=\boldsymbol{y}-\alpha\boldsymbol{u}$，则方程(6.1)可以满足. 那么 $\boldsymbol{y}-\hat{\boldsymbol{y}}$ 和 $\boldsymbol{u}$ 正交的充分必要条件是

$$(\boldsymbol{y}-\alpha\boldsymbol{u})\cdot\boldsymbol{u}=\boldsymbol{y}\cdot\boldsymbol{u}-(\alpha\boldsymbol{u})\cdot\boldsymbol{u}=\boldsymbol{y}\cdot\boldsymbol{u}-\alpha(\boldsymbol{u}\cdot\boldsymbol{u})=0$$

也就是满足方程(6.1)且 $\boldsymbol{z}$ 与 $\boldsymbol{u}$ 正交的充分必要条件是 $\alpha=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}$ 和 $\hat{\boldsymbol{y}}=\operatorname{proj}_L\boldsymbol{y}=\dfrac{\boldsymbol{y}\cdot\boldsymbol{u}}{\boldsymbol{u}\cdot\boldsymbol{u}}\cdot\boldsymbol{u}$. 向量 $\hat{\boldsymbol{y}}$ 称为 **$\boldsymbol{y}$ 在 $\boldsymbol{u}$ 上的正交投影**，向量 $\boldsymbol{z}$ 称为 **$\boldsymbol{y}$ 与 $\boldsymbol{u}$ 正交的分量**

如果集合 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是由单位向量构成的正交集，那么它是一个**单位正交集**

如果 W 是一个由单位正交集合生成的子空间，那么 $\{\boldsymbol{u}_1,\cdots,\boldsymbol{u}_p\}$ 是 W 的**单位正交基**

定理 6　一个 $m \times n$ 矩阵 U 具有单位正交列向量的充分必要条件是 $U^TU = I$

定理 7　假设 U 是一个具有单位正交列的 $m\times n$ 矩阵，且 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的向量，那么

a. $||U\boldsymbol{x}|| = ||\boldsymbol{x}||$

b. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$

c. $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = 0$ 的充分必要条件是 $\boldsymbol{x} \cdot \boldsymbol{y} = 0$

正交投影

对给定向量 $\boldsymbol{y}$ 和 $\mathbb{R}^n$ 中子空间 W，存在属于 W 的向量 $\hat{\boldsymbol{y}}$ 满足:

(1) W 中有唯一向量 $\hat{\boldsymbol{y}}$，使得 $\boldsymbol{y} - \hat{\boldsymbol{y}}$ 与 W 正交

(2) $\hat{\boldsymbol{y}}$ 是 W 中唯一最接近 $\boldsymbol{y}$ 的向量

定理 8 (正交分解定理)

若 W 是 $\mathbb{R}^n$ 的一个子空间，那么 $\mathbb{R}^n$ 中每一个向量 $\boldsymbol{y}$ 可以唯一表示为

$$\boldsymbol{y} = \hat{\boldsymbol{y}} + \boldsymbol{z}$$

其中 $\hat{\boldsymbol{y}}$ 属于 W 而 $\boldsymbol{z}$ 属于 $W^{\perp}$. 实际上，如果 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是 W 的任意正交基，那么

$$\hat{\boldsymbol{y}} = \frac{\boldsymbol{y} \cdot \boldsymbol{u_1}}{\boldsymbol{u_1} \cdot \boldsymbol{u_1}}\boldsymbol{u}_1 + \cdots + \frac{\boldsymbol{y} \cdot \boldsymbol{u_p}}{\boldsymbol{u_p} \cdot \boldsymbol{u_p}}\boldsymbol{u}_p$$

且 $\boldsymbol{z} = \boldsymbol{y} - \hat{\boldsymbol{y}}$

定理 9 (最佳逼近定理)

假设 W 是 $\mathbb{R}^n$ 的一个子空间，$\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的任意向量，$\hat{\boldsymbol{y}}$ 是 $\boldsymbol{y}$ 在 W 上的正交投影，那么 $\hat{\boldsymbol{y}}$ 是 W 中最接近 $\boldsymbol{y}$ 的点，也就是

$$||\boldsymbol{y} - \hat{\boldsymbol{y}}|| < ||\boldsymbol{y} - \boldsymbol{v}||$$

对所有属于 W 又异于 $\hat{\boldsymbol{y}}$ 的 $\boldsymbol{v}$ 成立

定理 10　如果 $\{\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p\}$ 是 $\mathbb{R}^n$ 中子空间 W 的单位正交基，那么

$$\mathrm{proj}_W\, \boldsymbol{y} = (\boldsymbol{y}\cdot\boldsymbol{u}_1)\boldsymbol{u}_1 + (\boldsymbol{y}\cdot\boldsymbol{u}_2)\boldsymbol{u}_2 + \cdots + (\boldsymbol{y}\cdot\boldsymbol{u}_p)\boldsymbol{u}_p$$

如果 $U = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \cdots \quad \boldsymbol{u}_p]$，则

$$\mathrm{proj}_W\, \boldsymbol{y} = UU^T\boldsymbol{y},\ \text{对所有}\boldsymbol{y}\in\mathbb{R}^n\text{成立}$$

格拉姆-施密特方法

定理 11 (格拉姆-施密特方法)
对 $\mathbb{R}^n$ 的子空间 W 的一个基 $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_p\}$，定义

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{x}_1\\
\boldsymbol{v}_2 &= \boldsymbol{x}_2 - \frac{\boldsymbol{x}_2\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1\\
\boldsymbol{v}_3 &= \boldsymbol{x}_3 - \frac{\boldsymbol{x}_3\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2 - \frac{\boldsymbol{x}_3\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1\\
&\vdots\\
\boldsymbol{v}_p &= \boldsymbol{x}_p - \frac{\boldsymbol{x}_p\cdot\boldsymbol{v}_{p-1}}{\boldsymbol{v}_{p-1}\cdot\boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1} - \cdots - \frac{\boldsymbol{x}_p\cdot\boldsymbol{v}_2}{\boldsymbol{v}_2\cdot\boldsymbol{v}_2}\boldsymbol{v}_2 - \frac{\boldsymbol{x}_p\cdot\boldsymbol{v}_1}{\boldsymbol{v}_1\cdot\boldsymbol{v}_1}\boldsymbol{v}_1
\end{aligned}$$

那么 $\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_p\}$ 是 W 的一个正交基. 此外，

$$\mathrm{Span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\} = \mathrm{Span}\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_k\},\ \text{其中}1 \leqslant k \leqslant p$$

定理 12 (QR 分解)
如果 $m \times n$ 矩阵 A 的列线性无关，那么 A 可以分解为 A=QR，其中 Q 是一个 $m \times n$ 矩阵，其列形成 $\mathrm{Col}\, A$ 的一个标准正交基，R 是一个 $n \times n$ 上三角可逆矩阵且在对角线上的元素为正数

最小二乘问题

当方程组的解不存在但又需要求解时，最好的方法是寻求 $\boldsymbol{x}$，使得 $A\boldsymbol{x}$ 尽可能接近 $\boldsymbol{b}$

定义　如果 $m\times n$ 矩阵 A 和向量 $\boldsymbol{b}$ 属于 $\mathbb{R}^m$，则 $A\boldsymbol{x}=\boldsymbol{b}$ 的**最小二乘解**是 $\mathbb{R}^n$ 中的 $\hat{\boldsymbol{x}}$，使得

$$||\boldsymbol{b}-A\hat{\boldsymbol{x}}||\leqslant||\boldsymbol{b}-A\boldsymbol{x}||$$

对所有 $\boldsymbol{x}\in\mathbb{R}^n$ 成立

对给定的 A 和 $\boldsymbol{b}$，应用最佳逼近定理于子空间 $\mathrm{Col}\,A$. 取

$$\hat{\boldsymbol{b}}=\mathrm{proj}_{\mathrm{Col}\,A}\,\boldsymbol{b}$$

由于 $\hat{\boldsymbol{b}}$ 属于 A 的列空间，故方程 $A\boldsymbol{x}=\hat{\boldsymbol{b}}$ 是相容的且存在一个属于 $\mathbb{R}^n$ 的 $\hat{\boldsymbol{x}}$ 使得

$$A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}\tag{6.2}$$

由于 $\hat{\boldsymbol{b}}$ 是 $\mathrm{Col}\,A$ 中最接近 $\boldsymbol{b}$ 的点，因此一个向量 $\hat{\boldsymbol{x}}$ 是 $A\boldsymbol{x}=\boldsymbol{b}$ 的一个最小二乘解的充分必要条件是 $\hat{\boldsymbol{x}}$ 满足(6.2). 这个属于 $\mathbb{R}^n$ 的 $\hat{\boldsymbol{x}}$ 是一系列由 A 的列构造的 $\hat{\boldsymbol{b}}$ 的权

若 $\hat{\boldsymbol{x}}$ 满足 $A\hat{\boldsymbol{x}}=\hat{\boldsymbol{b}}$, 则由正交分解定理, 投影 $\hat{\boldsymbol{b}}$ 具有性质 $\boldsymbol{b}-\hat{\boldsymbol{b}}$ 与 $\mathrm{Col}\,A$ 正交, 即 $\boldsymbol{b}-A\hat{\boldsymbol{x}}$ 正交于 A 的每一列. 如果 $\boldsymbol{a}_j$ 是 A 的任意列, 那么 $\boldsymbol{a}_j\cdot(\boldsymbol{b}-A\hat{\boldsymbol{x}})=0$. 由于每一个 $\boldsymbol{a}_j^T$ 是 A^T 的行，因此

$$A^T(\boldsymbol{b}-A\hat{\boldsymbol{x}})=\mathbf{0}$$

故有

$$A^T\boldsymbol{b}-A^TA\hat{\boldsymbol{x}}=\mathbf{0}$$

此计算表明 $A\boldsymbol{x}=\boldsymbol{b}$ 的每个最小二乘解满足方程

$$A^TA\boldsymbol{x}=A^T\boldsymbol{b}\tag{6.3}$$

矩阵方程(6.3)表示的线性方程组常称为 $A\boldsymbol{x} = \boldsymbol{b}$ 的**法方程**，(6.3)的解通常用 $\hat{\boldsymbol{x}}$ 表示

定理 13 *方程 $A\boldsymbol{x} = \boldsymbol{b}$ 的最小二乘解集和法方程 $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$ 的非空解集一致*

定理 14 *设 A 是 $m \times n$ 矩阵. 下面的条件是逻辑等价的:*
a) 对于 $\mathbb{R}^m$ 中的每个 $\boldsymbol{b}$, 方程 $A\boldsymbol{x} = \boldsymbol{b}$ 有唯一最小二乘解
b) A 的列是线性无关的
c) 矩阵 A^TA 是可逆的
当这些条件成立时, 最小二乘解 $\hat{\boldsymbol{x}}$ 有下面的表示:

$$\hat{\boldsymbol{x}} = (A^TA)^{-1}A^T\boldsymbol{b}$$

定理 15 *给定一个 $m \times n$ 矩阵 A, 它具有线性无关的列，取 $A = QR$ 是 A 类似定理 12 的 QR 分解, 那么对每一个属于 $\mathbb{R}^m$ 的 $\boldsymbol{b}$, 方程 $A\boldsymbol{x} = \boldsymbol{b}$ 有唯一的最小二乘解, 其解为*

$$\hat{\boldsymbol{x}} = R^{-1}Q^T\boldsymbol{b}$$

线性模型中的应用

为了更容易应用所讨论的实际问题，将 $A\boldsymbol{x} = \boldsymbol{b}$ 写成 $X\boldsymbol{\beta} = \boldsymbol{y}$，且称 X 为**设计矩阵**，$\boldsymbol{\beta}$ 为**参数向量**，$\boldsymbol{y}$ 为**观测向量**

变量 x 和 y 之间最简单的关系是线性方程 $y = \beta_0 + \beta_1 x$. 对应每一个数据点 (x_j, y_j)，有一个在直线上的点 $(x_j, \beta_0 + \beta_1 x_j)$ 具有同样的 x 坐标. 我们称 y_j 为 y 的**观测值**，而 $\beta_0 + \beta_1 x_j$ 位 y 的**预测值**. 观测 y 值和预测 y 值之

间的差称为**余差**

最小二乘直线 $y=\beta_0+\beta_1 x$ 是余差平方之和最小的，这条直线也被称为 $\boldsymbol{y}$ 对 $\boldsymbol{x}$ 的**回归直线**. 直线的系数 β_0,β_1 被称为（线性）**回归系数**

计算 $X\boldsymbol{\beta}=\boldsymbol{y}$ 的最小二乘问题等价于找出 $\boldsymbol{\beta}$

在计算最小二乘直线之前，常见的联系是计算原来 x 值的平均 $\bar{x}$，并形成一个新变量 $x^*=x-\bar{x}$. 新的 x 数据被称为**平均偏差形式**

在一些应用中，必须将数据点拟合为非直线形式. 引入**余差向量** $\boldsymbol{\varepsilon}$，定义为 $\boldsymbol{\varepsilon}=\boldsymbol{y}-X\boldsymbol{\beta}$，并且记住

$$\boldsymbol{y}=X\boldsymbol{\beta}+\boldsymbol{\varepsilon}$$

任何具有这种形式的方程称为**线性模型**. 一旦 X 和 $\boldsymbol{y}$ 被确定，使 $\boldsymbol{\varepsilon}$ 长度达到最小化相当于找出 $X\boldsymbol{\beta}=\boldsymbol{y}$ 的最小二乘解. 在每种情形下，最小二乘解 $\hat{\boldsymbol{\beta}}$ 是下面法方程的解:

$$X^TX\boldsymbol{\beta}=X^T\boldsymbol{y}$$

内积空间

定义　向量空间 V 上的内积是一个函数，对每一对属于 V 的向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$，存在一个实数 $\langle\boldsymbol{u},\boldsymbol{v}\rangle$ 满足下面公理，其中 $\boldsymbol{u}$,$\boldsymbol{v}$,$\boldsymbol{w}$ 属于 V，c 为所有数:

1. $\langle\boldsymbol{u},\boldsymbol{v}\rangle=\langle\boldsymbol{v},\boldsymbol{u}\rangle$

2. $\langle\boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle=\langle\boldsymbol{u},\boldsymbol{w}\rangle+\langle\boldsymbol{v},\boldsymbol{w}\rangle$

3. $\langle c\boldsymbol{u}, \boldsymbol{v}\rangle = c\langle \boldsymbol{u}, \boldsymbol{v}\rangle$

4. $\langle \boldsymbol{u}, \boldsymbol{u}\rangle \geqslant 0$且$\langle \boldsymbol{u}, \boldsymbol{u}\rangle = 0$的充分必要条件是$\boldsymbol{u} = 0$

一个赋予上面内积的向量空间称为**内积空间**

设 V 是一个内积空间，其内积记作 $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$. 像 $\mathbb{R}^n$ 中一样，我们定义一个向量 $\boldsymbol{v}$ 的**长度**或**范数**是数

$$||\boldsymbol{v}|| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}$$

即 $||\boldsymbol{v}|| = \langle \boldsymbol{v}, \boldsymbol{v}\rangle$

一个单位向量是长度为 1 的向量，向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 之间的距离是 $||\boldsymbol{u} - \boldsymbol{v}||$

向量 $\boldsymbol{u}$ 和向量 $\boldsymbol{v}$ 正交，如果 $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$ 成立

定理 16 (柯西-施瓦茨不等式)

对 V 中任意向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$, 有

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \leqslant ||\boldsymbol{u}||\ ||\boldsymbol{b}||$$

定理 17 (三角不等式)

对属于 V 的所有向量 $\boldsymbol{u}$,$\boldsymbol{v}$, 有

$$||\boldsymbol{u} + \boldsymbol{v}|| = ||\boldsymbol{u}|| + ||\boldsymbol{v}||$$

7

对称矩阵和二次型

对称矩阵的对角化

一个对称矩阵是一个满足 $A^T = A$ 的矩阵 A，这种矩阵当然是方阵，它的主对角线元素是任意的，但其他元素在主对角线的两边成对出现

定理 1 如果 A 是对称矩阵，那么不同特征空间的任意两个特征向量是正交的

一个矩阵 A 称为可**正交对角化**, 如果存在一个正交矩阵 P (满足 $P^{-1} = P^T$) 和一个对角矩阵 D 使得

$$A = PDP^T = PDP^{-1}$$

定理 2 一个 $n \times n$ 矩阵 A 可正交对角化的充分必要条件是 A 是对称矩阵

矩阵 A 的特征值的集合有时称为 A 的**谱**

定理 3 (对称矩阵的谱定理)
一个对称的 $n \times n$ 矩阵 A 具有下述性质:
a. A 有 n 个实特征值, 包含重复的特征值
b. 对每一个特征值 λ, 对应的特征空间的维数等于 λ 作为特征方程的根的重数
c. 特征空间相互正交, 这种正交性是在特征向量对应于不同特征值的意义下成立的
d. A 可正交对角化

二次型

计算 $\boldsymbol{x}^T\boldsymbol{x}$ 时的平方和及更一般形式的表达式称为**二次型**

$\mathbb{R}^n$ 上的一个二次型是一个定义在 $\mathbb{R}^n$ 上的函数，它在向量 $\boldsymbol{x}$ 处的值可由表达式 $Q(\boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x}$ 计算，其中 A 是一个 $n \times n$ 对称矩阵. 矩阵 A 称为**关于二次型的矩阵**

如果 $\boldsymbol{x}$ 表示 $\mathbb{R}^n$ 中的向量变量，那么**变量代换**是下面形式的等式:

$$\boldsymbol{x} = P\boldsymbol{y} \text{ 或 } \boldsymbol{y} = P^{-1}\boldsymbol{x} \tag{7.1}$$

其中 P 是可逆矩阵且 $\boldsymbol{y}$ 是 $\mathbb{R}^n$ 中的一个新的向量变量. 这里 P 的列可确定 $\mathbb{R}^n$ 的一个基，$\boldsymbol{y}$ 是相对于该基的向量 $\boldsymbol{x}$ 的坐标向量.
如果用变量代换(7.1)处理二次型 $\boldsymbol{x}^T A\boldsymbol{x}$，那么

$$\boldsymbol{x}^T A\boldsymbol{x} = (P\boldsymbol{y})^T A(P\boldsymbol{y}) = \boldsymbol{y}^T P^T AP\boldsymbol{y} = \boldsymbol{y}^T(P^T AP)\boldsymbol{y} \tag{7.2}$$

且新的二次型矩阵是 $P^T AP$. 因为 A 是对称的，故由定理 2，存在正交矩阵 P，使得 $P^T AP$ 是对角矩阵 D，(7.2)中的二次型变为 $\boldsymbol{y}^T D\boldsymbol{y}$

定理 4 (主轴定理)
设 A 是一个 $n \times n$ 对称矩阵, 那么存在一个正交变量代换 $bmx = P\boldsymbol{y}$, 它将二次型 $\boldsymbol{x}^T A\boldsymbol{x}$ 变换为不含交叉乘积项的二次型 $\boldsymbol{y}^T D\boldsymbol{y}$

矩阵 P 的列称为二次型 $\boldsymbol{x}^T A\boldsymbol{x}$ 的主轴

定义　一个二次型 Q 是:
a. 正定的，如果对所有 $x \neq 0$，有 $Q(\boldsymbol{x}) > 0$
b. 半正定的，如果对所有 $\boldsymbol{x}$，有 $Q(\boldsymbol{x}) \geqslant 0$
c. 负定的，如果对所有 $x \neq 0$，有 $Q(\boldsymbol{x}) < 0$
d. 半负定的，如果对所有 $\boldsymbol{x}$，有 $Q(\boldsymbol{x}) \leqslant 0$
e. 不定的，如果 $Q(\boldsymbol{x})$ 既有正值又有负值

定理 5 (二次型与特征值)
设 A 是 $n \times n$ 对称矩阵, 那么一个二次型 $\boldsymbol{x}^T A\boldsymbol{x}$ 是:
a. 正定的, 当且仅当 A 的所有特征值是正数
b. 负定的, 当且仅当 A 的所有特征值是负数
c. 不定的, 当且仅当 A 既有正特征值, 又有负特征值