高等工程數學 (3)

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黃 煒
博士，副教授
Dr. Wei Huang

矩阵的初等变换

矩阵论（3）

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄 炜 博士·副教授 Dr. Wei Huang

1.3 矩阵的初等变换及其应用

■ 1.3.1 矩阵的等价

矩阵的初等变换在矩阵论中有重要的作用. 所谓矩阵 $\boldsymbol{A}$ 的初等变换是指对 $\boldsymbol{A}$ 施行如下三种变换:

(1) 交换 $\boldsymbol{A}$ 的两行(列);

(2) 用一个非零数乘 $\boldsymbol{A}$ 的某一行(列);

(3) 用任一数乘以 $\boldsymbol{A}$ 的某一行(列)加到 $\boldsymbol{A}$ 的另一行(列).

在初等变换中如果仅限于行(列)的变换,称为**行(列)初等变换**.

对单位矩阵施行一次初等变换后得到的矩阵称为**初等矩阵**.

显然,初等矩阵都是方阵,每一种初等变换都有一个与之对应的初等矩阵.

1.3.1 矩阵的等价

第一种初等变换所对应的初等矩阵是

$$\boldsymbol{I}_{ij}=\begin{array}{l}\left[\begin{array}{ccccccc}1&&&&&&\\&\ddots&&&&&\\&&0&\cdots&1&&\\&&\vdots&\ddots&\vdots&&\\&&1&\cdots&0&&\\&&&&&\ddots&\\&&&&&&1\end{array}\right]\begin{array}{c}\\ \\ i\\ \\ j\\ \\ \\ \end{array}\\ \begin{array}{ccccccc}\quad&\quad&i&\quad&j&\quad&\quad\end{array}\end{array},$$

它表示把 $\boldsymbol{I}$ 中第 i、j 两行(列)进行交换.

第二种初等变换所对应的初等矩阵是

$$\boldsymbol{I}_i(k)=\begin{array}{l}\left[\begin{array}{ccccc}1&&&&\\&\ddots&&&\\&&k&&\\&&&\ddots&\\&&&&1\end{array}\right]\begin{array}{c}\\ \\ i\\ \\ \\ \end{array}\\ \begin{array}{ccccc}\quad&\quad&i&\quad&\quad\end{array}\end{array},$$

它表示把 $\boldsymbol{I}$ 中第 i 行(列)乘以非零数 k.

信息学院（国家示范性软件学院）School of Informatics 黄炜 博士/副教授 Dr. Wei Huang 厦門大學 XIAMEN UNIVERSITY

1.3.1 矩阵的等价

第三种初等变换所对应的初等矩阵是

$$
I_{ij}(k)=\begin{array}{cc}
\left[\begin{array}{ccccccc}
1&&&&&&\\
&\ddots&&&&&\\
&&1&&&&\\
&&\vdots&\ddots&&&\\
&&k&\cdots&1&&\\
&&&&&\ddots&\\
&&&&&&1
\end{array}\right] & \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array}\\
\begin{array}{ccccccc} \ \ &\ \ & i &\ \ & j &\ \ &\ \ \end{array} &
\end{array},\quad
I_{ij}^{\mathrm{T}}(k)=\begin{array}{cc}
\left[\begin{array}{ccccccc}
1&&&&&&\\
&\ddots&&&&&\\
&&1&\cdots&k&&\\
&&&\ddots&\vdots&&\\
&&&&1&&\\
&&&&&\ddots&\\
&&&&&&1
\end{array}\right] & \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array}\\
\begin{array}{ccccccc} \ \ &\ \ & i &\ \ & j &\ \ &\ \ \end{array} &
\end{array},
$$

其中 $I_{ij}(k)$ 表示把 I 第 i 行的 k 倍加到第 j 行上，或表示把 I 中第 j 列的 k 倍加到第 i 列上. 而 $I_{ij}^{\mathrm{T}}(k)$ 表示把 I 第 i 列的 k 倍加到第 j 列上，或表示把 I 中第 j 行的 k 倍加到第 i 行上.

初等矩阵都是可逆矩阵，且逆矩阵仍是初等矩阵. 事实上，有

$$I_{ij}^{-1}=I_{ij},\quad (I_i(k))^{-1}=I_i\left(\frac{1}{k}\right),\quad (I_{ij}(k))^{-1}=I_{ij}(-k).$$

用初等矩阵左(右)乘矩阵 A 相当于对 A 作一次相应的行(列)初等变换.

1.3.1 矩阵的等价

如果矩阵 $\boldsymbol{B}$ 可以由矩阵 $\boldsymbol{A}$ 经过有限次初等变换而得到，则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价，记作 $\boldsymbol{A} \cong \boldsymbol{B}$.

同型矩阵之间的等价，具有下列性质：

(1) 自反性　$\forall \boldsymbol{A}$(即对任意的 $\boldsymbol{A}$)，$\boldsymbol{A} \cong \boldsymbol{A}$；

(2) 对称性　若 $\boldsymbol{A} \cong \boldsymbol{B}$，则 $\boldsymbol{B} \cong \boldsymbol{A}$；

(3) 传递性　若 $\boldsymbol{A} \cong \boldsymbol{B}$，$\boldsymbol{B} \cong \boldsymbol{C}$，则 $\boldsymbol{A} \cong \boldsymbol{C}$.

形如

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

的矩阵称为**阶梯形矩阵**. 这种矩阵有如下特征：

(1) 零行(即元素全为零的行)全都位于非零行的下方；

(2) 各非零行左起第一个非零元素的列指标由上至下严格增大.

利用行初等变换可以把任一矩阵化成阶梯形矩阵.

厦門大學 XIAMEN UNIVERSITY 信息学院（国家示范性软件学院）School of Informatics 黄炜 博士/副教授 Dr. Wei Huang

1.3.1 矩阵的等价

例 1　将 $\boldsymbol{A}$ 化成阶梯形矩阵，其中

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 0 & 1\\ 0 & 2 & 1 & 2\\ -1 & 3 & 2 & 3\end{bmatrix}.$$

解　对 $\boldsymbol{A}$ 施行行初等变换，

$$\boldsymbol{A}\xrightarrow[\text{上第一行}]{\text{第三行加}}\begin{bmatrix}1 & 1 & 0 & 1\\ 0 & 2 & 1 & 2\\ 0 & 4 & 2 & 4\end{bmatrix}\xrightarrow[\text{第二行的两倍}]{\text{第三行减去}}\begin{bmatrix}1 & 1 & 0 & 1\\ 0 & 2 & 1 & 2\\ 0 & 0 & 0 & 0\end{bmatrix}. \quad ■$$

还可进一步对 $\boldsymbol{A}$ 继续施行行初等变换将其化成所谓 Hermite 阶梯形矩阵.

定义 1.3-1　矩阵 $\boldsymbol{A}$ 称为 Hermite **阶梯形矩阵**，如果它是阶梯形矩阵，且 $\boldsymbol{A}$ 的非零行左起第一个非零元素为 1，而且这些 1 分别是它们所在列的唯一非零元素.

例如，对于例 1 所得到的阶梯形矩阵再施行行初等变换，便得与 $\boldsymbol{A}$ 等价的 Hermite 阶梯形矩阵：

$$\begin{bmatrix}1 & 1 & 0 & 1\\ 0 & 2 & 1 & 2\\ 0 & 0 & 0 & 0\end{bmatrix}\xrightarrow[\text{第二行}]{\text{用}\frac{1}{2}\text{乘}}\begin{bmatrix}1 & 1 & 0 & 1\\ 0 & 1 & \frac{1}{2} & 1\\ 0 & 0 & 0 & 0\end{bmatrix}\xrightarrow[\text{第二行}]{\text{第一行减去}}\begin{bmatrix}1 & 0 & -\frac{1}{2} & 0\\ 0 & 1 & \frac{1}{2} & 1\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

1.3.1 矩阵的等价

上述对 $\boldsymbol{A}$ 施行行初等变换化成 Hermite 阶梯形矩阵可以概述为：存在可逆矩阵 $\boldsymbol{P}$ 使 $\boldsymbol{PA}$ 为 Hermite 阶梯形矩阵，且 $\boldsymbol{P}$ 是一系列初等矩阵的乘积. 事实上，若对分块矩阵 $[\boldsymbol{A} \vdots \boldsymbol{I}]$ 施行行初等变换化成如下形式：

$$\boldsymbol{P}[\boldsymbol{A} \vdots \boldsymbol{I}]=[\boldsymbol{PA} \vdots \boldsymbol{P}]=[\boldsymbol{H} \vdots \boldsymbol{P}],$$

其中 $\boldsymbol{H}=\boldsymbol{PA}$ 是 Hermite 阶梯形矩阵，则 $\boldsymbol{I}$ 所在的子块就把变换矩阵 $\boldsymbol{P}$ 记录了下来.

例 2 求可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{PA}$ 为 Hermite 阶梯形矩阵，其中

$$\boldsymbol{A}=\begin{bmatrix} 0 & 1 & 2 & 2 & 3 & 4 \\ 0 & 2 & 4 & -2 & 0 & -4 \\ 0 & 3 & 6 & 0 & 3 & 0 \end{bmatrix}.$$

解

$$\left[\begin{array}{cccccc:ccc} 0 & 1 & 2 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 2 & 4 & -2 & 0 & -4 & 0 & 1 & 0 \\ 0 & 3 & 6 & 0 & 3 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccccc:ccc} 0 & 1 & 2 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & -6 & -6 & -12 & -2 & 1 & 0 \\ 0 & 0 & 0 & -6 & -6 & -12 & -3 & 0 & 1 \end{array}\right]$$

$$\to \left[\begin{array}{cccccc:ccc} 0 & 1 & 2 & 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 & \frac{1}{3} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right] \to \left[\begin{array}{cccccc:ccc} 0 & 1 & 2 & 0 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 & \frac{1}{3} & -\frac{1}{6} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 \end{array}\right].$$

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄炜 博士·副教授 Dr. Wei Huang

1.3.1 矩阵的等价

因此，与 $\boldsymbol{A}$ 等价的 Hermite 阶梯形矩阵和所求的可逆矩阵 $\boldsymbol{P}$ 分别是

$$\begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{3} & -\frac{1}{6} & 0 \\ -1 & -1 & 1 \end{bmatrix}.$$ ■

如果允许使用列初等变换，则可将 $\boldsymbol{A}$ 化成 $\left[\begin{array}{c:c} \boldsymbol{I}_r & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O} \end{array}\right]$，称其为 $\boldsymbol{A}$ 的**等价标准形**，其中 r 为与 $\boldsymbol{A}$ 等价的 Hermite 阶梯形矩阵中非零行个数.

例如，在例 2 所得到的 Hermite 阶梯形矩阵中，把第二列与第一列互换，再把第四列与第二列互换，则得到如下矩阵

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

对此矩阵施行列初等变换，便得 $\boldsymbol{A}$ 的等价标准形：

$$\left[\begin{array}{cc:cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

1.3.1 矩阵的等价

这就是说，任何一个 $m\times n$ 矩阵 $\boldsymbol{A}$ 都存在 m 阶可逆矩阵 $\boldsymbol{P}$ 和 n 阶可逆矩阵 $\boldsymbol{Q}$，使

$$\boldsymbol{PAQ}=\begin{bmatrix}\boldsymbol{I}_r & \boldsymbol{O}\\ \boldsymbol{O} & \boldsymbol{O}\end{bmatrix}, \tag{1.3-1}$$

且通过对分块矩阵

$$\begin{bmatrix}\boldsymbol{A} & \boldsymbol{I}_m\\ \boldsymbol{I}_n & \boldsymbol{O}\end{bmatrix}$$

中子块 $\boldsymbol{A}$ 所在的 m 行施行行初等变换，$\boldsymbol{A}$ 所在的 n 列施行列初等变换，使 $\boldsymbol{PAQ}$ 化成形如(1.3-1)式的分块矩阵，那么 $\boldsymbol{I}_m$ 所在的子块记录下 $\boldsymbol{P}$，$\boldsymbol{I}_n$ 所在的子块记录下 $\boldsymbol{Q}$。

值得指出的是，矩阵 $\boldsymbol{A}$ 的等价标准形是唯一的。可逆矩阵 $\boldsymbol{A}$ 的等价标准形是单位矩阵，从而 $\boldsymbol{A}$ 为可逆矩阵的充要条件是，$\boldsymbol{A}$ 是一系列初等矩阵的乘积。

因此，$m\times n$ 矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 等价的充要条件是存在 m 阶可逆矩阵 $\boldsymbol{P}$ 和 n 阶可逆矩阵 $\boldsymbol{Q}$ 使 $\boldsymbol{PAQ}=\boldsymbol{B}$。

1.3.2 矩阵的秩

在 $m\times n$ 矩阵 $\boldsymbol{A}$ 中任取 k 行 k 列，位于这 k 行 k 列交叉处的 k^2 个元素按其原来的顺序构成的一个 k 阶行列式，称为 $\boldsymbol{A}$ 的一个 k **阶子式**.

矩阵 $\boldsymbol{A}$ 的不为零的子式的最高阶数称为 $\boldsymbol{A}$ 的**秩**，记为 $r(\boldsymbol{A})$.

零矩阵的秩规定为 0.

容易看出，若 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，则 $r(\boldsymbol{A})$ 满足

(1) $0\leqslant r(\boldsymbol{A})\leqslant \min\{m,n\}$；

(2) $r(\boldsymbol{A}^{\mathrm{T}})=r(\boldsymbol{A}),r(\boldsymbol{A}^{\mathrm{H}})=r(\boldsymbol{A})$；

(3) $r(k\boldsymbol{A})=\begin{cases}0, & k=0,\\ r(\boldsymbol{A}), & k\neq 0.\end{cases}$

设 $\boldsymbol{A}$ 是 $m\times n$ 矩阵，若 $r(\boldsymbol{A})=m$，则称 $\boldsymbol{A}$ 为**行满秩矩阵**；若 $r(\boldsymbol{A})=n$，则称 $\boldsymbol{A}$ 为**列满秩矩阵**. 若 $\boldsymbol{A}$ 为 n 阶方阵，且 $r(\boldsymbol{A})=n$，则称 $\boldsymbol{A}$ 为**满秩矩阵**.

显然，阶梯形矩阵的秩等于其非零行的个数.

可以证明[①]，初等变换不改变矩阵的秩，因此求矩阵 $\boldsymbol{A}$ 的秩时，只需将 $\boldsymbol{A}$ 施行行初等变换化成阶梯形矩阵，其非零行的个数即是 $\boldsymbol{A}$ 的秩.

信息学院（国家示范性软件学院） School of Informatics 黄炜 博士·副教授 Dr. Wei Huang
厦門大學 XIAMEN UNIVERSITY

1.3.2 矩阵的秩

例 3　设 a、b 为参数，讨论矩阵

$$A=\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 1 & -1 & b\\ 2 & 3 & a & 4\\ 3 & 5 & 1 & 7\end{bmatrix}$$

的秩.

解

$$A\rightarrow\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 1 & -1 & b\\ 0 & 1 & a-2 & 2\\ 0 & 2 & -2 & 4\end{bmatrix}\rightarrow\begin{bmatrix}1 & 1 & 1 & 1\\ 0 & 1 & -1 & b\\ 0 & 0 & a-1 & 2-b\\ 0 & 0 & 0 & 4-2b\end{bmatrix}.$$

因此，当 $a\neq 1$ 且 $b\neq 2$ 时，$r(A)=4$；当 $a=1,b\neq 2$ 或 $a\neq 1,b=2$时，$r(A)=3$；当 $a=1$ 且 $b=2$ 时，$r(A)=2$。 ■

对于 $m\times n$ 矩阵 A 来说，它的每一行可以看成一个 n 维行向量，用 $A^{(i)}$ 表示 A 的第 i 行；A 的每一列可以看成一个 m 维列向量，用 A_j 表示它的第 j 列，则 A 乘一个列向量 $x=[x_1,x_2,\cdots,x_n]^T$，其结果是将 A 的列向量进行线性组合，组合系数是 x 的各个分量. 事实上，

$$Ax=[A_1 \,\vdots\, \cdots \,\vdots\, A_n][x_1,x_2,\cdots,x_n]^T=x_1A_1+\cdots+x_nA_n.$$

1.3.2 矩阵的秩

同样，将一个行向量 $\boldsymbol{y}=[y_1, y_2, \cdots, y_m]$左乘 $\boldsymbol{A}$，则有

$$\boldsymbol{yA} = [y_1, y_2, \cdots, y_m]\begin{bmatrix} \boldsymbol{A}^{(1)} \\ \hdashline \vdots \\ \hdashline \boldsymbol{A}^{(m)} \end{bmatrix} = y_1\boldsymbol{A}^{(1)} + \cdots + y_m\boldsymbol{A}^{(m)}.$$

它表明 $\boldsymbol{yA}$ 是将 $\boldsymbol{A}$ 的行向量进行线性组合，且组合系数是 $\boldsymbol{y}$ 的各个分量.

可以证明②，任一矩阵 $\boldsymbol{A}$ 的秩等于 $\boldsymbol{A}$ 的列向量组的秩，也等于 $\boldsymbol{A}$ 的行向量组的秩. 从而求向量组的秩的一个有效方法是把向量组按列排成矩阵 $\boldsymbol{A}$，对 $\boldsymbol{A}$ 施行行初等变换化成阶梯形矩阵，则这个阶梯形矩阵的非零行的个数即为所求向量组的秩，而 $\boldsymbol{A}$ 的与非零行左起第一个非零元素所在列对应的向量构成的向量组，即为该向量组的一个极大线性无关组.

1.3.2 矩阵的秩

例 4 求向量组 $\boldsymbol{\alpha}_1=[1,0,1,0]^T$，$\boldsymbol{\alpha}_2=[1,1,0,0]^T$，$\boldsymbol{\alpha}_3=[2,1,1,0]^T$，$\boldsymbol{\alpha}_4=[0,0,1,1]^T$ 的秩及其极大线性无关组.

解 设

$$\boldsymbol{A}=\begin{bmatrix}1&1&2&0\\0&1&1&0\\1&0&1&1\\0&0&0&1\end{bmatrix},$$

对 $\boldsymbol{A}$ 施行行初等变换：

$$\boldsymbol{A}\to\begin{bmatrix}1&1&2&0\\0&1&1&0\\0&-1&-1&1\\0&0&0&1\end{bmatrix}\to\begin{bmatrix}1&1&2&0\\0&1&1&0\\0&0&0&1\\0&0&0&1\end{bmatrix}\to\begin{bmatrix}1&1&2&0\\0&1&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix},$$

由此可知 $r(\boldsymbol{A})=3$，即向量组 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4$ 的秩是 3. 又因非零行左起第一个非零元素所在的列分别是第一、第二和第四列，故 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$ 是所给向量组的一个极大线性无关组. ■

厦门大学 XIAMEN UNIVERSITY 信息学院（国家示范性软件学院）School of Informatics 黄炜 博士副教授 Dr. Wei Huang

1.3.2 矩阵的秩

定理 1.3-1 对矩阵 $\boldsymbol{A}$ 施行行(列)初等变换不会改变其列(行)向量之间的线性组合关系.

证 记 $\boldsymbol{PA}=\boldsymbol{B}$,则有 $$(\boldsymbol{PA})_j=\boldsymbol{PA}_j=\boldsymbol{B}_j.$$

如果 $\boldsymbol{A}$ 的某些列向量之间有某个线性组合关系,例如

$$a_1\boldsymbol{A}_{j_1}+\cdots+a_k\boldsymbol{A}_{j_k}=\boldsymbol{0},$$

用 $\boldsymbol{P}$ 左乘上式的两边得

$$\boldsymbol{P}(a_1\boldsymbol{A}_{j_1}+\cdots+a_k\boldsymbol{A}_{j_k})=a_1\boldsymbol{PA}_{j_1}+\cdots+a_k\boldsymbol{PA}_{j_k}=a_1\boldsymbol{B}_{j_1}+\cdots+a_k\boldsymbol{B}_{j_k}=\boldsymbol{0}.$$

由于 $\boldsymbol{P}$ 是一系列初等矩阵的乘积,故它是可逆矩阵,所以同理可得:若矩阵 $\boldsymbol{B}$ 的某些列向量之间有某个线性组合关系,则 $\boldsymbol{A}$ 中对应的列向量之间也有这个线性组合关系. ∎

由于 Hemite 阶梯形矩阵的特殊结构,所以它的其余列向量与极大线性无关组之间的关系是容易看出的. 因此,将矩阵 $\boldsymbol{A}$ 施行行初等变换化成 Hermite 阶梯形矩阵不仅能够得出 $\boldsymbol{A}$ 的列向量组的极大线性无关组,而且还可得出 $\boldsymbol{A}$ 的其余列向量与此极大线性无关组之间的线性组合关系.

厦門大學 XIAMEN UNIVERSITY 信息学院（国家示范性软件学院）School of Informatics 黄炜 博士副教授 Dr. Wei Huang

1.3.2 矩阵的秩

例 5 求向量组 $\boldsymbol{\alpha}_1=[2,1,4,3]^{\mathrm{T}},\boldsymbol{\alpha}_2=[-1,1,-6,6]^{\mathrm{T}},\boldsymbol{\alpha}_3=[-1,-2,2,-9]^{\mathrm{T}}$, $\boldsymbol{\alpha}_4=[1,1,-2,7]^{\mathrm{T}},\boldsymbol{\alpha}_5=[2,4,4,9]^{\mathrm{T}}$ 的一个极大线性无关组，并将其余向量用这个极大线性无关组线性表出.

解 以 $\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\alpha}_4,\boldsymbol{\alpha}_5$ 构造矩阵

$$\boldsymbol{A}=[\boldsymbol{\alpha}_1 \mid \boldsymbol{\alpha}_2 \mid \boldsymbol{\alpha}_3 \mid \boldsymbol{\alpha}_4 \mid \boldsymbol{\alpha}_5]=\begin{bmatrix}2&-1&-1&1&2\\1&1&-2&1&4\\4&-6&2&-2&4\\3&6&-9&7&9\end{bmatrix},$$

对 $\boldsymbol{A}$ 施行行初等变换化成 Hermite 阶梯形矩阵：

$$\boldsymbol{A}\to\begin{bmatrix}1&1&-2&1&4\\2&-1&-1&1&2\\4&-6&2&-2&4\\3&6&-9&7&9\end{bmatrix}\to\begin{bmatrix}1&1&-2&1&4\\0&-3&3&-1&-6\\0&-4&4&-4&0\\0&3&-3&4&-3\end{bmatrix}\to\begin{bmatrix}1&1&-2&1&4\\0&-4&4&-4&0\\0&-3&3&-1&-6\\0&3&-3&4&-3\end{bmatrix}$$

$$\to\begin{bmatrix}1&1&-2&1&4\\0&1&-1&1&0\\0&0&0&2&-6\\0&0&0&1&-3\end{bmatrix}\to\begin{bmatrix}1&1&-2&1&4\\0&1&-1&1&0\\0&0&0&1&-3\\0&0&0&0&0\end{bmatrix}\to\begin{bmatrix}1&0&-1&0&4\\0&1&-1&0&3\\0&0&0&1&-3\\0&0&0&0&0\end{bmatrix}.$$

因此，$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_4$ 是该向量组的一个极大线性无关组，且

$$\boldsymbol{\alpha}_3=-\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2,\quad \boldsymbol{\alpha}_5=4\boldsymbol{\alpha}_1+3\boldsymbol{\alpha}_2-3\boldsymbol{\alpha}_4.$$

信息学院 (国家示范性软件学院) School of Informatics 黄炜 博士 副教授 Dr. Wei Huang 厦門大學 XIAMEN UNIVERSITY

1.3.2 矩阵的秩

定义 1.3-2　设 A 是秩为 $r>0$ 的 $m\times n$ 矩阵，如果存在列满秩的 $m\times r$ 矩阵 F 和行满秩的 $r\times n$ 矩阵 G，使

$$A = FG, \tag{1.3-2}$$

则称(1.3-2)式为 A 的**满秩分解**.

定理 1.3-2　任何秩为 $r>0$ 的 $m\times n$ 矩阵 A 必有形如(1.3-2)式的满秩分解.

证　由于 $r(A)=r$，所以存在 m 阶可逆矩阵 P 和 n 阶可逆矩阵 Q，使

$$A=P\left[\begin{array}{c:c} I_r & O \\ \hdashline O & O \end{array}\right]Q=P\left[\begin{array}{c} I_r \\ \hdashline O \end{array}\right]\left[\begin{array}{c:c} I_r & O \end{array}\right]Q.$$

令 $F=P\left[\begin{array}{c} I_r \\ \hdashline O \end{array}\right]$，它是 P 的前 r 列组成的 $m\times r$ 矩阵，显然 F 是列满秩的；令 $G=\left[\begin{array}{c:c} I_r & O \end{array}\right]Q$，它是 Q 的前 r 行组成的 $r\times n$ 矩阵，显然 G 是行满秩的. 这样便得(1.3-2)式. ■

需要指出的是，A 的满秩分解不是唯一的. 因为若 D 是任一 r 阶可逆矩阵，则有

$$A=FG=FDD^{-1}G=(FD)(D^{-1}G)=F_1G_1,$$

其中 $F_1=FD, G_1=D^{-1}G$，它也是 A 的满秩分解.

求矩阵 A 的满秩分解的一种方法是利用行初等变换将 A 变成 Hermite 阶梯形矩阵 H. A 中与 H 之非零行左起第一个非零元素所在列对应的列所构成的矩阵可为 F，而 G 可取 H 中非零行所构成的矩阵.

厦門大學 XIAMEN UNIVERSITY 信息学院（国家示范性软件学院）School of Informatics 黄 炜 博士·副教授 Dr. Wei Huang

1.3.2 矩阵的秩

例 6 求矩阵 $\boldsymbol{A}$ 的满秩分解，其中

$$\boldsymbol{A}=\begin{bmatrix} 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 2 & 0 & 0 & 0 & -14 \\ 2 & -1 & 2 & -4 & 0 & 1 \\ -2 & 1 & -2 & 2 & 10 & 25 \end{bmatrix}.$$

解 用行初等变换把 $\boldsymbol{A}$ 变成 Hermite 阶梯形矩阵：

$$\boldsymbol{A}\to\begin{bmatrix} 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 2 & 0 & 0 & 0 & -14 \\ 2 & 0 & 2 & -5 & 5 & 7 \\ 0 & 0 & 0 & -2 & 10 & 26 \end{bmatrix}\to\begin{bmatrix} 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 0 & 0 & 2 & -10 & -26 \\ 2 & 0 & 2 & -5 & 5 & 7 \\ 0 & 0 & 0 & -2 & 10 & 26 \end{bmatrix}\to\begin{bmatrix} 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 0 & 0 & 1 & -5 & -13 \\ 2 & 0 & 2 & -5 & 5 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\to\begin{bmatrix} 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 0 & 0 & 1 & -5 & -13 \\ 1 & 0 & 1 & 0 & -10 & -29 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\to\begin{bmatrix} 1 & 0 & 1 & 0 & -10 & -29 \\ 0 & 1 & 0 & -1 & 5 & 6 \\ 0 & 0 & 0 & 1 & -5 & -13 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\to\begin{bmatrix} 1 & 0 & 1 & 0 & -10 & -29 \\ 0 & 1 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 1 & -5 & -13 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

信息学院（国家示范性软件学院）School of Informatics 黄炜 博士 副教授 Dr. Wei Huang
厦門大學 XIAMEN UNIVERSITY

1.3.2 矩阵的秩

因此，取 $\boldsymbol{A}$ 的第 1、第 2 和第 4 列构成矩阵 $\boldsymbol{F}$：

$$\boldsymbol{F}=\begin{bmatrix} 0 & 1 & -1 \\ 0 & 2 & 0 \\ 2 & -1 & -4 \\ -2 & 1 & 2 \end{bmatrix},$$

取 $\boldsymbol{G}$ 为

$$\boldsymbol{G}=\begin{bmatrix} 1 & 0 & 1 & 0 & -10 & -29 \\ 0 & 1 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 1 & -5 & -13 \end{bmatrix},$$

则有 $\boldsymbol{A}=\boldsymbol{F}\boldsymbol{G}$，即

$$\boldsymbol{A}=\begin{bmatrix} 0 & 1 & -1 \\ 0 & 2 & 0 \\ 2 & -1 & -4 \\ -2 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 & 0 & -10 & -29 \\ 0 & 1 & 0 & 0 & 0 & -7 \\ 0 & 0 & 0 & 1 & -5 & -13 \end{bmatrix}.$$

1.3.3 应用举例

（1）求逆矩阵和解矩阵方程

设 $\boldsymbol{A}$ 是 n 阶可逆矩阵，为求 $\boldsymbol{A}$ 的逆矩阵 $\boldsymbol{A}^{-1}$，我们对分块矩阵 $[\boldsymbol{A} \vdots \boldsymbol{I}]$（它是由 $\boldsymbol{A}$ 和 n 阶单位矩阵 $\boldsymbol{I}$ 合成的 $n\times 2n$ 矩阵）施行行初等变换，则当子块 $\boldsymbol{A}$ 化成 $\boldsymbol{I}$ 时，子块 $\boldsymbol{I}$ 就化成了 $\boldsymbol{A}^{-1}$，这是因为

$$\boldsymbol{A}^{-1}[\boldsymbol{A} \vdots \boldsymbol{I}]=[\boldsymbol{A}^{-1}\boldsymbol{A} \vdots \boldsymbol{A}^{-1}\boldsymbol{I}]=[\boldsymbol{I} \vdots \boldsymbol{A}^{-1}].$$

同样，我们也可以对分块矩阵 $\begin{bmatrix}\boldsymbol{A}\\ \cdots\\ \boldsymbol{I}\end{bmatrix}$ 施行列初等变换，当把子块 $\boldsymbol{A}$ 化成 $\boldsymbol{I}$ 时，则子块 $\boldsymbol{I}$ 就化成了 $\boldsymbol{A}^{-1}$.

对于求解矩阵方程 $\boldsymbol{AX}=\boldsymbol{B}$ （或 $\boldsymbol{XA}=\boldsymbol{B}$），

其中 $\boldsymbol{A}$ 是 n 阶可逆矩阵，$\boldsymbol{B}$ 是 $n\times m$（或 $m\times n$）矩阵，那么对分块矩阵 $[\boldsymbol{A} \vdots \boldsymbol{B}]$ $\left(\begin{bmatrix}\boldsymbol{A}\\ \cdots\\ \boldsymbol{B}\end{bmatrix}\right)$ 施行行（列）初等变换，当把子块 $\boldsymbol{A}$ 化成单位矩阵时，则子块 $\boldsymbol{B}$ 就化成了矩阵方程的解 $\boldsymbol{X}=\boldsymbol{A}^{-1}\boldsymbol{B}$ （$\boldsymbol{X}=\boldsymbol{BA}^{-1}$）.

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院） School of Informatics
黄炜 博士·副教授 Dr. Wei Huang

1.3.3 应用举例

例 7 求解矩阵方程 $\boldsymbol{X}=\boldsymbol{AX}+\boldsymbol{B}$，其中

$$\boldsymbol{A}=\begin{bmatrix}0 & 1 & 0\\ -1 & 1 & 1\\ -1 & 0 & -1\end{bmatrix},\quad \boldsymbol{B}=\begin{bmatrix}1 & -1\\ 2 & 0\\ 5 & -3\end{bmatrix}.$$

解 原方程等价于 $$(\boldsymbol{I}-\boldsymbol{A})\boldsymbol{X}=\boldsymbol{B},$$

对分块矩阵 $[\boldsymbol{I}-\boldsymbol{A} \mid \boldsymbol{B}]$ 施行行初等变换：

$$[\boldsymbol{I}-\boldsymbol{A} \mid \boldsymbol{B}]=\left[\begin{array}{ccc|cc}1 & -1 & 0 & 1 & -1\\ 1 & 0 & -1 & 2 & 0\\ 1 & 0 & 2 & 5 & -3\end{array}\right]\to\left[\begin{array}{ccc|cc}1 & -1 & 0 & 1 & -1\\ 0 & 1 & -1 & 1 & 1\\ 0 & 0 & 1 & 1 & -1\end{array}\right]$$

$$\to\left[\begin{array}{ccc|cc}1 & -1 & 0 & 1 & -1\\ 0 & 1 & 0 & 2 & 0\\ 0 & 0 & 1 & 1 & -1\end{array}\right]\to\left[\begin{array}{ccc|cc}1 & 0 & 0 & 3 & -1\\ 0 & 1 & 0 & 2 & 0\\ 0 & 0 & 1 & 1 & -1\end{array}\right].$$

于是得 $$\boldsymbol{X}=\begin{bmatrix}3 & -1\\ 2 & 0\\ 1 & -1\end{bmatrix}.$$ ∎

1.3.3 应用举例

(2) 关于矩阵之积与矩阵之和的秩

由于任一矩阵的秩等于其列秩(即该矩阵的列向量组的秩),也等于其行秩(即该矩阵的行向量组的秩),所以显然有下列不等式:

$$\max\{r(\boldsymbol{A}),r(\boldsymbol{B})\}\leqslant r([\boldsymbol{A}\ \vdots\ \boldsymbol{B}])\leqslant r(\boldsymbol{A})+r(\boldsymbol{B}), \tag{1.3-3}$$

$$\max\{r(\boldsymbol{A}),r(\boldsymbol{B})\}\leqslant r\left(\begin{bmatrix}\boldsymbol{A}\\ \hdashline \boldsymbol{B}\end{bmatrix}\right)\leqslant r(\boldsymbol{A})+r(\boldsymbol{B}). \tag{1.3-4}$$

定理 1.3-3 设 $\boldsymbol{A}$ 是 $m\times p$ 矩阵,$\boldsymbol{B}$ 是 $p\times n$ 矩阵,则

$$r(\boldsymbol{A})+r(\boldsymbol{B})-p\leqslant r(\boldsymbol{AB})\leqslant\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}. \tag{1.3-5}$$

证 对分块矩阵 $[\boldsymbol{A}\ \vdots\ \boldsymbol{O}]$ 施行列初等变换,化成 $[\boldsymbol{A}\ \vdots\ \boldsymbol{AB}]$,则由初等变换不改变矩阵的秩和(1.3-3)式,得

$$r(\boldsymbol{A})=r([\boldsymbol{A}\ \vdots\ \boldsymbol{O}])=r([\boldsymbol{A}\ \vdots\ \boldsymbol{AB}])\geqslant r(\boldsymbol{AB}).$$

若对分块矩阵 $\begin{bmatrix}\boldsymbol{B}\\ \hdashline \boldsymbol{O}\end{bmatrix}$ 施行行初等变换,化成 $\begin{bmatrix}\boldsymbol{B}\\ \hdashline \boldsymbol{AB}\end{bmatrix}$,则由(1.3-4)式知,

$$r(\boldsymbol{B})=r\left(\begin{bmatrix}\boldsymbol{B}\\ \hdashline \boldsymbol{O}\end{bmatrix}\right)=r\left(\begin{bmatrix}\boldsymbol{B}\\ \hdashline \boldsymbol{AB}\end{bmatrix}\right)\geqslant r(\boldsymbol{AB}).$$

因此得
$$r(\boldsymbol{AB})\leqslant\min\{r(\boldsymbol{A}),r(\boldsymbol{B})\}.$$

黄炜 博士·副教授 Dr. Wei Huang
信息学院（国家示范性软件学院）School of Informatics
厦門大學 XIAMEN UNIVERSITY

1.3.3 应用举例

下证前一个不等式. 设 $r(\boldsymbol{A})=k, r(\boldsymbol{B})=l$，则有 m 阶可逆矩阵 $\boldsymbol{P}$，p 阶可逆矩阵 $\boldsymbol{Q}$ 和 $\boldsymbol{F}$，以及 n 阶可逆矩阵 $\boldsymbol{G}$，使得

$$\boldsymbol{PAQ}=\left[\begin{array}{c:c}\boldsymbol{I}_k & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right], \quad \boldsymbol{FBG}=\left[\begin{array}{c:c}\boldsymbol{I}_l & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right].$$

由此得
$$\boldsymbol{PA}=\left[\begin{array}{c:c}\boldsymbol{I}_k & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right]\boldsymbol{Q}^{-1}, \quad \boldsymbol{BG}=\boldsymbol{F}^{-1}\left[\begin{array}{c:c}\boldsymbol{I}_l & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right],$$

从而
$$\boldsymbol{PABG}=\left[\begin{array}{c:c}\boldsymbol{I}_k & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right]\boldsymbol{Q}^{-1}\boldsymbol{F}^{-1}\left[\begin{array}{c:c}\boldsymbol{I}_l & \boldsymbol{O} \\ \hdashline \boldsymbol{O} & \boldsymbol{O}\end{array}\right].$$

令 p 阶可逆矩阵 $\boldsymbol{Q}^{-1}\boldsymbol{F}^{-1}$ 为

$$\begin{bmatrix} c_{11} & \cdots & c_{1l} & \cdots & c_{1p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_{k1} & \cdots & c_{kl} & \cdots & c_{kp} \\ \vdots & \cdots & \vdots & & \vdots \\ c_{p1} & \cdots & c_{pl} & \cdots & c_{pp} \end{bmatrix},$$

1.3.3 应用举例

则
$$PABG = \begin{bmatrix} c_{11} & \cdots & c_{1l} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ c_{k1} & \cdots & c_{kl} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

由于把一个矩阵的任一行(列)改为零行(列)，该矩阵的秩之减少不会超过 1，因此

$$\begin{aligned} r(\boldsymbol{AB}) &= r(\boldsymbol{PABG}) \geqslant p-(p-k)-(p-l) \\ &= k+l-p, \end{aligned}$$

即
$$r(\boldsymbol{AB}) \geqslant r(\boldsymbol{A})+r(\boldsymbol{B})-p. \quad \blacksquare$$

1.3.3 应用举例

例 8 设 $\boldsymbol{A}_i(i=1,\cdots,k)$ 是 n 阶方阵，证明当 $\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k=\boldsymbol{O}$ 时有不等式

$$\sum_{i=1}^{k}(n-r(\boldsymbol{A}_i))\geqslant n. \tag{1.3-6}$$

证 (1.3-6)式等价于 $\displaystyle\sum_{i=1}^{k}r(\boldsymbol{A}_i)\leqslant(k-1)n.$

根据(1.3-5)式，有

$$\begin{aligned}0=r(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_k)&\geqslant r(\boldsymbol{A}_1\cdots\boldsymbol{A}_{k-1})+r(\boldsymbol{A}_k)-n\\&\geqslant r(\boldsymbol{A}_1\cdots\boldsymbol{A}_{k-2})+r(\boldsymbol{A}_{k-1})+r(\boldsymbol{A}_k)-2n\\&\geqslant\cdots\geqslant\sum_{i=1}^{k}r(\boldsymbol{A}_i)-(k-1)n,\end{aligned}$$

即 $\displaystyle\sum_{i=1}^{k}r(\boldsymbol{A}_i)\leqslant(k-1)n$，亦即(1.3-6)式成立. ∎

1.3.3 应用举例

定理 1.3-4 设 A、B 都是 $m\times n$ 矩阵，则

$$r(A+B)\leqslant r(A)+r(B). \tag{1.3-7}$$

证 由于
$$A+B=[I_m \mid I_m]\begin{bmatrix}A\\ \hdashline B\end{bmatrix},$$

所以有
$$r(A+B)\leqslant r\left(\begin{bmatrix}A\\ \hdashline B\end{bmatrix}\right)\leqslant r(A)+r(B).$$ ■

(3) 分块矩阵的求逆公式

设 A 是 n 阶可逆矩阵，将其分块为

$$A=\begin{array}{c}\left[\begin{array}{c:c}A_{11} & A_{12}\\ \hdashline \underbrace{A_{21}}_{k} & \underbrace{A_{22}}_{n-k}\end{array}\right]\begin{array}{l}\}k\\ \}n-k\end{array}\end{array},$$

其中 A_{11} 是 k 阶方阵，A_{22} 是 $n-k$ 阶方阵，A_{12} 和 A_{21} 分别是 $k\times(n-k)$ 和 $(n-k)\times k$ 矩阵.

若 A_{11} 可逆，则有

$$\begin{array}{l}k\{\\ n-k\{\end{array}\left[\begin{array}{c:c}A_{11}^{-1} & O\\ \hdashline \underbrace{-A_{21}A_{11}^{-1}}_{k} & \underbrace{I}_{n-k}\end{array}\right]\left[\begin{array}{c:c}A_{11} & A_{12}\\ \hdashline A_{21} & A_{22}\end{array}\right]=\left[\begin{array}{c:c}I & A_{11}^{-1}A_{12}\\ \hdashline \underbrace{O}_{k} & \underbrace{A_{22.1}}_{n-k}\end{array}\right]\begin{array}{l}\}k\\ \}n-k\end{array},$$

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄 炜 博士·副教授 Dr. Wei Huang

1.3.3 应用举例

其中 $A_{22.1}=A_{22}-A_{21}A_{11}^{-1}A_{12}$ 是$(n-k)$阶可逆矩阵(为什么?). 容易验证,

$$\left[\begin{array}{c:c} I & A_{11}^{-1}A_{12} \\ \hdashline O & A_{22.1} \end{array}\right]^{-1}=\left[\begin{array}{c:c} I & -A_{11}^{-1}A_{12}A_{22.1}^{-1} \\ \hdashline O & A_{22.1}^{-1} \end{array}\right].$$

由此可得

$$A^{-1}=\left[\begin{array}{c:c} A_{11}^{-1}+A_{11}^{-1}A_{12}A_{22.1}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22.1}^{-1} \\ \hdashline \underbrace{-A_{22.1}^{-1}A_{21}A_{11}^{-1}}_{k} & \underbrace{A_{22.1}^{-1}}_{n-k} \end{array}\right]\begin{matrix}\}k \\ \}n-k\end{matrix}. \quad (1.3\text{-}8)$$

若 A_{22} 可逆,同理可得

$$A^{-1}=\left[\begin{array}{c:c} A_{11.2}^{-1} & -A_{11.2}^{-1}A_{12}A_{22}^{-1} \\ \hdashline \underbrace{-A_{22}^{-1}A_{21}A_{11.2}^{-1}}_{k} & \underbrace{A_{22}^{-1}+A_{22}^{-1}A_{21}A_{11.2}^{-1}A_{12}A_{22}^{-1}}_{n-k} \end{array}\right]\begin{matrix}\}k \\ \}n-k\end{matrix}, \quad (1.3\text{-}9)$$

其中 $A_{11.2}=A_{11}-A_{12}A_{22}^{-1}A_{21}$ 是 k 阶可逆矩阵.

类似地,当 A_{11} 和 A_{22} 都可逆时,则有

$$A^{-1}=\left[\begin{array}{c:c} A_{11.2}^{-1} & -A_{11}^{-1}A_{12}A_{22.1}^{-1} \\ \hdashline \underbrace{-A_{22}^{-1}A_{21}A_{11.2}^{-1}}_{k} & \underbrace{A_{22.1}^{-1}}_{n-k} \end{array}\right]\begin{matrix}\}k \\ \}n-k\end{matrix}. \quad (1.3\text{-}10)$$

因为可逆矩阵的逆矩阵是唯一的,所以(1.3-8)式与(1.3-9)式的对应元素相等,从而得到下列矩阵恒等式:

$$(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}=A_{11}^{-1}+A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1}, \quad (1.3\text{-}11)$$

$$A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}=(A_{11}-A_{12}A_{22}^{-1}A_{21})^{-1}A_{12}A_{22}^{-1}, \quad (1.3\text{-}12)$$

只要式中出现的逆矩阵都存在.

(1.3-11)式和(1.3-12)式在多变量线性系统理论中被称为**矩阵求逆引理**.

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院） School of Informatics
黄炜 博士 副教授 Dr. Wei Huang

作业（第1部分）

习题 1.3

3. 求可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{PA}$ 为 Hermite 阶梯形矩阵，其中

$$\boldsymbol{A}=\begin{bmatrix}1&1&2&2&3\\2&2&3&1&4\\1&0&1&1&5\\2&3&5&5&4\end{bmatrix}.$$

5. 求矩阵 $\boldsymbol{A}$ 的一个满秩分解，其中

$$\boldsymbol{A}=\begin{bmatrix}1&-1&2&0&-1\\1&1&-2&2&3\\3&-1&8&5&1\\1&3&-6&4&7\end{bmatrix}.$$

作业（第1部分）

习题 1.3

3. 求可逆矩阵 $\boldsymbol{P}$，使 $\boldsymbol{PA}$ 为 Hermite 阶梯形矩阵，其中

$$\boldsymbol{A}=\begin{bmatrix}1 & 1 & 2 & 2 & 3\\ 2 & 2 & 3 & 1 & 4\\ 1 & 0 & 1 & 1 & 5\\ 2 & 3 & 5 & 5 & 4\end{bmatrix}.$$

5. 求矩阵 $\boldsymbol{A}$ 的一个满秩分解，其中

$$\boldsymbol{A}=\begin{bmatrix}1 & -1 & 2 & 0 & -1\\ 1 & 1 & -2 & 2 & 3\\ 3 & -1 & 8 & 5 & 1\\ 1 & 3 & -6 & 4 & 7\end{bmatrix}.$$

误差的基本知识

数值计算（1）

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院） School of Informatics
黄 煒 博士·副教授 Dr. Wei Huang

第二部分　数值计算方法

电子数字计算机的高速运算和逻辑功能，使它成为数学应用于科学技术和社会经济各个领域的最强有力的工具．随着电子数字计算机广泛地用来解决各种工程实际和科学研究中的计算问题，数值计算方法作为一门学科迅速地发展起来，已经成为应用数学的一个最有活力的重要分支．

解决工程实际中的科学计算问题，一般需要经过建立数学模型、选用计算方法，用电子数字计算机计算问题的数值结果、验证所得结果是否符合客观实际等主要步骤．其中选用数值计算方法是一个重要的环节，它是使用电子数字计算机计算数值结果，并对这些数值结果进行分析的依据和基础．数值计算方法的任务是研究各种适用于科学计算的数值方法及有关的理论分析，亦即针对特定的数学模型，提供各种能用电子数字计算机有效地计算数值结果的数学方法，并对这些方法进行必要的理论分析和论证．

由于电子数字计算机只能根据人们给定的指令，完成加、减、乘、除等算术运算和一些逻辑运算，因此应用电子数字计算机求解各种数学问题，例如线性方程组的解、方阵的特征值及对应的特征向量、函数的最大值和最小值、微分方程的解等，都必须把求解过程分解为按一定规则进行的一系列四则算术运算．这种把数学问题的解法归结为由加、减、乘、除等有限步基本运算以及联系它们的一些逻辑运算，并明确规定运算顺序的完整计算方案，称为**数值算法**，简称**算法**．

第二部分　数值计算方法

对数值算法进行分析，是为了评价算法的优缺点，以便从各种算法中选用好的算法。这就涉及计算工作量、存贮量大小、逻辑结构繁简程度以及收敛性、稳定性、误差分析等问题.

这一部分的目的是介绍科学计算中常用的方法，阐述数值计算中基本的概念和原理，掌握适用于电子数字计算机的一些算法，使读者具有一定的理论分析能力，以便能将所学方法应用于解决实际计算问题.

第一章 误差的基本知识

用电子数字计算机进行实际问题的数值计算，计算误差是不可避免的．误差的来源主要有四个方面．

(1) 模型误差

用数学模型描述实际问题，一般都要作一定的简化，由此产生的数学模型的解与实际问题的解之间会有差异，这种差异称为**模型误差**．

(2) 观测误差

数学模型中包含的某些参数或常数，往往是通过仪器观测或实验获得其数值的，这样得到的观测数值与实际数值之间会有误差，这种误差称为**观测误差**．

(3) 截断误差

求解数学模型所用的数值计算方法往往是近似方法，从而只能得到数学模型的近似解，由此产生的误差称为**方法误差**．由于近似方法一般都要用有限的四则算术运算步骤来代替无穷的极限运算，这种由截断一个无穷过程而引起的误差，就是**截断误差**．因而方法误差也称为截断误差．

(4) 舍入误差

由于电子数字计算机只能将数表示成有限位进行运算，所以对超过位数的数字要按一定的规则作舍入，由此产生的误差称为**舍入误差**．

第一章　误差的基本知识

数值计算方法主要研究截断误差和舍入误差对计算结果的影响，一般不考虑模型误差和观测误差. 分析参数或常数的观测误差在数值计算中的影响的方法与分析舍入误差的影响所用的方法大致相同，而控制观测误差和模型误差则不是数学计算工作者所能独立解决的.

本章介绍误差的基本知识，包括绝对误差、相对误差、有效数字的概念，误差分析的基本方法，以及数值运算中应注意的原则.

1.1 绝对误差、相对误差及有效数字

数值计算中处理的数据和计算的结果，通常都是近似值，它们与准确值之间存在着误差.

设 x^* 是准确值 x 的一个近似值，则称

$$e = x^* - x \tag{1.1-1}$$

为近似值 x^* 的**绝对误差**. 由于准确值 x 一般是未知的，所以 e 也是未知的，往往只能根据实际情况估计其绝对值 $|e|$ 的一个上界 ε，即

$$|e| = |x^* - x| \leqslant \varepsilon,$$

ε 称为近似值 x^* 的**绝对误差限**. 如果估计出 ε，则可以确定 x 的数值范围：

$$x^* - \varepsilon \leqslant x \leqslant x^* + \varepsilon.$$

上述不等式在工程技术上常表示为 $x = x^* \pm \varepsilon.$

1.1 绝对误差、相对误差及有效数字

例 1 “四舍五入”的绝对误差限

实数 x 的十进制标准表示式是

$$x = \pm 0.a_1a_2\cdots a_na_{n+1}\cdots \times 10^m,$$

其中 m 为整数，$a_1, a_2, \cdots, a_n, a_{n+1}, \cdots$ 都是 0 至 9 中的一个数，且 $a_1 \neq 0$. 如果按“四舍五入”对 a_{n+1} 作舍入，得到 x 的近似值

$$x^* = \begin{cases} \pm 0.a_1a_2\cdots a_n \times 10^m, & a_{n+1} \leqslant 4, \\ \pm 0.a_1a_2\cdots (a_n+1) \times 10^m, & a_{n+1} \geqslant 5, \end{cases}$$

那么四舍时

$$|e| = |x^* - x| = (0.a_1a_2\cdots a_na_{n+1}\cdots - 0.a_1a_2\cdots a_n) \times 10^m$$

$$\leqslant 0.\underbrace{00\cdots 05}_{n个} \times 10^m = \frac{1}{2} \times 10^{m-n},$$

五入时 $|e| = |x^* - x| = (0.a_1a_2\cdots (a_n+1) - 0.a_1a_2\cdots a_na_{n+1}\cdots) \times 10^m$

$$\leqslant 0.\underbrace{00\cdots 05}_{n个} \times 10^m = \frac{1}{2} \times 10^{m-n}.$$

也就是说，对 x 进行四舍五入后，其绝对误差限为被保留的数字中最后数位的半个单位. ∎

1.1 绝对误差、相对误差及有效数字

单用绝对误差的概念不足以表征近似值的好坏程度. 要全面地反映近似值的精确程度,不仅要看绝对误差的大小,还要考虑绝对误差在原数值中所占的比例.

把近似值 x^* 的绝对误差与准确值 x 之比称为近似值 x^* 的**相对误差**,记为 e_r,即

$$e_r = \frac{e}{x} = \frac{x^* - x}{x}. \tag{1.1-2}$$

实际上,由于准确值 x 是未知的,所以通常把相对误差改取为

$$e_r = \frac{e}{x^*} = \frac{x^* - x}{x^*}. \tag{1.1-2$'$}$$

当 $e_r = \dfrac{e}{x^*}$ 很小时,两者的差为

$$\frac{e}{x} - \frac{e}{x^*} = \frac{e(x^* - x)}{xx^*} = \frac{1}{1 - e/x^*} \cdot \left(\frac{e}{x^*}\right)^2 \approx \left(\frac{e}{x^*}\right)^2,$$

它与 e_r^2 同阶,故可以忽略不计.

值得指出的是,相对误差是一个无量纲的数,而绝对误差是有量纲的.

类似地,相对误差的绝对值的上界称为近似值 x^* 的**相对误差限**,记为 ε_r,即 $|e_r| \leqslant \varepsilon_r$.

如果 ε 是近似值 x^* 的绝对误差限,那么近似值 x^* 的相对误差限为 $\dfrac{\varepsilon}{|x^*|}$.

1.1 绝对误差、相对误差及有效数字

例 2 $\pi=3.1415926\cdots$，按四舍五入取五位数字作为它的一个近似值 x^*，则 $x^*=3.1416$.

$$e=x^*-\pi=0.0000073\cdots,$$

$$e_r=\frac{x^*-\pi}{x^*}\approx 0.2324\times 10^{-5}. \qquad \blacksquare$$

设 x^* 是 x 的一个近似值，如果近似值 x^* 的绝对误差限是它的某一位的半个单位，而该位向左到 x^* 的第一位非零数字共有 n 位，则称 x^* 作为 x 的近似值具有 n 位**有效数字**，或者说近似值 x^* 准确到该位.

如果 x^* 表示为 $\qquad x^*=\pm 0.a_1a_2\cdots a_n\times 10^m, a_1\neq 0,$

那么，当

$$|e|=|x^*-x|\leqslant\frac{1}{2}\times 10^{m-n}, \tag{1.1-3}$$

则 x^* 作为 x 的近似值具有 n 位有效数字 $a_1,a_2,\cdots,a_n$.

从有效数字的概念可知，由准确值经过四舍五入得到的近似值，从它的末位数字向左到第一位非零数字都是有效数字. 同一个准确值的不同近似值，有效数字越多的近似值，它的绝对误差和相对误差越小. 因此，1.42032 的近似值 1.42 与 1.420 所表示的精度是不同的，前者有三位有效数字，而后者有四位有效数字.

1.1 绝对误差、相对误差及有效数字

例 3 下列近似值的绝对误差限都是 0.5×10^{-2}；$a=1.38$，$b=-0.0312$，$c=0.86\times10^{-4}$，问各个近似值有几位有效数字？

解 a 有三位有效数字 1，3，8；b 只有一位有效数字 3；而 c 没有有效数字. ■

同样可以给出相对误差限与有效数字位数之间的关系. 设近似值 x^* 具有 n 位有效数字，则

$$\varepsilon_r \leqslant \frac{1}{2a_1}\times10^{-(n-1)}. \tag{1.1-4}$$

例 4 求$\sqrt{5}$的近似值，使其相对误差小于 1%，问应取几位有效数字？

解 设应取 n 位有效数字，则因 $2<\sqrt{5}<3$，故按(1.1-4)式，n 应满足不等式

$$\frac{1}{2\times2}10^{-(n-1)}<0.01.$$

解之得 $n>2.4$，因此可取 $n=3$，即$\sqrt{5}$的近似值取为 2.24 时，其相对误差小于 1%. ■

作业（第2部分）

习题 1

2. 设 x 的相对误差为 1%，求 x^n 的相对误差.

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄炜 博士·副教授 Dr. Wei Huang

数理统计的常用分布

数理统计（1）

厦門大學
XIAMEN UNIVERSITY
信息学院
（国家示范性软件学院）
School of Informatics
黄 炜
博士·副教授
Dr. Wei Huang

1.3 数理统计中常见的几个分布

下列三种分布在数理统计中起着重要的作用.

1. χ^2 分布

定义 1.3-1 设随机变量 $X_1, X_2, \cdots, X_n$ 相互独立,且都服从标准正态分布 $N(0;1)$,则称随机变量

$$\chi^2 = \sum_{i=1}^{n} X_i^2 \tag{1.3-1}$$

所服从的分布为**自由度 n 的 χ^2 分布**,记为 $\chi^2(n)$.

定理 1.3-1 设随机变量 $\chi^2 \sim \chi^2(n)$,则其概率密度为

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, & x > 0, \\ 0, & x \leqslant 0. \end{cases} \tag{1.3-2}$$

证 因为 $X_1, X_2, \cdots, X_n$ 相互独立,且都服从 $N(0;1)$,所以 $\boldsymbol{X} = [X_1, X_2, \cdots, X_n]^{\mathrm{T}}$ 的概率密度为

$$f(x_1, x_2, \cdots, x_n) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} = (2\pi)^{-\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right).$$

于是,当 $x \geqslant 0, \Delta x > 0$ 时,随机事件 $\{x < \chi^2 \leqslant x + \Delta x\}$ 的概率为

$$P\{x < \chi^2 \leqslant x + \Delta x\} = (2\pi)^{-\frac{n}{2}} \iint_G \cdots \int \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) dx_1 dx_2 \cdots dx_n,$$

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院）School of Informatics
黄炜 博士·副教授 Dr. Wei Huang

1.3 数理统计中常见的几个分布

其中 G 是外半径为 $\sqrt{x+\Delta x}$、内半径为 $\sqrt{x}$ 的 n 维球壳. 记 G 的体积为 ΔV，则有

$$(2\pi)^{-\frac{n}{2}}\mathrm{e}^{-\frac{1}{2}(x+\Delta x)}\Delta V \leqslant P\{x<\chi^2\leqslant x+\Delta x\} \leqslant (2\pi)^{-\frac{n}{2}}\mathrm{e}^{-\frac{1}{2}x}\Delta V. \qquad (1.3\text{-}3)$$

由于半径为 r 的 n 维球体的体积是 Cr^n，这里的 C 是常数，故 $\Delta V = C[(x+\Delta x)^{\frac{n}{2}} - x^{\frac{n}{2}}]$，从而

$$\lim_{\Delta x\to 0}\frac{\Delta V}{\Delta x} = C\lim_{\Delta x\to 0}\frac{1}{\Delta x}[(x+\Delta x)^{\frac{n}{2}} - x^{\frac{n}{2}}] = C\,\frac{n}{2}x^{\frac{n}{2}-1}. \qquad (1.3\text{-}4)$$

因而，由(1.3-3)式和(1.3-4)式，得

$$\lim_{\Delta x\to 0}\frac{1}{\Delta x}P\{x<\chi^2\leqslant x+\Delta x\} = kx^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}, \qquad (1.3\text{-}5)$$

其中 $k=(2\pi)^{-\frac{n}{2}}\frac{n}{2}C$ 是常数. 这就是说，当 $x>0$ 时，χ^2 的概率密度是 $kx^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}$. 又因 $\chi^2\geqslant 0$，故当 $x\leqslant 0$ 时，χ^2 的概率密度显然为 0. 再由概率密度的性质知，

$$\int_{-\infty}^{+\infty}f(x)\mathrm{d}x = \int_0^{+\infty}kx^{\frac{n}{2}-1}\mathrm{e}^{-\frac{x}{2}}\mathrm{d}x = k2^{\frac{n}{2}}\int_0^{+\infty}u^{\frac{n}{2}-1}\mathrm{e}^{-u}\mathrm{d}u = k2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right),$$

即 $k=\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}$. 于是，(1.3-2) 式成立. ∎

1.3 数理统计中常见的几个分布

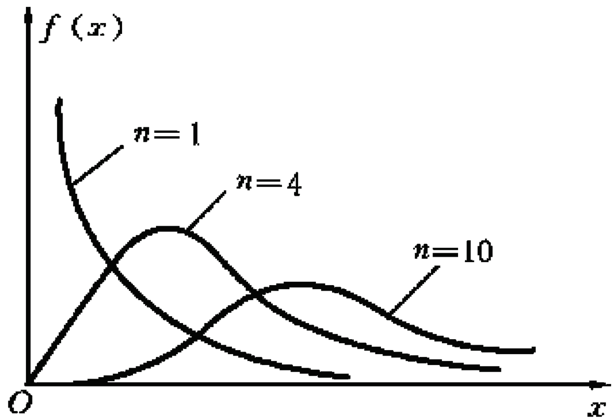

图 1.3-1　$\chi^2(n)$ 的概率密度函数

χ^2 分布的概率密度函数的图形如图 1.3-1 所示，它随自由度 n 的不同有所变动. 当 $n=1$ 时，$\lim\limits_{x\to 0^+} f(x)=+\infty$；而当 $n\geqslant 2$ 时，$\lim\limits_{x\to 0^+} f(x)=0$.

定理 1.3-2　设 $X\sim\chi^2(n)$，即 X 是自由度为 n 的 χ^2 随机变量，则

(1) X 的特征函数为

$$\varphi(t)=E(\mathrm{e}^{\mathrm{i}tX})=(1-2\mathrm{i}t)^{-\frac{n}{2}}. \tag{1.3-6}$$

(2) X 的数学期望与方差分别为

$$E(X)=n,\ D(X)=2n. \tag{1.3-7}$$

1.3 数理统计中常见的几个分布

证 (1)由特征函数的定义得

$$\varphi(t)=\int_0^{+\infty}\mathrm{e}^{\mathrm{i}tx}\,\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\mathrm{e}^{-\frac{x}{2}}x^{\frac{n}{2}-1}\mathrm{d}x=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}x^{\frac{n}{2}-1}\mathrm{e}^{-\left(\frac{1}{2}-\mathrm{i}t\right)x}\mathrm{d}x$$

$$=\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\left(\frac{1}{2}-\mathrm{i}t\right)^{-\frac{n}{2}}\int_0^{+\infty}u^{\frac{n}{2}-1}\mathrm{e}^{-u}\mathrm{d}u=(1-2\mathrm{i}t)^{-\frac{n}{2}}.$$

(2)
$$E(X)=\frac{1}{\mathrm{i}}\varphi'(0)=\frac{1}{\mathrm{i}}\left[-\frac{n}{2}(1-2\mathrm{i}t)^{-\frac{n}{2}-1}(-2\mathrm{i})\right]_{t=0}=n;$$

$$E(X^2)=\frac{1}{\mathrm{i}^2}\varphi''(0)=\frac{1}{\mathrm{i}^2}\left[n\mathrm{i}\left(-\frac{n}{2}-1\right)(1-2\mathrm{i}t)^{-\frac{n}{2}-2}(-2\mathrm{i})\right]_{t=0}=n(n+2),$$

从而
$$D(X)=E(X^2)-[E(X)]^2=2n.$$
$\blacksquare$

1.3 数理统计中常见的几个分布

定理 1.3-3 设 $X_1 \sim \chi^2(n_1)$，$X_2 \sim \chi^2(n_2)$，且 X_1 与 X_2 相互独立，则

$$X_1 + X_2 \sim \chi^2(n_1 + n_2). \tag{1.3-8}$$

证 由于 X_1、X_2 的特征函数分别为

$$\varphi_1(t) = (1-2\mathrm{i}t)^{-\frac{n_1}{2}}, \quad \varphi_2(t) = (1-2\mathrm{i}t)^{-\frac{n_2}{2}},$$

又 X_1 与 X_2 相互独立，故 $X_1 + X_2$ 的特征函数为

$$\varphi(t) = \varphi_1(t)\varphi_2(t) = (1-2\mathrm{i}t)^{-\frac{1}{2}(n_1+n_2)},$$

从而 $X_1 + X_2$ 服从自由度为 $n_1 + n_2$ 的 χ^2 分布. $\blacksquare$

定理 1.3-3 表明，χ^2 分布具有可加性. 更一般地，若 $X_i \sim \chi^2(n_i)(i = 1,2,\cdots,k)$，且 $X_1,X_2,\cdots,X_k$ 相互独立，则有 $\sum\limits_{i=1}^{k} X_i \sim \chi^2\left(\sum\limits_{i=1}^{k} n_i\right)$.

1.3 数理统计中常见的几个分布

2. t 分布

定义 1.3-2 设 $X \sim N(0;1)$，$Y \sim \chi^2(n)$，且 X 与 Y 相互独立，则称随机变量

$$T = \frac{X}{\sqrt{Y/n}} \tag{1.3-9}$$

所服从的分布为**自由度 n 的 t 分布**，记为 $t(n)$.

定理 1.3-4 设随机变量 $T \sim t(n)$，则 T 的概率密度为

$$f(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < +\infty. \tag{1.3-10}$$

证 记 $Z = \sqrt{\frac{Y}{n}}$，即 $Y = nZ^2$，则 Z 的概率密度为

$$f_Z(z) = \begin{cases} \dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}(nz^2)^{\frac{n}{2}-1}e^{-nz^2/2}2nz, & z > 0, \\ 0, & z \leqslant 0. \end{cases}$$

由于 X 与 Y 相互独立，故 X 与 Z 也相互独立，从而 $[X,Z]^{\mathrm{T}}$ 的概率密度为

$$f(x,z) = \begin{cases} \dfrac{n^{\frac{n}{2}}}{\sqrt{2\pi}2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}z^{n-1}e^{-(x^2+nz^2)/2}, & -\infty < x < +\infty, \; z > 0, \\ 0, & \text{其他}. \end{cases}$$

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院） School of Informatics
黄 炜 博士·副教授 Dr. Wei Huang

1.3 数理统计中常见的几个分布

现 $T=\dfrac{X}{Z}$，因而 T 的分布函数

$$F(t)=P\{T\leqslant t\}=P\left\{\frac{X}{Z}\leqslant t\right\}=P\{X\leqslant tZ\}=\int_{0}^{+\infty}\mathrm{d}z\int_{-\infty}^{tz}f(x,z)\mathrm{d}x.$$

于是，T 的概率密度为

$$f(t)=F'(t)=\int_{0}^{+\infty}f(tz,z)z\mathrm{d}z=\int_{0}^{+\infty}\frac{n^{\frac{n}{2}}}{\sqrt{2\pi}2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}z^{n}\mathrm{e}^{-(n+t^{2})z^{2}/2}\mathrm{d}z$$

$$=\frac{1}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^{2}}{n}\right)^{-(n+1)/2}\int_{0}^{+\infty}x^{\frac{n+1}{2}-1}\mathrm{e}^{-x}\mathrm{d}x=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^{2}}{n}\right)^{-(n+1)/2}. \quad \blacksquare$$

从(1.3-10)式看出，t 分布的概率密度函数 $f(t)$ 是偶函数，且 $\lim\limits_{t\to\infty}f(t)=0$. $f(t)$ 的图形如图 1.3-2 所示.

我们不加证明地给出，若 $T\sim t(n)$，且 $n\geqslant 2$，则对正整数 $r(r<n)$，$E(T^{r})$ 存在，且

$$E(T^{r})=\begin{cases}0, & r\text{ 为奇数},\\ n^{\frac{r}{2}}\dfrac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{n-r}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}, & r\text{ 为偶数}.\end{cases}$$

1.3 数理统计中常见的几个分布

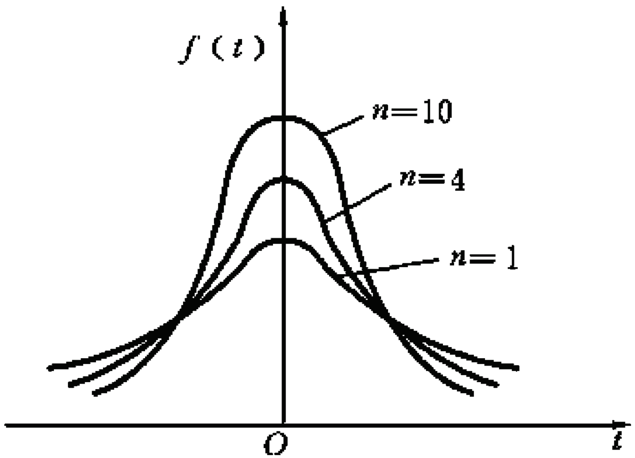

图 1.3-2　$t(n)$ 的概率密度函数

特别地，当 $n \geqslant 2$ 时，有 $E(T) = 0$；当 $n \geqslant 3$ 时，有 $D(T) = \dfrac{n}{n-2}$.

当 $n \to +\infty$ 时，利用 Γ 函数的 Stirling 公式，可得

$$\lim_{n\to\infty} f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}},$$

即当 n 很大（一般只要 $n > 30$）时，t 分布非常近似于 $N(0;1)$.

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院） School of Informatics
黄 炜 博士/副教授 Dr. Wei Huang

1.3 数理统计中常见的几个分布

3. F 分布

定义 1.3-3 设随机变量 X 与 Y 相互独立，且 $X\sim\chi^2(m)$，$Y\sim\chi^2(n)$，则称随机变量

$$F=\frac{X}{m}\Big/\frac{Y}{n}=\frac{n}{m}\,\frac{X}{Y} \tag{1.3-11}$$

所服从的分布为**自由度为(m,n)的 F 分布**，记为 $F(m,n)$，其中 m 称为第一自由度，n 称为第二自由度.

由 F 分布定义可知，当 $F\sim F(m,n)$ 时，则随机变量 $\dfrac{1}{F}$ 所服从的分布为自由度为 (n,m) 的 F 分布，即 $\dfrac{1}{F}\sim F(n,m)$.

定理 1.3-5 设随机变量 $F\sim F(m,n)$，则 F 的概率密度为

$$f(x)=\begin{cases}\dfrac{\Gamma\left(\dfrac{m+n}{2}\right)}{\Gamma\left(\dfrac{m}{2}\right)\Gamma\left(\dfrac{n}{2}\right)}\left(\dfrac{m}{n}\right)\left(\dfrac{m}{n}x\right)^{\frac{m}{2}-1}\left(1+\dfrac{m}{n}x\right)^{-\frac{m+n}{2}}, & x>0\\ 0, & x\leqslant 0.\end{cases} \tag{1.3-12}$$

证 令 $Z=\dfrac{X}{m}$，则由 X 的概率密度得到 Z 的概率密度为

$$f_W(w)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}(mw)^{\frac{n}{2}-1}\mathrm{e}^{-\frac{mw}{2}}, & w>0,\\ 0, & w\leqslant 0.\end{cases}$$

1.3 数理统计中常见的几个分布

由于 X 与 Y 相互独立，故 Z 与 W 相互独立，从而$[Z,W]^{\mathrm{T}}$ 的概率密度为

$$f(z,w)=\begin{cases}\dfrac{m^{\frac{m}{2}}\cdot n^{\frac{n}{2}}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}z^{\frac{m}{2}-1}w^{\frac{n}{2}-1}\mathrm{e}^{-(mz+nw)/2}, & z>0,\ w>0,\\ 0, & \text{其他}.\end{cases}$$

于是，$F=\dfrac{Z}{W}$ 的分布函数为

$$P\{F\leqslant x\}=P\{Z\leqslant xW\}=\begin{cases}\displaystyle\int_0^{+\infty}\mathrm{d}w\int_0^{xw}f(z,w)\mathrm{d}z, & x>0\\ 0, & x\leqslant 0.\end{cases}$$

因此，当 $x>0$ 时，F 的概率密度为

$$f(x)=\frac{\mathrm{d}}{\mathrm{d}x}P\{F\leqslant x\}=\int_0^{+\infty}wf(xw,w)\mathrm{d}w$$

$$=\frac{m^{\frac{m}{2}}\cdot n^{\frac{n}{2}}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\int_0^{+\infty}w(xw)^{\frac{m}{2}-1}w^{\frac{n}{2}-1}\mathrm{e}^{-(mx+n)w/2}\mathrm{d}w$$

$$=\frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)\left(\frac{m}{n}x\right)^{\frac{m}{2}-1}\left(1+\frac{m}{n}x\right)^{-\frac{m+n}{2}}.$$

而当 $x\leqslant 0$ 时，F 的概率密度显然为 0. 这就证得(1.3-12)式. ▌

1.3 数理统计中常见的几个分布

F 分布的概率密度函数的图形如图 1.3-3 所示.

χ^2 分布、t 分布及 F 分布都是由标准正态分布 $N(0;1)$ 诱导出来的. 因此,它们与正态总体下统计量的分布有着密切的联系.

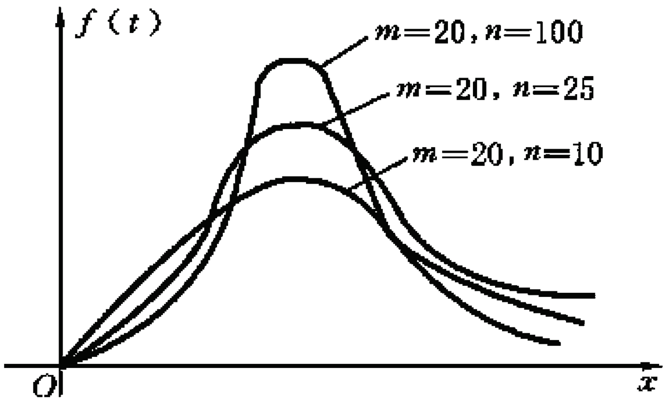

图 1.3-3　$F(m,n)$ 的概率密度函数

1.4 抽样分布

在统计推断中，经常要用到统计量的分布. 在一般情况下，要确定一个统计量的精确分布是很困难的，不过，如果总体服从正态分布，则可以求出一些重要统计量的分布. 正态总体在数理统计中占有特别重要的地位，一方面是因为在某些限制条件下假定总体服从正态分布是合理的，另一方面，数理统计中有许多总体的分布，都可以认为近似于正态分布. 因此研究正态总体下统计量的分布是很重要的.

定理 1.4-1 设 $(X_1, X_2, \cdots, X_n)$ 是取自正态总体 $N(\mu;\sigma^2)$ 的一个样本，$a_1, a_2, \cdots, a_n$ 是已知常数，则统计量 $U = \sum_{i=1}^{n} a_i X_i$ 服从正态分布 $N\left(\mu \sum_{i=1}^{n} a_i; \sigma^2 \sum_{i=1}^{n} a_i^2\right)$.

证 由于 $X_1, X_2, \cdots, X_n$ 相互独立，所以 U 是正态随机变量.

又因为

$$E(U) = \sum_{i=1}^{n} a_i E(X_i) = \mu \sum_{i=1}^{n} a_i,$$

$$D(U) = \sum_{i=1}^{n} D(a_i X_i) = \sum_{i=1}^{n} a_i^2 D(X_i) = \sigma^2 \sum_{i=1}^{n} a_i^2,$$

所以有

$$U \sim N\left(\mu \sum_{i=1}^{n} a_i; \sigma^2 \sum_{i=1}^{n} a_i^2\right). \qquad \blacksquare$$

信息学院 （国家示范性软件学院） School of Informatics
黄炜 博士 副教授 Dr. Wei Huang
厦門大學 XIAMEN UNIVERSITY

1.4 抽样分布

定理 1.4-2 设 $(X_1, X_2, \cdots, X_n)$ 是取自正态总体 $N(\mu;\sigma^2)$ 的一个样本，$\boldsymbol{A}=[a_{ij}]$ 是已知的 $p\times n$ 常数矩阵，又

$$[Y_1, Y_2, \cdots, Y_p]^{\mathrm{T}} = \boldsymbol{A}[X_1, X_2, \cdots, X_n]^{\mathrm{T}},$$

则

$$Y_i \sim N\left(\mu\sum_{j=1}^{n} a_{ij}, \sigma^2\sum_{j=1}^{n} a_{ij}^2\right), \quad i = 1, 2, \cdots, p;$$

$$\operatorname{cov}(Y_i, Y_j) = \sigma^2\sum_{k=1}^{n} a_{ik}a_{jk}, \quad i, j = 1, 2, \cdots, p.$$

请读者自证.

特别，当 $\mu=0$，且 $\boldsymbol{A}$ 为 n 阶正交矩阵时，$\operatorname{cov}(Y_i, Y_j) = \begin{cases}\sigma^2, i=j\\ 0, i\neq j\end{cases}$. 因此，统计量 Y_1，$Y_2, \cdots, Y_n$ 相互独立，并有相同的分布 $N(0,\sigma^2)$. 这就是说，均值为零且相互独立的同分布正态随机变量 $X_1, X_2, \cdots, X_n$，通过正交变换所得的随机变量 $Y_1, Y_2, \cdots, Y_n$，也是均值为零且相互独立的同分布正态随机变量.

1.4 抽样分布

定理 1.4-3 设$(X_1,X_2,\cdots,X_n)$是取自正态总体$N(\mu;\sigma^2)$的一个样本，则

(1) $\overline{X}\sim N\left(\mu;\frac{1}{n}\sigma^2\right)$, $\sqrt{n}\frac{\overline{X}-\mu}{\sigma}\sim N(0;1)$, (1.4-1)

(2) $\overline{X}$与S^2相互独立，且

$$\frac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}\frac{(X_i-\overline{X})^2}{\sigma^2}\sim\chi^2(n-1). \qquad (1.4\text{-}2)$$

证 (1) 由于$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，所以由定理1.4-1令$a_i=\frac{1}{n}(i=1,2,\cdots,n)$，即得$\overline{X}\sim N\left(\mu;\frac{1}{n}\sigma^2\right)$，从而

$$\frac{\overline{X}-\mu}{\sqrt{\frac{1}{n}\sigma^2}}=\sqrt{n}\frac{\overline{X}-\mu}{\sigma}\sim N(0;1).$$

(2) 令$Y_i=\frac{X_i-\mu}{\sigma}(i=1,2,\cdots,n)$，则$Y_1,Y_2,\cdots,Y_n$相互独立，且具有相同的分布$N(0;1)$. 作正交变换

$$\begin{cases} Z_1=\frac{1}{\sqrt{n}}Y_1+\frac{1}{\sqrt{n}}Y_2+\cdots+\frac{1}{\sqrt{n}}Y_n, \\ Z_i=c_{i1}Y_1+c_{i2}Y_2+\cdots+c_{in}Y_n, \quad i=2,3,\cdots,n, \end{cases}$$

黄 炜 博士/副教授 Dr. Wei Huang 信息学院（国家示范性软件学院）School of Informatics 厦門大學 XIAMEN UNIVERSITY

1.4 抽样分布

其中常数 $c_{ij}(i=2,3,\cdots,n,\quad j=1,2,\cdots,n)$ 使 n 阶方阵

$$\begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix}$$

为正交矩阵，这可以通过求 $n-1$ 维欧氏空间 $V=\left\{\boldsymbol{x}\in\mathbf{R}^n \mid \sum_{i=1}^{n} x_i = 0\right\}$ 的标准正交基，得到 c_{ij} 而找到这种正交矩阵. 于是 $Z_1, Z_2, \cdots, Z_n$ 相互独立，且都服从 $N(0;1)$. 记 $\overline{Y}=\frac{1}{n}\sum_{i=1}^{n} Y_i=\frac{\overline{X}-\mu}{\sigma}$，则 $Z_1=\sqrt{n}\overline{Y}=\sqrt{n}\frac{\overline{X}-\mu}{\sigma}\sim N(0;1)$. 又因为正交变换不改变向量的长度，故有

$$\sum_{i=1}^{n} Y_i^2 = \sum_{i=1}^{n} Z_i^2.$$

1.4 抽样分布

从而

$$\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}-\frac{\overline{X}-\mu}{\sigma}\right)^2=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2+\sum_{i=1}^{n}\left(\frac{\overline{X}-\mu}{\sigma}\right)^2$$

$$-2\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)\left(\frac{\overline{X}-\mu}{\sigma}\right)$$

$$=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2+n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2-2n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2$$

$$=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2-n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2=\sum_{i=1}^{n}Y_i^2-n\overline{Y}^2$$

$$=\sum_{i=1}^{n}Z_i^2-Z_1^2=\sum_{i=2}^{n}Z_i^2.$$

因此按 χ^2 分布的定义知

$$\frac{(n-1)S^2}{\sigma^2}=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2=\sum_{i=2}^{n}Z_i^2\sim\chi^2(n-1),$$

并且由于 $Z_1,Z_2,\cdots,Z_n$ 相互独立，故 Z_1 与 $\sum_{i=2}^{n}Z_i^2$ 相互独立，即$\sqrt{n}\frac{\overline{X}-\mu}{\sigma}$ 与 $\frac{(n-1)S^2}{\sigma^2}$ 相互独立，所以 $\overline{X}$ 与 S^2 相互独立. ∎

信息学院（国家示范性软件学院）School of Informatics 黄炜 博士·副教授 Dr. Wei Huang
厦門大學 XIAMEN UNIVERSITY

1.4 抽样分布

定理 1.4-4 设$(X_1,X_2,\cdots,X_n)$是取自正态总体$N(\mu;\sigma^2)$的一个样本，则

$$\sqrt{n}\,\frac{\bar{X}-\mu}{S}\sim t(n-1). \tag{1.4-3}$$

证 因为$\bar{X}$与S^2相互独立，且

$$\sqrt{n}\,\frac{\bar{X}-\mu}{\sigma}\sim N(0,1),\quad \frac{(n-1)S^2}{\sigma^2}\sim\chi^2(n-1),$$

所以由t分布的定义知 $\sqrt{n}\,\dfrac{\bar{X}-\mu}{S}=\dfrac{\sqrt{n}\,\dfrac{\bar{X}-\mu}{\sigma}}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2(n-1)}}}\sim t(n-1).$ $\blacksquare$

定理 1.4-5 设$(X_1,X_2,\cdots,X_m)$是取自正态总体$N(\mu_1;\sigma^2)$的样本，$(Y_1,Y_2,\cdots,Y_n)$是取自正态总体$N(\mu_2;\sigma^2)$的样本，且这两个样本相互独立，则

$$\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{S_w\sqrt{\dfrac{1}{m}+\dfrac{1}{n}}}\sim t(m+n-2), \tag{1.4-4}$$

其中 $$S_w^2=\frac{(m-1)S_1^2+(n-1)S_2^2}{m+n-2},\quad S_1^2=\frac{1}{m-1}\sum_{i=1}^{m}(X_i-\bar{X})^2,$$

$$S_2^2=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2.$$

厦門大學 XIAMEN UNIVERSITY
信息学院（国家示范性软件学院） School of Informatics
黄炜 博士副教授 Dr. Wei Huang

1.4 抽样分布

证 由 $\bar{X}$ 与 $\bar{Y}$ 相互独立以及 $\bar{X} \sim N\left(\mu_1; \dfrac{\sigma^2}{m}\right)$，$\bar{Y} \sim N\left(\mu_2; \dfrac{\sigma^2}{n}\right)$，得 $\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2; \sigma^2\left(\dfrac{1}{m} + \dfrac{1}{n}\right)\right)$，即

$$U = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sigma\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}} \sim N(0;1).$$

又因 $\dfrac{(m-1)S_1^2}{\sigma^2} \sim \chi^2(m-1)$，$\dfrac{(n-1)S_2^2}{\sigma^2} \sim \chi^2(n-1)$，

且它们相互独立，故由 χ^2 分布的可加性得

$$V = \frac{(m-1)S_1^2}{\sigma^2} + \frac{(n-1)S_2^2}{\sigma^2} \sim \chi^2(m+n-2).$$

从而由 U 与 V 相互独立和 t 分布的定义知

$$\frac{U}{\sqrt{\dfrac{V}{m+n-2}}} = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_w\sqrt{\dfrac{1}{m} + \dfrac{1}{n}}} \sim t(m+n-2). \quad \blacksquare$$

1.4 抽样分布

定理1.4-6 设$(X_1,X_2,\cdots,X_m)$是取自正态总体$N(\mu_1;\sigma_1^2)$的样本,$(Y_1,Y_2,\cdots,Y_n)$是取自正态总体$N(\mu_1;\sigma_2^2)$的样本,且这两个样本相互独立,则

$$\frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}=\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}\sim F(m-1,n-1). \qquad (1.4\text{-}5)$$

证 由于$\frac{(m-1)S_1^2}{\sigma_1^2}\sim\chi^2(m-1)$,$\frac{(n-1)S_2^2}{\sigma_2^2}\sim\chi^2(n-1)$,且它们相互独立,故由$F$分布的定义知

$$\frac{(m-1)S_1^2}{\sigma_1^2(m-1)}\Big/\frac{(n-1)S_2^2}{\sigma_2^2(n-1)}=\frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_2^2}\sim F(m-1,n-1). \quad \blacksquare$$

值得注意的是,上述两个定理中,定理1.4-5要求两个正态总体的方差相同,称为方差齐次;而定理1.4-6并不要求方差是相同的.

对于非正态总体,要求出统计量的分布是比较困难的,即使有时理论上可以求出精确的分布,但其形式复杂而难以应用,这时一般利用中心极限定理推出$n\to+\infty$下的极限分布.从而当样本容量很大时,总是应用有关统计量的极限分布作为其近似分布.

厦門大學 XIAMEN UNIVERSITY 信息学院(国家示范性软件学院) School of Informatics 黄炜 博士·副教授 Dr. Wei Huang

信息学院（国家示范性软件学院）School of Informatics
黄炜 博士/副教授 Dr. Wei Huang
厦門大學 XIAMEN UNIVERSITY

1.4 抽样分布

定理 1.4-7 设总体 X 的一阶矩、二阶矩存在：$E(X)=\mu, 0<D(X)=\sigma^2<+\infty$，$(X_1, X_2, \cdots, X_n)$ 是取自总体 X 的样本，则当 n 充分大时，

(1) 样本均值 $\overline{X}_n$ 近似地服从 $N\left(\mu;\dfrac{\sigma^2}{n}\right)$，

(2) 样本方差 S_n^2 依概率收敛于 σ^2，

(3) $\sqrt{n}\dfrac{\overline{X}_n-\mu}{S_n}$ 近似地服从 $N(0;1)$，其中

$$\overline{X}_n=\frac{1}{n}\sum_{i=1}^{n}X_i,\quad S_n^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2.$$

(请参阅何迎晖、闵华玲《数理统计》，北京：高等教育出版社，1989) ∎

信息学院（国家示范性软件学院）School of Informatics 黄炜 博士·副教授 Dr. Wei Huang 厦门大学 XIAMEN UNIVERSITY

1.5 分位数

数理统计中经常要用到分位数. 分位数是随机变量的一种数字特征，其定义如下.

定义 1.5-1 设随机变量 X 的分布函数为 $F(x)$，p 是一实数，$0<p<1$，若 x_p 使

$$P\{X \leqslant x_p\} = F(x_p) = p, \tag{1.5-1}$$

则称 x_p 为此分布的 p 分位数.

如果 $X \sim N(0;1)$，记它的 p 分位数为 u_p，即 u_p 满足

$$\Phi(u_p) = \int_{-\infty}^{u_p} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{u^2}{2}} \mathrm{d}u = p.$$

对于给定的 p，u_p 的值可由标准正态分布表（见附表一）查得，例如，$u_{0.95} = 1.645$，$u_{0.975} = 1.96$.

由于正态分布的概率密度函数是偶函数，所以有

$$\begin{aligned} P\{X \leqslant -u_{1-p}\} &= P\{X \geqslant u_{1-p}\} = 1 - P\{X < u_{1-p}\} \\ &= 1 - P\{X \leqslant u_{1-p}\} = 1 - (1-p) = p. \end{aligned}$$

从而由分位数的定义知 $\quad u_p = -u_{1-p} \quad$ 或 $\quad u_{1-p} = -u_p. \qquad (1.5\text{-}2)$

若 $X \sim \chi^2(n)$，用 $\chi_p^2(n)$ 表示此分布的 p 分位数，即

$$P\{X \leqslant \chi_p^2(n)\} = p.$$

$\chi_p^2(n)$ 的值可由 χ^2 分布表查得. 附表二对于某些 p 和较小的 n 给出了 $\chi_p^2(n)$ 的值. 例如，$\chi_{0.01}^2(8) = 1.646$，$\chi_{0.95}^2(10) = 18.307$. 当 $n > 45$ 时，可以用近似公式

$$\chi_p^2(n) \approx \frac{1}{2}(u_p + \sqrt{2n-1})^2 \tag{1.5-3}$$

得到 χ^2 分布的 p 分位数.

1.5 分位数

$t(n)$ 分布的 p 分位数 $t_p(n)$ 可由 t 分布表查得. 附表三对于某些大于 $\frac{1}{2}$ 的 p 和较小的 n 给出了 $t_p(n)$ 的值. 当 $p<\frac{1}{2}$ 时, 可以利用

$$t_p(n)=-t_{1-p}(n) \quad 或 \quad t_{1-p}(n)=-t_p(n) \tag{1.5-4}$$

得到 p 分位数. 例如, $t_{0.975}(14)=2.1448$, $t_{0.05}(7)=-t_{0.95}(7)=-1.8946$. 当 $n>45$ 时, 由于 $t(n)$ 近似于 $N(0;1)$, 所以可以用近似公式

$$t_p(n)\approx u_p \tag{1.5-5}$$

得到 t 分布的 p 分位数.

$F(m,n)$ 分布的 p 分位数记为 $F_p(m,n)$. 对于某些大于 $\frac{1}{2}$ 的 p 及某些 m、n, $F_p(m,n)$ 的值可以通过查附表四得到. 当 $p<\frac{1}{2}$ 时, 可以利用公式

$$F_p(m,n)=\frac{1}{F_{1-p}(n,m)} \tag{1.5-6}$$

得到 p 分位数. 这是因为

$$p=P\{F\leqslant F_p(m,n)\}=P\left\{\frac{1}{F}\geqslant\frac{1}{F(m,n)}\right\}=1-P\left\{\frac{1}{F}\leqslant\frac{1}{F_p(m,n)}\right\},$$

1.5 分位数

故由$\frac{1}{F} \sim F(n,m)$及分位数定义知，$\frac{1}{F_p(m,n)} = F_{1-p}(n,m)$，亦即(1.5-6)式成立. 例如，$F_{0.95}(5,9) = 3.48$，$F_{0.05}(5,9) = 1/F_{0.95}(9,5) = 1/4.77 = 0.21$.

对于分位数还需注意下列两种表示方法.

(1) 若λ使$P\{X > \lambda\} = p$，则称λ为X的上侧p分位数. 显然有

$$P\{X > \lambda\} = 1 - P\{X \leqslant \lambda\},$$

故X的上侧p分位数即为X的$1-p$分位数.

(2) 若λ_1、λ_2使

$$P\{X \leqslant \lambda_1\} = \frac{p}{2}, P\{X > \lambda_2\} = \frac{p}{2},$$

则称λ_1、λ_2为X的双侧p分位数. 显然，λ_1为X的$\frac{p}{2}$分位数，λ_2为X的$1-\frac{p}{2}$分位数.

作业（第3部分）

习题 1

7. 从总体 $N(20;(\sqrt{3})^2)$ 中抽取容量分别为 10 与 15 的两个相互独立的样本，求这两个样本均值之差的绝对值小于 0.3 的概率.

13. 设 $(X_1,\cdots,X_5)$ 是取自总体 $N(0;1)$ 的样本，

(1) 求常数 c_1、d_1，使 $c_1(X_1+X_2)^2+d_1(X_3+X_4+X_5)^2$ 服从 χ^2 分布，并指出其自由度.

(2) 求常数 c_2、d_2，使 $\dfrac{c_2(X_1^2+X_2^2)}{d_2(X_3+X_4+X_5)^2}$ 服从 F 分布，并指出其自由度.

謝謝觀看！

厦門大學
XIAMEN UNIVERSITY

信息学院
（国家示范性软件学院）
School of Informatics

黃 煒
博士·副教授
Dr. Wei Huang