高等工程數學 (2)

廈門大學
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

矩阵及其运算

矩阵论（2）

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

1.2 矩阵及其运算

数域 F 中的 $m \times n$ 个数 $a_{ij}(i=1,\cdots,m;j=1,\cdots,n)$ 排成 m 行 n 列的数表

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

叫做 F 上的 $m \times n$ 矩阵，简称矩阵，其中 a_{ij} 叫做此矩阵的第 i 行第 j 列元素. 元素都是实数的矩阵称为**实矩阵**；元素都是复数的矩阵称为**复矩阵**. 除非特别声明，一般讨论的都是实矩阵. 我们用大写英文字母 $\boldsymbol{A},\boldsymbol{B},\cdots$ 表示矩阵，如上面的矩阵可记作 $\boldsymbol{A}$. 为了方便也把此矩阵表示成

$$\boldsymbol{A}=[a_{ij}]_{m\times n} \quad 或 \quad \boldsymbol{A}=[a_{ij}].$$

特别地，当 $m=n$ 时，称 $\boldsymbol{A}$ 为 n **阶方阵**. n 阶方阵 $\boldsymbol{A}$ 的从左上角到右下角那条线叫做**主对角线**，简称**对角线**，其上的元素 $a_{11},a_{22},\cdots,a_{nn}$ 叫做 $\boldsymbol{A}$ 的**主对角线元素**.

主对角线元素全是 1，其余元素全为 0 的方阵称为**单位矩阵**，记做 $\boldsymbol{I}$.

除主对角线上的元素外，其余元素全为 0 的方阵称为**对角矩阵**.

满足 $\boldsymbol{A}^{\mathrm{T}}=\boldsymbol{A}$ 的实方阵 $\boldsymbol{A}$ 称为**实对称矩阵**，简称对称矩阵；满足 $\overline{\boldsymbol{A}^{\mathrm{T}}}=\boldsymbol{A}$ 的复方阵称为 **Hermite 矩阵**，通常把 $\overline{\boldsymbol{A}^{\mathrm{T}}}$ 记作 $\boldsymbol{A}^{\mathrm{H}}$，它表示 $\boldsymbol{A}$ 的转置再取共轭. 显然有 $\boldsymbol{A}^{\mathrm{H}}=(\bar{\boldsymbol{A}})^{\mathrm{T}}$，即 $\boldsymbol{A}^{\mathrm{H}}$ 是 $\boldsymbol{A}$ 的共轭转置.

1.2 矩阵及其运算

所有元素全是 0 的矩阵称为**零矩阵**，记作 $\boldsymbol{O}$.

设 $\boldsymbol{A}=[a_{ij}]$ 和 $\boldsymbol{B}=[b_{ij}]$ 都是 $m\times n$ 矩阵，如果它们的对应元素相等，即

$$a_{ij}=b_{ij},\quad i=1,2,\cdots,m;\ j=1,2,\cdots,n,$$

则称矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ **相等**，记作 $\boldsymbol{A}=\boldsymbol{B}$.

1.2.1 矩阵的运算

设 $\boldsymbol{A}=[a_{ij}]$，$\boldsymbol{B}=[b_{ij}]$ 都是 $m\times n$ 矩阵，定义 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的和为 $m\times n$ 矩阵 $[a_{ij}+b_{ij}]$，记为 $\boldsymbol{A}+\boldsymbol{B}$. 这种运算叫做矩阵的加法.

若 k 为一个数，定义 k 与 $\boldsymbol{A}$ 的乘积为 $m\times n$ 矩阵 $[ka_{ij}]$，记为 $k\boldsymbol{A}$ 或 $\boldsymbol{A}k$. 这种运算叫做矩阵的数乘.

矩阵的乘法是矩阵的重要运算之一. 设 $\boldsymbol{A}=[a_{ij}]$ 是 $m\times s$ 矩阵，$\boldsymbol{B}=[b_{ij}]$ 是 $s\times n$ 矩阵，则称 $m\times n$ 矩阵 $\boldsymbol{C}=[c_{ij}]$，其中 $c_{ij}=\sum_{k=1}^{s}a_{ik}b_{kj}$，$i=1,\cdots,m$，$j=1,\cdots,n$，为矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的乘积，记做 $\boldsymbol{C}=\boldsymbol{AB}$.

1.2.1 矩阵的运算

例 1 设 $A=\begin{bmatrix}1 & 2 & 3\\ -1 & 0 & -2\end{bmatrix}$，$B=\begin{bmatrix}2 & 1\\ -1 & 2\\ 0 & -1\end{bmatrix}$，

则 $$AB=\begin{bmatrix}1\times2+2\times(-1)+3\times0 & 1\times1+2\times2+3\times(-1)\\ (-1)\times2+0\times(-1)+(-2)\times0 & (-1)\times1+0\times2+(-2)\times(-1)\end{bmatrix}$$

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

$$BA=\begin{bmatrix}2\times1+1\times(-1) & 2\times2+1\times0 & 2\times3+1\times(-2)\\ (-1)\times1+2\times(-1) & (-1)\times2+2\times0 & (-1)\times3+2\times(-2)\\ 0\times1+(-1)\times(-1) & 0\times2+(-1)\times0 & 0\times3+(-1)\times(-2)\end{bmatrix}$$

$$=\begin{bmatrix}1 & 4 & 4\\ -3 & -2 & -7\\ 1 & 0 & 2\end{bmatrix}.$$ ■

由上例可见，矩阵乘法不满足交换律，即一般来说 $AB\neq BA$.

1.2.1 矩阵的运算

例 2 满足 $AB=BA$ 的矩阵 B 称为与 A 可交换的. 由矩阵乘法规则可知，若 A 与 B 是可交换的，则 A 与 B 必是同阶方阵. 设

$$A=\begin{bmatrix}2 & 1\\0 & 3\end{bmatrix},$$

求与 A 可交换的所有矩阵.

解 设与 A 可交换的矩阵 B 为 $B=\begin{bmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{bmatrix}$,

则由 $AB=BA$，即

$$\begin{bmatrix}2 & 1\\0 & 3\end{bmatrix}\begin{bmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{bmatrix}=\begin{bmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}2 & 1\\0 & 3\end{bmatrix},$$

得到

$$\begin{cases}2b_{11}+b_{21}=2b_{11},\\2b_{12}+b_{22}=b_{11}+3b_{12},\\3b_{21}=2b_{21},\\3b_{22}=b_{21}+3b_{22}.\end{cases}$$

因此 $b_{21}=0$，$b_{22}=b_{11}+b_{12}$，于是，与 A 可交换的所有矩阵是

$$B=\begin{bmatrix}b_{11} & b_{12}\\0 & b_{11}+b_{12}\end{bmatrix}.$$

∎

1.2.1 矩阵的运算

对于方阵 $\boldsymbol{A}$，k 个 $\boldsymbol{A}$ 相乘称为 $\boldsymbol{A}$ 的 k 次幂，记为 $\boldsymbol{A}^k$. 规定 $\boldsymbol{A}^0=\boldsymbol{I}$，即 $\boldsymbol{A}$ 的零次幂是单位矩阵.

对于任意非负整数 k、l，方阵 $\boldsymbol{A}$ 的幂满足

1.2.1 矩阵的运算

$$\boldsymbol{A}^k\boldsymbol{A}^l=\boldsymbol{A}^{k+l},\quad (\boldsymbol{A}^k)^l=\boldsymbol{A}^{kl}.$$

由于矩阵乘法不满足交换律，一般地说，

$$(\boldsymbol{A}+\boldsymbol{B})^2\neq\boldsymbol{A}^2+2\boldsymbol{A}\boldsymbol{B}+\boldsymbol{B}^2,\quad (\boldsymbol{A}\boldsymbol{B})^k\neq\boldsymbol{A}^k\boldsymbol{B}^k.$$

1.2.1 矩阵的运算

例 3 设矩阵 $\boldsymbol{J}=\begin{bmatrix}-2 & 1 & 0 & 0\\ 0 & -2 & 1 & 0\\ 0 & 0 & -2 & 1\\ 0 & 0 & 0 & -2\end{bmatrix}$（在 2.1 节中称之为 Jordan 块），求 $\boldsymbol{J}^2$、$\boldsymbol{J}^3$ 和 $\boldsymbol{J}^k(k\geqslant 4)$.

解 令
$$\boldsymbol{U}=\begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\end{bmatrix},$$

不难算出
$$\boldsymbol{U}^2=\begin{bmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\quad \boldsymbol{U}^3=\begin{bmatrix}0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\quad \boldsymbol{U}^k=\boldsymbol{O}(k\geqslant 4).$$

由于 $\boldsymbol{J}=(-2)\boldsymbol{I}+\boldsymbol{U}$，且单位矩阵与任何方阵是可交换的，所以

$$\boldsymbol{J}^2=((-2)\boldsymbol{I}+\boldsymbol{U})^2=(-2)^2\boldsymbol{I}+2\times(-2)\boldsymbol{U}+\boldsymbol{U}^2=\begin{bmatrix}4 & -4 & 1 & 0\\ & 4 & -4 & 1\\ & & 4 & -4\\ & & & 4\end{bmatrix},$$

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

1.2.1 矩阵的运算

$$
\begin{aligned}
\boldsymbol{J}^k &= ((-2)\boldsymbol{I}+\boldsymbol{U})^k = (-2)^k\boldsymbol{I}+C_k^1(-2)^{k-1}\boldsymbol{U}+C_k^2(-2)^{k-2}\boldsymbol{U}^2+C_k^3(-2)^{k-3}\boldsymbol{U}^3 \\
&= (-2)^k\boldsymbol{I}+k(-2)^{k-1}\boldsymbol{U}+\frac{k(k-1)}{2!}(-2)^{k-2}\boldsymbol{U}^2+\frac{k(k-1)(k-2)}{3!}(-2)^{k-3}\boldsymbol{U}^3 \\
&= \begin{bmatrix}
(-2)^k & k(-2)^{k-1} & \frac{k(k-1)}{2!}(-2)^{k-2} & \frac{k(k-1)(k-2)}{3!}(-2)^{k-3} \\
 & (-2)^k & k(-2)^{k-1} & \frac{k(k-1)}{2!}(-2)^{k-2} \\
 & & (-2)^k & k(-2)^{k-1} \\
 & & & (-2)^k
\end{bmatrix},
\end{aligned}
$$

J^2、J^3 和 J^k 中未写出的元素皆为 0. ■

1.2.1 矩阵的运算

例 4　设 $\boldsymbol{A}$ 是 $m\times n$ 复矩阵，证明 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{O}$ 的充要条件是 $\boldsymbol{A}=\boldsymbol{O}$.

证　若 $\boldsymbol{A}=\boldsymbol{O}$，则显然有 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{O}$. 下证必要性. 设

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

则

$$\boldsymbol{A}^{\mathrm{H}}=\begin{bmatrix} \overline{a_{11}} & \overline{a_{21}} & \cdots & \overline{a_{m1}} \\ \overline{a_{12}} & \overline{a_{22}} & \cdots & \overline{a_{m2}} \\ \vdots & \vdots & & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & \overline{a_{mn}} \end{bmatrix},$$

从而 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$ 的主对角线元素是

$$|a_{1j}|^2+|a_{2j}|^2+\cdots+|a_{mj}|^2,\quad j=1,2,\cdots,n.$$

由 $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}=\boldsymbol{O}$ 推出 $\sum_{i=1}^{m}|a_{ij}|^2=0(j=1,2,\cdots,n)$，即有

$$a_{ij}=0,\quad i=1,2,\cdots,m;\ j=1,2,\cdots,n.$$

因此，$\boldsymbol{A}=\boldsymbol{O}$. ▌

1.2.2 可逆矩阵与逆矩阵

设 $\boldsymbol{A}$ 是一个 n 阶方阵，如果存在一个 n 阶方阵 $\boldsymbol{G}$，使得

$$\boldsymbol{AG}=\boldsymbol{GA}=\boldsymbol{I},$$

则称 $\boldsymbol{A}$ 是**可逆矩阵**，并称 $\boldsymbol{G}$ 是 $\boldsymbol{A}$ 的**逆矩阵**；记为 $\boldsymbol{G}=\boldsymbol{A}^{-1}$.

给定 n 阶方阵 $\boldsymbol{A}=[a_{ij}]$，如何判定它是否可逆呢？为此引进 $\boldsymbol{A}$ 的伴随矩阵. 令 A_{ij} 是元素 a_{ij} 的代数余子式，则 $\boldsymbol{A}$ 的**伴随矩阵** $\mathrm{adj}\boldsymbol{A}$ 定义为

$$\mathrm{adj}\boldsymbol{A}=[A_{ij}]^{\mathrm{T}}. \tag{1.2-1}$$

方阵 $\boldsymbol{A}$ 与其伴随矩阵 $\mathrm{adj}\boldsymbol{A}$ 之间有下述关系：

$$\boldsymbol{A}\cdot\mathrm{adj}\boldsymbol{A}=\mathrm{adj}\boldsymbol{A}\cdot\boldsymbol{A}=|\boldsymbol{A}|\boldsymbol{I}, \tag{1.2-2}$$

其中 $|\boldsymbol{A}|$ 是 $\boldsymbol{A}$ 的行列式（有时也把方阵 $\boldsymbol{A}$ 的行列式记为 $\det\boldsymbol{A}$）. 这是伴随矩阵的一个重要性质.

由(1.2-2)式可知，$\boldsymbol{A}$ 可逆的充要条件是 $\boldsymbol{A}$ 的行列式不等于零. 当 $\boldsymbol{A}$ 可逆时，

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\mathrm{adj}\boldsymbol{A}.$$

1.2.2 可逆矩阵与逆矩阵

例 5 设 $\boldsymbol{A}=\begin{bmatrix}4&3&2\\3&2&1\\2&1&1\end{bmatrix}$，证明 $\boldsymbol{A}$ 可逆，并求 $\boldsymbol{A}^{-1}$.

解 由于

$$|\boldsymbol{A}|=\begin{vmatrix}4&3&2\\3&2&1\\2&1&1\end{vmatrix}=8+6+6-8-4-9=-1\neq 0,$$

故 $\boldsymbol{A}$ 可逆，且

$$\boldsymbol{A}^{-1}=\frac{1}{|\boldsymbol{A}|}\mathrm{adj}\boldsymbol{A}=\frac{1}{-1}\begin{bmatrix}1&-1&-1\\-1&0&2\\-1&2&-1\end{bmatrix}=\begin{bmatrix}-1&1&1\\1&0&-2\\1&-2&1\end{bmatrix}.\quad\blacksquare$$

1.2.2 可逆矩阵与逆矩阵

可逆矩阵有许多好性质，比较重要的叙述如下.

(1) 若 $\boldsymbol{A}$ 可逆，数 $k\neq 0$，则 $k\boldsymbol{A}$ 可逆，且

$$(k\boldsymbol{A})^{-1}=\frac{1}{k}\boldsymbol{A}^{-1};$$

(2) 若 $\boldsymbol{A}$、$\boldsymbol{B}$ 是同阶可逆方阵，则 $\boldsymbol{AB}$ 也可逆，且

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

(3) 若 $\boldsymbol{A}$ 可逆，则 $\boldsymbol{A}^{\mathrm{H}}$ 可逆，且

$$(\boldsymbol{A}^{\mathrm{H}})^{-1}=(\boldsymbol{A}^{-1})^{\mathrm{H}}.$$

对于高阶矩阵，用伴随矩阵求其逆矩阵是比较麻烦的，较好的方法是利用矩阵的初等变换.

1.2.3 分块矩阵

对于行数和列数较大的矩阵，经常采用“矩阵分块法”，即把一个大矩阵看成是以一些小矩阵为元素的矩阵. 在运算中，常把这些小矩阵当作“元素”来处理.

将一个矩阵 $\boldsymbol{A}$ 用若干条纵线和横线分成许多小矩阵，每个小矩阵叫做 $\boldsymbol{A}$ 的子块. 以这些子块为元素的形式上的矩阵称为**分块矩阵**.

分块矩阵的运算规则与普通矩阵的运算规则类似，例如若矩阵 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的行数、列数分别相同，并且 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 的分块方式也相同，则 $\boldsymbol{A}$ 与 $\boldsymbol{B}$ 相加时可将对应子块相加；数 k 与 $\boldsymbol{A}$ 相乘时可将 k 乘到各子块上去. 对于分块矩阵的乘法，只要前一个矩阵列的分块方式与后一矩阵行的分块方式一致，就可以将子块看成元素而按矩阵乘法规则进行运算.

形如
$$\begin{bmatrix} \boldsymbol{A}_1 & & \\ & \ddots & \\ & & \boldsymbol{A}_p \end{bmatrix}$$
的分块矩阵称为**分块对角矩阵**或**准对角矩阵**，它的特点是不在主对角线上的子块都是零矩阵，而主对角线上的子块均为方阵.

显然，当位于主对角线上的方阵 $\boldsymbol{A}_i(i=1,\cdots,p)$ 均为一阶矩阵时，就是对角矩阵.

作业（第1部分）

习题 1.2

4. 已知 $\boldsymbol{AP}=\boldsymbol{PB}$，其中

$$\boldsymbol{B}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&-1\end{bmatrix},\quad \boldsymbol{P}=\begin{bmatrix}1&0&1\\-1&1&1\\0&-1&1\end{bmatrix},$$

求 $\boldsymbol{A}$ 及 $\boldsymbol{A}^8$.

6. 设 $\boldsymbol{A}=\begin{bmatrix}1&2&0&0\\2&6&0&0\\0&0&3&5\\0&0&1&2\end{bmatrix}$，求 $|\boldsymbol{A}|$ 和 $|\boldsymbol{A}|^{-1}$.

数理统计的基本概念

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

前言

- 数理统计学是应用数学的一个重要分支， 它应用概率论的基本理论，根据试验或观测所获得的数据，对所研究的随机现象的客观规律性作出适当的、合理的估计和推断. 由于随机现象的统计规律性是在大量重复试验中呈现出来的，因此从理论上说，只需对所研究的随机现象进行足够多次的试验或观测， 便能揭示其规律性. 但是，实际上所允许的试验或观测往往是有限的，甚至是少谥的. 例如，我们不可能把生产出的所有炮弹都进行试射， 以研究这批炮弹的性能， 而只能抽取其中的一部分或几枚炮弹进行试射，获取有关数据，并通过这些数据对所研究的该批炮弹的性能作出估计和推断.

前言

- 通常把从所研究对象的全体中抽取一部分进行试验或观测，并获取试验数据的工作称为抽样．由于抽样所得的数据不可能包含研究对象的全部信息，从而据此而作出的估计和推断必然含有不确定性，对这种不确定性的处理就需要应用概率论的基本理论．数理统计学的基本任务是，研究如何以有效的方式收集、整理和分析受到随机性影响的数据，并对所研究的问题作出推断和预测，直至为采取决策和行动提供依据或者建议．因此，数理统计包括两个方面的内容：一是怎样合理和有效地收集数据—抽样方法与试验设计；二是由收集到的数据怎样比较正确地分析和合理地推测所研究对象的整体情况——统计推断．

前言

- 我们主要讲述统计推断的基本原理和方法，并应用统计推断的基本原理具体处理线性模型中的问题，同时也介绍一下合理安排试验及对试验所得的数据进行统计分析的正交试验设计方法.
- 数理统计方法的应用非常广泛，几乎在人类活动的所有领域内都能程度不同地找到它的应用，其原因是，实验是科学研究的根本方法，随机性因素对实验结果的影响无所不在反过来，这些应用又推动了数理统计理论和方法的发展，丰富了数理统计学研究的内容.

第一章 数理统计的基本概念

- 本章介绍数理统计的一些基本概念，如总体、样本、统计量等， 并在此基础上应用概率论，讨论某些常用统计量的分布.

1.1 总体和样本

数理统计中把所研究对象的全体称为**总体**，而把组成总体的每个元素称为**个体**. 例如，某厂生产了一大批灯泡，如果规定一个灯泡的使用寿命低于1000小时为次品，那么问题是如何确定这批灯泡的次品率？对于这个数理统计问题，这批灯泡的全体构成一个总体，其中每个灯泡是个体. 显然，研究的问题不同，则总体也会不同. 例如，研究某半导体收音机厂本月份产品的质量问题与研究其全年产品的质量问题，总体当然是不同的. 但为了研究总体的性质，我们就要按一定规则从总体中抽取若干个个体，这些个体称为该总体的一个**样本**，样本中所含个体的数目称为**样本容量**或**样本大小**.

实际工作中所关心的常常是总体中个体的某项数量指标，用 X 表示，其值一般随个体的不同而不同. 由于在数理统计问题中，我们从总体中抽取个体是随机的，所以 X 是一个随机变量. 于是，总体可以看成为某个随机变量所能取得的值的全体，个体是其中的一个值. 今后，凡提到总体总是指一个具有一定概率分布的随机变量，用 X、Y 等表示. 如果表征总体的随机变量 X 的分布函数为 $F(x)$，则此时我们也说总体为 $F(x)$. 如果总体所含的个体数目是有限的，则称此总体为**有限总体**；否则称为**无限总体**. 有时为了计算上和讨论问题的方便，当有限总体的个体数目很大时，也把它作为无限总体来处理.

当所考察的数量指标多于一项时，例如，我们要同时研究某校学生的身高和体重情况，那么总体可以用（二维）随机向量 (X,Y) 来表示，其中 X 与 Y 分别表示该校学生的身高与体重，这样的总体称为**二维总体**.

1.1 总体和样本

在数理统计学中，总体 X 的分布(分布函数 $F(x)$或概率密度 $f(x)$)通常是不能确知的，但它是客观存在的，故称其为理论分布，并且这正是我们所要了解的. 为了推断总体分布，就必须对总体进行抽样观测. 从总体中抽取一个个体，实质上是做一次试验并记录其结果. 如果进行了 n 次抽样观测，则得到总体的一组观测值 $x_1,x_2,\cdots,x_n$，其中 x_i 是第 i 次抽样观测的结果. 自然希望这组抽样观测数据能够很好地反映总体的情况，这就对抽取个体的方法有一定的要求：一是抽取个体必须是随机的，即总体中每个个体被抽到的机会是均等的；二是每次抽取的结果既不影响其他各次抽取的结果，也不受其他各次抽取结果的影响，亦即抽取个体必须是独立的，并且抽取一个个体后总体的情况不改变. 满足这样两个要求的抽样称为**简单随机抽样**，简单随机抽样的结果$(x_1,x_2,\cdots,x_n)$称为总体 X 的容量为 n 的样本观测值.

抽样，就具体方法而论，好比在袋中摸球. 对于有限总体，必定是有放回抽样，即将每个个体看成一个球，每次随机地从袋中摸取一个球，经过观测后放回并拌匀，然后再摸下一个球. 对于无限总体，放回与否并不改变总体的情况，因而可用不放回抽样，这样更方便些. 在实用中，即使总体为有限数 N，只要抽取的个体数 n 与 N 之比很小(经验是$\frac{n}{N}\leqslant 0.1$)时，可以近似地把不放回抽样看为有放回抽样. 今后若无特别说明，抽样总是指简单随机抽样.

1.1 总体和样本

设$(x_1,x_2,\cdots,x_n)$是总体 X 的一个容量为 n 的样本观测值，由于抽样的随机性和独立性，每个 $x_i(i=1,2,\cdots,n)$都可看成某个随机变量 X_i 所取的观测值，这里，$X_1,X_2,\cdots,X_n$ 相互独立，且都与总体 X 有相同的分布. 因此$(x_1,x_2,\cdots,x_n)$又可以看成为 n 维随机向量$[X_1,X_2,\cdots,X_n]$的一个实现. 在作观测前，样本观测值是不确定的，容量为 n 的一个样本$(X_1,X_2,\cdots,X_n)$是一个 n 维随机向量，它可能取值的全体称为**样本空间**. 样本空间可以是 n 维欧氏空间，也可以是 n 维欧氏空间的一个子集合. 但在观测后，样本观测值$(x_1,x_2,\cdots,x_n)$是完全确定的，它是一组已知数据，是样本空间中的一个点. 现用数学定义表述这些概念.

定义 1.1-1 设总体 X 的分布函数为 $F(x)$，如果 n 个随机变量 $X_1,X_2,\cdots,X_n$ 相互独立，且具有与 X 相同的分布函数 $F(x)$，则称$(X_1,X_2,\cdots,X_n)$是取自总体 X 的一个容量为 n 的**简单随机样本**，简称为**样本**，其观测值$(x_1,x_2,\cdots,x_n)$称为**样本观测值**.

因此，若$(X_1,X_2,\cdots,X_n)$是取自总体 X 的样本，且 X 的分布函数为 $F(x)$，那么这个样本的分布函数为

$$F^*(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}F(x_i). \qquad (1.1\text{-}1)$$

1.1 总体和样本

又若 X 的概率密度为 $f(x)$，则样本 $(X_1,X_2,\cdots,X_n)$ 的概率密度为

$$f^*(x_1,x_2,\cdots,x_n)=\prod_{i=1}^{n}f(x_i). \tag{1.1-2}$$

例 1 设 $(X_1,X_2,\cdots,X_n)$ 是取自正态总体 $N(\mu;\sigma^2)$ 的样本，则这个样本的概率密度为

$$\begin{aligned}f^*(x_1,x_2,\cdots,x_n)&=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]\\&=(2\pi\sigma^2)^{-\frac{n}{2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]\\&-\infty<x_1,x_2,\cdots,x_n<+\infty.\end{aligned}$$ ■

例 2 某厂要检查一批产品的质量，每件产品可区分为合格品与不合格品，用数字“0”表示合格品，用数字“1”表示不合格品，这时，总体 X 服从 0-1 分布：

$$P\{X=x\}=\begin{cases}p, x=1,\\1-p, x=0,\end{cases}$$

其中 p 表示一件产品为不合格品的概率. 上式可改写为

$$P\{X=x\}=p^x(1-p)^{1-x},\quad x=0,1. \tag{1.1-3}$$

若 $(X_1,X_2,\cdots,X_n)$ 是取自这个总体 X 的一个样本，则有

$$P\{X_1=x_1,X_2=x_2,\cdots,X_n=x_n\}=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}=p^{\sum_{i=1}^{n}x_i}(1-p)^{n-\sum_{i=1}^{n}x_i},$$

$$x_i=0,1, i=1,2,\cdots,n. \tag{1.1-4}$$ ■

1.2 统计量与样本矩

数理统计的任务之一，是利用样本所提供的信息对总体的情况进行推断. 但样本常常表现为一大批数据，即样本观测值，很难直接用来解决所研究的具体问题，从而需要对这些数据进行加工整理，将所需的信息集中起来. 也就是说，针对所研究的具体问题对样本进行一些运算，构造出样本的某种函数. 这种样本的函数称为统计量，其确切的定义如下.

定义 1.2-1 设$(X_1,X_2,\cdots,X_n)$为总体 X 的一个样本，$g(X_1,X_2,\cdots,X_n)$是$(X_1,X_2,\cdots,X_n)$的连续函数，如果 g 中不包含任何未知参数，则称 $g(X_1,X_2,\cdots,X_n)$为一个**统计量**.

统计量的基本特征是，当我们取得样本观测值$(x_1,x_2,\cdots,x_n)$时，即可得到统计量 $g(X_1,X_2,\cdots,X_n)$的观测值 $g(x_1,x_2,\cdots,x_n)$.

例如，设$(X_1,X_2,\cdots,X_n)$是取自正态总体 $N(\mu,\sigma^2)$的一个样本，且 μ 已知而 σ^2 未知，则 X_1，$2X_1-X_2$，$X_3-\mu$，$\frac{1}{3}(X_1+X_2+X_3)$，$\max\{X_1,X_2,X_3\}$等都是统计量；但$\frac{X_3-\mu}{\sigma}$，$\frac{1}{\sigma^2}(X_1^2+X_2^2+X_3^2)$等都不是统计量，因为它们含有未知参数 σ 或 σ^2.

由于样本$(X_1,X_2,\cdots,X_n)$是随机向量，而统计量是样本的函数，所以统计量也是随机变量，因而有相应的确定的概率分布.

常用的统计量为样本矩.

1.2 统计量与样本矩

定义 1.2-2 设$(X_1,X_2,\cdots,X_n)$是取自总体 X 的一个容量为 n 的样本，对任一正整数 k，称

$$M_k=\frac{1}{n}\sum_{i=1}^{n}X_i^k \tag{1.2-1}$$

为样本的 k **阶原点矩**；称

$$M'_k=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^k \tag{1.2-2}$$

为样本的 k **阶中心矩**，其中

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \tag{1.2-3}$$

称为**样本均值**，它实际是样本的 1 阶原点矩 M_1.

我们把

$$S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right) \tag{1.2-4}$$

称为**样本方差**，它表示样本偏离 $\overline{X}$ 的离散性特征.

值得指出的是，样本的 2 阶中心矩

$$M'_2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

与(2.1-4)式所给的样本方差 S^2 略有不同，

$$M'_2=\frac{n-1}{n}S^2,$$

通常将 M'_2 记为 $\tilde{S}^2$，并称 $\tilde{S}^2$ 为**修正的样本方差**.

1.2 统计量与样本矩

若样本观测值为$(x_1,x_2,\cdots,x_n)$，则

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\quad s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

分别为 $\overline{X}$、S^2 的观测值，且称 $s=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$ 为**样本标准差** $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$ 的观测值.

另一类有用的统计量是次序统计量. 设$(X_1,X_2,\cdots,X_n)$是取自总体 X 的样本，$(x_1,x_2,\cdots,x_n)$是样本观测值，把样本观测值按从小到大的次序排列成

$$x_{(1)}\leqslant x_{(2)}\leqslant\cdots\leqslant x_{(n-1)}\leqslant x_{(n)},$$

其中 $x_{(1)}$是 $x_1,x_2,\cdots,x_n$ 中最小的一个，$x_{(2)}$是 $x_1,x_2,\cdots,x_n$ 中次小的一个，$\cdots$，$x_{(n)}$是 $x_1,x_2,\cdots,x_n$ 中最大的一个.

定义 1.2-3 设$(X_1,X_2,\cdots,X_n)$是取自总体 X 的样本，不论样本观测值$(x_1,x_2,\cdots,x_n)$取怎样的数值，如果 $X_{(i)}(i=1,2,\cdots,n)$总是以 $x_{(i)}$为其观测值，则称 $X_{(i)}$为**第 i 个次序统计量**.

由次序统计量的定义可知，n 个次序统计量 $X_{(1)},X_{(2)},\cdots,X_{(n)}$总满足关系式

$$X_{(1)}\leqslant X_{(2)}\leqslant\cdots\leqslant X_{(n-1)}\leqslant X_{(n)},$$

且

$$X_{(1)}=\min_{1\leqslant i\leqslant n}X_{(i)},\ X_{(n)}=\max_{1\leqslant i\leqslant n}X_{(i)},$$

$X_{(1)}$称为最小次序统计量，$X_{(n)}$称为最大次序统计量.

1.2 统计量与样本矩

定义 1.2-4 设$(X_1,X_2,\cdots,X_n)$是取自总体 X 的样本，则称统计量 $R=X_{(n)}-X_{(1)}$为**样本极差**，称统计量

$$\operatorname{med}(X_1,X_2,\cdots,X_n)=\begin{cases} X_{(\frac{n+1}{2})}, & n\text{ 为奇数}, \\ \dfrac{1}{2}\left(X_{(\frac{n}{2})}+X_{(1+\frac{n}{2})}\right), & n\text{ 为偶数} \end{cases}$$

为**样本中位数**.

样本极差与样本中位数分别反映了数据在数轴上分布的离散性特征与位置特征.

设$(X_1,X_2,\cdots,X_n)$是取自总体 X 的样本，$(x_1,x_2,\cdots,x_n)$是样本观测值，对任一 $x\in(-\infty,+\infty)$，定义

$$F_n(x)=\frac{1}{n}(X_1,X_2,\cdots,X_n\text{ 中小于等于 }x\text{ 的个数}) \tag{1.2-5}$$

为经验分布，而称

$$F_n^*(x)=\frac{1}{n}(x_1,x_2,\cdots,x_n\text{ 中小于等于 }x\text{ 的个数}) \tag{1.2-6}$$

为总体 X 的经验分布函数，不难看出

$$F_n^*(x)=\begin{cases} 0, & x<x_{(1)}, \\ \dfrac{k}{n}, & x_{(k)}\leqslant x<x_{(k+1)}, \quad k=1,2,\cdots,n-1, \\ 1, & x\geqslant x_{(n)}. \end{cases} \tag{1.2-6$'$}$$

显然，经验分布是一个随机变量，只有在取得样本观测值后，其观测值 $F_n^*(x)$才是分布函数，具有分布函数的性质：$0\leqslant F_n^*(x)\leqslant 1$，$F_n^*(x)$非减及右连续，且在每个观测值处有跳跃 $1/n$.

1.2 统计量与样本矩

(1.2-5)式表明，经验分布 $F_n(x)$ 是样本 $(X_1,X_2,\cdots,X_n)$ 中不大于 x 的个体数与 n 之比，亦即在 n 次重复独立试验中，事件 $\{X\leqslant x\}$ 发生的频率，而总体 X 的分布函数 $F(x)$ 是事件 $\{X\leqslant x\}$ 发生的概率. 自然会问，当 n 充分大时，$F_n(x)$ 是否趋向于 $F(x)$ 呢？Гливенко 对这个问题作了肯定的回答：$F_n(x)$ 以概率 1 关于 x 一致收敛于 $F(x)$，即

$$P\{\lim_{n\to+\infty}\sup_{-\infty<x<+\infty}|F_n(x)-F(x)|=0\}=1.$$

Гливенко 定理表明，$F_n(x)$ 与 $F(x)$ 之差的绝对值在 $(-\infty,+\infty)$ 上的上确界当 $n\to+\infty$ 时，极限为 0 的概率是 1. 因此，只要 n 充分大，$F_n^*(x)$ 是 $F(x)$ 的很好的近似，而 $F_n^*(x)$ 是可由样本观测值得到的. 这就是用样本推断总体的理论基础.

作业（第2部分）

习题 1

2. 设(X_1,X_2,X_3)是取自总体 X 的样本，就下列总体 X 给出这个样本的概率函数或概率密度：

(1) X 服从参数为λ 的 Poisson 分布；

(2) X 服从参数为λ 的指数分布，即概率密度为 $f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0,\\ 0, & x\leqslant 0;\end{cases}$

(3) X 在区间(a,b)上服从均匀分布；

(4) X 服从参数为a,b的β分布，即概率密度为 $f(x;a,b)=\begin{cases}\dfrac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}, & 0<x<1,\\ 0, & \text{其他},\end{cases}$ 记为 $X\sim B(a,b)$.

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

黄 煒
博士，副教授
Dr. Wei Huang