旅游研究定量分析方法预习材料

杨旸（yangy@temple.edu）

一. 线性代数部分

1.1 基本概念

定义 1 由 $m \times n$ 个数 a_{ij} $(i = 1, 2, \cdots, m;\ j = 1, 2, \cdots, n)$ 排成的 m 行、n 列的数表

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

称为 m **行** n **列矩阵**，简称 $m \times n$ **矩阵**。一般用大写字母 ***A*，*B*，*C*……** 等表示，并记为：

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

为了便于表示，常把 $m \times n$ 矩阵简记为：$\boldsymbol{A} = \left(\boldsymbol{a_{ij}} \right)_{\boldsymbol{m \times n}}$。其中，$a_{ij}$ 称为矩阵的**元素**，第一下标 i 称为**行标**，表示这个元素处在第 i 行；第二下标 j 称为**列标**，表示这个元素处在第 j 列。所有元素及其相应位置是个整体，所以要加上一个大括弧表示它。a_{ij} 也称为**矩阵 *A*** 的（$\boldsymbol{i, j}$）**元**。

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

这个矩阵的特点是：**它是 n 阶方阵，它的主对角线上的元素都是 1，其它元素都是零**。称这样的矩阵为**单位矩阵**，通常记作 I_n。

方阵有对角线，从左上角到右下角的直线称为**方阵的主对角线**，从右上角到左下角的直线称为**方阵的副对角线**。主对角线一侧所有元素都为零的方阵，称为**三角形矩阵**。三角形矩阵分为**上三角矩阵**与**下三角矩阵**：

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} 与 \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

$1 \times n$ 矩阵

$$A = (a_1, a_2, \cdots, a_n)$$

称为**行矩阵**（也称为 n **维行向量**）；

$m \times 1$ 矩阵

$$B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

称为**列矩阵**（也称为 m **维列向量**）。换句话说，只有一行的矩阵称为**行矩阵**，只有一列的矩阵称为**列矩阵**。

1.2 基本运算

矩阵的**加法**与**矩阵**的数乘两种运算统称为矩阵的线性运算. 它满足下列运算规律:

设 A, B, C, O 都是同型矩阵， k, l 是常数，则

（1） $A + B = B + A;$

（2） $(A + B) + C = A + (B + C)$ ；

（3） $1A = A;$

（4） $k(l)A = (klA);$

（5） $(k + l)A = kA + lA;$

（6） $k(A + B) = kA + kB.$

矩阵的**乘法**满足下列运算规律(假定运算都是可行的):

（1） $(AB)C = A(BC);$

（2） $(A + B)C = AC + BC;$

（3） $C(A + B) = CA + CB;$

（4） $k(AB)=(kA)B=A(kB)$.

注：矩阵的乘法一般不满足交换律，即 $AB\neq BA$;

定义 2　把矩阵 A 的行换成同序数的列得到的新矩阵，称为 A 的**转置矩阵**，记作 A^T（或 A'）. 即若

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

则

$$A^T=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{pmatrix}.$$

矩阵的转置满足以下运算规律(假设运算都是可行的)：

(1) $(A^T)^T=A$;

(2) $(A+B)^T=A^T+B^T$;

(3) $(kA)^T=kA^T$;

(4) $(AB)^T=B^TA^T$.

定义 3：设 $A=(a_{ij})_{n\times n}$，若存在 n 阶方阵 B，使得 $AB=BA=I$，则称 A 为**可逆矩阵**，并称 B 为 A 的**逆矩阵**，记 $B=A^{-1}$。

1.3 线性相关

定义 4：对 n 维向量 $\alpha_1,\alpha_2,\cdots,\alpha_s$，若存在一组不全为零的数 $k_1,k_2,\cdots,k_s$，使 $k_1\alpha_1+\cdots+k_s\alpha=0$，则称向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **线性相关**。否则，称 $\alpha_1,\alpha_2,\cdots,\alpha_s$ **线性无关**。

如果矩阵的列向量组或者行向量组线性相关，则该矩阵**不可逆**。

1.4 特征根与特征矩阵

定义 5：方阵 A, 若数 λ_0 和非零向量 ξ, 使得

$$A\xi=\lambda_0\xi,$$

则称 λ_0 是 A 的**特征值**, ξ 是 A 的属于 λ_0 的**特征向量**.

二. 统计与概率论部分

2.1 总体与样本

总体是具有一定的共同属性的研究对象全体。组成总体的每一个元素称为个体。在数理统计中，把随机变量（或向量）X 的分布称为**总体分布**。**总体参数**，通常是指描述总体分布的一些特征，例如总体的均值（数学期望）、方差或标准差等等．一般，如果由总体抽得一个容量为的样本，则一个不包含任何未知参数的样本函数就称为一个**统计量**．一般地说，如果θ是一个未知的总体参数，当要对θ进行估计时，总是去寻找一个合适的统计量$\hat{\theta}=\hat{\theta}(X_1，X_2,\cdots，X_n)$，在取得样本观察值后，就计算出统计量$\hat{\theta}$的值作为对未知参数的一个估计值．这样对未知参数给出一个具体数值的估计方法称为参数的点估计．用于估计未知参数θ的统计量$\hat{\theta}$称为θ的一个**估计量**，或者简称为θ的一个**估计**．

常用统计量:

样本均值 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n} X_i$

样本方差 $S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$

样本标准差 $S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2}$

2.2 常见统计分布

(1)常用的统计分布

(a) χ^2分布

设$X_1,X_2,\cdots,X_n$是取自总体$N(0,1)$的样本，则称统计量$\chi^2=\sum_{i=1}^{n}X_i^2$服从自由度为$n$的$\chi^2$分布，记为$\chi^2\sim\chi^2(n)$。

(b) F分布

设$X\sim\chi^2(m)$, $Y\sim\chi^2(n)$，且X与Y相互独立，则称$Z=\frac{X/m}{Y/n}$服从自由度为m,n的F分布，记为$Z\sim F(m,n)$。

(c) t 分布

设 $X \sim N(0,1)$, $Y \sim \chi^2(n)$ ，且 X 与 Y 相互独立，则称 $T = \dfrac{X}{\sqrt{Y/n}}$ 服从自由度为 n 的 t 分布，记为 $T \sim t(n)$ 。

(2) 来自正态总体样本的抽样分布

设 $(X_1, X_2, \cdots, X_n)$ 为正态总体 $X \sim N(\mu, \sigma^2)$ 的样本， $\overline{X}$ 与 S^2 分别为该样本的样本均值与样本方差，则有

(a) $\overline{X} \sim N(\mu, \dfrac{\sigma^2}{n})$;

(b) $U = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$;

(c) $\dfrac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1)$;

(d) $T = \dfrac{\overline{X} - \mu}{S/\sqrt{n}} \sim t(n-1)$;

(e) $\overline{X}$ 与 S^2 相互独立。

(3) 各种常见分布之间的关系

(a) 正态分布与 χ^2 分布

若 $X_1, X_2, \cdots, X_n$ 独立同分布且 $X_i \sim N(0,1)$，则 $\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$;

(b) 正态分布, χ^2 分布与 t 分布

若 $X \sim N(0,1)$, $Y \sim \chi^2(n)$ ， X, Y 相互独立，则 $\dfrac{X}{\sqrt{Y/n}} \sim t(n)$;

(c) χ^2 分布与 F 分布

若 $X \sim \chi^2(n_1)$, $Y \sim \chi^2(n_2)$, X, Y 相互独立，则 $\dfrac{X/n_1}{Y/n_2} \sim F(n_1, n_2)$;

（d） χ^2 分布与 χ^2 分布

若 $X_1 \sim \chi^2(n_1), Y \sim \chi^2(n_2)$，$X, Y$ 相互独立，则 $X+Y \sim \chi^2(n_1+n_2)$；

（e） t 分布与 F 分布

若 $T \sim t(n)$，则 $T^2 \sim F(1,n)$；

（f） F 分布与 F 分布

若 $F \sim F(n_1,n_2)$，则 $\frac{1}{F} \sim F(n_2,n_1)$。

2.3 假设检验

(1). 假设检验的基本思想：对总体分布中的未知参数作出某种假设，根据样本在假设为真的前提下构造一个小概率事件，基于“小概率事件”在一次试验中几乎不可能发生而对假设作出拒绝或接受.

(2). 假设检验两类错误：**第一类错误**：原假设 H_0 为真，但拒绝了 H_0.
第二类错误；原假设 H_0 为假，但接受到了 H_0.