温兆和. 1020550143. 统计方法与机器学习作业 2.

1. 解：设转换后的系数为 $\tilde{\beta_0}, \tilde{\beta_1}$.

则 $Q(\tilde{\beta_0}\ \tilde{\beta_1}) = \sum_{i=1}^{n}\left(\frac{y_i - c_1}{d_1} - \tilde{\beta_0} - \tilde{\beta_1}\frac{x_i - c_2}{d_2}\right)^2$.

$$\begin{cases} \frac{\partial Q}{\partial \beta_0} = -2\sum_{i=1}^{n}\left(\frac{y_i - c_1}{d_1} - \tilde{\beta_0} - \tilde{\beta_1}\frac{x_i - c_2}{d_2}\right) \\ \frac{\partial Q}{\partial \beta_1} = -\frac{1}{2d_2}\sum_{i=1}^{n}\left(\frac{y_i - c_1}{d_1} - \tilde{\beta_0} - \tilde{\beta_1}\frac{x_i - c_2}{d_2}\right)x_i \end{cases}$$

令 $\begin{cases} \frac{\partial Q}{\partial \beta_0} = 0 \\ \frac{\partial Q}{\partial \beta_1} = 0 \end{cases}$ 解得: $\begin{cases} \tilde{\beta_0} = \frac{\bar{y} - c_1}{d_1} - \frac{\bar{x} - c_2}{d_1}\frac{l_{xy}}{l_{xx}} \\ \tilde{\beta_1} = \frac{d_2}{d_1}\frac{l_{xy}}{l_{xx}} \end{cases}$

$$\tilde{SS_T} = \sum_{i=1}^{n}\left(\frac{y_i - c_1}{d_1} - \frac{\bar{y} - c_1}{d_1}\right)^2$$

$$= \frac{1}{d^2}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{d^2}SS_T.$$

$$\tilde{SS_R} = \sum_{i=1}^{n}(\tilde{\beta_0} + \tilde{\beta_1}\tilde{x_i} - \tilde{\beta_0} - \tilde{\beta_1}\tilde{\bar{x}})^2$$

$$= \sum_{i=1}^{n}\left(\tilde{\beta_1}\left(\frac{x_i - c_2}{d_2} - \frac{\bar{x} - c_2}{d_2}\right)\right)^2$$

$$= \frac{1}{d_1^2}\beta_1^2 l_{xx} = \frac{1}{d^2}SS_R.$$

直接把转换后的值代入课件中的计算过程

又: $\sum_{i=1}^{n}(\tilde{y_i} - \tilde{\bar{y}})^2 = \sum_{i=1}^{n}(\tilde{y_i} - \tilde{\hat{y}_i} + \tilde{\hat{y}_i} - \tilde{\bar{y}})^2$

$$= \sum_{i=1}^{n}(\tilde{y_i} - \tilde{\hat{y}_i})^2 + \sum_{i=1}^{n}(\tilde{\hat{y}_i} - \tilde{\bar{y}})^2$$

故 $\tilde{SS_T} = \tilde{SS_R} + \tilde{SS_E}$

故 $\tilde{SS_E} = \frac{1}{d^2}SS_E$.

由于调整后的 SS_T, SS_R, SS_E 均为调整前差 $\frac{1}{d^2}$ 倍，

故调整后的检验统计量 $\tilde{F_0} = \tilde{SS_R}(n-2)/\tilde{SS_E}$

$= \frac{SS_R}{SS_E/(n-2)} = F_0$ 没有变化，

故转换前后F统计量的值保持不变。

2. 解：已知，在线性回归模型 $\hat{y} = a + bx$ 中，

$\hat{b} = \frac{l_{xy}}{l_{xx}}$ $\hat{a} = \bar{y} - \hat{b}\bar{x} = \bar{y} - \frac{l_{xy}}{l_{xx}}\bar{x}$

则在线性回归模型 $\hat{x} = c + dy$ 中

$\hat{d} = \frac{l_{xy}}{l_{yy}}$ $\hat{c} = \bar{x} - \frac{l_{xy}}{l_{yy}}\bar{y}$

故 $x = \bar{x} - \frac{l_{xy}}{l_{yy}}\bar{y} + \frac{l_{xy}}{l_{yy}}y$

将其与 $y = \bar{y} - \frac{l_{xy}}{l_{xx}}\bar{x} + \frac{l_{xy}}{l_{xx}}y$ 联立，得

$\begin{cases} x = \bar{x} \\ y = \bar{y} \end{cases}$ 故这两条直线有交点 $(\bar{x}, \bar{y})$，

直接代入最小二乘估计。

3. 证：$(I-H)^T = [I - X(X^TX)^{-1}X^T]^T$

$= I - X[(X^TX)^{-1}]^TX^T$

$= I - X[(X^TX)^T]^{-1}X^T$

$= I - X(X^TX)^{-1}X^T = I - H$

故对称性成立，

$(I-H)^2 = (I-H)(I-H)$

$= I - 2H + H^2$

$= I - 2X(X^TX)^{-1}X^T + X(X^TX)^{-1}X^TX(X^TX)^{-1}X^T$

$= I - 2X(X^TX)^{-1}X^T + X(X^TX)^{-1}X^T$

$= I - X(X^TX)^{-1}X^T = I - H$

故 ~~幂等性~~ $I-H$ 是对称幂等矩阵。

特征值 0，1

高级计量经济学作业2/Q2

由于 $tr(H) = p+1$, 故 $tr(I-H) = n-p-1$.

下证: 对称幂等矩阵特征值只可能是0或1.

考虑对称幂等矩阵 P. $P^2 = P$.

设 $P\alpha = \lambda\alpha$. 则 $P^2\alpha = \lambda P\alpha = \lambda^2\alpha = \lambda\alpha$

故 $\lambda^2 = \lambda$. $\lambda = 0$ 或 1.

故 $I-H$ 的特征值只能是0或1.

又: $tr(I-H) = n-p-1$.

故 $(I-H) \sim diag(\underbrace{1 \cdots 1}_{n-p-1\text{个}1} \underbrace{0 \cdots 0}_{p+1\text{个}0})$

由于相似矩阵有相同的秩. $r(I-H) = n-p-1$.

4. 证: $\sum_{i=1}^{n}(y_i - \hat{y}_i)$

$= \sum_{i=1}^{n}[x_i(\beta - \hat{\beta}) + \varepsilon_i]$

$= \sum_{i=1}^{n}[x_i(\beta - (X'X)^{-1}X'y) + \varepsilon_i]$

$\underline{X'X\beta = X'y}$ $\sum_{i=1}^{n}[x_i(\beta - \boxed{(X'X)^{-1}X'X}\beta) + \varepsilon_i)$

$= \sum_{i=1}^{n}\varepsilon_i = 0.$

问题: 未证对应系数之和为0。

5. 解: 令 $L = diag\left\{\frac{1}{\sqrt{L_{11}}}, \cdots, \frac{1}{\sqrt{L_{pp}}}\right\}$

则 X_c 为中心化后的 X.

设 $X_S = X_c L$,

$\beta_{S-c,\,slope} = (X_S'X_S)^{-1}X_S'y^*$

$= (LX_c'X_cL)^{-1}LX_c'y^*$

$= L^{-1}(X_c'X_c)^{-1}X_c'y^*$.

而 $\hat{\beta}_{c.slope} = \hat{\beta}_{slope} = (X_c'X_c)^{-1}X_c'y^*$

故 $\hat{\beta}_{s.c.slope} = L^{-1}\hat{\beta}_{slope}$,

$\tilde{\beta} = \begin{pmatrix} 0, \\ L^{-1}\hat{\beta}_{slope} \end{pmatrix}$

$E(\tilde{\beta}) = \begin{pmatrix} 0, \\ L^{-1}E(\hat{\beta}_{slope}) \end{pmatrix} = \begin{pmatrix} 0, \\ L^{-1}\beta_{slope} \end{pmatrix}$

$= (0, \sqrt{\lambda_1}\beta_1, \cdots, \sqrt{\lambda_p}\beta_p)'$.

6. 解: 在单因素方差分析模型中 $\sum_{i=1}^{a}\alpha_i = 0$.

$y_{ij} = \mu + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_{a-1}x_{a-1} + \varepsilon_{ij}$

当 y 在第 a 个水平下, 则 $x_1 = \cdots = x_{a-1} = -1$

其余情况下, 取到哪个水平, 哪个水平对应的 x 取 1.

其余 x 取 0.

$y = \begin{pmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{aj} \end{pmatrix}$ $\beta = \begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_{a-1} \end{pmatrix}$ $X = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 1 & -1 & -1 & \cdots & -1 \end{pmatrix}$

$\varepsilon = \begin{pmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{aj} \end{pmatrix}$

相当于用来自 a 个不同水平的响应变量值来其对应水平下的 x 的取值来估计 μ 和诸 α_i.

$\varepsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$.

ε_{ij} 下面用最小二乘法求 μ 和诸 α_i 的估计:

$$\text{令 } Q = \sum_{i=1}^{a}\sum_{j=1}^{m}(y_{ij}-\mu-\alpha_1 x_1-\cdots-\alpha_{a-1}x_{a-1})^2.$$

$$\text{求解 } \begin{cases}\dfrac{\partial L}{\partial \mu} = -2\sum_{i=1}^{a}\sum_{j=1}^{m}(y_{ij}-\mu-\alpha_1 x_1-\cdots-\alpha_{a-1}x_{a-1})=0\\ \dfrac{\partial L}{\partial \alpha_i} = -2\sum_{i=1}^{a}\sum_{j=1}^{m}(y_{ij}-\mu-\alpha_1 x_1-\cdots-\alpha_{a-1}x_{a-1})x_i=0.\end{cases}$$

$$\text{得: } \begin{cases}\hat{\mu} = \dfrac{\sum_{i=1}^{a}\bar{y}_{i\cdot}}{a}\\ \hat{\alpha}_i = \bar{y}_{i\cdot}-\dfrac{\sum_{i=1}^{a}\bar{y}_{i\cdot}}{a}\end{cases}$$

下面对该线性回归模型进行显著性检验.

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{m}(y_{ij}-\bar{y}_{\cdot\cdot})^2$$

$$\hat{y} = \mu+\alpha_i = \bar{y}_{i\cdot}$$

$$SS_R = \sum_{i=1}^{a}\sum_{j=1}^{m}(\bar{y}_{i\cdot}-\bar{y}_{\cdot\cdot})^2$$

$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{m}(y_{ij}-\bar{y}_{i\cdot})^2$$

样本量 $n = am$. 变量数 $p = a-1$

$$F_0 = \frac{SS_R/p}{SS_E/(n-p-1)} \sim F(p, n-p-1)$$

而 $F(p, n-p-1)$ 相当于 $F(a-1, n-a)$

与单因素方差分析中检验 $\alpha_1 = \cdots = \alpha_a = 0$ 检验统计量是完全相同

综上: 单因素方差分析模型可看作多元线性回归模型.

方差分析: $\alpha_i = 0$? $\Rightarrow$ 因子是否显著

相当于回归分析: α_i 的系数

看回归方程是否显著.

其实 y、β、X、ε 应该写成这样：

$$y=\begin{pmatrix} y_{11} \\ \vdots \\ y_{1m} \\ y_{21} \\ \vdots \\ y_{a1} \\ \vdots \\ y_{am} \end{pmatrix} \quad X=\begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ & & \vdots & & \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ & & \vdots & & \\ 1 & -1 & \cdots & & -1 \\ & & \vdots & & \\ 1 & -1 & & \cdots & -1 \end{pmatrix}$$

$$\beta=\begin{pmatrix} \mu \\ \alpha_1 \\ \vdots \\ \alpha_{a-1} \end{pmatrix} \quad \varepsilon=\begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1m} \\ \vdots \\ \varepsilon_{a1} \\ \vdots \\ \varepsilon_{am} \end{pmatrix}$$

这样可以让所有的 y_{ij} 都参与到 μ 和诸 α_i 的估计中。