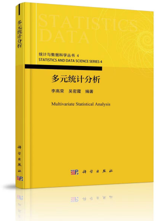
第8章
多元多重回归分析

李高荣

北京师范大学统计学院

E-mail: *ligaorong@bnu.edu.cn*



1. 多元方差分析模型
   - 单因子多元方差分析
   - 双因子多元方差分析

2. 多元多重回归
   - 多响应变量的多元多重回归模型
   - 模型参数的估计
   - 模型参数的检验
   - 多元多重线性回归模型的预测
   - 案例分析

3. 多元生长曲线模型

微信公众号：BNUlgr

- 扫二维码获取在线课件和相关教学电子资源
- 请遵守电子资源使用协议

学习目标与要求：

1. 掌握单因子多元方差分析和双因子多元方差分析的方法、理论和计算
2. 掌握多响应变量的多元多重回归模型的估计方法和理论，以及应用
3. 掌握多元生长曲线模型，以及应用

- 单因子多元方差分析模型的重要目的：检验r个相互独立的p元正态总体的均值向量是否有显著差异。
- 设$\boldsymbol{Y}_{k1}, \cdots, \boldsymbol{Y}_{kn_k}$ 是来自第k个p元正态总体$N_p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma})$ 的随机样本，其中$k = 1, \cdots, r$
- $\boldsymbol{Y}_{ki}$ 就可以表示为：

$$\boldsymbol{Y}_{ki} = \boldsymbol{\mu}_k + \varepsilon_{ki}, \quad k = 1, \cdots, r, \quad i = 1, \cdots, n_k,$$

  - 误差向量$\varepsilon_{ki} \sim N_p(\boldsymbol{0}, \boldsymbol{\Sigma})$，$k = 1, \cdots, r$，$i = 1, \cdots, n_k$，且相互独立

- 记

$$\boldsymbol{\mu} = \frac{1}{n}\sum_{k=1}^{r} n_k \boldsymbol{\mu}_k, \quad \boldsymbol{\tau}_k = \boldsymbol{\mu}_k - \boldsymbol{\mu}, \quad k = 1, \cdots, r,$$

  - ▷ 这里，$n = \sum_{k=1}^{r} n_k$

- $\boldsymbol{Y}_{ki}$ 可进一步用单向分类方差分析模型来表示，即

$$\boldsymbol{Y}_{ki} = \boldsymbol{\mu} + \boldsymbol{\tau}_k + \boldsymbol{\varepsilon}_{ki}.$$

- $\boldsymbol{\tau}_k$为第k个总体的效应，并满足条件：

$$\sum_{k=1}^{r} n_k \boldsymbol{\tau}_k = \boldsymbol{0}.$$

- 对r个多元正态总体均值向量的检验问题：

$$H_0: \boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_r, \qquad H_1: \boldsymbol{\mu}_1, \cdots, \boldsymbol{\mu}_r \text{不全相等}$$

- 等价于在单向分类方差分析模型下检验问题：

$$H_0: \boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_r = \mathbf{0}, \qquad H_1: \boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_r \text{不全为}\mathbf{0}.$$

- 为了完成假设检验，总离差阵$\mathbf{T}$有如下的分解：

$$\begin{aligned}\mathbf{T} &= \sum_{k=1}^{r}\sum_{i=1}^{n_k}\left(\boldsymbol{Y}_{ki} - \overline{\boldsymbol{Y}}\right)\left(\boldsymbol{Y}_{ki} - \overline{\boldsymbol{Y}}\right)' \\ &= \sum_{k=1}^{r}\sum_{i=1}^{n_k}\left(\boldsymbol{Y}_{ki} - \overline{\boldsymbol{Y}}_k\right)\left(\boldsymbol{Y}_{ki} - \overline{\boldsymbol{Y}}_k\right)' + \sum_{k=1}^{r} n_k\left(\overline{\boldsymbol{Y}}_k - \overline{\boldsymbol{Y}}\right)\left(\overline{\boldsymbol{Y}}_k - \overline{\boldsymbol{Y}}\right)' \\ &=: \mathbf{A} + \mathbf{B}.\end{aligned}$$

- $\mathbf{B}$称为组效应离差阵，刻画了效应大小的度量，并由效应所致
- $\mathbf{A}=\sum_{k=1}^{r}\mathbf{A}_k$ 称为误差离差阵，反映了随机误差ε_{ki}的影响
  - ▷ $\mathbf{A}_k=\sum_{i=1}^{n_k}\left(\boldsymbol{Y}_{ki}-\overline{\boldsymbol{Y}}_k\right)\left(\boldsymbol{Y}_{ki}-\overline{\boldsymbol{Y}}_k\right)'$
  - ▷ $\overline{\boldsymbol{Y}}=\frac{1}{n}\sum_{k=1}^{r}\sum_{i=1}^{n_k}\boldsymbol{Y}_{ki}$
  - ▷ $\overline{\boldsymbol{Y}}_k=\frac{1}{n_k}\sum_{i=1}^{n_k}\boldsymbol{Y}_{ki},\quad k=1,\cdots,r$

定理8.1.1

在单因子多元方差分析模型下，$\mathbf{A} \sim W_p(n-r, \Sigma)$。当原假设$H_0$成立时，则

1. $\mathbf{T} \sim W_p(n-1, \Sigma)$;
2. $\mathbf{B} \sim W_p(r-1, \Sigma)$;
3. $\mathbf{A}$和$\mathbf{B}$相互独立。

证明留作习题

- 当H_0成立时，$\mathbf{B}$ 相对于$\mathbf{A}$应该很小。
- 令$\lambda_1, \cdots, \lambda_g$为$\mathbf{BA}^{-1}$ 的所有非零特征值，其中$g = \min(p, r-1)$，考虑下面四个统计量：

Wilks-Λ 统计量：

$$\Lambda = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}|} = \frac{1}{|\mathbf{I}_p + \mathbf{BA}^{-1}|} = \prod_{k=1}^{r-1} \frac{1}{1+\lambda_k};$$

Hotelling-Lawley 迹统计量：

$$\mathrm{tr}(\mathbf{BA}^{-1}) = \sum_{k=1}^{g} \frac{1}{\lambda_k};$$

- **Pillai迹统计量：**

$$\mathrm{tr}(\mathbf{B}(\mathbf{A}+\mathbf{B})^{-1}) = \sum_{k=1}^{g} \frac{\lambda_k}{1+\lambda_k};$$

- **Roy最大特征值统计量：**

$$\lambda_1(\mathbf{B}\mathbf{A}^{-1}).$$

- 虽然以上四个统计量具有几乎相同的渐近分布，但在有限样本下的表现并不完全一致。

- 由第6章知：Wilks-Λ统计量的检验等价于似然比检验。当$p \leq 2$或$r \leq 3$时，Wilks-Λ统计量有精确的零分布，详细的结果见第4章和第6章。
- 当$n \to \infty$时，可用Bartlett (1938)提出的渐近χ^2分布近似：即在原假设H_0成立时，则

$$\chi^2 = -\left(n - 1 - \frac{p+r}{2}\right) \ln \Lambda \xrightarrow{d} \chi^2_{(r-1)p}.$$

- 因此，可得拒绝原假设H_0的拒绝域为：

$$-\left(n - 1 - \frac{p+r}{2}\right) \ln \Lambda > \chi^2_{(r-1)p}(\alpha).$$

- R语言中提供了多元方差分析的函数manova()，对输出结果可用函数summary()进行具体分析，调用格式为：

```
summary(object,
        test = c("Pillai", "Wilks",
        "Hotelling-Lawley", "Roy"),
        intercept = FALSE, tol = 1e-7, ...)
```

其中object为函数manova()生成的对象，test对应提到的四种统计量。

例：Wisconsin养老院数据

Wisconsin养老院数据，该数据选自Johnson 和Wichern (2008)。养老院分为私人养老院、非营利组织养老院和公立养老院。Wisconsin州卫生及公共服务部需要要根据护理水平、平均工资和州平均工资等因素来计算每个养老院的成本，从而提供合理的补助。这时需检验三种不同所有权下的养老院成本是否有显著差异，显著性水平取$\alpha = 0.01$。

解：

- 记Y_1 为雇佣护工的成本
- Y_2为饮食的成本
- Y_3为机构的植被种植和维护成本
- Y_4为房间保洁和清洗衣物的成本
- 共收集了$n=516$家机构的数据，其中
  1. 私人养老院$n_1=271$家
  2. 非营利组织养老院$n_2=138$家
  3. 公立养老院$n_3=107$家

- 三组数据的样本均值向量和样本协方差矩阵分别为：

$$\bar{y}_1 = \begin{pmatrix} 2.066 \\ 0.480 \\ 0.082 \\ 0.360 \end{pmatrix}, \quad \widehat{\Sigma}_1 = \begin{pmatrix} 0.291 & & & \\ -0.001 & 0.011 & & \\ 0.002 & 0.000 & 0.001 & \\ 0.010 & 0.003 & 0.000 & 0.010 \end{pmatrix};$$

$$\bar{y}_2 = \begin{pmatrix} 2.167 \\ 0.596 \\ 0.124 \\ 0.418 \end{pmatrix}, \quad \widehat{\Sigma}_2 = \begin{pmatrix} 0.561 & & & \\ 0.011 & 0.025 & & \\ 0.001 & 0.004 & 0.005 & \\ 0.037 & 0.007 & 0.002 & 0.019 \end{pmatrix};$$

$$\bar{y}_3 = \begin{pmatrix} 2.273 \\ 0.521 \\ 0.125 \\ 0.383 \end{pmatrix}, \quad \widehat{\Sigma}_3 = \begin{pmatrix} 0.261 & & & \\ 0.030 & 0.017 & & \\ 0.003 & -0.00 & 0.004 & \\ 0.018 & 0.006 & 0.001 & 0.013 \end{pmatrix}.$$

- 误差离差阵为：

$$\mathbf{A}=\mathbf{A}_1+\mathbf{A}_2+\mathbf{A}_3=(n_1-1)\widehat{\mathbf{\Sigma}}_1+(n_2-1)\widehat{\mathbf{\Sigma}}_2+(n_3-1)\widehat{\mathbf{\Sigma}}_3$$

$$=\begin{pmatrix} 182.962 & & & \\ 4.408 & 8.200 & & \\ 1.695 & 0.633 & 1.484 & \\ 9.581 & 2.428 & 0.394 & 6.538 \end{pmatrix},$$

- 总样本均值向量为

$$\bar{\mathbf{y}}=\frac{1}{n}\sum_{k=1}^{3}n_k\bar{\mathbf{y}}_k=(2.136,0.519,0.102,0.380)'.$$

- 进一步，计算可得组效应离差阵为：

$$\mathbf{B} = \sum_{k=1}^{3} n_k(\bar{\mathbf{y}}_k - \bar{\mathbf{y}})(\bar{\mathbf{y}}_k - \bar{\mathbf{y}})' = \begin{pmatrix} 3.475 & & & \\ 1.111 & 1.225 & & \\ 0.821 & 0.453 & 0.235 & \\ 0.584 & 0.610 & 0.230 & 0.304 \end{pmatrix}.$$

- 为了检验三种性质的养老院的成本是否相等，计算Wilks-Λ统计量，可得

$$\Lambda = \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}|} = 0.7714.$$

- 注意到$r=3$，可采用精确的F检验统计量，计算

$$
\begin{aligned}
F &= \left(\frac{n-p-2}{p}\right)\left(\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\right) \\
&= \left(\frac{516-4-2}{4}\right)\left(\frac{1-\sqrt{0.7714}}{\sqrt{0.7714}}\right) = 17.67.
\end{aligned}
$$

- 取显著性水平为$\alpha=0.01$，由于

$$
\begin{aligned}
F =& 17.67 > F_{2p,2(n-p-2)}(0.01) \\
=& F_{8,1020}(0.01) \approx \chi^2(0.01)/8 = 2.51.
\end{aligned}
$$

- 拒绝原假设H_0，认为三种性质的养老院的成本不相等，即养老院的所有权性质对养老院的成本有显著影响。

- 由于样本量$n = 516$很大，也可用Bartlett (1938)提出的χ^2近似检验，计算

$$\chi^2 = -\left(n - 1 - \frac{p+r}{2}\right)\ln\Lambda$$
$$= -511.5\ln(0.7714) = 132.76.$$

- 查表得$\chi^2_{(r-1)p}(0.01) = \chi^2_8(0.01) = 20.09 < 132.76$。
- 因此，拒绝原假设H_0，得到与精确检验下相同的结论。

- 目的：当拒绝原假设H_0时，有必要构造效应差$\boldsymbol{\tau}_k - \boldsymbol{\tau}_l$(或均值向量差$\boldsymbol{\mu}_k - \boldsymbol{\mu}_l$)的Bonferroni同时置信区间
- 令τ_{kj}表示$\boldsymbol{\tau}_k$的第j个分量，并得到$\boldsymbol{\tau}_k$的估计量为：

$$\hat{\boldsymbol{\tau}}_k = \overline{\boldsymbol{Y}}_k - \overline{\boldsymbol{Y}}$$

- 则τ_{kj}和$\tau_{kj} - \tau_{lj}$的估计量分别为

$$\hat{\tau}_{kj} = \overline{Y}_{kj} - \overline{Y}_j, \quad \hat{\tau}_{kj} - \hat{\tau}_{lj} = \overline{Y}_{kj} - \overline{Y}_{lj},$$

其中$l < k = 1, \cdots, r$和$j = 1, \cdots, p$。

- 注意到：

$$\mathrm{Var}(\widehat{\tau}_{kj} - \widehat{\tau}_{lj}) = \mathrm{Var}(\overline{Y}_{kj} - \overline{Y}_{lj}) = \left(\frac{1}{n_k} + \frac{1}{n_l}\right)\sigma_{jj},$$

其中σ_{jj}是Σ的第j个对角线元素

- 进一步，可得$\mathrm{Var}(\overline{Y}_{kj} - \overline{Y}_{lj})$的估计为

$$\widehat{\mathrm{Var}}(\overline{Y}_{kj} - \overline{Y}_{lj}) = \left(\frac{1}{n_k} + \frac{1}{n_l}\right)\frac{a_{jj}}{n-r},$$

其中a_{jj}为离差阵$\mathbf{A}$的第j个对角线元素

- 给定观测样本$\mathbf{y}_{ki}$，其中$k = 1, \cdots, r$; $i = 1, \cdots, n_k$
- $\tau_{kj} - \tau_{lj}$ 置信度至少为$1 - \alpha$ 的Bonferroni 同时置信区间为：

$$(\bar{y}_{kj} - \bar{y}_{lj}) \pm t_{n-r}\left(\frac{\alpha}{pr(r-1)}\right)\sqrt{\frac{a_{jj}}{n-r}\left(\frac{1}{n_k} + \frac{1}{n_l}\right)},$$

其中$j = 1, \cdots, p$和$l < k = 1, \cdots, r$。

Bonferroni同时置信区间

- 对Wisconsin养老院数据构造效应差的Bonferroni 同时置信区间
- 计算，可得

$$\widehat{\tau}_1 = (\bar{\boldsymbol{y}}_1 - \bar{\boldsymbol{y}}) = (-0.070, -0.039, -0.020, -0.020)',$$

$$\widehat{\tau}_3 = (\bar{\boldsymbol{y}}_3 - \bar{\boldsymbol{y}}) = (-0.137, 0.002, 0.023, 0.003)',$$

$$\mathbf{A} = \begin{pmatrix} 182.962 & & & \\ 4.408 & 8.200 & & \\ 1.695 & 0.633 & 1.484 & \\ 9.581 & 2.428 & 0.394 & 6.538 \end{pmatrix}.$$

- 可得：$\widehat{\tau}_{13}-\widehat{\tau}_{33}=-0.020-0.023=-0.043$，且$n=271+138+107=516$
- 这时，有

$$\sqrt{\frac{a_{33}}{n-r}\left(\frac{1}{n_1}+\frac{1}{n_3}\right)}=\sqrt{\frac{1.484}{516-3}\left(\frac{1}{271}+\frac{1}{107}\right)}=0.00614.$$

- 可知：$p=4$，$r=3$和$\alpha=0.05$，及

$$t_{n-r}(\alpha/pr(r-1))=t_{513}(0.05/24)\approx 2.87$$

- 则$\tau_{13}-\tau_{33}$的95%的Bonferroni同时置信区间为：

$$(\bar{y}_{13}-\bar{y}_{33})\pm 2.87\times 0.00614=-0.043\pm 0.018.$$

- 考虑如下双因子多元方差分析模型：

$$\boldsymbol{Y}_{kli} = \boldsymbol{\mu} + \boldsymbol{\tau}_k + \boldsymbol{\beta}_l + \boldsymbol{\gamma}_{kl} + \boldsymbol{\varepsilon}_{kli},$$

  - 误差向量$\boldsymbol{\varepsilon}_{kli} \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$，且$\boldsymbol{\varepsilon}_{kli}$ 相互独立
  - $k = 1, \cdots, r,\ \ l = 1, \cdots, m,\ \ i = 1, \cdots, n$
- $\boldsymbol{\tau}_k$和$\boldsymbol{\beta}_l$为固定效应向量，$\boldsymbol{\gamma}_{kl}$为交互固定效应
- 满足条件：

$$\sum_{k=1}^{r} \boldsymbol{\tau}_k = \sum_{l=1}^{m} \boldsymbol{\beta}_l = \sum_{k=1}^{r} \boldsymbol{\gamma}_{kl} = \sum_{l=1}^{m} \boldsymbol{\gamma}_{kl} = \mathbf{0}.$$

- 在双因子多元方差分析模型下，考虑如下的三个假设检验问题：

$$H_{10}: \boldsymbol{\tau}_1 = \cdots = \boldsymbol{\tau}_r = \mathbf{0}, \qquad H_{11}: \boldsymbol{\tau}_1, \cdots, \boldsymbol{\tau}_r \text{不全为} \mathbf{0};$$

$$H_{20}: \boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_m = \mathbf{0}, \qquad H_{21}: \boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_m \text{不全为} \mathbf{0};$$

$$H_{30}: \boldsymbol{\gamma}_{11} = \boldsymbol{\gamma}_{12} = \cdots = \boldsymbol{\gamma}_{rm} = \mathbf{0}, \quad H_{31}: \boldsymbol{\gamma}_{11}, \boldsymbol{\gamma}_{12}, \cdots, \boldsymbol{\gamma}_{rm} \text{不全为} \mathbf{0}.$$

- 给定观测样本$\{y_{kli},\ k=1,\cdots,r;\ l=1,\cdots,m;\ i=1,\cdots,n\}$，由双因子多元方差分析模型，则有

$$y_{kli}=\bar{y}+(\bar{y}_{k.}-\bar{y})+(\bar{y}_{.l}-\bar{y})+(\bar{y}_{kl}-\bar{y}_{k.}-\bar{y}_{.l}+\bar{y})+(y_{kli}-\bar{y}_{kl}),$$

其中

$$\bar{y}=\frac{1}{rmn}\sum_{k=1}^{r}\sum_{l=1}^{m}\sum_{i=1}^{n}y_{kli},\qquad \bar{y}_{k.}=\frac{1}{mn}\sum_{l=1}^{m}\sum_{i=1}^{n}y_{kli},$$

$$\bar{y}_{.l}=\frac{1}{rn}\sum_{k=1}^{r}\sum_{i=1}^{n}y_{kli},\qquad \bar{y}_{kl}=\frac{1}{n}\sum_{i=1}^{n}y_{kli}.$$

- 总的离差阵可以分解为：

$$\mathbf{T}=\sum_{k=1}^{r}\sum_{l=1}^{m}\sum_{i=1}^{n}(\boldsymbol{y}_{kli}-\bar{\boldsymbol{y}})(\boldsymbol{y}_{kli}-\bar{\boldsymbol{y}})'=\mathbf{A}+\mathbf{B}_1+\mathbf{B}_2+\mathbf{B}_{1\times 2},$$

其中

$$\mathbf{A}=\sum_{k=1}^{r}\sum_{l=1}^{m}\mathbf{A}_{kl}=\sum_{k=1}^{r}\sum_{l=1}^{m}\sum_{i=1}^{n}(\boldsymbol{y}_{kli}-\bar{\boldsymbol{y}}_{kl})(\boldsymbol{y}_{kli}-\bar{\boldsymbol{y}}_{kl})',$$

$$\mathbf{B}_1=\sum_{k=1}^{r}mn(\boldsymbol{y}_{k.}-\bar{\boldsymbol{y}})(\boldsymbol{y}_{k.}-\bar{\boldsymbol{y}})',\qquad \mathbf{B}_2=\sum_{l=1}^{m}rn(\boldsymbol{y}_{.l}-\bar{\boldsymbol{y}})(\boldsymbol{y}_{.l}-\bar{\boldsymbol{y}})',$$

$$\mathbf{B}_{1\times 2}=\sum_{k=1}^{r}\sum_{l=1}^{m}n(\bar{\boldsymbol{y}}_{kl}-\bar{\boldsymbol{y}}_{k.}-\bar{\boldsymbol{y}}_{.l}+\bar{\boldsymbol{y}})(\bar{\boldsymbol{y}}_{kl}-\bar{\boldsymbol{y}}_{k.}-\bar{\boldsymbol{y}}_{.l}+\bar{\boldsymbol{y}})'.$$

Table 1: 双因子多元方差分析模型的多元方差分析表

变差来源	离差阵	自由度
因子1	$\mathbf{B}_1$	$r-1$
因子2	$\mathbf{B}_2$	$m-1$
交叉作用	$\mathbf{B}_{1\times 2}$	$(r-1)(m-1)$
误差	$\mathbf{A}$	$rm(n-1)$
总离差阵	$\mathbf{T}$	$rmn-1$

- $\mathbf{A}$称为误差离差阵
- $\mathbf{B}_1$称为因子1的效应离差阵
- $\mathbf{B}_2$ 称为因子2 的效应离差阵
- $\mathbf{B}_{1\times 2}$称为因子1和因子2的交叉效应离差阵

可以证明：

1. $\mathbf{A} \sim W_p(rm(n-1), \Sigma)$
2. 当原假设H_{10}成立时，$\mathbf{B}_1 \sim W_p(r-1, \Sigma)$，且$\mathbf{A}$和$\mathbf{B}_1$ 相互独立
3. 当原假设H_{20} 成立时，$\mathbf{B}_2 \sim W_p(m-1, \Sigma)$，且$\mathbf{A}$和$\mathbf{B}_2$ 相互独立
4. 当原假设H_{30} 成立时，$\mathbf{B}_{1\times 2} \sim W_p((r-1)(m-1), \Sigma)$，且$\mathbf{A}$ 和$\mathbf{B}_{1\times 2}$ 相互独立

- 对检验因子1是否存在效应差异，定义Wilks-Λ 统计量如下：

$$\Lambda_1 = \frac{|\mathbf{A}|}{|\mathbf{A}+\mathbf{B}_1|}.$$

- 可知，当原假设H_{01}成立时，

$$\Lambda_1 \sim \Lambda(p, rm(n-1), r-1).$$

- 当$p > 2$ 和$r > 3$时，且$n \to \infty$时，可得拒绝原假设H_{01}的拒绝域为：

$$-\left[rm(n-1) - \frac{p+1-(r-1)}{2}\right] \ln \Lambda_1 > \chi^2_{(r-1)p}(\alpha).$$

- 对检验因子2是否存在效应差异，定义Wilks-Λ 统计量如下：

$$\Lambda_2 = \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}_2|}.$$

- 当$p > 2$ 和$m > 3$时，且当样本量$n \to \infty$时，类似地，可得拒绝原假设H_{02}的拒绝域为：

$$-\left[rm(n-1) - \frac{p+1-(m-1)}{2}\right] \ln \Lambda_2 > \chi^2_{(m-1)p}(\alpha).$$

- 对检验因子1和因子2是否存在交互效应作用，定义Wilks-Λ 统计量如下：

$$\Lambda_{1\times 2} = \frac{|\mathbf{A}|}{|\mathbf{A} + \mathbf{B}_{1\times 2}|}.$$

- 当$n \to \infty$时，可得拒绝原假设H_{03}的拒绝域为：

$$-\left[rm(n-1) - \frac{p+1-(r-1)(m-1)}{2}\right] \ln \Lambda_{1\times 2}$$
$$>\chi^2_{(r-1)(m-1)p}(\alpha).$$

- 在忽略交互效应作用的情况下，分别构造效应差$\boldsymbol{\tau}_k-\boldsymbol{\tau}_s$和$\boldsymbol{\beta}_l-\boldsymbol{\beta}_q$的Bonferroni同时置信区间
- $\tau_{kj}-\tau_{sj}$ 置信水平为$1-\alpha$的Bonferroni同时置信区间：

$$(\bar{y}_{k.j}-\bar{y}_{s.j}) \pm t_\nu\left(\frac{\alpha}{pr(r-1)}\right)\sqrt{\frac{a_{jj}}{\nu}\frac{2}{mn}}, \qquad s<k=1,\cdots,r,$$

其中$\nu=rm(n-1)$，a_{jj}是离差阵$\mathbf{A}$的第j个对角线元素。
- $\beta_{lj}-\beta_{qj}$置信水平为$1-\alpha$的Bonferroni同时置信区间：

$$(\bar{y}_{.lj}-\bar{y}_{.qj}) \pm t_\nu\left(\frac{\alpha}{pm(m-1)}\right)\sqrt{\frac{a_{jj}}{\nu}\frac{2}{rn}}, \qquad q<l=1,\cdots,m,$$

其中$\nu=rm(n-1)$，a_{jj}是离差阵$\mathbf{A}$的第j个对角线元素。

例

考虑Johnson和Wichern (2008)中例6.13中塑料薄膜数据的双因子多元方差分析。该数据考虑两个因子：挤出率和纤维含量；每个因子有两个水平：低(Low)和高(High)；每个水平包含三个变量：耐撕裂性(tear resistance, X_1)、光泽度(gloss, X_2)和不透明性(opacity, X_3)。对两个因子，在每种水平组合下重复测量$n = 5$次。在显著性水平$\alpha = 0.05$下，试研究该数据中是否存在因子的效应差异和交互效应作用？

Table 2: 塑料薄膜数据

		因子2：纤维含量					
		Low(1.0%)			High(1.5%)		
		X_1	X_2	X_3	X_1	X_2	X_3
因子1：挤出率	Low(-10%)	6.5	9.5	4.4	6.9	9.1	5.7
		6.2	9.9	6.4	7.2	10.0	2.0
		5.8	9.6	3.0	6.9	9.9	3.9
		6.5	9.6	4.1	6.1	9.5	1.9
		6.5	9.2	0.8	6.3	9.4	5.7
	High(10%)	6.7	9.1	2.8	7.1	9.2	8.4
		6.6	9.3	4.1	7.0	8.8	5.2
		7.2	8.3	3.8	7.2	9.7	6.9
		7.1	8.4	1.6	7.5	10.1	2.7
		6.8	8.5	3.4	7.6	9.2	1.9

- **解：** 可知：$r = 2$，$m = 2$和$n = 5$
- 用函数manova()进行双因子多元方差分析，并用函数summary() 输出结果

```
Plastic = read.table("Plastic.DAT"); Y = Plastic[, 3:5];
colnames(Y) = c("tear", "gloss", "opacity"); Y = as.matrix(Y)
rate = factor(Plastic[,1], labels = c("Low", "High"))
additive = factor(Plastic[,2], labels = c("Low", "High"))
fit = manova(Y ~ rate * additive)
summary(fit, test = "Wilks")  ## ANOVA table of Wilks' Lambda
#### 输出结果:
              Df   Wilks approx F num Df den Df   Pr(>F)
rate           1 0.38186   7.5543      3     14 0.003034 **
additive       1 0.52303   4.2556      3     14 0.024745 *
rate:additive  1 0.77711   1.3385      3     14 0.301782
Residuals     16
---
Signif.codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- 三个假设检验的Wilks-Λ 统计量分别为：

$$\Lambda_1 = 0.38186, \quad \Lambda_2 = 0.52303, \quad \Lambda_{1\times 2} = 0.77711.$$

- p值分别为：

$$p_1 = 0.003034, \quad p_2 = 0.024745, \quad p_{1\times 2} = 0.301782.$$

- 从p值可以判断得到：
  1. 因为$p_1 = 0.003034 < \alpha = 0.05$，则拒绝$H_{01} : \boldsymbol{\tau}_1 = \boldsymbol{\tau}_2 = \mathbf{0}$，即存在因子1 的效应差异
  2. 因为$p_2 = 0.024745 < \alpha = 0.05$，则拒绝$H_{02} : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \mathbf{0}$，即存在因子2的效应差异
  3. 因为$p_{1\times 2} = 0.301782 > \alpha = 0.05$，则接受$H_{03}$，即因子1 和因子2的交互效应作用不显著

- 当响应变量向量$\boldsymbol{Y}=(Y_1,\cdots,Y_q)'$与协变量向量$\boldsymbol{X}=(X_1,\cdots,X_{p-1})'$的联合分布为多元正态分布时，即

$$\begin{pmatrix} \boldsymbol{Y} \\ \boldsymbol{X} \end{pmatrix} \sim N_{q+p-1}\left(\begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{D}_{11} & \mathbf{D}_{12} \\ \mathbf{D}_{21} & \mathbf{D}_{22} \end{pmatrix}\right),$$

这里，

- ▷ $\boldsymbol{\mu}_1=E(\boldsymbol{Y})$
- ▷ $\boldsymbol{\mu}_2=E(\boldsymbol{X})$
- ▷ $\mathbf{D}_{11}=\mathrm{Cov}(\boldsymbol{Y})$
- ▷ $\mathbf{D}_{12}=\mathbf{D}_{21}'=\mathrm{Cov}(\boldsymbol{Y},\boldsymbol{X})$
- ▷ $\mathbf{D}_{22}=\mathrm{Cov}(\boldsymbol{X})$

- 由多元正态分布的条件分布相关理论知：

$$E(\boldsymbol{Y}|\boldsymbol{X}) = (\boldsymbol{\mu}_1 - \mathbf{D}_{12}\mathbf{D}_{11}^{-1}\boldsymbol{\mu}_2) + \mathbf{D}_{12}\mathbf{D}_{11}^{-1}\boldsymbol{X},$$

$$\mathrm{Cov}(\boldsymbol{Y}|\boldsymbol{X}) = \mathbf{D}_{11.2} = \mathbf{D}_{11} - \mathbf{D}_{12}\mathbf{D}_{22}^{-1}\mathbf{D}_{21}.$$

- 记

$$\boldsymbol{\Theta} = \begin{pmatrix} (\boldsymbol{\mu}_1 - \mathbf{D}_{12}\mathbf{D}_{11}^{-1}\boldsymbol{\mu}_2)' \\ (\mathbf{D}_{12}\mathbf{D}_{11}^{-1})' \end{pmatrix}, \quad \boldsymbol{\Sigma} = (\sigma_{ij}) = \mathbf{D}_{11.2}.$$

- 于是，$\boldsymbol{Y}=(Y_1,\cdots,Y_q)'$对于$\boldsymbol{X}=(X_1,\cdots,X_{p-1})'$有如下线性表达：

$$\boldsymbol{Y}=\boldsymbol{\Theta}'\begin{pmatrix}1\\ \boldsymbol{X}\end{pmatrix}+\varepsilon,$$

其中$E(\varepsilon)=\mathbf{0}$和$\mathrm{Cov}(\varepsilon)=\boldsymbol{\Sigma}$。

- 为了估计未知的$p\times q$回归参数矩阵$\boldsymbol{\Theta}$和$q\times q$的协方差矩阵$\boldsymbol{\Sigma}$，对$\boldsymbol{Y}=(Y_1,\cdots,Y_q)'$和$\boldsymbol{X}=(X_1,\cdots,X_{p-1})'$作$n$次独立观测，得到观测数据：

$$(y_{i1},\cdots,y_{iq}),\quad (x_{i1},\cdots,x_{i,p-1}),\quad i=1,\cdots,n.$$

- 记

$$\mathbf{Y}=\begin{pmatrix} y_{11} & \cdots & y_{1q} \\ y_{21} & \cdots & y_{2q} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{nq} \end{pmatrix}=(\boldsymbol{Y}_1,\cdots,\boldsymbol{Y}_q),\quad \mathbf{X}=\begin{pmatrix} 1 & x_{11} & \cdots & x_{1(p-1)} \\ 1 & x_{21} & \cdots & x_{2(p-1)} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{n(p-1)} \end{pmatrix},$$

$$\boldsymbol{\Theta}=\begin{pmatrix} \beta_{01} & \beta_{02} & \cdots & \beta_{0q} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1q} \\ \vdots & \vdots & & \vdots \\ \beta_{(p-1)1} & \beta_{(p-1)2} & \cdots & \beta_{(p-1)q} \end{pmatrix}=(\boldsymbol{\beta}_0,\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_q),$$

$$\mathbf{E}=\begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \cdots & \varepsilon_{1q} \\ \varepsilon_{21} & \varepsilon_{22} & \cdots & \varepsilon_{2q} \\ \vdots & \vdots & & \vdots \\ \varepsilon_{n1} & \varepsilon_{n2} & \cdots & \varepsilon_{nq} \end{pmatrix}=(\boldsymbol{\varepsilon}_1,\boldsymbol{\varepsilon}_2,\cdots,\boldsymbol{\varepsilon}_q),$$

- 注意到：$\mathbf{E}$的不同行对应于不同次观测，它们相互独立，均值均为零，协方差矩阵均为$\boldsymbol{\Sigma}$
- 可得Gauss-Markov多元多重线性回归模型：

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\Theta} + \mathbf{E},$$

其中$E(\mathbf{E}) = \mathbf{0}$和$\mathrm{Cov}(\mathrm{Vec}(\mathbf{E})) = \boldsymbol{\Sigma} \otimes \mathbf{I}_n$。

- 模型的第j列满足如下的多元线性回归模型：

$$Y_j = \mathbf{X}\beta_j + \varepsilon_j, \qquad j = 1, \cdots, q,$$

其中$E(\varepsilon_j) = \mathbf{0}$, $\mathrm{Cov}(\varepsilon_j) = \sigma_{jj}\mathbf{I}_n$, $\mathrm{Cov}(\varepsilon_k, \varepsilon_j) = \sigma_{kj}\mathbf{I}_n$。

- 如何估计未知参数$\mathbf{\Theta}$和Σ?
- 若$\mathbf{X}$为列满秩矩阵，则$\mathbf{\Theta}$ 可估。$\mathbf{\Theta}$的最小二乘估计$\widehat{\mathbf{\Theta}}$ 可定义为：

$$\mathrm{tr}\left[(\mathbf{Y}-\mathbf{X}\widehat{\mathbf{\Theta}})'(\mathbf{Y}-\mathbf{X}\widehat{\mathbf{\Theta}})\right]=\min_{\mathbf{\Theta}}\left\{\mathrm{tr}\left[(\mathbf{Y}-\mathbf{X}\mathbf{\Theta})'(\mathbf{Y}-\mathbf{X}\mathbf{\Theta})\right]\right\}.$$

- 注意到：

$$\mathrm{tr}\left[(\mathbf{Y}-\mathbf{X}\mathbf{\Theta})'(\mathbf{Y}-\mathbf{X}\mathbf{\Theta})\right]=\sum_{j=1}^{q}(\boldsymbol{Y}_j-\mathbf{X}\boldsymbol{\beta}_j)'(\boldsymbol{Y}_j-\mathbf{X}\boldsymbol{\beta}_j).$$

- 对每个最小二乘目标函数$(Y_j - \mathbf{X}\beta_j)'(Y_j - \mathbf{X}\beta_j)$ 进行极小化
- 可得

$$\widehat{\beta}_j = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y_j, \quad j = 1, \cdots, q.$$

- $\mathbf{\Theta}$的最小二乘估计定义为：

$$\widehat{\mathbf{\Theta}} = (\widehat{\beta}_1, \cdots, \widehat{\beta}_q) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

- 将$\widehat{\mathbf{\Theta}}$代入$\mathbf{Q}(\mathbf{\Theta}) = (\mathbf{Y} - \mathbf{X\Theta})'(\mathbf{Y} - \mathbf{X\Theta})$，可得

$$\mathbf{Q}(\widehat{\mathbf{\Theta}}) = \mathbf{Y}' \left[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] \mathbf{Y}.$$

- 注意到：$\mathrm{Cov}(\boldsymbol{Y}_j, \boldsymbol{Y}_k) = \sigma_{jk}\mathbf{I}_n$，则有

$$
\begin{aligned}
E[\boldsymbol{Y}_k'(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\boldsymbol{Y}_j] =& \mathrm{tr}\left[(\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathrm{Cov}(\boldsymbol{Y}_j, \boldsymbol{Y}_k)\right] \\
=&(n-q)\sigma_{kj}
\end{aligned}
$$

- 即 $E[\mathbf{Q}(\widehat{\boldsymbol{\Theta}})] = (n-q)\boldsymbol{\Sigma}$。
- 故证得

$$
\widehat{\boldsymbol{\Sigma}} = \frac{1}{n-q}\mathbf{Y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}
$$

为$\boldsymbol{\Sigma}$的一个无偏估计，也把$\widehat{\boldsymbol{\Sigma}}$称为$\boldsymbol{\Sigma}$的最小二乘估计。

定理8.2.1

最小二乘估计$\widehat{\boldsymbol{\Theta}}$为Gauss-Markov多元多重线性回归模型中参数$\boldsymbol{\Theta}$的最佳线性无偏估计，即对于$\boldsymbol{\Theta}$的任意线性无偏估计$\mathbf{TY}$，在非负定意义下，总有

$$\mathrm{Cov}(\mathrm{Vec}(\widehat{\boldsymbol{\Theta}})) \leq \mathrm{Cov}(\mathrm{Vec}(\mathbf{TY})), \quad \mathrm{Vec}(\boldsymbol{\Theta}) \in \mathbb{R}^{pq}.$$

证明：由于$\mathbf{TY}$为$\boldsymbol{\Theta}$的无偏估计，即对任意的$\boldsymbol{\Theta}$ 都有

$$E(\mathbf{TY}) = \mathbf{TX}\boldsymbol{\Theta} = \boldsymbol{\Theta}.$$

因此，矩阵$\mathbf{T}$要满足条件：$\mathbf{TX} = \mathbf{I}_p$。

结合事实：$\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \geq \mathbf{0}$ (即正交投影阵是半正定的)，于是

$$\begin{aligned}
\mathrm{Cov}(\mathrm{Vec}(\mathbf{TY})) &= \mathrm{Cov}((\mathbf{I}_q \otimes \mathbf{T})\mathrm{Vec}(\mathbf{Y})) \\
&= (\mathbf{I}_q \otimes \mathbf{T})(\boldsymbol{\Sigma} \otimes \mathbf{I}_n)(\mathbf{I}_q \otimes \mathbf{T}') \\
&= \boldsymbol{\Sigma} \otimes (\mathbf{TT}') \geq \boldsymbol{\Sigma} \otimes (\mathbf{TX}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{T}') \\
&= \boldsymbol{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1} = \mathrm{Cov}(\mathrm{Vec}(\widehat{\boldsymbol{\Theta}})).
\end{aligned}$$

- 应用矩阵拉直运算Vec(·)，模型转化为如下单响应变量的线性回归模型：

$$\text{Vec}(\mathbf{Y}) = (\mathbf{I}_q \otimes \mathbf{X})\text{Vec}(\boldsymbol{\Theta}) + \text{Vec}(\mathbf{E})$$

- 记$\boldsymbol{\beta} = \text{Vec}(\boldsymbol{\Theta})$，$\boldsymbol{\beta}$最佳线性无偏估计为

$$\begin{aligned}
\boldsymbol{\beta}^* &= \text{Vec}(\boldsymbol{\Theta}^*) = \left[(\mathbf{I}_q \otimes \mathbf{X})'(\boldsymbol{\Sigma} \otimes \mathbf{I}_n)^{-1}(\mathbf{I}_q \otimes \mathbf{X})\right]^{-1}(\mathbf{I}_q \otimes \mathbf{X})(\boldsymbol{\Sigma} \otimes \mathbf{I}_n)^{-1}\text{Vec}(\mathbf{Y}) \\
&= (\boldsymbol{\Sigma} \otimes \mathbf{X}'\mathbf{X})^{-1}(\boldsymbol{\Sigma}^{-1} \otimes \mathbf{X}')\text{Vec}(\mathbf{Y}) = [\mathbf{I}_q \otimes (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\text{Vec}(\mathbf{Y}) \\
&= \text{Vec}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}) = \text{Vec}(\widehat{\boldsymbol{\Theta}}).
\end{aligned}$$

定理8.2.2

若Gauss-Markov多元多重线性回归模型的随机误差$\mathbf{E}$服从多元正态分布，即$\mathrm{Vec}(\mathbf{E}) \sim N_{nq}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I}_n)$，则$\mathbf{\Theta}$和$\mathbf{\Sigma}$的极大似然估计分别为$\widehat{\mathbf{\Theta}}$和$(n-p)\widehat{\mathbf{\Sigma}}/n$。

- 详细证明见教材

- 当$\mathbf{X}$列不满秩，即$\text{rank}(\mathbf{X}) = r < p$时，则模型变为降秩模型，这时$\boldsymbol{\Theta}$是不可估的
- 考虑$\boldsymbol{\Theta}$的线性函数$\mathbf{C}\boldsymbol{\Theta}$的可估性问题

定义8.1：可估函数

如果存在线性估计$\mathbf{TY}$，使得$E(\mathbf{TY}) = \mathbf{C}\boldsymbol{\Theta}$，则称$\mathbf{C}\boldsymbol{\Theta}$为可估函数。

定理8.2.3

设$\mathbf{C}$为$m \times p$常数矩阵，则在多元多重线性回归模型下，$\mathbf{C}\boldsymbol{\Theta}$ 可估的充要条件为：

$$\mathbf{C} = \mathbf{T}\mathbf{X},$$

这里，$\mathbf{T}$为$n \times q$的常数矩阵。

- 详细证明见教材

- 可估函数的直观解释为：当$\mathbf{X}$列不满秩，即rank$(\mathbf{X}) = r < p$时，可通过$\mathbf{X}$ 进行满秩分解$\mathbf{X} = \mathbf{GM}$，其中
  - $\mathbf{G}$为$n \times r$的列满秩矩阵
  - $\mathbf{M}$为$r \times p$的行满秩矩阵
- 令$\mathbf{X}^* = \mathbf{G}$和$\mathbf{\Theta}^* = \mathbf{M\Theta}$，则模型(2.1)可被重新参数化为：

$$\mathbf{Y} = \mathbf{X}^*\mathbf{\Theta}^* + \mathbf{E}$$

- 由于$\mathbf{X}^*$为列满秩矩阵，因此$\mathbf{\Theta}^* = \mathbf{M\Theta}$ 可估
- 从而$\mathbf{C\Theta}$可估，本质上等价于$\mathbf{C\Theta}$可表示为$\mathbf{\Theta}^*$的一个线性函数，即存在矩阵$\mathbf{T}^*$，使得$\mathbf{C\Theta} = \mathbf{T}^*\mathbf{\Theta}^*$。

- 当$\mathbf{X}$列不满秩时，经过上述重新参数化，可以求得正态分布假设下$\boldsymbol{\Theta}^*$ 和$\boldsymbol{\Sigma}$ 的极大似然估计，分别为

$$\widetilde{\boldsymbol{\Theta}}^* = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{Y}, \qquad \widetilde{\boldsymbol{\Sigma}} = \frac{1}{n}\mathbf{Y}'[\mathbf{I}_n - \mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}']\mathbf{Y}.$$

- 注意：$\mathcal{M}(\mathbf{G}) = \mathcal{M}(\mathbf{X})$，两矩阵的列向量空间的正交投影阵相等，即$\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}' = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'$，其中$(\mathbf{X}'\mathbf{X})^{-}$表示矩阵$\mathbf{X}'\mathbf{X}$的广义逆。
- 故$\boldsymbol{\Sigma}$的极大似然估计可重新表示为：

$$\widetilde{\boldsymbol{\Sigma}} = \frac{1}{n}\mathbf{Y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}']\mathbf{Y}.$$

- 由$\mathbf{C\Theta}$ 的可估性表明：存在矩阵$\mathbf{T}$，使得

$$\mathbf{C\Theta} = \mathbf{TX\Theta} = \mathbf{TG\Theta}^*.$$

- 因此，应用极大似然估计的变换不变性，得到可估函数$\mathbf{C\Theta}$的极大似然估计：

$$\mathbf{TG}\widetilde{\mathbf{\Theta}^*} = \mathbf{TG}(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{Y} = \mathbf{TX}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y} = \mathbf{C}\widehat{\mathbf{\Theta}}$$

- 这里，$\widehat{\mathbf{\Theta}} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$ 为当$\mathbf{X}$列不满秩时的一个最小二乘解。

- 另外，由最小二乘解$\widehat{\boldsymbol{\Theta}}$，得可估函数$\mathbf{C}\boldsymbol{\Theta}$的最小二乘估计$\mathbf{C}\widehat{\boldsymbol{\Theta}}$，以及$\boldsymbol{\Sigma}$的最小二乘估计为：

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n-r}(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\Theta}})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\Theta}}) = \frac{1}{n-r}\mathbf{Y}'[\mathbf{I}_n - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}']\mathbf{Y}.$$

定理8.2.4

若模型的随机误差$\mathbf{E}$服从正态分布，即$\mathrm{Vec}(\mathbf{E}) \sim N_{nq}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n)$，则

1. $\mathbf{C}\widehat{\boldsymbol{\Theta}} \sim N_{m\times q}(\mathbf{C}\boldsymbol{\Theta}, \boldsymbol{\Sigma} \otimes (\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}'))$；
2. $(n-r)\widehat{\boldsymbol{\Sigma}} \sim W_p(n-r, \boldsymbol{\Sigma})$；
3. $\mathbf{C}\widehat{\boldsymbol{\Theta}}$和$\widehat{\boldsymbol{\Sigma}}$相互独立。

- 考虑正态假设下多元多重线性回归模型的线性假设检验问题：

$$H_0 : \mathbf{C}\boldsymbol{\Theta} = \mathbf{D}, \qquad H_1 : \mathbf{C}\boldsymbol{\Theta} \neq \mathbf{D},$$

这里

  - $\mathbf{C}\boldsymbol{\Theta}$为可估函数
  - $\mathbf{D}$为某个给定的$m \times q$ 常数矩阵
  - $\mathbf{C}$为$m \times p$的行满秩常数矩阵

- 假设$\mathrm{Vec}(\mathbf{E}) \sim N_{nq}(\mathbf{0}, \boldsymbol{\Sigma} \otimes \mathbf{I}_n)$，采用似然比构造检验统计量。
- 记 $\widehat{\boldsymbol{\Theta}}_C = \widehat{\boldsymbol{\Theta}} - (\mathbf{X}'\mathbf{X})^{-}\mathbf{C}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\boldsymbol{\Theta}} - \mathbf{D})$。

- 可见，$\widehat{\mathbf{\Theta}}_C$为$\mathbf{\Theta}$的约束最小二乘解。
- 由$\mathbf{C\Theta}$的可估性，易证满足$\mathbf{C}\widehat{\mathbf{\Theta}}_C = \mathbf{D}$。
- 当原假设H_0成立时，注意到

$$
\begin{aligned}
&(\mathbf{X}\widehat{\mathbf{\Theta}}_C - \mathbf{X\Theta})'(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{\Theta}}_C)\\
=&(\widehat{\mathbf{\Theta}}_C - \mathbf{\Theta})(\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\mathbf{\Theta}} - \mathbf{D})\\
=&(\widehat{\mathbf{\Theta}}_C - \mathbf{\Theta})'\mathbf{C}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\mathbf{\Theta}} - \mathbf{D})\\
=&(\mathbf{C}\widehat{\mathbf{\Theta}}_C - \mathbf{D})'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\mathbf{\Theta}} - \mathbf{D}) = \mathbf{0}.
\end{aligned}
$$

- 因此，有

$$\begin{aligned}
&(\mathbf{Y}-\mathbf{X}\boldsymbol{\Theta})'(\mathbf{Y}-\mathbf{X}\boldsymbol{\Theta})\\
=&(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C+\mathbf{X}\widehat{\boldsymbol{\Theta}}_C-\mathbf{X}\boldsymbol{\Theta})'(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C+\mathbf{X}\widehat{\boldsymbol{\Theta}}_C-\mathbf{X}\boldsymbol{\Theta})\\
=&(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C)'(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C)+(\widehat{\boldsymbol{\Theta}}_C-\boldsymbol{\Theta})'\mathbf{X}'\mathbf{X}(\widehat{\boldsymbol{\Theta}}_C-\boldsymbol{\Theta})\\
\geq&(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C)'(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C).
\end{aligned}$$

- 在H_0成立时，可证得对数似然函数有如下不等式：

$$\begin{aligned}
\ln L(\boldsymbol{\Theta},\boldsymbol{\Sigma}) &= -\frac{n}{2}\ln|2\pi\boldsymbol{\Sigma}|-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\mathbf{X}\boldsymbol{\Theta})'(\mathbf{Y}-\mathbf{X}\boldsymbol{\Theta})\right]\\
&\leq -\frac{n}{2}\ln|2\pi\boldsymbol{\Sigma}|-\frac{1}{2}\mathrm{tr}\left[\boldsymbol{\Sigma}^{-1}(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C)'(\mathbf{Y}-\mathbf{X}\widehat{\boldsymbol{\Theta}}_C)\right]
\end{aligned}$$

- 因此，在原假设H_0下，$\mathbf{X\Theta}$的约束极大似然估计为：$\mathbf{X}\widehat{\mathbf{\Theta}}_C$
- $\mathbf{\Sigma}$的极大似然估计为：

$$\widetilde{\mathbf{\Sigma}}_C = \frac{1}{n}(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{\Theta}}_C)'(\mathbf{Y} - \mathbf{X}\widehat{\mathbf{\Theta}}_C).$$

- 证得

$$\max_{H_0} L(\mathbf{\Theta}, \mathbf{\Sigma}) = L(\widehat{\mathbf{\Theta}}_C, \widehat{\mathbf{\Sigma}}_C) = |2\pi\widetilde{\mathbf{\Sigma}}_C|^{-n/2}.$$

- 在备择假设H_1下，$\mathbf{X\Theta}$ 和$\mathbf{\Sigma}$的极大似然估计为：

$$\mathbf{X}\widetilde{\mathbf{\Theta}} = \mathbf{X}\widehat{\mathbf{\Theta}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$$

和

$$\widetilde{\mathbf{\Sigma}} = \frac{n-r}{n}\widehat{\mathbf{\Sigma}} = \frac{1}{n}(\mathbf{Y}-\mathbf{X}\widehat{\mathbf{\Theta}})'(\mathbf{Y}-\mathbf{X}\widehat{\mathbf{\Theta}}).$$

- 故有

$$\max_{H_1} L(\mathbf{\Theta}, \mathbf{\Sigma}) = L(\widetilde{\mathbf{\Theta}}, \widetilde{\mathbf{\Sigma}}) = |2\pi\widetilde{\mathbf{\Sigma}}|^{-n/2}.$$

模型参数的检验

- 似然比为

$$\lambda = \frac{\max\limits_{H_0} L(\Theta, \Sigma)}{\max\limits_{H_1} L(\Theta, \Sigma)} = \left(\frac{|\widetilde{\Sigma}|}{|\widetilde{\Sigma}_C|} \right)^{n/2}.$$

- 由于似然比为$|\widetilde{\Sigma}|/|\widetilde{\Sigma}_C|$ 的单调增函数
- 故由似然比可导出Wilks-Λ检验统计量：

$$\Lambda = \frac{|\widetilde{\Sigma}|}{|\widetilde{\Sigma}_C|} = \frac{|n\widetilde{\Sigma}|}{|n\widetilde{\Sigma} + n(\widetilde{\Sigma}_C - \widetilde{\Sigma})|}.$$

定理8.2.5

若原假设$H_0 : \mathbf{C}\boldsymbol{\Theta} = \mathbf{D}$成立，则

$$\Lambda = \frac{|\widetilde{\Sigma}|}{|\widetilde{\Sigma}_C|} \sim \Lambda(q, n-r, m),$$

其中r为矩阵$\mathbf{X}$的秩，$\mathbf{C}\boldsymbol{\Theta}$为可估函数，且$\mathbf{C}$为行满秩的$m \times p$ 矩阵。

证明：记$\mathbf{S}_0 = n\widetilde{\Sigma}$和$\mathbf{S}_C = n(\widetilde{\Sigma}_C - \widetilde{\Sigma})$。注意到：$\mathbf{S}_0 = (\mathbf{Y} - \mathbf{X}\widehat{\Theta})'(\mathbf{Y} - \mathbf{X}\widehat{\Theta})$和

$$\begin{aligned} \mathbf{S}_C &= (\mathbf{X}\widehat{\Theta}_C - \mathbf{X}\widehat{\Theta})'(\mathbf{X}\widehat{\Theta}_C - \mathbf{X}\widehat{\Theta}) \\ &= (\mathbf{C}\widehat{\Theta} - \mathbf{D})'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-}\mathbf{C}']^{-1}(\mathbf{C}\widehat{\Theta} - \mathbf{D}). \end{aligned}$$

由定理8.2.4，有$\mathbf{S}_0 \sim W_q(n-r, \Sigma)$。

在原假设H_0下，$\mathbf{S}_C \sim W_q(m, \Sigma)$，且两者独立。

因此，由Wilks-Λ 统计量的定义，可得

$$\Lambda = \frac{|\mathbf{S}_0|}{|\mathbf{S}_0 + \mathbf{S}_C|} \sim \Lambda(q, n-r, m).$$

■ 当$q=1$时，模型退化为线性回归模型：$\boldsymbol{Y}_{n\times 1}=\mathbf{X}_{n\times p}\boldsymbol{\beta}_{p\times 1}+\boldsymbol{\varepsilon}_{n\times 1}$。

- 进一步，检验问题退化为：$H_0:\boldsymbol{C\beta}=d$，其中$\boldsymbol{C}$为给定的$1\times p$向量，且$d$为一个给定的常数。
- 这时，F检验统计量为：

$$F_1=\frac{n-r}{m}\frac{1-\Lambda}{\Lambda}=\frac{n-r}{m}\frac{S_C}{S_0},$$

其中

- ▷ $S_0=(\boldsymbol{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}})'(\boldsymbol{Y}-\mathbf{X}\widehat{\boldsymbol{\beta}})$
- ▷ $S_C=(\boldsymbol{C}\widehat{\boldsymbol{\beta}}-d)'[\boldsymbol{C}(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{C}']^{-1}(\boldsymbol{C}\widehat{\boldsymbol{\beta}}-d)$
- ▷ $\boldsymbol{\beta}=(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\boldsymbol{Y}$

- 结合定理8.2.5，得到$H_0 : C\beta = d$ 的拒绝域：

$$F_1 = \frac{n-r}{m}\frac{S_C}{S_0} \geq F_{m,n-r}(\alpha).$$

- 当$m = 1$时，则$\mathbf{C}$和$\mathbf{D}$分别为$1 \times p$和$1 \times q$的行向量，记$\mathbf{C} = \boldsymbol{c}'$ 和$\mathbf{D} = \boldsymbol{d}'$。

- 于是，Wilks-Λ统计量定义为：

$$\Lambda = \frac{|\mathbf{S}_0|}{|\mathbf{S}_0 + \mathbf{S}_C|} = \frac{1}{|\mathbf{I}_m + \mathbf{S}_0^{-1}\mathbf{S}_C|} = \frac{1}{1 + \mathrm{tr}(\mathbf{S}_0^{-1}\mathbf{S}_C)},$$

  - ▷ $\mathbf{S}_0 = (\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\Theta}})'(\mathbf{Y} - \mathbf{X}\widehat{\boldsymbol{\Theta}})$和$\mathbf{S}_C = \frac{1}{\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta})'(\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta})$

- 进一步，计算可得

$$\mathrm{tr}(\mathbf{S}_0^{-1}\mathbf{S}_C) = \mathrm{tr}(\mathbf{S}_0^{-1}\mathbf{S}_C) = \frac{1}{\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}\mathrm{tr}\left[\mathbf{S}_0^{-1}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta})'(\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta})\right]$$
$$= \left(\frac{\boldsymbol{c}'\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta}}{\sqrt{\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}}\right)\mathbf{S}_0^{-1}\left(\frac{\boldsymbol{c}'\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta}}{\sqrt{\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}}\right)'.$$

- 由定理8.2.5，在原假设H_0成立时，有

$$F_2 = \frac{n-r-q+1}{q}\frac{1-\Lambda}{\Lambda}$$
$$= \frac{n-r-q+1}{q(n-r)}\left(\frac{\boldsymbol{c}'\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta}}{\sqrt{\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}}\right)\widehat{\boldsymbol{\Sigma}}^{-1}\left(\frac{\boldsymbol{c}'\widehat{\boldsymbol{\Theta}} - \boldsymbol{c}'\boldsymbol{\Theta}}{\sqrt{\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}}\right)'$$
$$\sim F_{q,n-r-q+1}.$$

定理8.2.6

当$m=1$时，拒绝原假设$H_0:\boldsymbol{c}'\widehat{\boldsymbol{\Theta}}=\boldsymbol{d}'$的拒绝域为：

$$(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}}-\boldsymbol{d}')\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}}-\boldsymbol{d}')'>\frac{q(n-r)\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{c}}{n-r-q+1}F_{q,n-r-q+1}(\alpha), \quad (2.1)$$

其中$\boldsymbol{c}$和$\boldsymbol{d}$分别为$p\times 1$和$q\times 1$的常数向量，$\boldsymbol{c}$属于$\mathbf{X}$的行向量空间。

定理8.2.7

当$q=2$时，拒绝原假设$H_0:\mathbf{C}\Theta=\mathbf{D}$的拒绝域为：

$$\frac{n-r-1}{m}\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}>F_{2m,2(n-r-1)}(\alpha).$$

定理8.2.8

当$m=2$ 时，拒绝原假设$H_0:\mathbf{C}\widehat{\Theta}=\mathbf{D}$ 的拒绝域为：

$$\frac{n-r-q+1}{m}\frac{1-\sqrt{\Lambda}}{\sqrt{\Lambda}}\geq F_{2q,2(n-r-q+1)}(\alpha).$$

- 当$q \neq 1,2$且$m \neq 1,2$时，通常基于Wilks-Λ统计量的渐近分布得到H_0的拒绝域，即

$$-2\ln\lambda = -n\ln\Lambda > \chi^2_f(\alpha),$$

其中自由度$f = mq$。

- 当n很大时，也可采用调整的检验统计量，提高拒绝域或p值的精度，相应的拒绝域为：

$$-\left[n - r - \frac{1}{2}(q - m + 1)\right]\ln\Lambda > \chi^2_f(\alpha).$$

- 基于随机矩阵$\mathbf{S}_0$和$\mathbf{S}_C$还可构造其他检验统计量，如：

$$\text{Roy最大特征值检验统计量} = \lambda_1\left(\mathbf{S}_C\mathbf{S}_0^{-1}\right),$$

$$\text{Pillai迹检验统计量} = \mathrm{tr}\left[\mathbf{S}_C(\mathbf{S}_0+\mathbf{S}_C)^{-1}\right],$$

$$\text{Hotelling-Lawley 迹统计量} = \mathrm{tr}\left(\mathbf{S}_C\mathbf{S}_0^{-1}\right).$$

- 目的：考虑q个因变量$\boldsymbol{Y}_0 = (Y_{01}, \cdots, Y_{0q})'$ 在$\boldsymbol{X}_0 = (1, x_{01}, \cdots, x_{0(p-1)})'$处的预测问题
- 假设$\boldsymbol{Y}_0$与历史数据$\mathbf{Y}$ 来源于同一线性回归模型：$\boldsymbol{Y}_0' = \boldsymbol{X}_0'\boldsymbol{\Theta} + \boldsymbol{E}_0'$
- 令$\mathcal{M}(\mathbf{A})$表示由矩阵$\mathbf{A}$的列向量张成的空间

定义8.2：可被预测

称$\boldsymbol{Y}_0 = (Y_{01}, \cdots, Y_{0q})'$可被预测，若$\boldsymbol{X}_0'\boldsymbol{\Theta}$ 可估，即$\boldsymbol{X}_0 \in \mathcal{M}(\mathbf{X}')$。

- 由于$E(\boldsymbol{Y}_0') = \boldsymbol{X}_0'\boldsymbol{\Theta}$，则$\boldsymbol{Y}_0$的点预测为：

$$\widehat{\boldsymbol{Y}}_0 = (\boldsymbol{X}_0'\widehat{\boldsymbol{\Theta}})' = \widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0$$

- 下面考虑$E(\boldsymbol{Y}_0)$的置信域和$\boldsymbol{Y}_0$ 的预测域问题
- 假设随机误差$\mathrm{Vec}(\mathbf{E}) \sim N_{nq}(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I}_n)$。令$\boldsymbol{c} = \boldsymbol{X}_0$，可得：

$$\frac{n-r-q+1}{q(n-r)} \left(\frac{\boldsymbol{X}_0'\widehat{\boldsymbol{\Theta}} - \boldsymbol{X}_0'\boldsymbol{\Theta}}{\sqrt{\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0}} \right) \widehat{\mathbf{\Sigma}}^{-1} \left(\frac{\boldsymbol{X}_0'\widehat{\boldsymbol{\Theta}} - \boldsymbol{X}_0'\boldsymbol{\Theta}}{\sqrt{\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0}} \right)'$$

$$\sim F_{q,n-r-q+1}.$$

- $E(\boldsymbol{Y}_0)$的$1-\alpha$置信超椭球为：

$$(\boldsymbol{X}_0'\widehat{\boldsymbol{\Theta}}-\boldsymbol{X}_0'\boldsymbol{\Theta})\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{X}_0'\widehat{\boldsymbol{\Theta}}-\boldsymbol{X}_0'\boldsymbol{\Theta})'$$
$$\leq\frac{q(n-r)[\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(\alpha).$$

- q个响应变量均值$E(Y_{0j}),\cdots,E(Y_{0q})$的$1-\alpha$ 同时置信区间为：

$$\boldsymbol{X}_0'\widehat{\boldsymbol{\beta}}_j\pm\sqrt{\frac{q(n-r)[\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(\alpha)}\cdot\widehat{\sigma}_{jj},\quad j=1,\cdots,q,$$

  - ▷ $\widehat{\sigma}_{jj}$为$\widehat{\boldsymbol{\Sigma}}$的第$j$个对角线元素

- 目的： 考虑响应变量$\boldsymbol{Y}_0' = \boldsymbol{X}_0'\widehat{\boldsymbol{\Theta}} + \boldsymbol{E}_0'$的预测置信域问题
- 假设$\boldsymbol{E}_0 \sim N_q(\mathbf{0}, \boldsymbol{\Sigma})$，与模型的随机误差阵$\mathbf{E}$独立
- 注意到：

$$\boldsymbol{Y}_0 - \widehat{\boldsymbol{Y}}_0 = \boldsymbol{Y}_0 - \widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0 \sim N_q\left(\mathbf{0}, [1 + \boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]\boldsymbol{\Sigma}\right)$$

与$\widehat{\boldsymbol{\Sigma}}$独立。

- 因此，$\boldsymbol{Y}_0$的$1-\alpha$预测域为：

$$(\boldsymbol{Y}_0-\widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0)'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{Y}_0-\widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0) \le \frac{q(n-r)[1+\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(\alpha).$$

- 响应变量$Y_{01},\cdots,Y_{0q}$的$1-\alpha$ 同时预测区间为：

$$\boldsymbol{X}_0'\widehat{\boldsymbol{\beta}}_j \pm \sqrt{\frac{q(n-r)[1+\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(\alpha)\cdot\widehat{\sigma}_{jj}},$$

其中$j=1,\cdots,q$。

- 注意到：$(n-r)\widehat{\sigma}_{jj}/\sigma_{jj} \sim \chi^2_{n-r}$，则有

$$\frac{X_0'\widehat{\beta}_j - X_0'\beta_j}{\sqrt{\widehat{\sigma}_{jj}[X_0'(\mathbf{X}'\mathbf{X})^- X_0]}} \sim t_{n-r}, \qquad \frac{Y_{0j} - X_0'\widehat{\beta}_j}{\sqrt{\widehat{\sigma}_{jj}[1 + X_0'(\mathbf{X}'\mathbf{X})^- X_0]}} \sim t_{n-r}.$$

- $E(Y_{0j})$ 的$1-\alpha$置信区间为：

$$X_0'\widehat{\beta}_j \pm \sqrt{\widehat{\sigma}_{jj} X_0'(\mathbf{X}'\mathbf{X})^- X_0}\, t_{n-r}(\alpha/2), \quad j = 1, \cdots, q.$$

- 响应变量Y_{0j} 的$1-\alpha$预测区间为：

$$X_0'\widehat{\beta}_j \pm \sqrt{\widehat{\sigma}_{jj}[1 + X_0'(\mathbf{X}'\mathbf{X})^- X_0]}\, t_{n-r}(\alpha/2), \; j = 1, \cdots, q.$$

例

各公司/各单位在采购计算机前，首先会先根据自己公司/单位需求对计算机的各项性能进行评估。因此，类似的公司，对计算机硬件设备的要求也是类似的。采用多元多重线性回归模型来刻画订单信息与计算机硬件设备性能之间的关系。记

$X_1 =$ 客户的订单量(单位：千台)；

$X_2 =$ 添加-删除项数(单位：千项)；

$Y_1 =$ CPU(central processing unit)时间(单位：小时)；

$Y_2 =$ 磁盘输入/输出能力(disk I/O)(单位：小时)。

Table 3: 计算机数据

X_1 (订单量)	X_2 (添加-删除项数)	Y_1 (CPU时间)	Y_2 (磁盘输入/输出能力)
123.5	2.108	141.5	301.8
146.1	9.213	168.9	296.1
133.9	1.905	154.8	328.2
128.5	0.815	146.5	307.4
151.1	1.061	172.8	362.4
136.2	8.603	160.1	369.5
92.0	1.125	108.5	229.1

- 考虑下面的模型：

$$Y_j = \beta_{0j} + X_1\beta_{1j} + X_2\beta_{2j} + \varepsilon_j, \qquad j = 1, 2.$$

- 已知$n = 7$，$q = 2$，$p = r = 3$
- 记$\boldsymbol{\Theta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$和$\boldsymbol{\Sigma} = (\sigma_{ij})$，其中$\boldsymbol{\beta}_j = (\beta_{0j}, \beta_{1j}, \beta_{2j})'$，$j = 1, 2$，且$\sigma_{ij} = \mathrm{Cov}(\varepsilon_i, \varepsilon_j)$，$1 \le i, j \le 2$
- 计算得

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 8.17969 & & \\ -0.06411 & 0.00052 & \\ 0.08831 & -0.00107 & 0.01440 \end{pmatrix}.$$

- 回归系数矩阵Θ和协方差矩阵Σ的最小二乘估计分别为：

$$\widehat{\Theta} = (\widehat{\beta}_1, \widehat{\beta}_2) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \begin{pmatrix} 8.42 & 14.14 \\ 1.08 & 2.25 \\ 0.42 & 5.67 \end{pmatrix}$$

和

$$\widehat{\Sigma} = \frac{1}{7-2}(\mathbf{Y} - \mathbf{X}\widehat{\Theta})'(\mathbf{Y} - \mathbf{X}\widehat{\Theta}) = \begin{pmatrix} 1.160 & 1.060 \\ 1.060 & 2.626 \end{pmatrix}.$$

- 得到Y_1和Y_2的经验回归方程分别为：

$$Y_1 = 8.42 + 1.08X_1 + 0.42X_2,$$

$$Y_2 = 14.14 + 2.25X_1 + 5.67X_2.$$

- 首先，进行回归方程的显著性检验，即检验响应变量(Y_1, Y_2)是否与协变量(X_1, X_2)存在线性关系，检验问题可表达为：

$$H_0: \quad \beta_{11} = \beta_{12} = \beta_{21} = \beta_{22} = 0 \Leftrightarrow H_1: \quad \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}\text{不全为零}$$

- 等价于

$$H_0: \quad \mathbf{C}\mathbf{\Theta} = \mathbf{0}, \qquad H_1: \quad \mathbf{C}\widehat{\mathbf{\Theta}} \neq \mathbf{0},$$

其中$\mathbf{C} = (\mathbf{0}, \mathbf{I}_2)$。

- 应用定理8.2.5，当原假设H_0成立时，有

$$\Lambda = \frac{|\mathbf{S}_0|}{|\mathbf{S}_0 + \mathbf{S}_C|} \sim \Lambda(2, 4, 2).$$

- 注意$m=q=2$，由定理8.2.7 或定理8.2.8可得$H_0:\mathbf{C\Theta}=\mathbf{0}$的拒绝域为：

$$\frac{3(1-\sqrt{\Lambda})}{2\sqrt{\Lambda}}\geq F_{4,6}(\alpha).$$

- 查表得$F_{4,6}(0.05)=4.534$。计算得

$$\mathbf{S}_0=(n-r)\widehat{\Sigma}=5\times\begin{pmatrix}1.16 & 1.06\\ 1.06 & 2.626\end{pmatrix}=\begin{pmatrix}5.80 & 5.30\\ 5.30 & 13.13\end{pmatrix},$$

$$\begin{aligned}\mathbf{S}_C&=(\mathbf{C}\widehat{\mathbf{\Theta}})'[\mathbf{C}(\mathbf{X}'\mathbf{X})^-\mathbf{C}']^{-1}(\mathbf{C}\widehat{\mathbf{\Theta}})\\ &=\begin{pmatrix}1.08 & 2.25\\ 0.42 & 5.67\end{pmatrix}\begin{pmatrix}0.00052 & -0.00107\\ -0.00107 & 0.01440\end{pmatrix}^{-1}\begin{pmatrix}1.08 & 2.25\\ 0.42 & 5.67\end{pmatrix}'\\ &=\begin{pmatrix}3882.780 & 3267.983\\ 3267.983 & 3839.414\end{pmatrix}.\end{aligned}$$

- 计算Wilks-Λ统计量为：

$$\Lambda = \frac{|\mathbf{S}_0|}{|\mathbf{S}_0 + \mathbf{S}_C|} = 1.12653 \times 10^{-5},$$

且

$$\frac{3(1-\sqrt{\Lambda})}{2\sqrt{\Lambda}} = 445.4092 \geq F_{4,6}(0.05) = 4.534.$$

- 因此，拒绝H_0，认为(Y_1, Y_2) 关于(X_1, X_2)的线性回归显著。
- 接下来，依次检验模型中X_1的回归系数的显著性问题：

$$H_{01}: \quad \beta_{11} = \beta_{21} = 0, \qquad H_{11}: \quad \beta_{11}\text{和}\beta_{21}\text{不全为零。}$$

令$\boldsymbol{C} = \boldsymbol{c}' = (0, 1, 0)$，则$H_{01}: \beta_{11} = \beta_{21} = 0$ 等价于$H_{01}: (0, 1, 0)\boldsymbol{\Theta} = \mathbf{0}$。

- 注意到：$m=1$，所以应用定理8.2.6得到H_{01}的拒绝域为：

$$(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}})\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}})' \geq \frac{q(n-r)\boldsymbol{c}'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{c}}{n-r-q+1}F_{q,n-r-q+1}(\alpha).$$

- 计算得$n-r-q+1=3$，计算得

$$\begin{aligned}
&(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}})\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}})' \\
=&(\widehat{\beta}_{11},\widehat{\beta}_{21})\widehat{\boldsymbol{\Sigma}}^{-1}(\widehat{\beta}_{11},\widehat{\beta}_{21})' \\
=&(1.08,0.42)\begin{pmatrix} 1.16 & 1.06 \\ 1.06 & 2.626 \end{pmatrix}^{-1}\begin{pmatrix} 1.08 \\ 0.42 \end{pmatrix} \\
=&1.199
\end{aligned}$$

$$
\begin{aligned}
\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c} &= (0,1,0)\begin{pmatrix} 8.17969 & & \\ -0.06411 & 0.00052 & \\ 0.08831 & -0.00107 & 0.01440 \end{pmatrix}(0,1,0)' \\
&= 0.00052.
\end{aligned}
$$

- 由上面的结果，可得

$$
\begin{aligned}
\frac{q(n-r)\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}{(n-r-q+1)}F_{q,n-r-q+1}(\alpha) &= \frac{8\times 0.00052}{3}F_{2,3}(0.05) \\
&= \frac{8\times 0.00052\times 9.552}{3} \\
&= 0.0132 < 1.199.
\end{aligned}
$$

- 因此，拒绝原假设H_{01}，认为模型中X_1的回归系数显著。

- 同样的方法可检验对模型中X_2的回归系数显著性，即检验

$$H_{02}: \ \beta_{12}=\beta_{22}=0, \Leftrightarrow H_{11}: \ \beta_{12}\text{和}\beta_{22}\text{不全为零}$$

- 令$\boldsymbol{c}=(0,0,1)'$，则$H_{02}:\boldsymbol{c}'\boldsymbol{\Theta}=\mathbf{0}$
- 计算得

$$\begin{aligned}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}})\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{c}'\widehat{\boldsymbol{\Theta}})' &= (\widehat{\beta}_{12},\widehat{\beta}_{22})\widehat{\boldsymbol{\Sigma}}^{-1}(\widehat{\beta}_{12},\widehat{\beta}_{22})' \\ &= (2.25, 5.67)\begin{pmatrix} 1.16 & 1.06 \\ 1.06 & 2.626 \end{pmatrix}^{-1}\begin{pmatrix} 2.25 \\ 5.67 \end{pmatrix} \\ &= 12.245\end{aligned}$$

- 以及

$$
\begin{aligned}
\frac{q(n-r)\mathbf{c}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}}{(n-r-q+1)}F_{q,n-r-q+1}(\alpha) &= \frac{8\times 0.0144\times 9.552}{3} \\
&= 0.367 \\
&< 12.245.
\end{aligned}
$$

- 因此，拒绝原假设H_{02}，认为模型中X_2的回归系数显著。

- 预测： 下面以$\boldsymbol{X}_0=(1,130,7.5)'$ 为例，给出响应变量$\boldsymbol{Y}_0=(Y_{01},Y_{02})'$的预测
- 首先，$\boldsymbol{Y}_0$的点预测就是$\boldsymbol{\Theta}'\boldsymbol{X}_0$ 的点估计，即为

$$\widehat{\boldsymbol{Y}}_0=\widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0=\begin{pmatrix}\boldsymbol{X}_0'\widehat{\boldsymbol{\beta}}_1\\ \boldsymbol{X}_0'\widehat{\boldsymbol{\beta}}_2\end{pmatrix}=\begin{pmatrix}151.97\\ 349.17\end{pmatrix}.$$

- 目标： 接着分别计算$E(\boldsymbol{Y}_0)$和$\boldsymbol{Y}_0$及其分量的95% 置信域和预测域
- 计算，得

$$\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{X}_0=0.34725$$

案例分析

$$\frac{q(n-r)[\mathbf{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(0.05)=\frac{2\times 4\times 0.34725}{3}F_{2,3}(0.05)=8.845$$

$$\frac{q(n-r)[1+\mathbf{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(0.05)=34.317$$

$$\sqrt{\widehat{\sigma}_{11}\frac{q(n-r)[\mathbf{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(0.05)}=\sqrt{1.16\times 8.845}=3.203$$

$$\sqrt{\widehat{\sigma}_{22}\frac{q(n-r)[\mathbf{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(0.05)}=\sqrt{2.626\times 8.845}=4.819$$

$$\sqrt{\widehat{\sigma}_{11}\frac{q(n-r)[1+\mathbf{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(0.05)}=\sqrt{1.16\times 34.317}=6.309$$

$$\sqrt{\widehat{\sigma}_{22}\frac{q(n-r)[1+\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-1}\boldsymbol{X}_0]}{n-r-q+1}F_{q,n-r-q+1}(0.05)}=\sqrt{2.626\times 34.317}=9.493$$

$$\sqrt{\widehat{\sigma}_{11}[\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}t_{n-r}(\alpha/2)=\sqrt{1.16\times 0.34725}\times t_4(0.025)=1.762$$

$$\sqrt{\widehat{\sigma}_{22}[\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}t_{n-r}(\alpha/2)=\sqrt{2.626\times 0.34725}\times t_4(0.025)=2.651,$$

$$\sqrt{\widehat{\sigma}_{11}[1+\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}t_{n-r}(\alpha/2)=\sqrt{1.16\times 1.34725}\times t_4(0.025)=3.471$$

$$\sqrt{\widehat{\sigma}_{22}[1+\boldsymbol{X}_0'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{X}_0]}t_{n-r}(\alpha/2)=\sqrt{2.626\times 1.34725}\times t_4(0.025)=5.222$$

- 由上面的结果，可得$E(\boldsymbol{Y}_0) = (\boldsymbol{X}_0'\boldsymbol{\beta}_1, \boldsymbol{X}_0'\boldsymbol{\beta}_1)'$ 的95%置信椭圆为：

$$\begin{pmatrix} \boldsymbol{X}_0'\boldsymbol{\beta}_1 - 151.97 \\ \boldsymbol{X}_0'\boldsymbol{\beta}_2 - 349.17 \end{pmatrix}' \begin{pmatrix} 1.16 & 1.06 \\ 1.06 & 2.626 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{X}_0'\boldsymbol{\beta}_1 - 151.97 \\ \boldsymbol{X}_0'\boldsymbol{\beta}_2 - 349.17 \end{pmatrix} \leq 8.845.$$

- $\boldsymbol{Y}_0 = (Y_{01}, Y_{02})'$ 的95%预测椭圆为：

$$(\boldsymbol{Y}_0 - \widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0)'\widehat{\boldsymbol{\Sigma}}^{-1}(\boldsymbol{Y}_0 - \widehat{\boldsymbol{\Theta}}'\boldsymbol{X}_0) \leq 34.317.$$

- 进一步，可以计算每个分量的预测区间，课下练习

- 多元多重线性回归模型的第j个响应变量仅依赖于参数矩阵Θ的第j列参数β，但在生物医学中，往往所研究的q个响应变量可能会共享参数，表示为

$$\Theta = \Upsilon \mathbf{H},$$

  - ▷ Υ为$q \times k$的参数矩阵
  - ▷ $\mathbf{H}$为$k \times p$ 的已知矩阵

- 代入模型，得

$$\mathbf{Y} = \mathbf{X}\Upsilon\mathbf{H} + \mathbf{E}, \qquad \mathrm{Cov}(\mathrm{Vec}(\mathbf{E})) = \mathbf{\Sigma} \otimes \mathbf{I}_n$$

- 称该模型为生长曲线模型(growth-curve model)，这源于此模型的不少例子来自生物生长问题。

- 例：生物学家欲研究白鼠的某个特征随时间变化情况，随机选用n只小白鼠做试验。在时刻$t_1, \cdots, t_p$对每只小白鼠观测该特征的值。
- 设第i只小白鼠的p次观测值为：

$$y_{i1}, \cdots, y_{ip}, \qquad i = 1, \cdots, n.$$

- 假定不同白鼠的观测值是不相关的，而同一只白鼠的p次观测却是相关的，且协方差阵为Σ。

- 从理论分析认为，这些观测值与观测时间t的关系为$k-1$阶多项式：

$$Y = f(t) = \beta_0 + \beta_1 t + \cdots + \beta_{k-1} t^{k-1},$$

这就是所谓理论生长曲线。

- 生物学家的目的是估计$\beta_0, \beta_1, \cdots, \beta_{k-1}$，以得到经验生长曲线。
- 记$\Upsilon = (\beta_0, \beta_1, \cdots, \beta_{k-1})$，且

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_p \\ \vdots & \vdots & & \vdots \\ t_1^{k-1} & t_2^{k-1} & \cdots & t_p^{k-1} \end{pmatrix}.$$

- 因此，可写成下面矩阵形式的模型：

$$\mathbf{Y} = \mathbf{1}_n \Upsilon \mathbf{H} + \mathbf{E}.$$

- 例：现在欲建立m个经验生长曲线。假设对n只小白鼠依品种或其他指标分成m个小组，第i组有n_i只，$n = \sum_{i=1}^{m} n_i$。
- 在时刻$t_1, \cdots, t_p$ 对每只小白鼠的特征进行观测。在理论上，对每小组有一条生长曲线，即

$$E(Y_i) = f_i(t) = \beta_{i0} + \beta_{i1} t + \cdots + \beta_{ik-1} t^{k-1}, \quad i = 1, \cdots, m.$$

- 记$y_{ij\,l}$为在时刻t_l对第i组的第j只小白鼠的观测值，

$$\mathbf{Y}_i = (y_{ij\,l})_{n_i \times p}, \quad i = 1, \cdots, m.$$

- 因此，观测矩阵为

$$\mathbf{Y}_{n\times p} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_m \end{pmatrix}.$$

- 这时，可用生长曲线模型来表达，其中

$$\mathbf{X} = \begin{pmatrix} \mathbf{1}_{n_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{n_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1}_{n_m} \end{pmatrix}, \qquad \boldsymbol{\Upsilon}_{p\times q} = \begin{pmatrix} \beta_{10} & \beta_{11} & \cdots & \beta_{1k-1} \\ \beta_{20} & \beta_{21} & \cdots & \beta_{2k-1} \\ \vdots & \vdots & & \vdots \\ \beta_{m0} & \beta_{m1} & \cdots & \beta_{mk-1} \end{pmatrix},$$

$$\mathbf{H} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ t_1 & t_2 & \cdots & t_p \\ \vdots & \vdots & & \vdots \\ t_1^{k-1} & t_2^{k-1} & \cdots & t_p^{k-1} \end{pmatrix}.$$

- 事实上，有

$$\begin{cases} \mathsf{Vec}(\mathbf{Y}) = (\mathbf{H}' \otimes \mathbf{X})\mathsf{Vec}(\boldsymbol{\Upsilon}) + \mathsf{Vec}(\mathbf{E}), \\ \mathsf{Cov}(\mathsf{Vec}(\mathbf{E})) = \boldsymbol{\Sigma} \otimes \mathbf{I}_n. \end{cases}$$

- 不难证明，线性函数$\varphi = \mathrm{tr}(\mathbf{C}\boldsymbol{\Upsilon})$ 可估的充要条件是，存在矩阵$\mathbf{T}_{n\times q}$，使得

$$\mathbf{C}' = \mathbf{X}'\mathbf{T}\mathbf{H}.$$

- 若$\boldsymbol{\Sigma}$已知，则$\beta^* = \mathrm{Vec}(\boldsymbol{\Upsilon}^*)$ 的广义最小二乘解为：

$$\beta^* = \mathrm{Vec}[(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}\boldsymbol{\Sigma}^{-1}\mathbf{H}'(\mathbf{H}\boldsymbol{\Sigma}^{-1}\mathbf{H}')^{-}],$$

- 等价地$\boldsymbol{\Upsilon}^* = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}\boldsymbol{\Sigma}^{-1}\mathbf{H}'(\mathbf{H}\boldsymbol{\Sigma}^{-1}\mathbf{H}')^{-}$。
- 故任一可估函数$\varphi = \mathrm{tr}(\mathbf{C}\boldsymbol{\Upsilon})$ 的广义最小二乘估计为$\varphi = \mathrm{tr}(\mathbf{C}\boldsymbol{\Upsilon}^*)$。

- 当Σ未知时，Σ的无偏估计为：

$$\mathbf{S} = \frac{1}{n - \mathsf{rank}(\mathbf{X})}\mathbf{Y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}')\mathbf{Y}$$

- 当$n - \mathsf{rank}(\mathbf{X}) > q$时，它以概率为1正定。将无偏估计$\mathbf{S}$代入，得到

$$\widetilde{\Upsilon} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}\mathbf{S}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{S}^{-1}\mathbf{H}')^{-}.$$

- 把所得估计称为Υ的两步广义最小二乘解。对任一可估函数$\varphi = \mathsf{tr}(\mathbf{C}\Upsilon)$，其两步估计为

$$\widetilde{\varphi} = \mathsf{tr}(\mathbf{C}\widetilde{\Upsilon}).$$


STATISTICS
DATA
统计与数据科学丛书 4
STATISTICS AND DATA SCIENCE SERIES 4
多元统计分析
李高荣 吴密霞 编著
Multivariate Statistical Analysis
科学出版社

谢谢，请多提宝贵意见！