作业 **10**

1. 对于一般线性模型 $\mathbf{y}=X\boldsymbol{\beta}+\boldsymbol{\epsilon},\boldsymbol{\epsilon}\sim(\mathbf{0},\sigma^2 I_n)$, 我们知道 $(I_n-P_X)X=0$, 其中 $P_X=X(X^\top X)^- X^\top$ 是投影阵，所以
$$RSS=\mathbf{y}^\top(I_n-P_X)\mathbf{y}=\boldsymbol{\epsilon}^\top(I_n-P_X)\boldsymbol{\epsilon},$$
即 RSS 不依赖于 $\boldsymbol{\beta}$。基于上述观点重新证明第 6 讲引理 1(2)：对于简单线性模型，残差平方和 $RSS=s_{\epsilon\epsilon}-s_{\epsilon x}^2/s_{xx}$。

2. 假设数据 $(y_i,x_i,z_i),i=1,...,n$ 满足线性模型
$$y_i=a+bx_i+cz_i+\epsilon_i,\epsilon_i\sim(0,\sigma^2),\epsilon_i\text{与}x_i,z_i\text{独立}$$
或等价地 ($\mathbf{x}=(x_1,...,x_n)^\top$,$\mathbf{z}=(z_1,...,z_n)^\top$)
$$\mathbf{y}=\mathbb{1}a+\mathbf{x}b+\mathbf{z}c+\boldsymbol{\epsilon},\quad \boldsymbol{\epsilon}\sim(0,\sigma^2 I_n),\quad \boldsymbol{\epsilon}\perp\!\!\!\perp(\mathbf{x},\mathbf{z}).$$
   (a) 记 $\hat{c}_0=S_{xz}/S_{zz}$, $S_{xz}=\frac{1}{n-1}\sum(x_i-\bar{x})(z_i-\bar{z}), S_{zz}=\frac{1}{n-1}\sum(z_i-\bar{z})^2$, 试用第 11 讲命题 4 证明 1 的投影方法（即先将三个向量 $\mathbb{1},\mathbf{x},\mathbf{z}$ 正交化，再投影）而不是直接应用命题 4 的结论，证明 b 的 LS 估计
$$\hat{b}=\frac{\sum[(x_i-\bar{x}-(z_i-\bar{z})\hat{c}_0]y_i}{\sum[(x_i-\bar{x}-(z_i-\bar{z})\hat{c}_0]^2}.$$
   （注：该表达式表明，为了求解 b 的 LS 估计，我们可以分两步回归：(1) $lm(x\sim z)$: $x_i=a_0+c_0z_i+\delta_i$, 求得回归系数 $\hat{c}_0=s_{xz}/s_{zz}$ 以及 $\hat{a}_0=\bar{x}-\bar{z}\hat{c}_0$, 得到残差 $x_i^\perp=x_i-\bar{x}-(z_i-\bar{z})\hat{c}_0$. (2) $y\sim x^\perp$: $y_i=\alpha+\beta x_i^\perp+error$，由此得到的斜率 LS 估计 $\hat{\beta}$ 就是 $\hat{b}$）

   (b) 证明
$$R^2=\frac{r_{yx}^2+r_{yz}^2-2r_{yx}r_{yz}r_{xz}}{1-r_{xz}^2},$$
   其中 r_{xz} 为两个自变量之间的样本相关系数。

3. 假设线性模型 $\mathbf{y}=X_1\boldsymbol{\beta}_1+X_2\boldsymbol{\beta}_2+\boldsymbol{\epsilon}$, 其中 X_1 为 $n\times(p-q)$ 矩阵，X_2 为 $n\times q$ 矩阵, $\boldsymbol{\beta}_2$ 为 $q\times 1$ 参数向量。记 $X_2^\perp=X_2-P_{X_1}X_2$。假设 W 是任何 $n\times q$ 矩阵使得 $W^\top X_1=0,W^\top X_2$ 可逆，证明 $\tilde{\boldsymbol{\beta}}_2=(W^\top X_2)^{-1}W^\top\mathbf{y}$ 是 $\boldsymbol{\beta}_2$ 的无偏估计，并证明 $\mathrm{var}(\tilde{\boldsymbol{\beta}}_2|X)\geq\sigma^2(X_2^{\perp\top}X_2^\perp)^{-1}$, 当 $W=X_2^\perp$ 时等号成立。

4. 假设 A 为任一 $n\times p$ 列满秩矩阵，其各列为 $A=(\mathbf{a}_1,...,\mathbf{a}_p)$, 我们知道 $A^\top A$ 的 (i,j) 元为 $\mathbf{a}_i^\top\mathbf{a}_j$。记 $\mathbf{a}_i^\perp=\mathbf{a}_i-P_{A_{(-i)}}\mathbf{a}_i$, 其中 $A_{(-i)}$ 为 A 删除第 i 列后剩余的列组成的矩阵。证明 $(A^\top A)^{-1}$ 的 (i,j) 元等于 $\frac{\mathbf{a}_i^{\perp\top}\mathbf{a}_j^\perp}{\|\mathbf{a}_i^\perp\|^2\|\mathbf{a}_j^\perp\|^2}$.

   提示：假设线性模型 $\mathbf{y}=A\boldsymbol{\beta}+\boldsymbol{\epsilon}$,$\boldsymbol{\epsilon}\sim(\mathbf{0},\sigma^2 I_n)$, $\boldsymbol{\beta}=(\beta_1,...,\beta_p)^\top$ ，注意这里 A 的第一列未必是 $\mathbb{1}$，但 LS 估计仍具有形式 $\hat{\boldsymbol{\beta}}=(A^\top A)^{-1}A^\top\mathbf{y}$, 方差矩阵为 $var(\hat{\boldsymbol{\beta}}|A)=\sigma^2(A^\top A)^{-1}$, 其 (i,j) 元为 $\hat{\beta}_i,\hat{\beta}_j$ 的协方差, 因此只需计算 $cov(\hat{\beta}_i,\hat{\beta}_j|A)$ （利用 11 讲命题 4 给出 $\hat{\beta}_i$，$\hat{\beta}_j$ 的表达）.