Exercise 5.4

Let $X = \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{bmatrix}$ Then we have

$$
\begin{aligned}
X^TX &= \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{p1} & x_{p2} & x_{p3} & \dots & x_{pn} \end{bmatrix} \begin{bmatrix} 1 & x_{11} & x_{21} & \dots & x_{p1} \\ 1 & x_{12} & x_{22} & \dots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \dots & x_{pn} \end{bmatrix} \\
&= \begin{bmatrix} n & \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{2i} & \dots & \sum_{i=1}^n x_{pi} \\ \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{1i}^2 & \sum_{i=1}^n x_{1i}x_{2i} & \dots & \sum_{i=1}^n x_{1i}x_{pi} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n x_{pi} & \sum_{i=1}^n x_{pi}x_{1i} & \sum_{i=1}^n x_{pi}x_{2i} & \dots & \sum_{i=1}^n x_{pi}^2 \end{bmatrix}
\end{aligned}
$$

From this, let $A_{11} = [n]$ be the 1×1 matrix, $A_{12} = \begin{bmatrix} \sum_{i=1}^n x_{1i} & \sum_{i=1}^n x_{2i} & \dots & \sum_{i=1}^n x_{pi} \end{bmatrix}$, and let $A_{22} = (\sum_{i=1}^n x_{ki}x_{ji})_{1\leq k,j\leq p}$.

First note that

$$
\begin{aligned}
\mathcal{X}^T\mathcal{X} &= \begin{bmatrix} (x_{11}-\overline{x}_1) & (x_{21}-\overline{x}_1) & \dots & (x_{n1}-\overline{x}_1) \\ (x_{12}-\overline{x}_2) & (x_{22}-\overline{x}_2) & \dots & (x_{n2}-\overline{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{1p}-\overline{x}_p) & (x_{2p}-\overline{x}_p) & \dots & (x_{np}-\overline{x}_p) \end{bmatrix} \begin{bmatrix} (x_{11}-\overline{x}_1) & (x_{12}-\overline{x}_2) & \dots & (x_{1p}-\overline{x}_1) \\ (x_{21}-\overline{x}_1) & (x_{22}-\overline{x}_2) & \dots & (x_{2p}-\overline{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{n1}-\overline{x}_1) & (x_{n2}-\overline{x}_2) & \dots & (x_{np}-\overline{x}_p) \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^n(x_{i1}-\overline{x}_1)^2 & \sum_{i=1}^n(x_{i1}-\overline{x}_1)(x_{i2}-\overline{x}_2) & \dots & \sum_{i=1}^n(x_{i1}-\overline{x}_1)(x_{ip}-\overline{x}_p) \\ \sum_{i=1}^n(x_{i1}-\overline{x}_1)(x_{i2}-\overline{x}_2) & \sum_{i=1}^n(x_{i2}-\overline{x}_2)^2 & \dots & \sum_{i=1}^n(x_{i2}-\overline{x}_2)(x_{ip}-\overline{x}_p) \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n(x_{i1}-\overline{x}_1)(x_{ip}-\overline{x}_p) & \sum_{i=1}^n(x_{i2}-\overline{x}_2)(x_{ip}-\overline{x}_p) & \dots & \sum_{i=1}^n(x_{ip}-\overline{x}_p)^2 \end{bmatrix}
\end{aligned}
$$

Now, for *any* $1\leq j\leq p$ and $1\leq k\leq p$

$$
\begin{aligned}
\sum_{i=1}^n(x_{ik}-\overline{x}_k)(x_{ij}-\overline{x}_j) &= \sum_{i=1}^n x_{ik}x_{ij} - \overline{x}_j\sum_{i=1}^n x_{ik} - \overline{x}_k\sum_{i=1}^n x_{ij} + n\overline{x}_j\overline{x}_k \\
&= \sum_{i=1}^n x_{ik}x_{ij} - 2n\overline{x}_j\overline{x}_k + n\overline{x}_j\overline{x}_k \\
&= \sum_{i=1}^n x_{ik}x_{ik} - n\overline{x}_j\overline{x}_k
\end{aligned}
$$

We will now show that our $A_{22} - A_{12}^TA_{11}^{-1}A_{12} = \mathcal{X}^T\mathcal{X}$. First note that $A_{12}^TA_{11}^{-1}A_{12} = \frac{1}{n}A_{12}^TA_{12}$. Moreover, the entries of $A_{12}^TA_{12}$ are given by

$$A_{12}^TA_{12} = \big(\sum_{i=1}^n x_{ki}\sum_{i=1}^n x_{ji}\big)_{1\leq k,j\leq p} = \big(n^2\overline{x}_k\overline{x}_k\big)_{1\leq j,k\leq p}$$

Therefore we see

$$A_{12}^TA_{11}^{-1}A_{12} = (n\overline{x}_k\overline{x}_j)_{1\leq k,j\leq p}$$

Combining this result with the definition of $A_{22} = (\sum_{i=1}^n x_{ki}x_{ji})_{1\leq k,j\leq p}$. Thus we see

$$A_{22} - A_{12}^TA_{11}^{-1}A_{12} = \Big(\sum_{i=1}^n x_{ki}x_{ji} - n\overline{x}_k\overline{x}_k\Big)_{1\leq k,j\leq p} = \mathcal{X}^T\mathcal{X}$$

Now, notice that $A_{11}^{-1}A_{12} = (\frac{1}{n}\sum_{i=1}^n x_{ik})1 \leq k \leq p = \overline{\mathbf{x}}^T$ and $A_{11}^{-1}A_{12}^T = (\frac{1}{n}\sum_{i=1}^n x_{ik})1 \leq k \leq p = \overline{\mathbf{x}}$. Having shown these relationshops hold we have

$$(\mathbf{X}^T\mathbf{X})^{-1} = \begin{bmatrix} \frac{1}{n} + (\overline{\mathbf{x}}^T)(\mathcal{X}^T\mathcal{X})^{-1}\overline{\mathbf{x}} & -\overline{\mathbf{x}}(\mathcal{X}^T\mathcal{X})^{-1}\overline{\mathbf{x}} \\ -(\mathcal{X}^T\mathcal{X})^{-1}\overline{\mathbf{x}} & (\mathcal{X}^T\mathcal{X})^{-1} \end{bmatrix}$$