Notes in RED are not part of solution.

Technical Exercise 1

(1) $\sum_{t=3}^{6} X_t$ and $\sum_{j=3}^{6} X_t$ are the same, while $\sum_{j=3}^{6} X_t$ is not equal to either.

Observe that

$$\sum_{t=3}^{6} X_t = X_3 + X_4 + X_5 + X_6$$

$$\sum_{j=3}^{6} X_j = X_3 + X_4 + X_5 + X_6$$

j and t are really placeholders

$$\sum_{j=3}^{6} X_t = X_t + X_t + X_t + X_t = 4X_t$$

(2)

$$\sum_{i=1}^{3} \sum_{j=2}^{3} X_{ij} = \sum_{i=1}^{3} \left(X_{i2} + X_{i3} \right) = \sum_{i=1}^{3} X_{i2} + \sum_{i=1}^{3} X_{i3} = X_{12} + X_{22} + X_{32} + X_{13} + X_{23} + X_{33}$$

property (3) in ExSet02

(3) Case of $n=2$:

$$\begin{aligned}\sum_{t=1}^{2} (X_t - \bar{X})(Y_t - \bar{Y}) &= (X_1 - \bar{X})(Y_1 - \bar{Y}) + (X_2 - \bar{X})(Y_2 - \bar{Y}) \\ &= \left(X_1 - \frac{X_1 + X_2}{2}\right)\left(Y_1 - \frac{Y_1 + Y_2}{2}\right) + \left(X_2 - \frac{X_1 + X_2}{2}\right)\left(Y_2 - \frac{Y_1 + Y_2}{2}\right) \\ &= \frac{1}{2}(X_1 - X_2) \cdot \frac{1}{2}(Y_1 - Y_2) + \frac{1}{2}(X_2 - X_1)\frac{1}{2}(Y_2 - Y_1)\end{aligned}$$

$$= \frac{1}{4}(X_2-X_1)(Y_2-Y_1) + \frac{1}{4}(X_2-X_1)(Y_2-Y_1)$$

$$= \frac{1}{2}(X_2-X_1)(Y_2-Y_1)$$

$$\sum_{i=1}^{2}\sum_{j=1}^{2}(X_j-X_i)(Y_j-Y_i) = \sum_{i=1}^{2}\Big[(X_1-X_i)(Y_1-Y_i) + (X_2-X_i)(Y_2-Y_i)\Big]$$

$$= (X_1-X_1)(Y_1-Y_1)+(X_2-X_1)(Y_2-Y_1)+(X_1-X_2)(Y_1-Y_2) + (X_2-X_2)(Y_2-Y_2)$$

$$= 2(X_2-X_1)(Y_2-Y_1)$$

NOTE that I have a very big typo here in what I asked you to prove. I was too excited to share this result and I wanted to reduce the amount of work by excluding the denominator of the slope. Eventually, those constants 1/2 and 2 would get canceled out. The correct equality is

$$\sum_{t=1}^{n}(X_t-\bar{X})(Y_t-\bar{Y}) = \boxed{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_j-X_i)(Y_j-Y_i)$$

The missing part.

Technical Exercise 2

① slope of regression of Y on X_1 $= \dfrac{\sum_{t=1}^{n}(X_{1t}-\bar{X}_1)(Y_t-\bar{Y})}{\sum_{t=1}^{n}(X_{1t}-\bar{X}_1)^2}$

② slope of regression of Z on W $= \dfrac{\sum_{t=1}^{n}(W_t-\bar{W})(Z_t-\bar{Z})}{\sum_{t=1}^{n}(W_t-\bar{W})^2}$

Tech Exer 2
Ex Set 02

③ observe that $\sum_{t=1}^{n}(W_t-\bar{W})(Z_t-\bar{Z}) = \sum_{t=1}^{n}\left(aX_{1t}+b-(a\bar{X}_1+b)\right)\left(cY_t+d-(c\bar{Y}+d)\right)$

$$= \sum_{t=1}^{n} a(X_{1t}-\bar{X}_1)c(Y_t-\bar{Y})$$

$$= ac\sum_{t=1}^{n}(X_{1t}-\bar{X}_1)(Y_t-\bar{Y})$$

Next, $\sum_{t=1}^{n}(W_t-\bar{W})^2 = \sum_{t=1}^{n}\left(aX_{1t}+b-(a\bar{X}+b)\right)^2 = \sum_{t=1}^{n}a^2(X_{1t}-\bar{X}_1)^2 = a^2\sum_{t=1}^{n}(X_{1t}-\bar{X}_1)^2$

Therefore,

slope of regression of Z on W $= \dfrac{ac\sum_{t=1}^{n}(X_{1t}-\bar{X}_1)(Y_t-\bar{Y})}{a^2\sum_{t=1}^{n}(X_{1t}-\bar{X}_1)^2} = \dfrac{c}{a}\cdot$ slope of regression of Y on X_1

Technical Exercise 3

(1) Observe that $\sum_{t=1}^{n}(Y_t-\hat\beta_1 X_{1t})^2=\sum_{t=1}^{n}\left(Y_t^2-2\hat\beta_1 Y_t X_{1t}+\hat\beta_1^2 X_{1t}^2\right)$

$$=\sum_{t=1}^{n}Y_t^2-2\hat\beta_1\sum_{t=1}^{n}Y_tX_{1t}+\hat\beta_1^2\sum_{t=1}^{n}X_{1t}^2$$

The first-order condition is

$$\frac{d}{d\hat\beta_1}\left[\sum_{t=1}^{n}Y_t^2-2\hat\beta_1\sum_{t=1}^{n}Y_tX_{1t}+\hat\beta_1^2\sum_{t=1}^{n}X_{1t}^2\right]=0$$

$$-2\sum_{t=1}^{n}Y_tX_{1t}+2\hat\beta_1\sum_{t=1}^{n}X_{1t}^2=0$$

$$\hat\beta_1=\frac{\sum_{t=1}^{n}X_{1t}Y_t}{\sum_{t=1}^{n}X_{1t}^2}.$$

The second-order condition is given by

$$\frac{d}{d\hat\beta_1}\left[-2\sum_{t=1}^{n}Y_tX_{1t}+2\hat\beta_1\sum_{t=1}^{n}X_{1t}^2\right]=2\sum_{t=1}^{n}X_{1t}^2.$$

$\hat\beta_1$ will be optimal (minimize sum of squared residuals) if $\sum_{t=1}^{n}X_{1t}^2>0$ which is always true (Sum of positive numbers is positive) with one exception. The exception happens when all observations of X_1 are all equal to zero meaning $X_{1t} = 0$ for all $t = 1, \dots, n$. But this would never happen because $\hat{\beta}_1$ will not even exist in this situation.

(2) The residuals are $e_t = Y_t - \hat{\beta}_1 X_{1t} = Y_t - \boxed{\left(\frac{\sum_{t=1}^{n} X_{1t} Y_t}{\sum_{t=1}^{n} X_{1t}^2} \right)} X_{1t}$.

→ really an expression which no longer depends on t

Taking averages, we have

$$\frac{1}{n} \sum_{t=1}^{n} e_t = \frac{1}{n} \sum_{t=1}^{n} Y_t - \left(\frac{\sum_{t=1}^{n} Y_t X_{1t}}{\sum_{t=1}^{n} X_{1t}^2} \right) \frac{1}{n} \sum_{t=1}^{n} X_{1t}$$

$$= \bar{Y} - \left(\frac{\sum_{t=1}^{n} Y_t X_{1t}}{\sum_{t=1}^{n} X_{1t}^2} \right) \bar{X}_1$$

There is no way to guarantee that the average of the residuals is equal to zero. It will only happen when $\bar{Y}$ and $\bar{X}_1$ are <u>both</u> equal to zero.