Exercise 1.

Task 1.1. *It is equal.*

Task 1.2.

$$\hat{x} = \left(\sum \frac{1}{\sigma_i^2}\right)^{-1} \sum \frac{y_i}{\sigma_i^2}$$

Exercise 2.

Task 2.1.

$$\mathbf{G} = \begin{bmatrix} 1 & T & T^2 \\ 1 & 2T & (2T)^2 \\ \vdots & \vdots & \vdots \\ 1 & NT & (NT)^2 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^\top, \qquad \mathbf{r} = \begin{bmatrix} r_1 & \cdots & r_N \end{bmatrix}^\top$$

Task 2.2. *Just simulation.*

Exercise 3.

Task 3.1. *Yes, but proof it!.*

Task 3.2.

$$\mathbb{E}[y_i y_j] = \begin{cases} x^2, & i \neq j, \\ x^2 + \sigma^2, & i = j. \end{cases}$$

Task 3.3. *It is a biased estimate.*

$$\begin{aligned} \mathbb{E}\hat{\sigma^2} =& \mathbb{E}\frac{1}{N}\sum_i (y_i - \hat{x})^2 = \mathbb{E}\frac{1}{N}\sum_i (\frac{1}{n}\sum_j (y_i - y_j))^2 \\ =& \mathbb{E}\frac{1}{N}\sum_i (\frac{1}{N}\sum_j (r_i - r_j))^2 \\ =& \mathbb{E}\frac{1}{N}\sum_i \frac{1}{N^2}\sum_j \sum_{j'} (r_i - r_j)(r_i - r_{j'}) \\ =& \frac{1}{N}\sum_i \frac{1}{N^2}\sum_j \sum_{j'} \mathbb{E}r_i^2 - \mathbb{E}r_i r_j - \mathbb{E}r_i r_{j'} + \mathbb{E}r_j r_{j'} \\ =& \frac{1}{N}\sum_i \frac{1}{N^2}\left(N^2\sigma^2 - 2N\sigma^2 + N\sigma^2\right) \\ =& \frac{\sigma^2(N-1)}{N} \end{aligned}$$

Exercise 4. Stacking the euqations over n gives

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_N \end{bmatrix} \mathbf{x} + \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}. \tag{1}$$

Therefore we have

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \ \mathbf{G} = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_N \end{bmatrix}, \ \mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix}. \tag{2}$$

Since r_n are an i.i.d sequence the covariance matrix of $\mathbf{r}$ is block diagonal with each block having the entry R. That is,

$$\mathbf{R} = \begin{bmatrix} R & 0 & \dots & \dots & 0 \\ 0 & R & 0 & \dots & 0 \\ \vdots & 0 & \ddots & & \vdots \end{bmatrix}. \tag{3}$$

From the previous result we have

$$\mathbf{r} = \mathbf{Y} - \mathbf{G}\mathbf{x}, \tag{4}$$

hence

$$\|\mathbf{r}\|^2_{\mathbf{R}^{-1}} = (\mathbf{Y} - \mathbf{G}\mathbf{x})^\mathsf{T}\mathbf{R}^{-1}(\mathbf{Y} - \mathbf{G}\mathbf{x}). \tag{5}$$

Differentiation with respect to $\mathbf{x}$ gives the gradient as

$$\nabla_\mathbf{x}\|\mathbf{r}\|^2_{\mathbf{R}^{-1}} = -2\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}(\mathbf{Y} - \mathbf{G}\mathbf{x}). \tag{6}$$

Setting the gradient to zero and solving for $\mathbf{x}$ gives the least squares estimator as

$$\hat{\mathbf{x}} = [\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}]^{-1}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{Y} \tag{7}$$

Substituting $\mathbf{Y}$ for $\mathbf{G}\mathbf{x} + \mathbf{r}$ gives

$$\begin{aligned} \hat{\mathbf{x}} &= [\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}]^{-1}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}(\mathbf{G}\mathbf{x} + \mathbf{r}) \\ &= \mathbf{x} + [\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}]^{-1}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{r}. \end{aligned} \tag{8}$$

Using the standard rules for expectation and covariance matrix, and the statistics of $\mathbf{r}$ gives

$$\mathbb{E}[\hat{\mathbf{x}}] = \mathbf{x}, \tag{9a}$$

$$\begin{aligned} \mathbb{V}[\hat{\mathbf{x}}] &= [\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}]^{-1}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbb{V}[\mathbf{r}][[\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}]^{-1}\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}]^\mathsf{T} \\ &= [\mathbf{G}^\mathsf{T}\mathbf{R}^{-1}\mathbf{G}]^{-1}. \end{aligned} \tag{9b}$$

Exercise 5. Writing the equations in matrix form gives

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{bmatrix} \mathbf{x} + \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}. \tag{10}$$

Therefore, the model matrices for the least squares problem using the k first measurements are given by

$$\mathbf{Y}_k = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} \tag{11}$$

$$\mathbf{G}_k = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{bmatrix} \tag{12}$$

$$\mathbf{r}_k = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_k \end{bmatrix}. \tag{13}$$

The least squares estimator have already been computed in the previous exercise and is give by (for the k first measurements):

$$\hat{\mathbf{x}}_k = [\mathbf{G}_k^\mathsf{T}\mathbf{G}_k]^{-1}\mathbf{G}_k^\mathsf{T}\mathbf{Y}_k. \tag{14}$$

Writing down $\mathbf{G}_k^\mathsf{T}\mathbf{G}_k$ according to how $\mathbf{G}_k$ is defined gives

$$\begin{aligned} \mathbf{G}_k^\mathsf{T}\mathbf{G}_k &= \begin{bmatrix} G_1^\mathsf{T} & G_2^\mathsf{T} & \dots & G_k^\mathsf{T} \end{bmatrix} \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{bmatrix} \\ &= \sum_{l=1}^{k} G_l^\mathsf{T} G_l \\ &= \sum_{l=1}^{k-1} G_l^\mathsf{T} G_l + G_k^\mathsf{T} G_k. \end{aligned} \tag{15}$$

Similarly, for $\mathbf{G}_k^\mathsf{T}\mathbf{Y}_k$ we have

$$\begin{aligned} \mathbf{G}_k^\mathsf{T}\mathbf{Y}_k &= \begin{bmatrix} G_1^\mathsf{T} & G_2^\mathsf{T} & \dots & G_k^\mathsf{T} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} \\ &= \sum_{l=1}^{k} G_l^\mathsf{T} y_l \\ &= \sum_{l=1}^{k-1} G_l^\mathsf{T} y_l + G_k^\mathsf{T} y_k. \end{aligned} \tag{16}$$

Therefore,

$$U_k = G_k^\mathsf{T} G_k, \tag{17}$$

$$u_k = G_k, \tag{18}$$

$$\xi_k = G_k^\mathsf{T} y_k. \tag{19}$$

Using the expression for $\hat{\mathbf{x}}_k$, the previously derived recursion, and the definition of P_k we get

$$\begin{aligned}\hat{\mathbf{x}}_k &= [\mathbf{G}_k^\top \mathbf{G}_k]^{-1}\mathbf{G}_k^\top \mathbf{Y}_k \\ &= [\mathbf{G}_{k-1}^\top \mathbf{G}_{k-1} + U_k]^{-1}(\mathbf{G}_{k-1}^\top \mathbf{Y}_{k-1} + \xi_k) \\ &= [\mathbf{G}_{k-1}^\top \mathbf{G}_{k-1} + G_k^\top G_k]^{-1}(\mathbf{G}_{k-1}^\top \mathbf{Y}_{k-1} + G_k^\top y_k) \\ &= [P_{k-1}^{-1} + G_k^\top G_k]^{-1}(\mathbf{G}_{k-1}^\top \mathbf{Y}_{k-1} + G_k^\top y_k).\end{aligned} \tag{20}$$

The matrix inversion lemma states that:

$$(A + CBC^\top)^{-1} = A^{-1} - A^{-1}C(B + C^\top A^{-1}C)^{-1}C^\top A^{-1}. \tag{21}$$

Taking $A = P_{k-1}^{-1}$, $B = 1$, and $C = G_k^\top$ gives:

$$[P_{k-1}^{-1} + G_k^\top G_k]^{-1} = P_{k-1} - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k P_{k-1}. \tag{22}$$

Furthermore, using the defition of $\hat{\mathbf{x}}_k$ gives:

$$P_{k-1}\mathbf{G}_{k-1}^\top \mathbf{Y}_{k-1} = \hat{\mathbf{x}}_{k-1}. \tag{23}$$

Making these substitutions in the expression for $\hat{\mathbf{x}}_k$ gives

$$\begin{aligned}\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k-1} - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k \hat{\mathbf{x}}_{k-1} \\ &\quad + P_{k-1}G_k^\top y_k - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k P_{k-1} G_k^\top y_k \\ &= \hat{\mathbf{x}}_{k-1} - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k \mathbf{x}_{k-1} \\ &\quad + P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} (G_k P_{k-1} G_k^\top + 1) y_k \\ &\quad - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k P_{k-1} G_k^\top y_k \\ &= \hat{\mathbf{x}}_{k-1} - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k \hat{\mathbf{x}}_{k-1} \\ &\quad + P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} \Big((G_k P_{k-1} G_k^\top + 1) y_k - G_k P_{k-1} G_k^\top y_k\Big) \\ &= \hat{\mathbf{x}}_{k-1} - P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} G_k \hat{\mathbf{x}}_{k-1} \\ &\quad + P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} y_k \\ &= \hat{\mathbf{x}}_{k-1} + P_{k-1}G_k^\top (G_k P_{k-1} G_k^\top + 1)^{-1} (y_k - G_k \hat{\mathbf{x}}_{k-1}).\end{aligned} \tag{24}$$

Note that the least squares estimate at $k = 1$ is given by

$$\hat{\mathbf{x}}_1 = [G_1^\top G_1]^{-1} G_1^\top y_1. \tag{25}$$

However, in this exercise $G_1^\top G_1$ is not invertible, hence the recursion needs at the normal least squares estimate for some $k > 1$, (in this case $k = 2$).

The rest is just coding.