# 1 Static Linear Models

**Exercise 1.** From Exercise 1 at the previous session, we know that a resistor with unknown resistance $x$ was found at Sensor Fusion lab. A student then tried to estimate what was the value of $x$ from a set of $k$ noisy multimeter measurements.

**Task 1.1.** *If the student intuitively estimate the resistance as an average of k measurements, what will be the difference between his result and the least square estimate?*

**Task 1.2.** *Consider Exercise 2 from the previous session. What will be the closed form of the weighted least square estimate of $x$?*

**Exercise 2.** Suppose that we know from some theories that measurement records that we have collected are quadratic function of time. That is

$$y_k = f(x_1, x_2, x_3, kT) + r_k,$$

where $x_1, x_2, x_3$ are the unknown variables, $k$ is the time index, $T$ is the sampling time, and $r_k$ is the measurement noise which is assumed to be $\mathscr{N}(0, \sigma^2)$.

**Task 2.1.** *If we collected $N$ measurements, construct the measurement equations in the matrix form*

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{r}.$$

**Task 2.2** (Simulation Exercise)**.** *Let $x_1, x_2 = 0$, $x_3 = 9.8$, and $T = 0.01$. Generate a set of data from $t = 0$ up to $t = 1$ with variance $\sigma^2 = 1e-2$. Using the result of Task 2.1 compute the least square estimate of $x_1, x_2, x_3$. Now reduce the variance $r$ to half of the original value. Generate again the data with the new variance. Using the weighted cost function:*

$$J = \sum_{k=1}^{N} (y_k - f(\hat{x}))\lambda^{N-k},$$

*compute the estimate of the unknown variables. Select $\lambda \in (0, 1)$, and compare the estimate $\hat{\mathbf{x}}$ with the estimate obtained from the least square method.*

**Exercise 3.** Suppose that $\{y_1, \cdots, y_N\}$ is a set of measurements of the unknown resistor in Task 1.1. Let us assume that the expectation of $\mathbb{E}[(y_i - x)(y_j - x)] = 0$ for $i \neq j$. Since we dont know what $x$ and $\sigma^2$ are, we estimate these two quantities as follows:

$$\hat{x} = \frac{1}{n} \sum_{i=1}^{N} y_i,$$
$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{N} (y_i - \hat{x})^2.$$

**Task 3.1.** *Is $\hat{x}$ an unbiased estimate of $x$? I.e., is $\mathbb{E}(\hat{x}) = x$?*

**Task 3.2.** *Find $\mathbb{E}(y_i y_j)$ in terms of $x$ and $\sigma^2$ for $i = j$ and $i \neq j$.*

**Task 3.3.** *Is $\hat{\sigma}^2$ an unbiased estimate of $\sigma^2$? If not, how should we change $\hat{\sigma}^2$ to make it an unbiased estimate?*

**Exercise 4.**
Let $\mathbf{y}_n \in \mathbb{R}^d$, $\mathbf{G}_n \in \mathbb{R}^{d\times d}$, $\mathbf{x} \in \mathbb{R}^d$ and $\mathbf{r}_n$ be i.i.d $d$-dimensional random vector with mean $\mathbb{E}[r_n] = 0$ and covariance matrix $\mathbb{V}[\mathbf{r}_n] = \mathbb{E}[(\mathbf{r}_n - \mathbb{E}[\mathbf{r}_n])(\mathbf{r}_n - \mathbb{E}[\mathbf{r}_n])^\top] = \mathbf{R}$, where $n = 1, \ldots, N$. In this exercise the following model will be considered:

$$\mathbf{y}_n = \mathbf{G}_n\mathbf{x} + \mathbf{r}_n \tag{1}$$

**Task 4.1** (Matrix formulaton of least squares)**.** *Write the system Equation* (1) *in matrix form, that is*

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{r}. \tag{2}$$

*How do the elemnts of* $\mathbf{y}$ *and* $\mathbf{r}$ *look? How does the rows in* $\mathbf{G}$ *look? What is the structure of the covariance matrix,* $\mathbf{R}$*, of* $\mathbf{r}$*?*

**Task 4.2** (Least squares estimation)**.** *Write down an expression of the weighted (Euclidean) norm* $\|\mathbf{r}\|^2_{\mathbf{R}^{-1}}$ *in terms of just* $\mathbf{G}$*,* $\mathbf{x}$*, and* $\mathbf{y}$*. Find the least squares estimator of* $\mathbf{x}$*,* $\hat{\mathbf{x}}$ *by computing the gradient of* $\|\mathbf{r}\|^2_{\mathbf{R}^{-1}}$ *with respect to* $\mathbf{x}$ *and solving for zero.*

**Task 4.3** (Examining the statistics of the least squares estimator)**.** $\hat{\mathbf{x}}$ *is a random variable as it depends* $\mathbf{y}$ *which in turn depends on the stochastic vector* $\mathbf{r}$*. Compute the expected value and covariance matrix of* $\hat{\mathbf{x}}$*.*

**Exercise 5.** Consider the linear regression model:

$$y_n = \mathbf{G}_n x + r_n, \quad n = 1, \ldots, N, \tag{3}$$

where $y_n \in \mathbb{R}$, $r_n$ is a Gaussian random variable with zero mean and variance $\sigma^2$, $\mathbf{G}_n = \begin{bmatrix} 1 & n/N \end{bmatrix}$, $\mathbf{x}^\mathsf{T} = \begin{bmatrix} -1 & 0.2 \end{bmatrix}$, and $N = 2^{10}$.

**Task 5.1.** *Consider least squares estimation using only the first* $k$ *measurements of Equation* (3) *(*$1 \leq k \leq N$*). Write down the model matrices* $\mathbf{y}_k$*,* $\mathbf{G}_k$ *for solving the least squares problem using only the* $k$ *first measurements and write down the least squares estimator of* $\mathbf{x}$*.*

**Task 5.2.** *Let* $\mathbf{y}_k$ *and* $\mathbf{G}_k$ *be defined as in the previous exercise. Find a recursive expression for* $\mathbf{G}_k^\mathsf{T}\mathbf{G}_k$ *and* $\mathbf{G}_k^\mathsf{T}\mathbf{y}_k$*, according to*

$$\mathbf{G}_k^\mathsf{T}\mathbf{G}_k = \mathbf{G}_{k-1}^\mathsf{T}\mathbf{G}_{k-1} + \mathbf{U}_k, \tag{4a}$$
$$\mathbf{G}_k^\mathsf{T}\mathbf{y}_k = \mathbf{G}_{k-1}^\mathsf{T}\mathbf{y}_{k-1} + \boldsymbol{\xi}_k. \tag{4b}$$

*That is, find* $\mathbf{U}_k$ *and* $\boldsymbol{\xi}_k$*.*

**Task 5.3.** *The matrix* $\mathbf{U}_k$ *from the previous task can be written as an outer product:*

$$\mathbf{U}_k = u_k u_k^\mathsf{T}. \tag{5}$$

*Give an expression for* $u_k$*.*

**Task 5.4.** *Let* $\mathbf{P}_k^{-1} = \mathbf{G}_k^\mathsf{T}\mathbf{G}_k$*. Use the results of the previous tasks to derive an expression for a recursive estimator,* $\hat{x}_k$*, of* $x$*. Hint:* $(\mathbf{P}_{k-1}^{-1} + u_k u_k^\mathsf{T})^{-1}$ *can be computed using the Woodbury identity (also known as Matrix inversion lemma). Google matrix cookbook for formula. How do you initialise the recursion? Any problems?*

**Task 5.5.** *Simuate from the model in Equation* (3)*. Implement the least squares estimator and the recursive least squares estimator for the* $k$ *first samples,* $k = 1, \ldots, N$*. Is there any difference? Also compare them in terms of computational speed, what's the conclusion?*