HW 2

ARNAB DEY
Student ID: 5563169
Email: dey00011@umn.edu

Problem 1

Problem 2.5 The probability density function of an exponential random variable, X, is

$$f_X(x) = \begin{cases} ae^{-ax}, & x \geq 0 \\ 0 & x < 0, \end{cases}$$

where $a \geq 0$.

(a) Probability distribution function of the exponentially distributed random variable, X, is:

$$F_X(x) = \int_{-\infty}^{x} f_X(z)dz$$

case 1: $x < 0$:

$$F_X(x) = \int_{-\infty}^{x} 0dz = 0.$$

case 2: $x \geq 0$:

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(z)dz \\ &= \int_{-\infty}^{0} 0dz + \int_{0}^{x} ae^{-az}dz \\ &= -e^{-az}\Big|_{0}^{x} \\ &= 1 - e^{-ax}. \end{aligned}$$

Therefore,

$$F_X(x) = \begin{cases} 1 - e^{-ax}, & x \geq 0 \\ 0 & x < 0. \end{cases}$$

(b) The mean of X is given by (assuming $a > 0$),

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} zf_X(z)dz \\ &= a\int_{0}^{\infty} ze^{-az}dz \\ &= a\left[\left[-\frac{z}{a}e^{-az}\right]_0^{\infty} + \frac{1}{a}\int_0^{\infty} e^{-az}dz\right] \\ &= -\frac{1}{a}\left[e^{-az}\right]_0^{\infty} \\ &= \frac{1}{a}. \end{aligned} \tag{1}$$

(c) The second moment of X is,

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} z^2 f_X(z)dz \\
&= a\int_0^{\infty} z^2 e^{-az}dz \\
&= a\left[\left[-\frac{z^2}{a}e^{-az}\right]_0^{\infty} + \frac{2}{a}\int_0^{\infty} ze^{-az}dz\right] \\
&= a\left[\frac{2}{a}\frac{1}{a^2}\right] \qquad [\text{ from } (1)] \\
&= \frac{2}{a^2}. \qquad (2)
\end{aligned}$$

(d) The variance of X is,

$$\begin{aligned}
\mathbb{E}[(X-\mathbb{E}[X])^2] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
&= \frac{2}{a^2} - \frac{1}{a^2} \qquad [\text{ from } (1) \text{ and } (2)] \qquad (3) \\
&= \frac{1}{a^2}. \qquad (4)
\end{aligned}$$

(e) The probability that X takes on a value within one standard deviation of its mean is given by,

$$\begin{aligned}
P(\frac{1}{a} - \frac{1}{a} \leq X \leq \frac{1}{a} + \frac{1}{a}) &= P(0 \leq X \leq \frac{2}{a}) \\
&= F_X\left(\frac{2}{a}\right) - F_X(0) \\
&= 1 - e^{-2} \\
&= 0.865.
\end{aligned}$$

Problem 2

For a valid joint pdf, it needs to satisfy,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(u,z)dudz = 1.$$

Substituting the expression of $f_{XY}(x,y)$ and assuming $a > 0$, we get,

$$\begin{aligned}
&\int_0^{\infty}\int_0^{\infty} ae^{-2u}e^{-3z}dudz = 1 \\
\implies &\int_0^{\infty} e^{-2u}\left[\int_0^{\infty} e^{-3z}dz\right]du = \frac{1}{a} \\
\implies &\int_0^{\infty} e^{-2u}\left[-\frac{1}{3}e^{-3z}\right]_0^{\infty} du = \frac{1}{a} \\
\implies &\frac{1}{3}\int_0^{\infty} e^{-2u}du = \frac{1}{a} \\
\implies &\frac{1}{3}\left[-\frac{1}{2}e^{-2u}\right]_0^{\infty} = \frac{1}{a} \\
\implies &\frac{1}{3}\frac{1}{2} = \frac{1}{a} \\
\implies &a = 6.
\end{aligned}$$

(b) First let us calculate marginal pdf of X which is given by,

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{XY}(x,z)dz \\
&= 6e^{-2x}\int_0^{\infty} e^{-3z}dz \\
&= 6e^{-2x}\left[-\frac{1}{3}e^{-3z}\right]_0^{\infty} \\
&= 6e^{-2x}\frac{1}{3} \\
&= 2e^{-2x}, \qquad (5)
\end{aligned}$$

if $x > 0$ and 0 otherwise. Therefore, we can see that $X \sim exp(2)$, *i.e.* an exponential random variable with parameter value equals to 2. Now, to find $\bar{x}$, we can compute the expectation of X as follows,

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} uf_X(u)du \\
&= 2\int_0^{\infty} ue^{-2u}du \\
&= \frac{1}{2} \qquad [\text{ from (1)}]. \qquad (6)
\end{aligned}$$

Similarly, we can find the marginal pdf of Y as follows:

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{XY}(u,y)du \\
&= 6e^{-3y}\int_0^{\infty} e^{-2u}du \\
&= 6e^{-3y}\left[-\frac{1}{2}e^{-2u}\right]_0^{\infty} \\
&= 6e^{-3y}\frac{1}{2} \\
&= 3e^{-3y}, \qquad (7)
\end{aligned}$$

if $y > 0$ and 0 otherwise. Therefore, we can see that $Y \sim exp(3)$, *i.e.* an exponential random variable with parameter value equals to 3. Now, to find $\bar{y}$, we can compute the expectation of Y as follows,

$$\begin{aligned}
\mathbb{E}[Y] &= \int_{-\infty}^{\infty} zf_Y(z)dz \\
&= 3\int_0^{\infty} ze^{-3z}dz \\
&= \frac{1}{3} \qquad [\text{ from (1)}]. \qquad (8)
\end{aligned}$$

(c) The second moment of X is given by,

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} u^2 f_X(u)du \\
&= \frac{2}{2^2} \qquad [\text{ from (2)}] \\
&= \frac{1}{2}. \qquad (9)
\end{aligned}$$

Similarly, the second moment of Y is given by,

$$\mathbb{E}[Y^2] = \int_{-\infty}^{\infty} z^2 f_Y(z)dz$$

$$
\begin{aligned}
&= \frac{2}{3^2} \qquad [\text{ from } (2)] \\
&= \frac{2}{9}. \qquad (10)
\end{aligned}
$$

Now, from (5) and (7), we can see that,

$$f_{XY}(x,y) = f_X(x) f_Y(y). \qquad (11)$$

Therefore, by definition, X and Y are independent. Hence,

$$
\begin{aligned}
\mathbb{E}[XY] &= \mathbb{E}[X]\mathbb{E}[Y] \\
&= \frac{1}{2}\frac{1}{3} \\
&= \frac{1}{6}. \qquad (12)
\end{aligned}
$$

(d) The autocorrelation matrix of the random vector $[X\ Y]^T$ is given by,

$$
\begin{aligned}
\mathbb{E}[[X\ Y]^T[X\ Y]] &= \begin{bmatrix} \mathbb{E}[X^2] & \mathbb{E}[XY] \\ \mathbb{E}[YX] & \mathbb{E}[Y^2] \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{2} & \frac{1}{6} \\ \frac{1}{6} & \frac{2}{9} \end{bmatrix}. \qquad (13)
\end{aligned}
$$

(e) The variance of X is given by,

$$
\begin{aligned}
\mathbb{E}[(X - \mathbb{E}[X])^2] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\
&= \frac{1}{2} - \frac{1}{4} \\
&= \frac{1}{4}. \qquad (14)
\end{aligned}
$$

The variance of Y is given by,

$$
\begin{aligned}
\mathbb{E}[(Y - \mathbb{E}[Y])^2] &= \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\
&= \frac{2}{9} - \frac{1}{9} \\
&= \frac{1}{9}. \qquad (15)
\end{aligned}
$$

The covariance of X and Y is given by,

$$
\begin{aligned}
C_{XY} &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\
&= \mathbb{E}[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] - \mathbb{E}[X]\mathbb{E}[Y]] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\
&= \frac{1}{6} - \frac{1}{2}\frac{1}{3} \\
&= 0. \qquad (16)
\end{aligned}
$$

(f) The autocovariance of the random vector $[X\ Y]^T$ is given by,

$$
\begin{aligned}
\mathbb{E}[[(X - \mathbb{E}[X])\ (Y - \mathbb{E}[Y])]^T[(X - \mathbb{E}[X])\ (Y - \mathbb{E}[Y])]] &= \begin{bmatrix} \mathbb{E}[(X - \mathbb{E}[X])^2] & \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ \mathbb{E}[(Y - \mathbb{E}[Y])(X - \mathbb{E}[]X)] & \mathbb{E}[(Y - \mathbb{E}[Y])^2] \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{9} \end{bmatrix}. \qquad (17)
\end{aligned}
$$

(g) The correlation coefficient between X and Y is

$$\rho = \frac{C_{XY}}{\sigma_x \sigma_y} = 0.$$

(h) Yes, X and Y are independent as we have already derived in the previous sections and shown that $f_{XY}(x, y) = f_X(x) f_Y(y)$ in (11).

Problem 3

Total number of residents, $|M \cap I| + |M \cap C| + |F \cap I| + |F \cap C| = 340454 + 40854 + 359881 + 172742 = 913931$.
Total number of male residents, $|M| = |M \cap I| + |M \cap C| = 340454 + 40854 = 381308$.
Total number of female residents, $|F| = |F \cap I| + |F \cap C| = 359881 + 172742 = 532623$.
Therefore,

$$\begin{aligned}
P(M) &= \frac{381308}{913931} = 0.417 \\
P(F) &= \frac{532623}{913931} = 0.583 \\
P(C) &= \frac{|M \cap C| + |F \cap C|}{913931} = \frac{40854 + 172742}{913931} = 0.234 \\
P(I) &= \frac{|M \cap I| + |F \cap I|}{913931} = \frac{340454 + 359881}{913931} = 0.766 \\
P(M \cap I) &= \frac{|M \cap I|}{913931} = \frac{340454}{913931} = 0.373 \\
P(F \cap I) &= \frac{|F \cap I|}{913931} = \frac{359881}{913931} = 0.394 \\
P(M \cap C) &= \frac{|M \cap C|}{913931} = \frac{40854}{913931} = 0.045 \\
P(F \cap C) &= \frac{|F \cap C|}{913931} = \frac{172742}{913931} = 0.19.
\end{aligned}$$

(a) Now, $P(M|C) = \frac{P(M \cap C)}{P(C)} = \frac{0.045}{0.234} = 0.192$. Therefore, the probability that a male resident of the island picked at random will be a convicted felon is,

$$\begin{aligned}
P(C|M) &= \frac{P(M|C)P(C)}{P(M)} \\
&= \frac{0.192 \times 0.234}{0.417} \\
&= 0.108.
\end{aligned}$$

Note that we can obtain the same result from the given relative frequencies as $P(C|M) = \frac{40854}{40854+340454} = 0.107$.

(b) $P(F|C) = \frac{P(F \cap C)}{P(C)} = \frac{0.19}{0.234} = 0.812$. Therefore, the probability that a female resident of the island picked at random will be a convicted felon is

$$\begin{aligned}
P(C|F) &= \frac{P(F|C)P(C)}{P(F)} \\
&= \frac{0.812 \times 0.234}{0.583} \\
&= 0.326.
\end{aligned}$$

Note that we can obtain the same result from the given relative frequencies as $P(C|F) = \frac{172742}{172742+359881} = 0.324$.

(c)

$$
\begin{aligned}
P(F|C) &= \frac{P(C|F)P(F)}{P(C)} \\
&= \frac{0.326 \times 0.583}{0.234} \\
&= 0.81.
\end{aligned}
$$

Note that we can obtain the same result from the given relative frequencies as $P(F|C) = \frac{172742}{172742+40854} = 0.81$.

(d)

$$
\begin{aligned}
P(M|C) &= \frac{P(C|M)P(M)}{P(C)} \\
&= \frac{0.108 \times 0.417}{0.234} \\
&= 0.19.
\end{aligned}
$$

Note that we can obtain the same result from the given relative frequencies as $P(M|C) = \frac{40854}{40854+172742} = 0.19$.

Problem 4

I propose to optimally combine the present radar system with a new radar system that has a measurement variance of 6. The reason is explained in the next paragraphs.

To address this problem, either we can use iterative least square or batch weighted least square and both leads to the same solution. I am using batch weighted least square where I assume that all the data from each sensor are available and processed simultaneously. Let the measurements $y \in \mathbb{R}^m$ coming from different sensors and true position $x \in \mathbb{R}^n$ are related as follows:

$$
y = Hx + e,
$$

where, $H \in \mathbb{R}^{m \times n}$ is the measurement matrix and $e \in \mathbb{R}^m$ is error matrix associated with m measurements. Let the error covariance matrix is,

$$
R = \mathbb{E}[ee^T] = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \sigma_m^2 \end{bmatrix}.
$$

We know that weighted least square estimate of x is:

$$
\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} y.
$$

The error associated with estimation of true position is

$$
\begin{aligned}
\epsilon_x &= (x - \hat{x}) \\
&= x - (H^T R^{-1} H)^{-1} H^T R^{-1} y \\
&= x - (H^T R^{-1} H)^{-1} H^T R^{-1} (Hx + e) \\
&= x - x - (H^T R^{-1} H)^{-1} H^T R^{-1} e \\
&= -(H^T R^{-1} H)^{-1} H^T R^{-1} e.
\end{aligned}
$$

Therefore, we can find the variance of estimation error as follows:

$$
\begin{aligned}
\mathbb{E}[\epsilon_x \epsilon_x^T] &= (H^T R^{-1} H)^{-1} H^T R^{-1} \mathbb{E}[ee^T] R^{-1} H (H^T R^{-1} H)^{-1} && (18) \\
&= (H^T R^{-1} H)^{-1} H^T R^{-1} H (H^T R^{-1} H)^{-1} && (19) \\
&= (H^T R^{-1} H)^{-1}. && (20)
\end{aligned}
$$

Case 1: When we have only one sensor with measure variance of $\sigma_1^2 = 10$, we have the following model

$$y = \underbrace{[1]}_{H} x + e$$

and $R = [10]$. From (18),

$$\mathbb{E}[\epsilon_x \epsilon_x^T] = 10.$$

Case 2: When we combine a sensor having measurement variance of $\sigma_2^2 = 6$, with the current sensor with $\sigma_1^2 = 10$, we have the following model

$$y = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{H} x + e$$

and $R = \begin{bmatrix} 10 & 0 \\ 0 & 6 \end{bmatrix}$. From (18),

$$\begin{aligned}
\mathbb{E}[\epsilon_x \epsilon_x^T] &= \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{10} & 0 \\ 0 & \frac{1}{6} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^{-1} \\
&= \left(\frac{1}{10} + \frac{1}{6} \right)^{-1} \\
&= 3.75.
\end{aligned}$$

Case 3: When we combine two sensors having same measurement variance of 10, as the original system, along with the original one, we have the following model

$$y = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{H} x + e$$

and $R = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 10 \end{bmatrix} = 10 I_3$. From (18),

$$\begin{aligned}
\mathbb{E}[\epsilon_x \epsilon_x^T] &= \left(\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \frac{1}{10} I_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right)^{-1} \\
&= \frac{10}{3} \\
&= 3.33.
\end{aligned}$$

Therefore, we can see that case 2 gives the least variance in error of position estimation. Hence, a good choice of sensors would be the combination of current sensor having measurement variance 10 with another with measurement variance of 6.

Problem 5

(a) The system dynamics given are as follows:

$$\begin{aligned}
\alpha_k &= F \alpha_{k-1} + G u_{k-1} \\
y_k &= \tilde{y} + v_k = C \alpha_k + v_k,
\end{aligned} \tag{21}$$

where,

$$F = \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix}, \; G = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

From (21),

$$\begin{aligned}
\tilde{y}_k &= C\alpha_k \\
&= C[F\alpha_{k-1} + Gu_{k-1}] \\
&= CF\alpha_{k-1} + CGu_{k-1} \\
&= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix} \alpha_{k-1} + \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u_{k-1} \\
&= \begin{bmatrix} -a_1 & 1 \end{bmatrix} \alpha_{k-1} + b_1 u_{k-1} \\
&= -a_1 \begin{bmatrix} 1 & 0 \end{bmatrix} \alpha_{k-1} + \begin{bmatrix} 0 & 1 \end{bmatrix} \alpha_{k-1} + b_1 u_{k-1} \\
&= -a_1 C\alpha_{k-1} + \begin{bmatrix} 0 & 1 \end{bmatrix} [F\alpha_{k-2} + Gu_{k-2}] + b_1 u_{k-1} \\
&= -a_1 \tilde{y}_{k-1} + \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} -a_1 & 1 \\ -a_2 & 0 \end{bmatrix} \alpha_{k-2} + \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u_{k-2} + b_1 u_{k-1} \\
&= -a_1 \tilde{y}_{k-1} + \begin{bmatrix} -a_2 & 0 \end{bmatrix} \alpha_{k-2} + b_2 u_{k-2} + b_1 u_{k-1} \\
&= -a_1 \tilde{y}_{k-1} - a_2 \tilde{y}_{k-2} + b_1 u_{k-1} + b_2 u_{k-2}.
\end{aligned} \tag{22}$$

Therefore, we can write it in the following matrix form

$$\tilde{y}_k = \underbrace{\begin{bmatrix} -\tilde{y}_{k-1} & -\tilde{y}_{k-2} & u_{k-1} & u_{k-2} \end{bmatrix}}_{H_k} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix}}_{x}. \tag{23}$$

Stacking all the H_k for all $k \in \{2, 3, \ldots, N\}$, we get

$$\underbrace{\begin{bmatrix} \tilde{y}_2 \\ \tilde{y}_3 \\ \vdots \\ \tilde{y}_N \end{bmatrix}}_{\tilde{y}} = \underbrace{\begin{bmatrix} -\tilde{y}_1 & -\tilde{y}_0 & u_1 & u_0 \\ -\tilde{y}_2 & -\tilde{y}_1 & u_2 & u_1 \\ \vdots & \vdots & \vdots & \vdots \\ -\tilde{y}_{N-1} & -\tilde{y}_{N-2} & u_{N-1} & u_{N-2} \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix}}_{x},$$

where N is the total number of measurements.

(b) Using the ordinary least squares (OLS) method, we get the following estimate of x:

$$\hat{x}_{OLS} = \begin{bmatrix} 0.5 \\ 0.25 \\ 1.0 \\ 0 \end{bmatrix} \tag{24}$$

I have also implemented recursive least squares (RLS) method and got the following estimate of x:

$$\hat{x}_{RLS} = \begin{bmatrix} 0.52 \\ 0.27 \\ 1.0 \\ 0.02 \end{bmatrix}. \tag{25}$$

(c), (d) After adding the noise with $\sigma = 0.1$ and running the OLS and RLS methods I got the following estimates of state vector x:

$$\hat{x}_{OLS} = \begin{bmatrix} 0.036 \\ -0.019 \\ 0.927 \\ -0.383 \end{bmatrix}, \; \hat{x}_{RLS} = \begin{bmatrix} 0.036 \\ -0.019 \\ 0.927 \\ -0.383 \end{bmatrix}. \tag{26}$$

(e) After adding the noise with $\sigma = 0.001$ and running the OLS and RLS methods I got the following estimates of state vector x:

$$\hat{x}_{OLS} = \begin{bmatrix} 0.475 \\ 0.234 \\ 0.996 \\ -0.021 \end{bmatrix}, \; \hat{x}_{RLS} = \begin{bmatrix} 0.476 \\ 0.235 \\ 0.996 \\ -0.020 \end{bmatrix}. \tag{27}$$

Therefore, we can see that the system is very sensitive to noise. Reducing the noise variance greatly improved the estimates to produce values closer to the estimate calculated in part (b) with 0 noise.

(f) To validate the accuracy of the estimates we can run cross-validation where we choose $N_{train} < N$ number of time index randomly from the available data and construct the $H_{train} \in \mathbb{R}^{N_{train} \times 4}$ matrix based on those time index and then use OLS / RLS to estimate $\hat{x}_{train}$ based on those randomly chosen time index. Then we can use the rest $N_{val} = N - N_{train}$ number of points from the available data and construct the $H_{val} \in \mathbb{R}^{N_{val} \times 4}$ and then predict $y_{val} \in \mathbb{R}^{N_{val}}$ based on H_{val} and $\hat{x}_{train}$. Now we can compute the root mean square error from y_{val} and y values from the original dataset. We can perform the operation for many times choosing different time index for train and test and then take the mean and standard deviation to show the confidence of the estimates.

I have used 10-fold cross validation for this purpose. Fig. 1 shows the validation error plot for three different noise standard deviation, $0.1, 0.01, 0.001$. The line connect the mean values of the root mean square error collected for each fold for each noise variance. x-axis denotes the noise standard deviation. Thus we

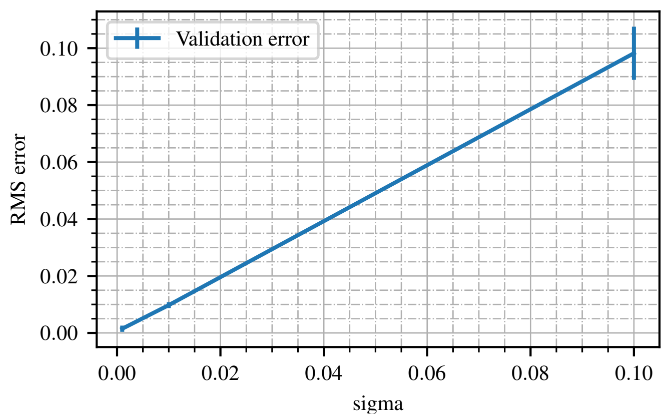

Figure 1: Q5.f: Validation error plot for $\sigma = 0.1, 0.01, 0.001$

can see that the root mean square error increases with increase in noise variance which indicates lower confidence in the estimates of x. Table 1 shows the RMS error for each folds of the cross validation and Table 2 shows the mean and standard deviation across all the folds for each noise standard deviation.

Git location of the code: The code file has been attached as a PDF file. Moreover, the code can be found at
`https://github.umn.edu/dey00011/EE5251_optimal_filter_estimation/tree/master/HW2/code`

Table 1: Q5.f: Error table for 10-fold cross validation

	fold 1	fold 2	fold 3	fold 4	fold 5	fold 6	fold 7	fold 8	fold 9	fold 10
$\sigma_{noise} = 0.001$	0.005	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
$\sigma_{noise} = 0.01$	0.012	0.009	0.011	0.009	0.01	0.01	0.008	0.011	0.008	0.009
$\sigma_{noise} = 0.1$	0.09	0.09	0.11	0.10	0.09	0.10	0.08	0.11	0.11	0.10

Table 2: Q5.f: Mean and Std of RMS error

σ_{noise}	Mean	Standard deviation
0.001	0.001	0.001
0.01	0.01	0.001
0.1	0.098	0.009