ASSIGNMENT - 1

1) (a) $f_{X,Y}(x,y) = \begin{cases} \dfrac{K e^{-\frac{x}{y}} e^{-y}}{y} & x>0,\ y>0 \\ 0 & \text{elsewhere} \end{cases}$

(i) $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$

$$= \int_0^{\infty}\int_0^{\infty} \frac{e^{-\frac{x}{y}} e^{-y}}{y} K\,dx\,dy = \int_0^{\infty} \frac{K \cdot e^{-y}}{y} \left. \frac{(-e^{-\frac{x}{y}})}{\frac{1}{y}} \right|_0^{\infty} dy$$

$$= \int_0^{\infty} K e^{-y} [0+1]\,dy.$$

$$= \int_0^{\infty} K e^{-y} dy = -K e^{-y}\Big|_0^{\infty} = \boxed{K=1}$$

$$\Rightarrow f_{X,Y}(x,y) = \begin{cases} \dfrac{e^{-\frac{x}{y}} e^{-y}}{y} & x>0,\ y>0 \\ 0 & \text{o.t} \end{cases}$$

(ii) $f_Y(Y) = \displaystyle\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$

$$= \int_0^{\infty} f_{X,Y}(x,y)\,dx = \cancel{K}\int_0^{\infty} \frac{e^{-\frac{x}{y}} e^{-y}}{y}\,dx$$

$$= \frac{e^{-y}}{y} \left. \frac{e^{-\frac{x}{y}}}{-\frac{1}{y}} \right|_{x=0}^{x=\infty}$$

$$f_Y(Y) = \frac{e^{-y}}{\cancel{y}}[0+1] = \begin{cases} e^{-y} & y>0 \\ 0 & \text{o.t} \end{cases}$$

(iii) $P(0<X<1;\ 0.2<Y<0.4) = \displaystyle\int_{0.2}^{0.4}\int_0^{1} f_{X,Y}(x,y)\,dx\,dy.$

Using Matlab numerical integration, we get :-

$P(0 < X < 0.1 ; 0.2 < Y < 0.4) = 0.1429.$

(iv) Conditional expectation : $E(X|Y)$

$$E(X|Y) = \int x f_{X|Y}(x)\,dx$$

$$= \int_{-\infty}^{\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)}\,dx$$

$$= \int_{0}^{\infty} x \frac{\frac{e^{-\frac{x}{y}}\cancel{e^{-y}}}{y}}{\cancel{e^{-y}}}\,dx$$

$$= \int_{0}^{\infty} \frac{x e^{-\frac{x}{y}}}{y}\,dx$$

Using Byparts ; $\underline{\underline{E(X|Y) = y}}.$

(b) X, Y are two joint RV with $\mu_x, \sigma_x, \mu_y, \sigma_y$.
Let X_1, Y_1 be $X-\mu_x, Y-\mu_x$ which are also RV with $0, \sigma_x, 0, \sigma_y$.

Let $\hat{X} = \frac{\rho \sigma_x Y_1}{\sigma_y}$; $\tilde{X} = X_1 - \hat{X}$

where ; $\rho = \frac{E(X_1 Y_1) - \overset{0}{\cancel{E(X_1)}}\overset{0}{\cancel{E(Y_1)}}}{\sigma_x \sigma_y}$

$\tilde{X} = X_1 - \hat{X} \rightarrow ①$

Since X_1 & Y_1 are joint normal random variable, they are linear combination of some U, V which are two independent RV.

$$X_1 = aU + bV$$
$$Y_1 = cU + dV$$

$\Rightarrow \tilde{X} = a'U + b'V$ [from $X_1 - \underset{\overset{\|}{\frac{\rho\sigma_x}{\sigma_y}Y_1}}{\hat{X}} = \tilde{X}$ (①)]

Hence Y_1 & $\tilde{X}$ are also jointly normal.

$$\downarrow$$
$(0, \sigma_Y)$

$$E(\tilde{X}) = E(X_1) - E(\hat{X})$$
$$= 0 - E\left(\frac{\rho\sigma_x}{\sigma_Y} Y\right)$$
$$= 0.$$

$$\Rightarrow \operatorname{cov}(\tilde{X} Y) = E(\tilde{X} Y_1)$$
$$= E(Y_1(X_1 - \hat{X})) = E(X_1 Y_1) - E(Y_1 \hat{X})$$
$$= \rho\sigma_X\sigma_Y - E\left(\frac{\rho\sigma_X}{\sigma_Y} Y_1^2\right)$$
$$= \rho\sigma_X\sigma_Y - \frac{\rho\sigma_X}{\sigma_Y}\sigma_Y^2$$
$$= 0.$$

Hence $\tilde{X}$ and Y_1 are uncorrelated $\Rightarrow$ Independent.

[This is true for a joint normal RV]

$$\Rightarrow X = \tilde{X} + \hat{X} \quad \text{and} \quad \hat{X} = \frac{\rho\sigma_X}{\sigma_Y} Y_1$$

Since $\tilde{X}$ is indep. Y_1

$\tilde{X}$ is indep cY_1

$\Rightarrow$ $\tilde{X}$ is independent of $\hat{X}$.

$$\Rightarrow E(X_1 | Y_1) = \frac{\rho\sigma_X}{\sigma_Y} E(Y_1 | Y_1) + \underset{\underset{E(\tilde{X})}{\parallel}}{E(\tilde{X} | Y_1)}$$

$$E(X_1 | Y_1) = \frac{\rho\sigma_X}{\sigma_Y} Y_1$$

$$\Rightarrow E(X|Y) - \mu_x = \frac{\rho\sigma_X}{\sigma_Y}(Y - \mu_y)$$

$$\Rightarrow \boxed{E(X|Y) = \mu_x + \frac{\rho\sigma_X}{\sigma_Y}(Y - \mu_y)}$$

= LINEAR FUNCTION OF X & Y.

2. $x[K]$ and $y[K]$

$$\hat{\sigma}_{yx} = \frac{1}{N} \sum_{K=1}^{N} (y[K] - \bar{y})(x[K] - \bar{x})$$

$\bar{x}, \bar{y}$ = sample means.

$N \equiv$ sample size.

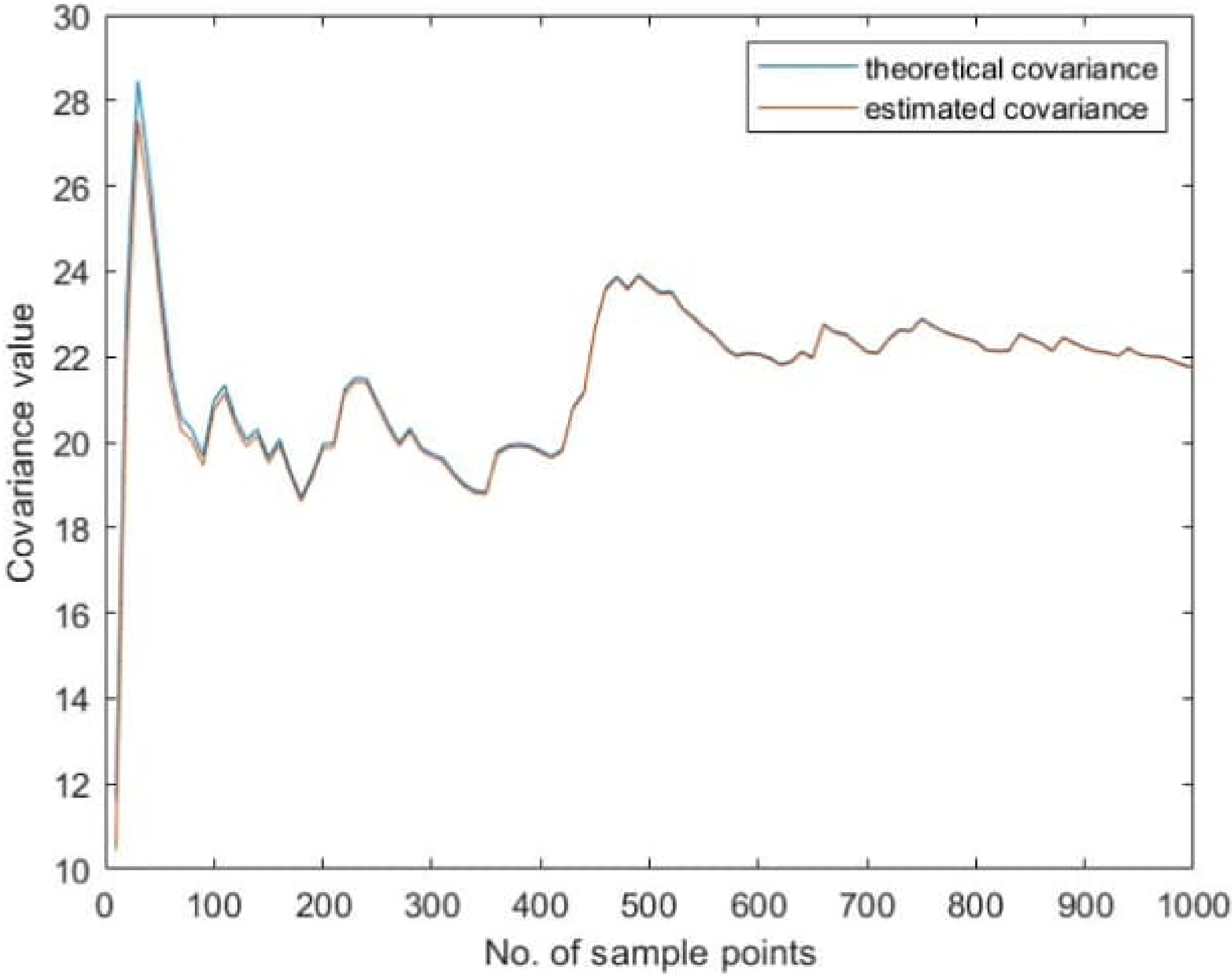
From graph, it is evident that as sample point increases, estimated $\hat{\sigma}_{yx}$ tends to theoretical value.

At 1000 sample points,

Estimated covariance = 21.735674

Theoretical Covariance = 21.757431

Absolute difference = 0.021757 = 0.099 % error



theoretical covariance
estimated covariance
Covariance value
No. of sample points

3. $\Sigma = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix}$

$\sigma_1 = 2$

$\sigma_2 = 3$

$\sigma_3 = 5$

Correlation Matrix $= \begin{bmatrix} 1 & \frac{1}{2.3} & \frac{2}{2.5} \\ \frac{1}{2.3} & 1 & \frac{-3}{3.5} \\ \frac{2}{2.5} & \frac{-3}{3.5} & 1 \end{bmatrix}$

$= \begin{bmatrix} 1 & 1/6 & 1/5 \\ 1/6 & 1 & -1/5 \\ 1/5 & -1/5 & 1 \end{bmatrix}$

(b) correlation btw X_1 ; $1/2\, X_2 + 1/3\, X_3$.

$\text{cov}(X_1, \frac{1}{2}X_2 + \frac{1}{3}X_3) = E\left(\frac{X_1X_2}{2} + \frac{X_1X_3}{3}\right) - E(X_1)\left[\frac{E(X_2)}{2} + \frac{E(X_3)}{3}\right]$

$= \frac{1}{2}\text{cov}(X_1,X_2) + \frac{1}{3}\text{cov}(X_1,X_3)$

$= \frac{1}{2}\cdot 1 + \frac{1}{3}\cdot 2 = \frac{1}{2} + \frac{2}{3} = \frac{7}{6}.$

$\sigma_{X_1} = 2$

$\sigma_{(\frac{1}{2}X_2 + \frac{1}{3}X_3)} \Rightarrow \left(\text{var}\left(\frac{1}{2}X_2 + \frac{1}{3}X_3\right)\right)^{1/2} = \left(\text{var}\left(\frac{1}{2}X_2\right) + \text{var}\left(\frac{1}{3}X_3\right) + 2\text{cov}\left(\frac{1}{2}X_2, \frac{1}{3}X_3\right)\right)^{1/2}$

$= \left(\frac{1}{4}\sigma_2^2 + \frac{1}{9}\sigma_3^2 + 2\left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\sigma_{23}\right)^{1/2}$

$= \left(\frac{1}{4}(9) + \frac{1}{9}(25) + \frac{\cancel{2}\,1}{\cancel{6}\,\cancel{3}}(\cancel{-3})^{-1}\right)^{1/2}$

$= \left(\frac{9}{4} + \frac{25}{9} - 1\right)^{1/2} = 2.0069$

$\Rightarrow \rho = \frac{7/6}{2 \times (2.0069)} = 0.29065.$

4. $X \sim \chi^2(10)$

(a) → Optimal MAE value is 9.339

→ Average absolute error = 3.432117 (cost function value)

(b) $Pr(0.9X^* < X < 1.1X^*) = 0.1724$

$Pr(0.9\mu_x < X < 1.1\mu_x) = 0.1746$. where $\mu_x = 10$.

Hence the first probability is lower than the second one.

POSSIBLE REASONS :-

X^* we estimated is the median as it is MAE.

In chi squared distribution, tail probabilities are very low.

Mean has more weightage as it is $\int x f_X(x)$ compared to median with rather depends on $f_X(x)$. Hence Mean is located more closer towards the peak. Closer to peak implies more density hence more probability.