BY: S. VISHAL

CHS115 : ASSIGNMENT - 1 CH18B020

1(a) i) We know that $\int_{\forall y}\int_{\forall x} f(x,y)\, dx\, dy = 1$.

$$\Rightarrow K\int_0^\infty\int_0^\infty \frac{e^{-x/y}e^{-y}}{y}\, dx\, dy = 1$$

$$\Rightarrow K\int_0^\infty \frac{y}{y}\left(-e^{-x/y}\right)\Big|_0^\infty e^{-y} dy = K\int_0^\infty e^{-y} dy = 1.$$

$$\Rightarrow K(1-0) = 1$$

$$\Rightarrow \boxed{K = 1}$$

ii) Marginal density of Y, $f_Y(y) = \int_{\forall x} f(x,y)\, dx$.

$$= \int_0^\infty \frac{e^{-x/y}e^{-y}}{y}\, dx = \left(\frac{e^{-y}}{y}\right)(y)(e^{-x})\Big|_0^\infty$$

$$\Rightarrow \boxed{f_Y(y) = e^{-y}}$$

iii) $Pr(0<X<1, 0.2<Y<0.4) = \int_{0.2}^{0.4}\int_0^1 f(x,y)\, dx\, dy$

$$= \int_{0.2}^{0.4}\int_0^1 \frac{e^{-x/y}e^{-y}}{y}\, dx\, dy = \int_{0.2}^{0.4} (1-e^{-1/y})\, e^{-y}\, dy$$

Numerically integrating using 'integral' in MATLAB,

$\rightarrow Pr(0<X<1, 0.2<Y<0.4) = \boxed{0.1429.}$

iv) $$f_{X|Y=y}(x) = \frac{f(x,y)}{f_Y(y)}$$

From part (ii) we know the marginal density of Y,

$$\Rightarrow f_{X|Y=y}(x) = \frac{e^{-x/y} e^{-y}}{y} \times \frac{1}{e^{-y}} = \frac{e^{-x/y}}{y}$$

$$E(X|Y=y) = \int_{\forall x} x f_{X|Y}(x)\, dx.$$

$$= \int_0^\infty \frac{x e^{-x/y}}{y} dx \, ;$$

Let $\frac{x}{y} = \beta \Rightarrow dx = y\, d\beta$.

and $x=0 \Rightarrow \beta = 0$,

$x \to \infty \Rightarrow \beta \to \infty$

$$\Rightarrow E(X|Y=y) = \int_0^\infty \beta e^{-\beta} y\, d\beta = y \int_0^\infty \beta e^{-\beta} d\beta.$$

$$= y\left(-\beta e^{-\beta}\Big|_0^\infty + \int_0^\infty e^{-\beta} d\beta \right).$$

$$\Rightarrow \boxed{E(X|Y=y) = y}$$

b) For a bivariate Gaussian,

$$\Sigma = \begin{bmatrix} \sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2 \end{bmatrix}$$

$$\Rightarrow |\Sigma| = \sigma_X^2 \sigma_Y^2 - \sigma_{XY}^2 = (1-\rho^2)\sigma_X^2 \sigma_Y^2$$

$$\& \quad \Sigma^{-1} = \begin{bmatrix} \sigma_Y^2 & -\sigma_{XY} \\ -\sigma_{XY} & \sigma_X^2 \end{bmatrix}$$

$$f(x,y) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2} \begin{bmatrix} X-\mu_X & Y-\mu_Y \end{bmatrix} \Sigma^{-1} \begin{bmatrix} X-\mu_X \\ Y-\mu_Y \end{bmatrix} \right.$$

$$\Rightarrow f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2\sigma_X^2\sigma_Y^2(1-\rho^2)}(x-\mu_X) \begin{pmatrix} \sigma_Y^2 & -\sigma_{XY} \\ -\sigma_{XY} & \sigma_X^2 \end{pmatrix} (\right.$$

Substituting the values of $|\Sigma|$ and Σ^{-1} and simplifying,

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2(x-\mu_x)(y-\mu_y)\rho}{\sigma_x\sigma_y}\right)\right)$$

Let $\frac{x-\mu_x}{\sigma_X} = \alpha$ and $\frac{y-\mu_y}{\sigma_Y} = \beta$.

$$\therefore \Rightarrow f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)}\left(\alpha^2+\beta^2-2\alpha\beta\rho\right)\right)$$

$$E(Y|X) = \int \left(f_{Y/X=x}(y)\right) y\, dy$$

$$= \int \frac{f(x,y)}{f_X(x)}\, y\, dy$$

we need $f_X(x)$ (marginal distribution of X)

$$f_X(x) = \int_{\forall y} f(x,y)\,dy = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2(1-\rho^2)}(\alpha^2+\beta^2-2\rho\alpha\beta)\right) dy$$

completing squares, $\alpha^2+\beta^2-2\rho\alpha\beta = \beta^2-2\rho\alpha\beta+\rho^2\alpha^2 + (1-\rho^2)\alpha^2$

$$= (\beta-\rho\alpha)^2 + (1-\rho^2)\alpha^2$$

Also $\frac{y-\mu_Y}{\sigma_Y} = \beta \Rightarrow \frac{dy}{\sigma_Y} = d\beta$.

~~$\Rightarrow f_X(x) = \frac{1}{2\pi\sigma_X}\int \exp\left(-\frac{(\alpha+\rho\beta)^2}{2(1-\rho^2)}\right)$~~

$$\Rightarrow f_X(x) = \frac{1}{2\pi\sigma_X}\int_{-\infty}^{\infty} \exp\left(-\left[\frac{(\beta-\rho\alpha)^2+(1-\rho^2)\alpha^2}{2(1-\rho^2)}\right]\right)\frac{d\beta}{\sqrt{1-\rho^2}}$$

$$= \frac{\exp\left(\frac{-\alpha^2}{2}\right)}{\sqrt{2\pi}\,\sigma_X} \int_{-\infty}^{\infty} \exp\left(-\left[\frac{(\beta-\rho\alpha)^2}{2(1-\rho^2)}\right]\right)\frac{d\beta}{\sqrt{2\pi}\sqrt{1-\rho^2}}$$

The given integral is the integral of a normal random variable with mean $\rho\alpha$ & $\sigma = \sqrt{1-\rho^2}$.

$$\Rightarrow f_X(x) = \frac{\exp\left(\frac{-\alpha^2}{2}\right)}{\sqrt{2\pi}\,\sigma_X}$$

in terms of x, $f_X(x) = \frac{\exp\left(-\frac{(x-\mu_X)^2}{2\sigma_X^2}\right)}{\sqrt{2\pi}\,\sigma_X}$

(A normal RV with mean μ_X and standard deviation σ_X)

$$\therefore E(Y|X) = \int y \frac{\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}\exp\left(\frac{-1}{2}\left(\frac{\alpha^2+\beta^2-2\alpha\beta\rho}{1-\rho^2}\right)\right)dy}{\frac{\exp\left(-\frac{\alpha^2}{2}\right)}{\sqrt{2\pi}\cdot\sigma_X}}$$

$$\Rightarrow E(Y|X) = \int_{\forall y} \frac{y}{\sqrt{2\pi}\,\sigma_Y(\sqrt{1-\rho^2})} \exp\left(-\frac{1}{2}\left(\frac{\alpha^2+\beta^2-2\alpha\beta\rho}{1-\rho^2} - \alpha^2\right)\right) dy$$

$$= \int_{\forall y} \frac{y}{\sqrt{2\pi}\,\sigma_Y(\sqrt{1-\rho^2})} \left(\exp\left(-\frac{1}{2}\frac{(\beta-\alpha\rho)^2}{(1-\rho^2)}\right)\right) dy$$

Let $\frac{y - \mu_Y}{\sigma_Y} = \beta \Rightarrow \frac{dy}{\sigma_Y} = d\beta$ and $y \to \infty \Rightarrow \beta \to \infty$, $y \to -\infty \Rightarrow \beta \to -\infty$

Adding and subtracting μ_Y in the Numerator,

$$= \int \left(\frac{y-\mu_Y}{\sigma_Y}\right) \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2}\left(\frac{(\beta-\alpha\rho)^2}{(1-\rho^2)}\right)\right) d\beta.\sigma_Y$$

$$+ \frac{\mu_Y}{\sigma_Y\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2}\left(\frac{(\beta-\alpha\rho)^2}{(1-\rho^2)}\right)\right) dy$$

$$= \frac{\sigma_Y}{\sqrt{2\pi}\sqrt{1-\rho^2}}\left[\int_{\forall\beta} \beta.\exp\left(\frac{-1}{2}\frac{(\beta-\alpha\rho)^2}{(1-\rho^2)}\right) d\beta + \int_{\forall\beta} \frac{\mu_Y}{\sigma_Y} \exp\left(\frac{-1}{2}\left(\frac{(\beta-\alpha\rho)^2}{(1-\rho^2)}\right)\right) d\beta\right]$$

first integral: $$\frac{\sigma_Y}{\sqrt{2\pi}\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \beta \exp\left(\frac{-1}{2}\left(\frac{(\beta-\alpha\rho)^2}{(1-\rho^2)}\right)\right) d\beta.$$

The integral is the first moment of a normal RV having $E(X) = \rho\alpha$ and $\sigma^2 = (1-\rho^2)$

$\Rightarrow$ first integral $= \sigma_Y (\rho\alpha) = \sigma_Y \left(\frac{\sigma_{XY}}{\sigma_X \sigma_Y} \alpha \right)$

$$= \frac{\sigma_{XY}}{\sigma_X} \alpha.$$

Second integral: $\frac{\mu_Y}{1} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left(\frac{-1}{2} \frac{(\beta - \rho\alpha)^2}{(1-\rho^2)} \right) d\beta.$

The integral is the integral of a PDF of normal RV with $E(X) = \rho\alpha$ & $\sigma = \sqrt{1-\rho^2}$ throughout its domain

$\therefore$ value of the integral is 1

$\rightarrow$ Second integral $= \mu_Y$

$$\therefore E(Y | X) = \frac{\sigma_{XY}}{\sigma_X} \alpha + \mu_Y$$

But $\alpha = \frac{X - \mu_X}{\sigma_X}$

$$\Rightarrow E(Y | X) = \frac{\sigma_{XY}}{\sigma_X^2} (X - \mu_X) + \mu_Y$$

$$\Rightarrow E(Y|X) = \left(\frac{\sigma_{XY}}{\sigma_X^2} \right) X - \frac{\mu_X \sigma_{XY}}{\sigma_X^2} + \mu_Y$$

Thus $E(Y|X)$ is a linear function of X

Slope $= \frac{\sigma_{XY}}{\sigma_X^2}$, intercept $= \mu_Y - \frac{\mu_X \sigma_{XY}}{\sigma_X^2}$

② → The given expression for sample covariance ~~matrix~~ has been used and a function for the evaluating the variance - covariance matrix has been written in MATLAB.

→ We note that the cov ( ) function in MATLAB normalises by N-1 rather than N by default. So I have passed an additional parameter cov([x y], 1) to ensure that it normalises by N and it can be compared with the user defined function.

→ X has been declared using randn as

X = randn([size, 1]) * $\sqrt{2}$ + 1;

($\sigma = \sqrt{2}$ and $\mu = 1$).

Y is simply ~~3 X.^4~~ .. 3X.^2 + 5X

(elementwise multiplication and elementwise addition in case of vector X)

→ Theoritical ~~expectation~~: $\sigma_{XY}: E(XY) - E(X)E(Y)$

$$= \left(E(X(3X^2+5X)) - E(X)E(3X^2+5X)\right)$$

$$= 3E(X^3) + 5E(X^2) - (E(X))(3)(E(X^2)) - 5(E(X))^2$$

$E(X) = \mu_X = 1.$

$E(X^2) = \sigma_X^2 + (E(X))^2 = \cancel{4}^{\,2} + 1 = \cancel{5}\ 3$

To find $E(X^3)$, lets find ~~E(...)~~ $E(A^3)$ where $A \sim N(1,0)$

$E(A^3) = \int_{\forall a} a^3 f(a)\, da$ Since Gaussian is symmetric, $f(a) = f(-a)$

$= \int_{-\infty}^{\infty} a^3 f(a)\, da$

$\Rightarrow a^3 f(a)$ is an odd function

$\Rightarrow E(A^3) = 0.$

That means $E\left(\left(\frac{X - \mu_X}{\sigma_X^2}\right)^3\right) = 0$

$\Rightarrow E((X-\mu)^3) = 0 \cdot \sigma^3$

$\Rightarrow E(X^3) - \mu^3 + 3E(X)\mu^2 - (E(X^2))(3\mu) = 0.$

We know $E(X) = \mu_X = 1$, $E(X^2) = 3$

$\Rightarrow E(X^3) = 7$

$\therefore \underline{\sigma_{XY} \text{ theoritical}} = 3 \times 7 + (5)(3) - 3 \times 3 \times 5 \times 1^2$

$= \boxed{22}$

To examine convergence:

i) I generated samples of size $1 - 10^4$, although the convergence wasn't monotonic there was convergence

ii) I then decided to try higher values of n and then plot ~~a curve~~ σ_{XY} for values between 10^5 and 10^7 in steps of 10^6 ~~again it wasn't~~

iii) convergence is monitored using deviation vs sample size plot. (deviation = |estimate - theoretical value|

iv) The convergence isn't monotonic but there was convergence. (Deviation goes closer to zero as we increase sample size). This displays consistency.

Sample size	avg	Deviation
10^1	15.9148	6.0852
10^2	29.9677	7.967
10^3	20.8897	1.1103
10^4	21.316	0.684
10^5	21.1117	0.112
10^6	22.0121	0.0121
2×10^6	21.9598	0.0402
3×10^6	21.9767	0.0233

We can see that as we ↑se n, the deviation is bound closer and closer to zero.

(or 'local mean' of deviation goes to zero)

(3) a) $\Sigma = \begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & -3 \\ 2 & -3 & 25 \end{bmatrix}$ Let ~~$X_2 \to$~~ Y, ~~X_2~~ $Z = X_3$

$Y = X_2$

variance - covariance matrix: $\begin{bmatrix} \sigma_X^2 & \sigma_{XY} & \sigma_{XZ} \\ \sigma_{XY} & \sigma_Y^2 & \sigma_{YZ} \\ \sigma_{XZ} & \sigma_{YZ} & \sigma_Z^2 \end{bmatrix}$

to convert it into correlation matrix: $\begin{bmatrix} \frac{\sigma_X^2}{\sigma_X^2} & \frac{\sigma_{XY}}{\sigma_X \sigma_Y} & \frac{\sigma_{XZ}}{\sigma_X \sigma_Z} \\ \frac{\sigma_{XY}}{\sigma_X \sigma_Y} & \frac{\sigma_Y^2}{\sigma_Y^2} & \frac{\sigma_{YZ}}{\sigma_Y \sigma_Z} \\ \frac{\sigma_{XZ}}{\sigma_X \sigma_Z} & \frac{\sigma_{YZ}}{\sigma_Y \sigma_Z} & \frac{\sigma_Z^2}{\sigma_Z^2} \end{bmatrix}$

(we need to normalise the values as depicted above)

S. ~~Calc~~. From Σ we can obtain,

$$\boxed{\begin{matrix} \sigma_X = 2 & \sigma_Y = 3 & \sigma_Z = 5 \\ \sigma_{XY} = 1 & \sigma_{XZ} = 2 & \sigma_{YZ} = -3 \end{matrix}}$$

$\therefore$ correlation matrix ρ : $\begin{bmatrix} 1 & 1/6 & 1/5 \\ 1/6 & 1 & -1/5 \\ 1/5 & -1/5 & 1 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0.1667 & 0.2 \\ 0.1667 & 1 & -0.2 \\ 0.2 & -0.2 & 1 \end{bmatrix}$$

③ b) $\rho\left(X_1, \frac{X_2}{2}+\frac{X_3}{2}\right) = \dfrac{E\left((X_1)\left(\frac{X_2}{2}+\frac{X_3}{2}\right)\right) - E(X_1)E\left(\frac{X_2}{2}+\frac{X_3}{2}\right)}{\sigma_{X_1}\ \sigma_{\left(\frac{X_2}{2}+\frac{X_3}{2}\right)}}$

$E\left(\frac{X_2}{2}+\frac{X_3}{2}\right) = E\left(\frac{X_2}{2}\right) + E\left(\frac{X_3}{2}\right) = \frac{E(X_2)+E(X_3)}{2}$ — ①

$var\left(\frac{X_2+X_3}{2}\right) = \frac{1}{4}E\left(\left(\frac{X_2+X_3}{2}\right)^2\right) - \left(E\left(\frac{X_2+X_3}{2}\right)\right)^2$

$= \frac{1}{4}\left[E(X_2^2) + E(X_3^2) + 2E(X_2X_3) - (E(X_2))^2 - (E(X_3))^2 - 2E(X_2)E(X_3)\right]$

$= \frac{1}{4}\left[\left(E(X_2^2) - (E(X_2))^2\right) + \left(E(X_3^2) - (E(X_3))^2\right) + 2\left(E(X_2X_3) - E(X_2)E(X_3)\right)\right]$

$= \frac{1}{4}\left[\sigma_{X_2}^2 + \sigma_{X_3}^2 + 2\sigma_{X_2X_3}\right]$

$\therefore \sigma_{\left(\frac{X_2+X_3}{2}\right)} = \frac{1}{2}\sqrt{\sigma_{X_2}^2 + \sigma_{X_3}^2 + 2\sigma_{X_2X_3}}$ — ②

$E\left(X_1\left(\frac{X_2}{2}+\frac{X_3}{2}\right)\right) - E(X_1)E\left(\frac{X_2+X_3}{2}\right)$

$= \frac{E(X_1X_2)}{2} - \frac{E(X_1)E(X_2)}{2} + \frac{E(X_1X_3) - E(X_1)E(X_3)}{2}$

$= \frac{1}{2}\left(\sigma_{X_1X_2} + \sigma_{X_1X_3}\right)$ — ③

From ①, ②, ③ we get,

$$\rho\left(X_1, \frac{X_2+X_3}{2}\right) = \frac{\frac{1}{2}(\sigma_{X_1X_2} + \sigma_{X_1X_3})}{\sigma_{X_1}\left(\frac{1}{2}\sqrt{\sigma_{X_2}^2 + \sigma_{X_3}^2 + 2\sigma_{X_2X_3}}\right)}$$

Substituting these values from variance - covariance matrix

$$\Rightarrow \rho\left(X_1, \frac{X_2+X_3}{2}\right) = \frac{(2+1)}{2\sqrt{(9+25-6)}}$$

$$\Rightarrow \boxed{\rho = +0.2835}$$

④ a). ◦ 200 records, each containing 10000 samples are generated from χ^2 distribution of dof 10.

◦ From the shape of the distribution the optimal MAE predictor x^* $\left(\min_{\hat{x}} E(|x - \hat{x}|)\right)$ we can safely say that x^* should lie between 8 and 11

◦ A vector containing guess values of x^* is declared. They are consecutive values with 0.001 spacing.

◦ the cost function for each record = ~~norm~~ $|x - \hat{x}^*|$, is found using '~~norm~~ sum' function and 'abs' function. (~~We use the 1-norm on the residual vector of each record~~)

◦. Each record is given $1/N$ weightage

J for a given record = $\sum_{\forall i} |x(i) - \hat{x}| \times \frac{1}{N}$

- The minimum cost is found using min( ).
- min_J = 3.469 ⇒ Average absolute error $= \boxed{3.469}$

  $x^* = 9.3360 = \boxed{9.336}$
- We note that x^* varies in each run of the program.
- J vs $\hat{X}$ curve is also plot.

b) μ_x is found using chi2stat( ) = 10

The required probabilities are found using the cdf of χ^2 distribution. $Pr(0.9x^* < X < 1.1x^*)$

$= F(x \leq 1.1x^*) - F(x \leq 0.9x^*)$

where F is the cumulative distribution function.

We find that

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

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

$\Rightarrow Pr(0.9\mu_x < X < 1.1\mu_x) > Pr(0.9x^* < X < 1.1x^*)$.

From the cdf plot we observe that,

μ_x is located nearer to the peak (in a more denser region) than x^* — ①

Also $\mu_x > x^* \Rightarrow 0.2\mu_x > x^*(0.2)$

⇒ Range of values is greater in case of μ_x — ②

Because of these 2 reasons we see

$$Pr(0.9\mu_X < X < 1.1\mu_X) > Pr(0.9 \cdot X^* < X < 1.1 X^*)$$

A more Qualitative reasoning:

→ x^* ~~is to~~ can be considered an estimate of the median, since median minimises $E(|X - x^*|)$

~~Definition~~ We can derive that $\int_m^\infty f(x)\,dx = \int_0^m f(x)\,dx$.
where m is the median. $(\Rightarrow F(X \le m) = 0.5)$

So for X^2 distribution which has a sharp rise followed by a sharper fall with the tail values really low compared to the left half of the peak, the median, by definition, being in the middle, will tend to the left side of the peak.

~~However~~ $\mu_X = E(X) = \int f(x)\,x\,dx$. Here, eventhough the tail probabilities are low, the ~~median~~ mean gives higher weightage to it than mean because of the fact that the 'x' continously increases in the right. This effect causes $E(X)$ to move a bit more to the right than the median.

So μ_X is close to peak than X^*

$\Rightarrow \mu_X$ is in a denser region, so the Probability of the RV taking a value around it is more than ~~$Pr(0.9X^* < X < 1.1X^*)$~~ the RV taking value around the median.