دانشگاه صنعتی امیرکبیر
دانشکده مهندسی کامپیوتر

گزارش تکلیف اول درس شناسائی آماری الگو

دانشجو:
سید احمد نقوی نوزاد
ش-د: ۹٤۱۳۱۰٦۰

استاد:
دکتر رحمتی

پائیز ۹۵

جواب سوال ۱

$$E(X)=2, E(X^2)=2, Y=-2X-5$$

$$a)\quad V(X)=E(X^2)-E(X)^2=2-2^2=-2$$

$$b)\quad V(Y)=V(-2X-5)=(-2)^2V(X)=4\times(-2)=-8$$

$$c)\quad E(Y^2)=?\rightarrow V(Y)=E(Y^2)-E(Y)^2\Rightarrow E(Y^2)=V(Y)+E(Y)^2\ <*>,$$

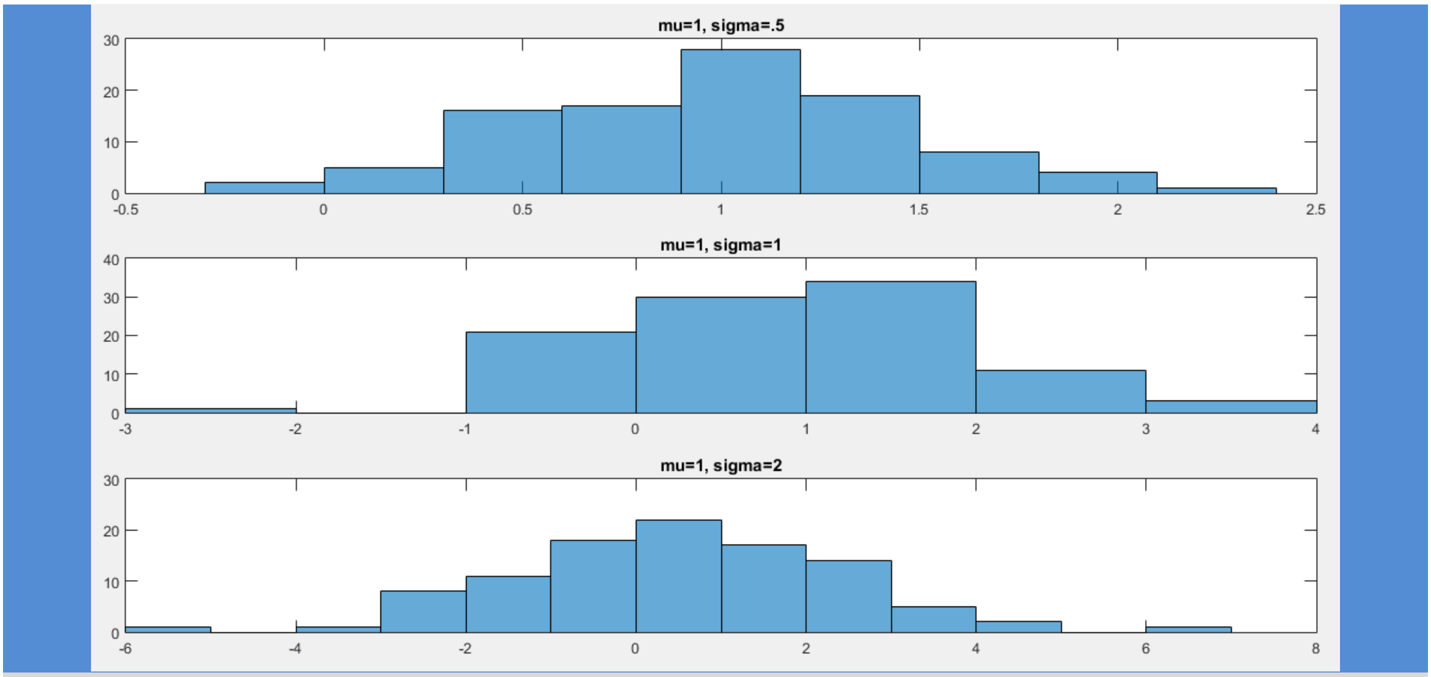
$$E(Y)=E(-2X-5)=-2E(X)-5=-2\times 2-5=-9$$

$$<*>\Rightarrow E(Y^2)=-8+(-9)^2=73$$

جواب سوال ۲

کد مربوط به این سوال در فایل HW01_ex02 قرار دارد.

در این‌جا با استفاده از تابع متلب با نام mvnrnd() ، تعداد ۱۰۰ عدد داده‌ی تصادفی دارای توزیع نرمال با مقادیر میانگین $\mu = 1$ و $\sigma =.5, 1, 2$ تولید نموده و نمودار هیستوگرام مربوطه را مطابق زیر رسم می‌نمائیم:

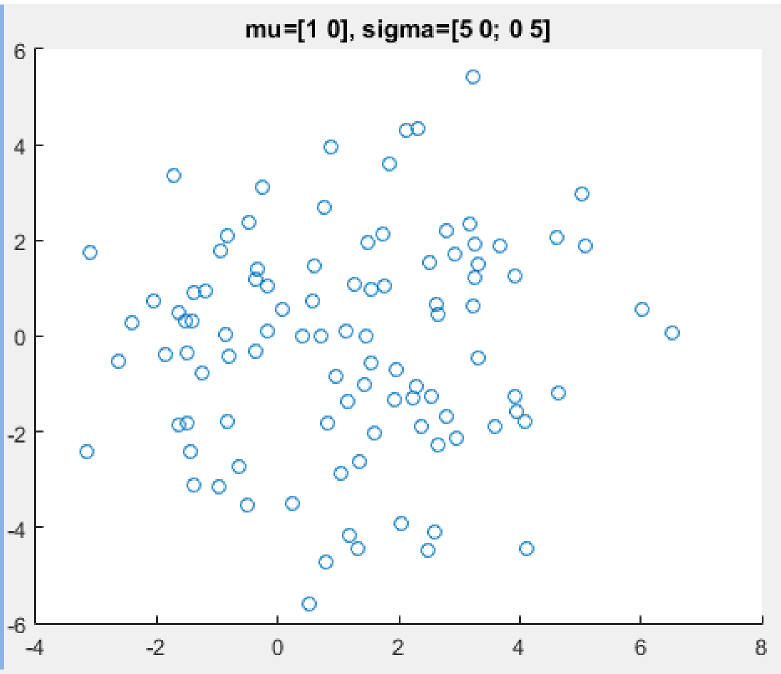


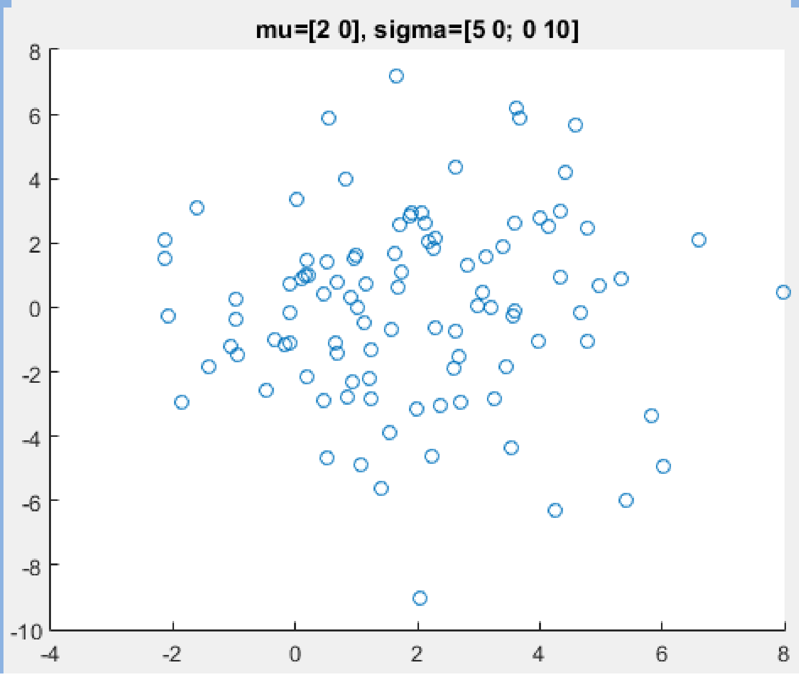
شکل ۱. نمودار هستوگرام مربوط به داده‌های با توزیع نرمال

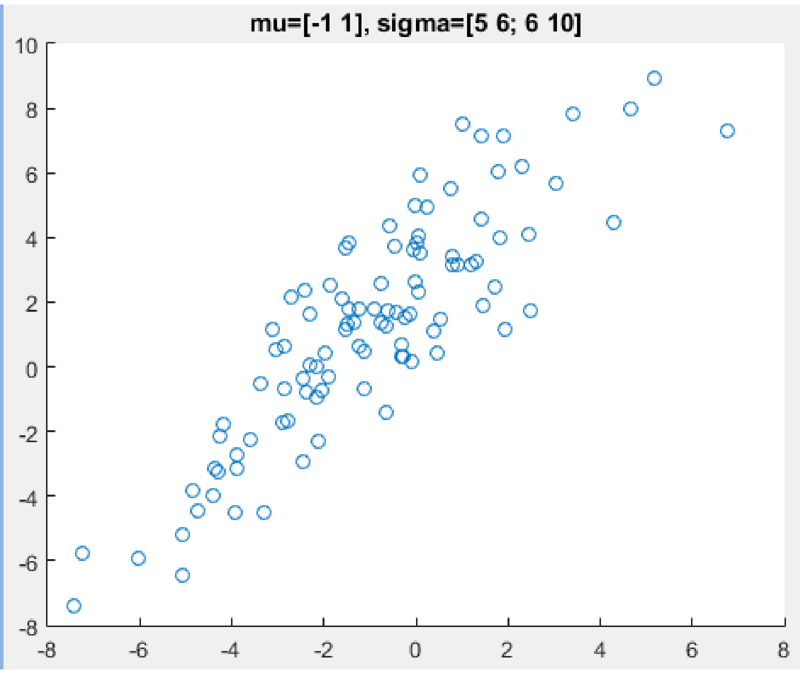
همانطور که قابل مشاهده است، میانگین توزیع نرمال در هر سه مورد برابر ۱ بوده و تجمع داده‌ها نیز مطابق نمودار هیستوگرام حول همان مقدار ۱ می‌باشد؛ اما به ازای افزایش مقدار انحراف از معیار σ پراکندگی داده‌ها افزایش پیدا کرده و توزیع داده‌ها در محدوده‌ی وسیع‌تری قرار می‌گیرد. به عنوان مثال برای مقدار انحراف از معیار برابر ۱، توزیع داده‌ها در محدوده‌ی [-.5,2.5] می‌باشد، در حالی که برای مقدار انحراف از معیار برابر ۲، توزیع داده‌ها در محدوده‌ی وسیع‌تر [-6,8] می‌باشد.

جواب سوال ۳

کد مربوط به این سوال در فایل HW01_ex03 قرار دارد.







جواب سوال ۴

کد مربوط به این سوال در فایل HW01_ex04 قرار دارد.

مقادیر اصلی مربوط به سوال ۳ به شرح ذیل می‌باشد:

$$\text{a.}\quad N = 100, M = \begin{bmatrix}1\\0\end{bmatrix}, \Sigma = \begin{bmatrix}5 & 0\\0 & 5\end{bmatrix}$$

$$\text{b.}\quad N = 100, M = \begin{bmatrix}2\\0\end{bmatrix}, \Sigma = \begin{bmatrix}5 & 0\\0 & 10\end{bmatrix}$$

$$\text{c.}\quad N = 100, M = \begin{bmatrix}-1\\1\end{bmatrix}, \Sigma = \begin{bmatrix}5 & 6\\6 & 10\end{bmatrix}$$

مقادیر حاصله از متلب نیز به شرح ذیل می‌باشد:

$$\boldsymbol{a.}\ \boldsymbol{N} = \mathbf{100}, \qquad \boldsymbol{M} = \begin{bmatrix}\mathbf{1.2123}\\ \mathbf{.1554}\end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix}\mathbf{4.3448} & \mathbf{-.3805}\\ \mathbf{-.3805} & \mathbf{4.4114}\end{bmatrix}$$

$$\boldsymbol{b.}\ \boldsymbol{N} = \mathbf{100}, \qquad \boldsymbol{M} = \begin{bmatrix}\mathbf{2.0270}\\ \mathbf{.0223}\end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix}\mathbf{4.8392} & \mathbf{.6865}\\ \mathbf{.6865} & \mathbf{10.6814}\end{bmatrix}$$

$$\boldsymbol{c.}\ \boldsymbol{N} = \mathbf{100}, \qquad \boldsymbol{M} = \begin{bmatrix}\mathbf{-.9474}\\ \mathbf{1.2152}\end{bmatrix}, \qquad \boldsymbol{\Sigma} = \begin{bmatrix}\mathbf{4.7443} & \mathbf{5.5196}\\ \mathbf{5.5196} & \mathbf{8.9793}\end{bmatrix}$$

همانطور که قابل مشاهده است، مقادیر حاصله از متلب با مقادیر اولیه چندان تفاوتی ندارند و علت این اختلاف نیز در آن است که تعداد نمونه‌های تصادفی اندک می‌باشند و در صورت تولید نمونه‌های بیشتر، این مقادیر به مقادیر واقعی نزدیک‌تر هم خواهند شد.

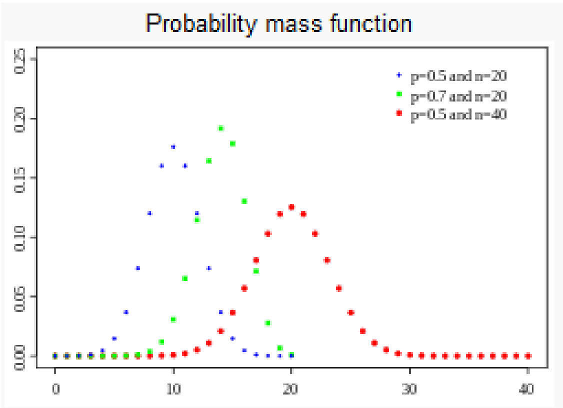
جواب سوال ۵

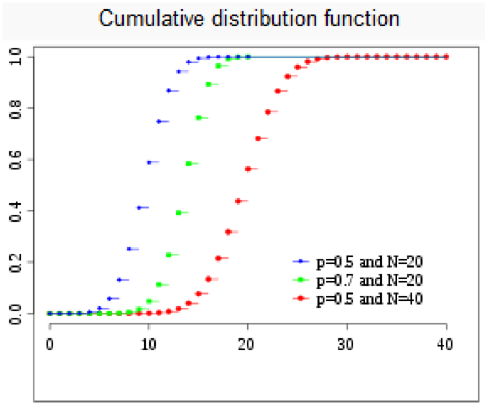
کد مربوط به این سوال در فایل HW01_ex05 قرار دارد.

توزیع Binomial دارای توزیع زیر می‌باشد:

$$f(k;n,p) = \Pr(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

که در آن k برابر تعداد دفعات شیرآمدن در n پرتاب یک سکه‌ی غیر مورب (fair coin) می‌باشد، و p نیز برابر احتمال شیرآمدن می‌باشد. شکل نمودار تابع توزیع جرم احتمال (pmf) و تابع توزیع تجمعی احتمال (cdf) برای توزیع دوجمله‌ای یا همان Binomial نیز به صورت زیر می‌باشد:





همانطور که قابل مشاهده است هر دو شکل مربوط به pdf و cdf برای این توزیع تا حدود زیادی شبیه به توابع مربوط به توزیع نرمال می‌باشند. حال اگر تعداد دفعات پرتاب سکه زیاد باشد (حداقل ۲۰ پرتاب) و مقدار احتمال p نیز نه زیاد نزدیک به ۰ و نه زیاد نزدیک به ۱ باشد (مثلا حدود ۰.۵)، در آن‌صورت میزان چولگی توزیع دوجمله‌ای زیاد نبوده و توزیع نرمال با پارامترهای $N(np, np(1-p))$ می‌تواند یک تخمین منطقی برای توزیع دوجمله‌ایِ $B(n,p)$ باشد.

حال در این تمرین مقدار n برابر ۲۰ بوده و مقدار p را نیز برابر ۰.۵ اختیار می‌نمائیم، و در نتیجه توزیع نرمال مربوطه به صورت $N(20*.5, 20*.5(1-.5)^2) = N(10, 5^2)$ خواهد بود. داریم:

$$B(20,.5) = N(10,25) = \frac{1}{5\sqrt{2\pi}} exp\left\{-\frac{(x-10)^2}{50}\right\}$$

که در آن x برابر تعداد دفعات شیرآمدن در n پرتاب سکه می‌باشد. حال در این تمرین از ما خواسته شده تا مقدار تابع توزیع جرم احتمال دو جمله‌ای را با استفاده از تابع توزیع چگالی احتمال نرمال و برای مقادیر $x = [12, 13, 14, 15]$ تخمین بزنیم. نتایج حاصله از کد متلب مربوطه به صورت زیر بوده و خود گویای تخمین نسبتا دقیق دلخواه ما می‌باشد:

x value	12	13	14	15
Binomial Dist.	0.1201	0.0739	0.0370	0.0148
Normal Dist.	0.0737	0.0666	0.0579	0.0484

جواب سوال ۶

کد مربوط به این سوال در فایل HW01_ex06 قرار دارد.

قسمت الف)

$$A=\begin{bmatrix}2&0&0\\-1&3&3\\6&-6&-6\end{bmatrix}\rightarrow \text{what are eigenvalues and eigenvectors of A?}$$

$$|A-\lambda I|=0\Rightarrow\left|\begin{bmatrix}2&0&0\\-1&3&3\\6&-6&-6\end{bmatrix}-\begin{bmatrix}\lambda&0&0\\0&\lambda&0\\0&0&\lambda\end{bmatrix}\right|=0\Rightarrow\begin{vmatrix}2-\lambda&0&0\\0&3-\lambda&0\\0&0&-6-\lambda\end{vmatrix}=0$$

$$\Rightarrow(2-\lambda)\begin{vmatrix}3-\lambda&3\\-6&-6-\lambda\end{vmatrix}=0\Rightarrow(2-\lambda)[(3-\lambda)(-6-\lambda)+18]=0$$

$$\Rightarrow\lambda(2-\lambda)(3+\lambda)=0\Rightarrow\qquad \lambda=-3,0,2\rightarrow \text{these are eigenvalues of A}$$

$$Av=\lambda v\ \rightarrow \text{and v is a eigenvector}$$

$$\lambda_1=-3\rightarrow\begin{bmatrix}2&0&0\\-1&3&3\\6&-6&-6\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=-3\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\Rightarrow\begin{cases}2v_1=-3v_1\Rightarrow v_1=0\\-v_1+3v_2+3v_3=-3v_2\\6v_1-6v_2-6v_3=-3v_3\end{cases}$$

$$\Rightarrow\begin{cases}6v_2+3v_3=0\\-6v_2-3v_3=0\end{cases}\Rightarrow v_2=-\frac{1}{2}v_3\Rightarrow v_3=-2,v_2=1\Rightarrow \underline{v_1}=c_1\begin{bmatrix}0\\1\\-2\end{bmatrix}$$

$$\lambda_2=0\rightarrow\begin{bmatrix}2&0&0\\-1&3&3\\6&-6&-6\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=0\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\Rightarrow\begin{cases}2v_1=0\Rightarrow v_1=0\\-v_1+3v_2+3v_3=0\\6v_1-6v_2-6v_3=0\end{cases}$$

$$\Rightarrow\begin{cases}3v_2+3v_3=0\\-6v_2-6v_3=0\end{cases}\Rightarrow v_2=-v_3\Rightarrow v_3=2,v_2=-2\Rightarrow \underline{v_2}=c_2\begin{bmatrix}0\\-2\\2\end{bmatrix}$$

$$\lambda_3=2\rightarrow\begin{bmatrix}2&0&0\\-1&3&3\\6&-6&-6\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}=2\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\Rightarrow\begin{cases}2v_1=2v_1\Rightarrow v_1=1\\-v_1+3v_2+3v_3=2v_2\\6v_1-6v_2-6v_3=2v_3\end{cases}$$

$$\Rightarrow\begin{cases}v_2+3v_3=1\\6v_2+8v_3=6\end{cases}\Rightarrow v_3=0,v_2=1\Rightarrow \underline{v_3}=c_3\begin{bmatrix}1\\1\\0\end{bmatrix}$$

در این‌جا با استفاده از تابع متلب با نام **eig** و در قالب **[V,D]=eig(A)** دو ماتریس **V** و **D** را به دست می‌آوریم، که **V** یک ماتریس با همان ابعاد ماتریس **A** می‌باشد، که هر ستون آن مطابق یک **eigenvector** مربوط به یکی از **eigenvalue**ها می‌باشد و البته چینش این **eigenvalue**ها نیز به صورت صعودی می‌باشد. ماتریس **D** نیز یک ماتریس قطری می‌باشد که درایه‌های روی قطر

اصلی آن همان eigenvalueها و به ترتیب صعودی از چپ به راست می‌باشند. نتایج حاصله به همراه نتایج محاسباتی بالا به شرح ذیل می‌باشند:

eigenValue1 = -3		eigenValue2 = 0		eigenValue3 = 2	
eigenVector1		**eigenVector2**		**eigenVector3**	
Comp. ($c_1 \times$)	**MATLAB**	**Comp. ($c_2 \times$)**	**MATLAB**	**Comp. ($c_3 \times$)**	**MATLAB**
0	0	0	0	1	.7071
1	.4472	-2	-.7071	1	.7071
-2	-.8944	2	.7071	0	-1.1776e-16

همانطور که قابل مشاهده است، با توجه به این‌که نتایج محاسباتی همگی ضریبی از یک مقدار خاص می‌باشند، با نتایج حاصله از متلب همخوانی تام داشته و در نتیجه خوشبختانه محاسبات ما صحیح بوده است.

قسمت ب)

$$Av_i = \lambda_i v_i$$

$$Av_1 = \lambda_1 v_1 \Rightarrow \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \begin{cases} a_{11} = 2 \\ a_{21} = 0 \end{cases}$$

$$Av_2 = \lambda_2 v_2 \Rightarrow \begin{bmatrix} 2 & a_{12} \\ 0 & a_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \begin{cases} a_{12} = 3 \\ a_{22} = 5 \end{cases}$$

$$\Rightarrow A = \begin{bmatrix} 2 & 3 \\ 0 & 5 \end{bmatrix}$$

جواب سوال ۷

$$f(x,y)=\begin{cases} c(x+y^2)\ ;\ 0<x<1, 0<y<1 \\ 0 \quad ; \quad \text{o.w} \end{cases} \rightarrow \begin{cases} c=? \\ f_X(x)=? \\ f_{X|Y}(x|y)=? \\ X\,\Pi\,Y \quad ? \\ \Pr(X<\tfrac{1}{2}|Y=\tfrac{1}{2})=? \end{cases}$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy=1 \Rightarrow \int_0^1\int_0^1 c(x+y^2)dxdy=1 \Rightarrow c\int_0^1 [\frac{1}{2}x^2+xy^2]_{x=0}^{x=1}dy=1 \Rightarrow$$

$$c\int_0^1 (\frac{1}{2}+y^2)dy=1 \Rightarrow c[\frac{1}{2}y+\frac{1}{3}y^3]_{y=0}^{y=1}=1 \Rightarrow c(\frac{1}{2}+\frac{1}{3})=1 \quad \Rightarrow \quad c=\frac{6}{5}$$

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^1 \frac{6}{5}(x+y^2)dy=\frac{6}{5}[xy+\frac{1}{3}y^3]_{y=0}^{y=1}=\frac{6}{5}(x+\frac{1}{3}) \quad \Rightarrow \quad f_X(x)=\frac{6}{5}x+\frac{2}{5}$$

$$f_{X|Y}(x|y)=\frac{f_{X,Y}(x,y)}{f_Y(y)} \rightarrow f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_0^1 \frac{6}{5}(x+y^2)dx=\frac{6}{5}[\frac{1}{2}x^2+xy^2]_{x=0}^{x=1}=$$

$$\frac{6}{5}[\frac{1}{2}x^2+xy^2]_{x=0}^{x=1}=\frac{6}{5}(y^2+\frac{1}{2}) \Rightarrow f_Y(y)=\frac{6}{5}y^2+\frac{3}{5}$$

$$\Rightarrow f_{X|Y}(x|y)=\frac{\frac{6}{5}(x+y^2)}{\frac{6}{5}(y^2+\frac{1}{2})} \quad \Rightarrow \quad f_{X|Y}(x|y)=\frac{x+y^2}{y^2+\frac{1}{2}}$$

If X and Y are independent ($X\Pi Y$) then: $f_{X,Y}(x,y)=f_X(x)f_Y(y)$

$\frac{6}{5}(x+y^2)\neq\frac{6}{5}(x+\frac{1}{3})\times\frac{6}{5}(y^2+\frac{1}{2}) \quad \Rightarrow$ so X and Y are NOT independent

$$\Pr(X<\tfrac{1}{2}|Y=\tfrac{1}{2})=\int_{-\infty}^{1/2} f_{X<1/2|Y=1/2}(x|\tfrac{1}{2})dx=\int_0^{1/2}\left(\frac{x+(\frac{1}{2})^2}{(\frac{1}{2})^2+\frac{1}{2}}\right)dx=\int_0^{1/2}\left(\frac{x+\frac{1}{4}}{\frac{3}{4}}\right)dx$$

$$=\frac{4}{3}\int_0^{1/2}\left(x+\frac{1}{4}\right)dx=\frac{4}{3}\left[\frac{x^2}{2}+\frac{x}{4}\right]_{x=0}^{x=1/2}=\frac{4}{3}(\frac{1}{8}+\frac{1}{8})=\frac{1}{3} \quad \Rightarrow \quad \Pr(X<\tfrac{1}{2}|Y=\tfrac{1}{2})=\frac{1}{3}$$

جواب سوال ۸

کد مربوط به این سوال در فایل HW01_ex08 قرار دارد.

$$\underline{x} \in R^n \quad \rightarrow \quad p(\underline{x} \mid \mu_{\underline{X}}, \Sigma_{\underline{X}}) = \frac{1}{(2\pi)^{(\frac{n}{2})} |\Sigma_{\underline{X}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\underline{x}-\mu_{\underline{X}})^T \Sigma_{\underline{X}}^{-1} (\underline{x}-\mu_{\underline{X}})\right\}$$

قسمت الف)

$y = \underline{v}_i^T \underline{x}$, where $\underline{v}_i^T$ is an eigenvector of $\Sigma_{\underline{X}}$ with an eigenvalue of λ_i. p(y)=?

so y becomes a constant! and we have:

$$\Sigma_{\underline{X}} \underline{v}_i = \lambda_i \underline{v}_i \Rightarrow (\Sigma_{\underline{X}} \underline{v}_i)^T = (\lambda_i \underline{v}_i)^T \Rightarrow \underline{v}_i^T \Sigma_{\underline{X}}^T = \lambda_i \underline{v}_i^T \Rightarrow \underline{v}_i^T \Sigma_{\underline{X}}^T \underline{x} = \lambda_i \underline{v}_i^T \underline{x} \Rightarrow$$

$$\underline{v}_i^T \underline{x} = \left(\frac{1}{\lambda_i} \underline{v}_i^T \Sigma_{\underline{X}}^T\right) \underline{x} \Rightarrow y = \left(\frac{1}{\lambda_i} \underline{v}_i^T \Sigma_{\underline{X}}^T\right) \underline{x} \Rightarrow y = \left(\frac{1}{\lambda_i} \Sigma_{\underline{X}} \underline{v}_i\right)^T \underline{x} \Rightarrow y = A^T \underline{x}$$

where $A = \frac{1}{\lambda_i} \Sigma_{\underline{X}} \underline{v}_i = \underline{v}_i$ and it's the linear transformation matrix.

$\underline{X}$ is a random variable, so $\underline{Y}$ is random variable too. So we have:

$$\mu_{\underline{Y}} = A^T \mu_{\underline{X}} = \left(\frac{1}{\lambda_i} \Sigma_{\underline{X}} \underline{v}_i\right)^T \mu_{\underline{X}} = \left(\frac{1}{\lambda_i} \lambda_i \underline{v}_i\right)^T \mu_{\underline{X}} \quad \Rightarrow \quad \mu_{\underline{Y}} = \underline{v}_i^T \mu_{\underline{X}}$$

$$\Sigma_{\underline{Y}} = A^T \Sigma_{\underline{X}} A = \left(\frac{1}{\lambda_i} \Sigma_{\underline{X}} \underline{v}_i\right)^T \Sigma_{\underline{X}} \left(\frac{1}{\lambda_i} \Sigma_{\underline{X}} \underline{v}_i\right) = \left(\frac{1}{\lambda_i} \lambda_i \underline{v}_i\right)^T \Sigma_{\underline{X}} \left(\frac{1}{\lambda_i} \lambda_i \underline{v}_i\right) \quad \Rightarrow$$

$$\Sigma_{\underline{Y}} = \underline{v}_i^T \Sigma_{\underline{X}} \underline{v}_i$$

$$p(y \mid \mu_{\underline{Y}}, \Sigma_{\underline{Y}}) = \frac{1}{(2\pi)^{(\frac{1}{2})} |\Sigma_{\underline{Y}}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(y-\mu_{\underline{Y}})^T \Sigma_{\underline{Y}}^{-1} (y-\mu_{\underline{Y}})\right\} \Rightarrow$$

$$p(y \mid \mu_{\underline{Y}}, \Sigma_{\underline{Y}}) = \frac{1}{(2\pi)^{(\frac{1}{2})} \left(\underline{v}_i^T \Sigma_{\underline{X}} \underline{v}_i\right)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(y-\underline{v}_i^T \mu_{\underline{X}})^T \left(\underline{v}_i^T \Sigma_{\underline{X}} \underline{v}_i\right)^{-1} (y-\underline{v}_i^T \mu_{\underline{X}})\right\}$$

قسمت ب)

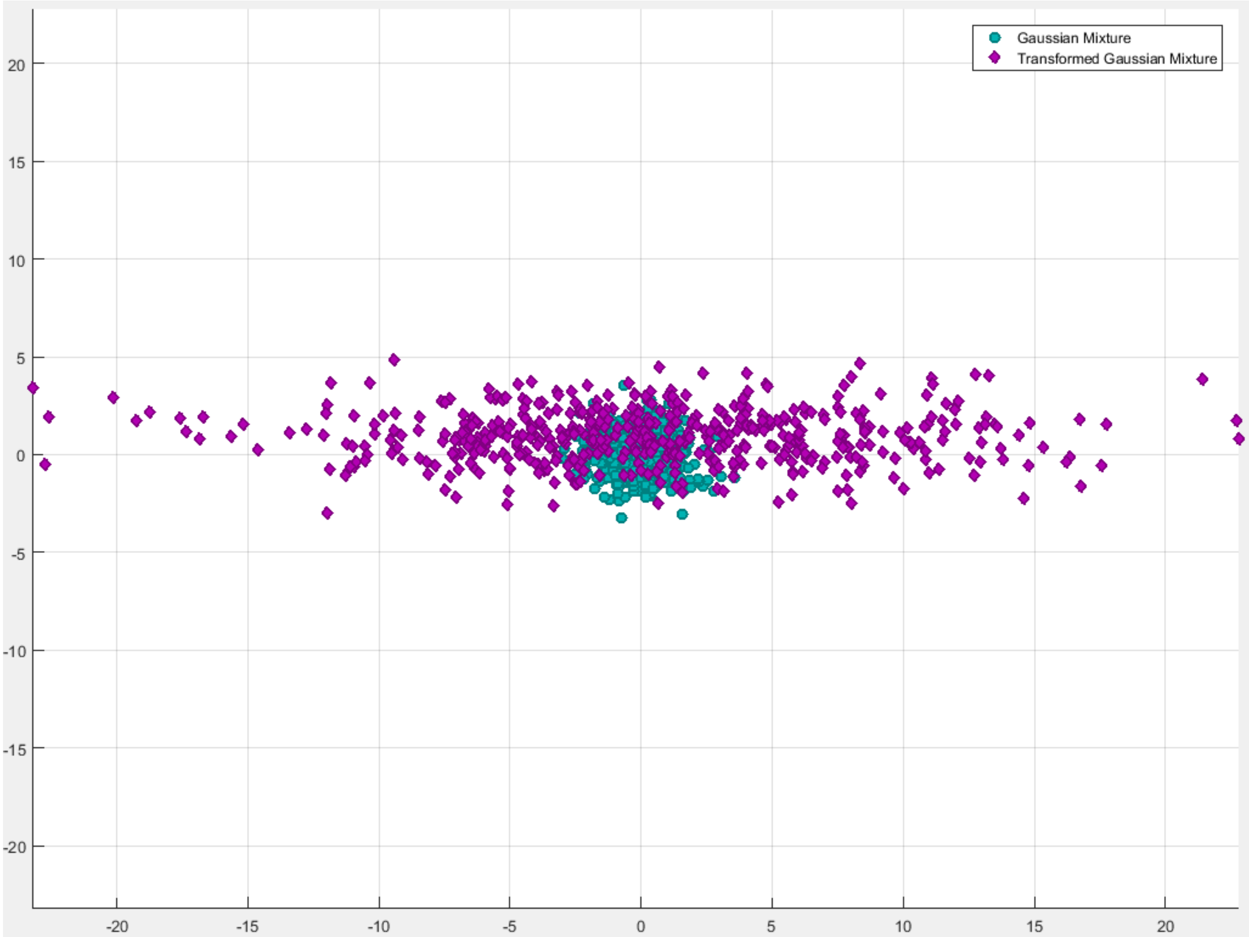
$\underline{X}$ is a gaussian random vector with zero mean and a isotropic covariance (Σ=I). Let $\underline{Y}$=A$\underline{X}$+b. Compute the mean and variance of $\underline{Y}$.

$$\mu_{\underline{X}} = 0 \rightarrow \mu_{\underline{Y}} = E\{A\underline{X} + b\} = AE\{\underline{X}\} + b = A\mu_{\underline{X}} + b \quad \Rightarrow \quad \mu_{\underline{Y}} = b$$

$$\Sigma_{\underline{Y}} = Var\{A\underline{X} + b\} = A^T \Sigma_{\underline{X}} A \quad \xrightarrow{\Sigma_{\underline{X}} = I} \quad \Sigma_{\underline{Y}} = A^T A$$

قسمت ج)

نتایج حاصله از نرم‌افزار متلب با توجه به مقادیر ارائه‌شده در صورت مسئله به شرح ذیل می‌باشد:



همانطور که از تصویر بالا مشخص است، تبدیل مربوطه سبب افزایش مقیاس یکی از بُعدها نسبت به حالت اولیه شده است.

جواب سوال ۹

$\mathrm{X\Pi Y}$ with mean λ and variance σ^2, $\mathrm{E}\left\{(X-Y)^2\right\}=?$

$$
\begin{aligned}
&E\left\{(X-Y)^2\right\}=E\left\{X^2-2XY+Y^2\right\}=E\{X^2\}-2E\{XY\}+E\{Y^2\}\\
&\rightarrow E\{X^2\}=E\{Y^2\}=\sigma^2+\lambda^2,\ \ \mathrm{X\Pi Y}\Rightarrow E\{XY\}=E\{X\}E\{Y\}=\lambda^2\\
&\Rightarrow E\left\{(X-Y)^2\right\}=2(\sigma^2+\lambda^2)-2\lambda^2\Rightarrow E\left\{(X-Y)^2\right\}=2\sigma^2
\end{aligned}
$$

جواب سوال ۱۰

برای این‌که بتوان دو ماتریس را به طور همزمان قطری نمود، می‌بایست هر دوماتریس از نوع **Hermittian** باشند. یعنی باید خود ماتریس با اصطلاحاً ماتریس ضمیمه‌ی[1] آن برابر باشند. یعنی:

$$a_{ij} = \overline{a_{ji}} \quad \text{or} \quad A = \overline{A^T}$$

حال اگر این مسئله را در مورد ماتریس‌های ارائه‌شده در مسئله بررسی نمائیم، خواهیم دید که رابطه‌ی مزبور در مورد آن‌ها صادق نمی‌باشد!؟ اما در عین حال ما به دنبال مطلوب مسئله خواهیم بود. حال برای آن‌که ماتریس‌های **Hermittian** را به طور همزمان قطری نمائیم، می‌بایست ابتدا مقادیر ویژه و بردارهای ویژه‌ی آن‌ها را به دست آورده و سپس با استفاده از این بردارهای ویژه‌ی مشترک، ماتریس جدیدی را با همان ابعاد ماتریس‌های اولیه بسازیم به طوری که هر ستون آن یکی از بردارهای ویژه‌ی مشترک باشد. حال این ماتریس جدید، ماتریسی است که ترانهاده‌ی آن با خود آن برابر است و آن را **U** می‌نامیم. در نهایت برای این‌که هر کدام از ماتریس‌های اولیه را قطری نمائیم، می‌بایست ترانهاده‌ی ماتریس **U** (که برابر با معکوس آن می‌باشد) را در ماتریس مربوطه ضرب نموده و سپس حاصل را در خود ماتریس **U** ضرب نمائیم.[2]

اما در نهایت با بررسی مقادیر ویژه و بردارهای ویژه‌ی دو ماتریس ارائه‌شده در صورت مسئله، مشاهده می‌نمائیم که هیچ بردار ویژه‌ی مشترکی را دارا نیستند!!؟ لذا در نهایت قادر نیستیم تا ماتریس **U** گفته‌شده در بالا را ساخته و ماتریس‌های مربوطه را همزمان قطری نمائیم.

ماتریس‌های ارائه‌شده در صورت مسئله و مقادیر ویژه و بردارهای ویژه‌ی آن‌ها همگی به قرار زیر می‌باشند:

$$\Sigma_1 = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} 1+\sqrt{3/4} & .5 \\ .5 & 1-\sqrt{3/4} \end{bmatrix}$$

Σ_1 eigenvalues and eigenvectors:

$$\Lambda_1 = \begin{bmatrix} .5 & 0 \\ 0 & 1.5 \end{bmatrix}, \quad V_1 = \begin{bmatrix} -.7071 & .7071 \\ .7071 & .7071 \end{bmatrix}$$

Σ_2 eigenvalues and eigenvectors:

$$\Lambda_2 = \begin{bmatrix} 5.5511e-17 \cong 0 & 0 \\ 0 & 2 \end{bmatrix}, \quad V_2 = \begin{bmatrix} .2588 & -.9659 \\ -.9659 & -.2588 \end{bmatrix}$$

همان‌طور که قابل مشاهده است، دو ماتریس مربوطه به هیچ عنوان دارای بردارهای ویژه‌ی مشترکی نمی‌باشند و در نتیجه ادامه‌ی عملیات نیز میسر نخواهد بود.

---

[1] Adjoint

[2] http://www.quantumsciencephilippines.com/216/simultaneous-diagonalization-hermitian-matrices/comment-page-1/

جواب سوال ۱۱

$$p(\underline{X}) = N_{\underline{X}}(\mu_{\underline{X}}, \Sigma_{\underline{X}}) \text{ where } \underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \mu_{\underline{X}} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma_{\underline{X}} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$p(x_1) = N_{x_1}(\mu_1, \sigma_1^2) \text{ (a marginal density)}$$

$$p(\underline{x} \mid \mu_{\underline{X}}, \Sigma_{\underline{X}}) = \frac{1}{(2\pi)^{(\frac{2}{2})} \left|\Sigma_{\underline{X}}\right|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\underline{x}-\mu_{\underline{X}})^T \Sigma_{\underline{X}}^{-1}(\underline{x}-\mu_{\underline{X}})\right\},$$

$$\text{where } \left|\Sigma_{\underline{X}}\right| = \sigma_1^2\sigma_2^2 - \rho^2\sigma_1^2\sigma_2^2 = \sigma_1^2\sigma_2^2(1-\rho^2), \quad \Sigma_{\underline{X}}^{-1} = \frac{1}{\det\left(\Sigma_{\underline{X}}\right)} adj(\Sigma_{\underline{X}}) =$$

$$\frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{bmatrix}$$

$$\Rightarrow (\underline{x}-\mu_{\underline{X}})^T \Sigma_{\underline{X}}^{-1}(\underline{x}-\mu_{\underline{X}}) = \begin{bmatrix} \mathrm{x}_1-\mu_1 & \mathrm{x}_2-\mu_2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2(1-\rho^2)} & \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} \\ \frac{-\rho}{\sigma_1\sigma_2(1-\rho^2)} & \frac{1}{\sigma_2^2(1-\rho^2)} \end{bmatrix} \begin{bmatrix} \mathrm{x}_1-\mu_1 \\ \mathrm{x}_2-\mu_2 \end{bmatrix} =$$

$$\frac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(\mathrm{x}_2-\mu_2)^2}{\sigma_2^2}$$

$$\mathrm{p}(\mathrm{x}_1,\mathrm{x}_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{z}{2\sqrt{1-\rho^2}}\right\},$$

$$\text{where } z \equiv \frac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(\mathrm{x}_2-\mu_2)^2}{\sigma_2^2}$$

$$p(\mathrm{x}_1) = \int_{-\infty}^{+\infty} \mathrm{p}(\mathrm{x}_1,\mathrm{x}_2) dx_2 =$$

$$\int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2\sqrt{1-\rho^2}}\left[\frac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(\mathrm{x}_2-\mu_2)^2}{\sigma_2^2}\right]\right\} dx_2 =$$

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2\sqrt{1-\rho^2}}\left[\frac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(\mathrm{x}_2-\mu_2)^2}{\sigma_2^2}\right]\right\} dx_2,$$

$$\mathrm{u} = \frac{\mathrm{x}_2-\mu_2}{\sigma_2} \Rightarrow du = \frac{\mathrm{dx}_2}{\sigma_2} \Rightarrow \mathrm{dx}_2 = \sigma_2 du$$

$$\Rightarrow p(\mathrm{x}_1) = \frac{\sigma_2}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{+\infty} \exp\left\{-\frac{1}{2\sqrt{1-\rho^2}}\left[\frac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(\mathrm{x}_1-\mu_1)}{\sigma_1} u + u^2\right]\right\} du = ???$$

Sorry! but it seems that I cannot solve this integral!!?? But the answer should be as follows:

$$p(\mathrm{x}_1) = \frac{1}{\sigma_1^2\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\mathrm{x}_1-\mu_1}{\sigma_1}\right)^2\right\}$$

$p(x_1 | x_2) = N_{x_1}(\mu_1 + \rho\sigma_1(\mathrm{x}_2 - \mu_2)/\sigma_2, \sigma_1^2(1-\rho^2))$ (a conditional density)

$$f(x_1 | x_2) = \frac{f(x_1, x_2)}{f(x_2)} = \frac{\dfrac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\dfrac{z}{2\sqrt{1-\rho^2}}\right\}}{\dfrac{1}{\sigma_2^2\sqrt{2\pi}} \exp\left\{-\dfrac{1}{2}\left(\dfrac{\mathrm{x}_2-\mu_2}{\sigma_2}\right)^2\right\}},$$

where $z \equiv \dfrac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \dfrac{2\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2} + \dfrac{(\mathrm{x}_2-\mu_2)^2}{\sigma_2^2}$

$$\Rightarrow f(x_1 | x_2) = \frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2}\left[\frac{z}{\sqrt{1-\rho^2}} - \left(\frac{\mathrm{x}_2-\mu_2}{\sigma_2}\right)^2\right]\right\} =$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2}\left[\frac{z}{\sqrt{1-\rho^2}} - \left(\frac{\mathrm{x}_2-\mu_2}{\sigma_2}\right)^2\right]\right\} =$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2\sqrt{1-\rho^2}}\left[\frac{(\mathrm{x}_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(\mathrm{x}_2-\mu_2)^2}{\sigma_2^2}\right] + \frac{1}{2}\left(\frac{\mathrm{x}_2-\mu_2}{\sigma_2}\right)^2\right\} =$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2\sqrt{1-\rho^2}}\left(\frac{\mathrm{x}_1-\mu_1}{\sigma_1}\right)^2 + \frac{\rho(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2)}{\sigma_1\sigma_2\sqrt{1-\rho^2}} - \frac{1-\sqrt{1-\rho^2}}{2\sqrt{1-\rho^2}}\left(\frac{\mathrm{x}_2-\mu_2}{\sigma_2}\right)^2\right\} =$$

$$\frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \times \ldots$$

$$\ldots \exp\left\{-\frac{1}{2}\left[\frac{(\mathrm{x}_1-\mu_1)^2\sqrt{1-\rho^2} - 2\sigma_1\sigma_2\rho\sqrt{1-\rho^2}(\mathrm{x}_1-\mu_1)(\mathrm{x}_2-\mu_2) + \left(\frac{\sigma_1}{\sigma_2}\right)^2\sqrt{1-\rho^2}\left(1-\sqrt{1-\rho^2}\right)(\mathrm{x}_2-\mu_2)^2}{\sigma_1^2(1-\rho^2)}\right]\right\} =$$

Sorry again! Honestly it seems that I cannot disintegrate this numerator of the fraction!

But the final result should be as follows:

$$p(x_1 | x_2) = \frac{1}{\sqrt{2\pi\sigma_1^2(1-\rho^2)}} \exp\left\{-\frac{1}{2}\left[\frac{x_1 - \left(\mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}(\mathrm{x}_2-\mu_2)\right)}{\sigma_1\sqrt{1-\rho^2}}\right]^2\right\}$$