**HW6, EE416 Autumn 2009**

**Project 5.1**

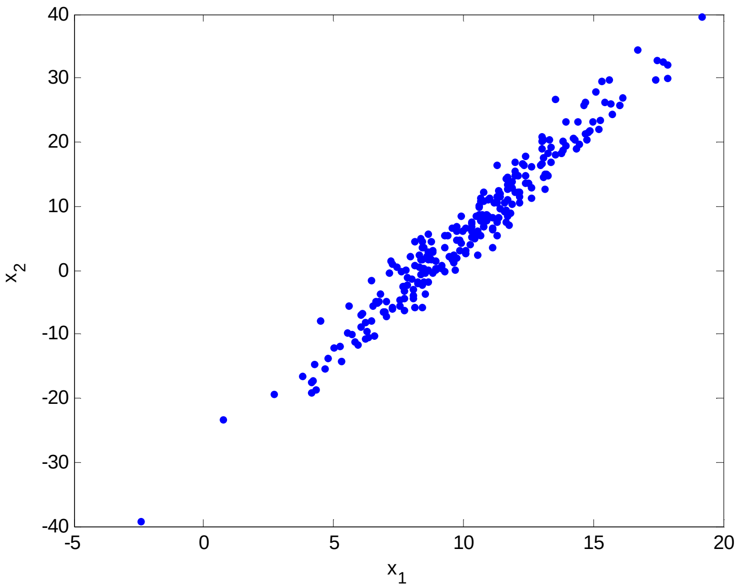
**1)** $\hat{\mathbf{m}}_X = \begin{bmatrix} 10.2014 \\ 6.4956 \end{bmatrix}$, $\hat{\mathbf{C}} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 10.2031 & 36.6203 \\ 36.6203 & 137.7496 \end{bmatrix}$

| $\hat{m}_{x_1}$ | $\hat{m}_{x_2}$ | $\hat{\sigma}_1$ | $\hat{\sigma}_2$ | $\hat{c}_{12}$ | $\hat{\rho}$ |
|---|---|---|---|---|---|
| 10.2014 | 6.4956 | 3.1942 | 11.7367 | 36.6203 | 0.9768 |

%matlab code

```
X=load('G2Data.txt');
X=X';  % size 2*256
mX=1/256*sum(X,2)
CX=1/256*X*X'-mX*mX';
sigma1=sqrt(CX(1,1));
sigma2=sqrt(CX(2,2));
rho=CX(1,2)/sigma1/sigma2;
```

**2)** Scatter plot.

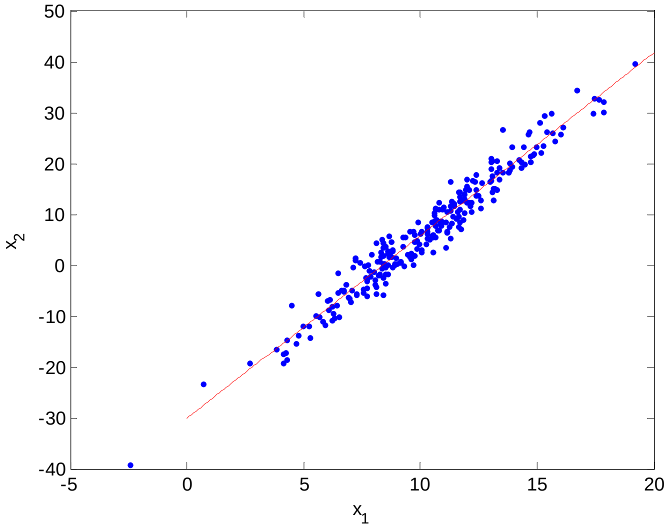



```
plot(X(1,:),X(2,:),'.')
xlabel('x_1');
ylabel('x_2');
```

**3a)** Plot the linear regression.

Linear regression is to find a line $x_2 = ax_1 + b$ so that $E\left\{\left|x_2 - (ax_1 + b)\right|^2\right\}$ is minimized.

It can be derived that the optimal line is $x_2 - m_{x_2} = \frac{\mathrm{cov}(X_1, X_2)}{\mathrm{cov}(X_1, X_1)}(x_1 - m_{x_1})$.




%matlab code

```
x1=0:0.1:20;
x2=mX(2)+rho*sigma2/sigma1* (x1-mX(1));
plot(X(1,:),X(2,:),'.',x1,x2,'r-')
xlabel('x_1');
ylabel('x_2');
```

**3b)** The ellipse on which the pdf has the same value is given by

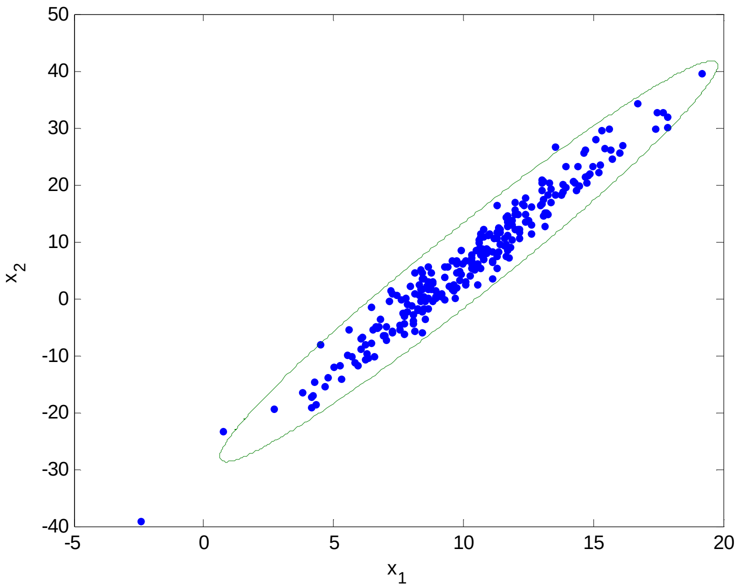
$$\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \mathbf{m}_X\right)^T \mathbf{C}^{-1} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \mathbf{m}_X\right) = \mathrm{const}.$$

I use polar coordinates to express the ellipse.

Diagonize $\mathbf{C}$ by eigen-decomposition so that $\mathbf{C} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^T = \mathbf{U}\begin{bmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{bmatrix}\mathbf{U}^T$, where $\mathbf{U}$ is a unitary matrix.

Then the ellipse is given by $\left(\mathbf{U}^T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \mathbf{m}_X\right)\right)^T \begin{bmatrix} \frac{1}{\lambda_1^2} & 0 \\ 0 & \frac{1}{\lambda_2^2} \end{bmatrix} \left(\mathbf{U}^T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \mathbf{m}_X\right)\right) = \mathrm{const}$.

Now points on the ellipse can by expressed as $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{m}_X + \mathrm{cosnt} \times \mathbf{U}\begin{bmatrix} \lambda_1 \cos(\theta) \\ \lambda_2 \sin(\theta) \end{bmatrix}, \theta \in [0, 2\pi]$.




%matlab code

```
[U,D]=eig(CX);
theta=0:0.01:2*pi
const=3;
ellipse=const*U*[sqrt(D(1,1))*cos(theta); sqrt(D(2,2))*sin(theta) ];
ellipse=repmat(mX,1,size(ellipse,2))+ellipse;
plot(X(1,:),X(2,:),'.',ellipse(1,:),ellipse(2,:))
xlabel('x_1');
ylabel('x_2');
```

**4)** The shape of the ellipse indicates the magnitude of correlation. More tighter the ellipse, higer correlation; more circular, lesser correlation. Two extreme cases: if $\rho = \pm 1$, the ellipse degrades to a line; $\rho = 0$ then the ellipse becomes a circle.

The orientation of the regression line indicates positive correlation ($\rho > 0$).

**5. 17**

$$\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-m_1)(x_2-m_2)\, f_{X_1X_2}(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2\right\}^2$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\underbrace{(x_1-m_1)\sqrt{f_{X_1X_2}(x_1,x_2)}}_{h}\underbrace{(x_2-m_2)\sqrt{f_{X_1X_2}(x_1,x_2)}}_{g}\mathrm{d}x_1\mathrm{d}x_2 ,$$

$$\le\left\{\int_{-\infty}^{\infty}(x_1-m_1)^2 f_{X_1X_2}(x_1,x_2)\mathrm{d}x_1\right\}\left\{\int_{-\infty}^{\infty}(x_2-m_2)^2 f_{X_1X_2}(x_1,x_2)\mathrm{d}x_2\right\}$$

Thus $c^2 \le \sigma_1^2\sigma_2^2$.

And $\rho^2 = \dfrac{c^2}{\sigma_1^2\sigma_2^2} \le 1$.

When $h$ and $g$ are linearly related, that is, h=ag, the equality holds. So

$$(x_1 - m_1)\sqrt{f_{X_1X_2}(x_1,x_2)} = a(x_2 - m_2)\sqrt{f_{X_1X_2}(x_1,x_2)}$$, or $x_1 = ax_2 + \text{const}$.

**5．20**

(a) $J(y) = \dfrac{1}{\left\|\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}\right\|} = 1$

(b) $X_1 = Y_1; \quad X_2 = Y_2 - aY_1$.

$$f_{Y_1Y_2}(y_1,y_2) = J(y)\frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{y_1^2 - 2\rho y_1(y_2 - ay_1) + (y_2 - ay_1)^2}{2(1-\rho^2)}}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{(1+2a\rho+a^2)y_1^2 - 2(a+\rho)y_1y_2 + y_2^2}{2(1-\rho^2)}}$$

(c) for $\rho = 1/2$, if $a = -1/2$, the cross term vanishes so

$$f_{Y_1Y_2}(y_1,y_2) = \frac{1}{\pi\sqrt{3}} e^{-\left(\frac{1}{2}y_1^2 + \frac{2}{3}y_2^2\right)} = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_1^2}}_{f_{Y_1}(y_1)} \underbrace{\frac{1}{\sqrt{2\pi}\sqrt{\frac{3}{4}}} e^{-\frac{y_2^2}{2*\frac{3}{4}}}}_{f_{Y_2}(y_2)}$$

So $Y_1$ has variance $\sigma^2 = 1$, $Y_2$ has variance $\sigma^2 = \dfrac{3}{4}$.

**5．21**

for general value of $\rho$, if $a = -\rho$, the cross term vanishes so

$$f_{Y_1Y_2}(y_1,y_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{(1-2\rho^2+\rho^2)y_1^2 + y_2^2}{2(1-\rho^2)}} = \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_1^2}}_{f_{Y_1}(y_1)} \underbrace{\frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{y_2^2}{2(1-\rho^2)}}}_{f_{Y_2}(y_2)}$$

So $Y_1$ has variance $\sigma^2 = 1$, $Y_2$ has variance $\sigma^2 = 1-\rho^2$.

**5．22**

$$f_{X_1X_2}(x_1,x_2) = \frac{1}{2\pi|\mathbf{C}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m}_X)^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m}_X)}$$

(a) $\mathbf{m}_X = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} 4 & 3.2 \\ 3.2 & 4 \end{bmatrix}$

$$f_{X_1X_2}(x_1,x_2) = \frac{1}{4.8\pi} e^{-\frac{1}{2.88}\left(x_1^2 - 1.6x_1x_2 + x_2^2\right)}$$

(b) $\mathbf{m}_X = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} s^2 & 0 \\ 0 & s^2 \end{bmatrix}$

$$f_{X_1X_2}(x_1,x_2)=\frac{1}{2\pi s^2}e^{-\frac{1}{2s^2}\left(x_1^2+x_2^2\right)}$$

(c) $\mathbf{m}_X=\begin{bmatrix}1\\1\end{bmatrix}$, $\mathbf{C}=\begin{bmatrix}9 & 0\\0 & 9\end{bmatrix}$

$$f_{X_1X_2}(x_1,x_2)=\frac{1}{18\pi}e^{-\frac{1}{18}\left((x_1-1)^2+(x_2-1)^2\right)}$$

**5．30**

(a) $M_{X_i}(s)=e^{\left(m_is+\frac{1}{2}\sigma_i^2s^2\right)}$

(b) $M_Y(s)=\prod_{i=1}^{n}M_{X_i}(s)=\prod_{i=1}^{n}e^{\left(m_is+\frac{1}{2}\sigma_i^2s^2\right)}=e^{\left(\left(\sum_{i=1}^{n}m_i\right)s+\frac{1}{2}\left(\sum_{i=1}^{n}\sigma_i^2\right)s^2\right)}$

(c) So $Y$ is a Gaussian r.v with mean $\left(\sum_{i=1}^{n}m_i\right)$ and variance $\left(\sum_{i=1}^{n}\sigma_i^2\right)$.

**5．36**

(a) $\mathbf{C}_X=\mathbf{R}_X-\mathbf{m}_X\mathbf{m}_X^T=\begin{bmatrix}4 & -1\\-1 & 3\end{bmatrix}-\begin{bmatrix}1\\1\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix}=\begin{bmatrix}3 & -2\\-2 & 2\end{bmatrix}$

(b) because $c_{12}=\rho_{X_1X_2}\sigma_{X_1}\sigma_{X_2}$, $\rho_{X_1X_2}=\frac{-2}{\sqrt{3}\sqrt{2}}=-\frac{\sqrt{6}}{3}$

(c) $\mathbf{C}_X^{-1}=\begin{bmatrix}1 & 1\\1 & \frac{3}{2}\end{bmatrix}$

$$f_{X_1X_2}(x_1,x_2)=\frac{1}{2\pi|\mathbf{C}|^{\frac{1}{2}}}e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m}_X)^T\mathbf{C}^{-1}(\mathbf{x}-\mathbf{m}_X)}=\frac{1}{2\pi\sqrt{2}}e^{-\left\{\frac{1}{2}(x_1-1)^2+(x_1-1)(x_2-1)+\frac{3}{4}(x_2-1)^2\right\}}$$

**5．37**

Using the definition,

$$\begin{aligned}\mathbf{C}_y&=E\left\{(\mathbf{y}-\mathbf{m}_y)(\mathbf{y}-\mathbf{m}_y)^T\right\}\\&=E\left\{(\mathbf{Ax}-\mathbf{Am}_x)(\mathbf{Ax}-\mathbf{Am}_x)^T\right\}\\&=E\left\{\mathbf{A}(\mathbf{x}-\mathbf{m}_x)(\mathbf{x}-\mathbf{m}_x)^T\mathbf{A}^T\right\}\\&=\mathbf{A}E\left\{(\mathbf{x}-\mathbf{m}_x)(\mathbf{x}-\mathbf{m}_x)^T\right\}\mathbf{A}^T\\&=\mathbf{AC}_x\mathbf{A}^T\end{aligned}$$

Repeat the proof using $\mathbf{C}_y=\mathbf{R}_y-\mathbf{m}_y\mathbf{m}_y^T$,

$$
\begin{aligned}
\mathbf{C}_y &= \mathbf{R}_y - \mathbf{m}_y\mathbf{m}_y^T \\
&= \mathbf{A}\mathbf{R}_x\mathbf{A}^T - (\mathbf{A}\mathbf{m}_x)(\mathbf{A}\mathbf{m}_x)^T \\
&= \mathbf{A}\left(\mathbf{R}_x - \mathbf{m}_x\mathbf{m}_x^T\right)\mathbf{A}^T \\
&= \mathbf{A}\mathbf{C}_x\mathbf{A}^T
\end{aligned}
$$

If $\mathbf{A}$ is a square matrix, then $\left|\mathbf{R}_y\right| = \left|\mathbf{A}\mathbf{R}_x\mathbf{A}^T\right| = |\mathbf{A}||\mathbf{R}_x|\left|\mathbf{A}^T\right| = |\mathbf{A}|^2|\mathbf{R}_x|$.

Grading:
25 for Project 5.1;
10 for 5.17;
15 for 5.20+5.21;
15 for 5.22;
10 for 5.30;
15 for 5.36
10 for 5.37.