1.

关于x轴方向的一般最小二乘：

$D_{LS}^{(1)}=\sum_{i=1}^{5}[m(x_i-\bar{x})+(y_i-\bar{y})]^2$，其中 $\bar{x}=4$，$\bar{y}=4.2$

故 $D_{LS}^{(1)}=(-3m-1.2)^2+(-m-3.2)^2+(m+2.8)^2+1.8^2+(3m-0.2)^2$

$\frac{\partial D_{LS}^{(1)}}{\partial m}=40m+18=0 \Rightarrow m=-0.45$

故直线方程为 $-0.45(x-4)+(y-4.2)=0$

距离平方和 $d_{LS}^{(1)}=\frac{D_{LS}^{(1)}}{1+m^2}=15.5925$

关于y轴方向的一般最小二乘：

$D_{LS}^{(2)}=\sum_{i=1}^{5}[(x_i-\bar{x})+m(y_i-\bar{y})]^2=(-3-1.2m)^2+(-1-3.2m)^2+(1+2.8m)^2+(1.8m)^2+(3-0.2m)^2$

$\frac{\partial D_{LS}^{(2)}}{\partial m}=45.6m+18=0 \Rightarrow m=-0.3947$，直线方程为 $(x-4)-0.3947(y-4.2)=0$

距离平方和 $d_{LS}^{(2)}=\frac{D_{LS}^{(2)}}{1+m^2}=14.2304$

总体最小二乘：

构造矩阵 $M=\begin{bmatrix}-3 & -1.2\\ -1 & -3.2\\ 1 & 2.8\\ 0 & 1.8\\ 3 & -0.2\end{bmatrix}$，对 M^TM 进行特征值分解，$M^TM=\begin{bmatrix}20 & 9\\ 9 & 22.8\end{bmatrix}$

$=\begin{bmatrix}0.6505 & -0.7595\\ 0.7595 & 0.6505\end{bmatrix}\begin{bmatrix}30.5082 & \\ & 12.2918\end{bmatrix}\begin{bmatrix}0.6505 & -0.7595\\ 0.7595 & 0.6505\end{bmatrix}^T$

故直线方程为：$-0.7595(x-4)+0.6505(y-4.2)=0$

距离平方和 $D_{TLS}=\left\| \begin{bmatrix} -3 & -1.2 \\ -1 & -3.2 \\ 1 & 2.8 \\ 0 & 1.8 \\ 3 & -0.2 \end{bmatrix} \begin{bmatrix} -0.7595 \\ 0.6505 \end{bmatrix} \right\|_2^2 = 12.2918$

2、

$J_W(x)=(y-Ax)^T W(y-Ax)$

$\frac{\partial J_W(x)}{\partial x}=0 \Rightarrow A^TWy=A^TWAx$

由此可得 $x_{WLS}=(A^TWA)^{-1}A^TWy$

$x_{WLS}=(A^TWA)^{-1}A^TW(Ax+e)=x+(A^TWA)^{-1}A^TWe$

估计误差的协方差矩阵可表示为

$K=E[(x-x_{WLS})(x-x_{WLS})^T]$

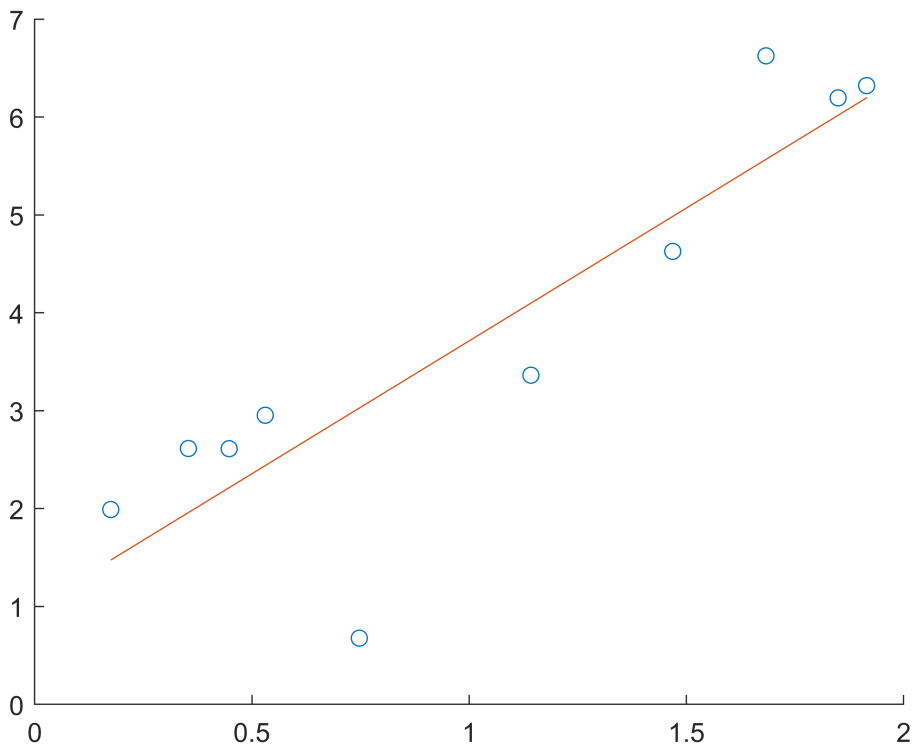
$=(A^TWA)^{-1}A^TWReW^TA((A^TWA)^{-1})^T=(A^TWA)^{-1}A^TWReWA(A^TWA)^{-1}$

$\frac{\partial tr(K)}{\partial W}=0 \Rightarrow A(A^TWA)^{-2}A^TWReWA(A^TWA)^{-1}A^T=ReWA(A^TWA)^{-2}A^T,\ W=W_{opt}=Re^{-1}$

```
% data size
n = 10;

% generate data
a = 3;
b = 1;
x = 2*rand(n,1);
y = a*x + b + randn(n,1);
x_mean = mean(x);
y_mean = mean(y);

% LS
syms m
D_LS = sum((m*(x-x_mean) + (y-y_mean)).^2) ;
delta = diff(D_LS, m);
m_value = double(solve(delta==0, m));
y_LS = -m_value*(x-x_mean) + y_mean;
scatter(x,y), hold on
plot(x, y_LS), hold off
```

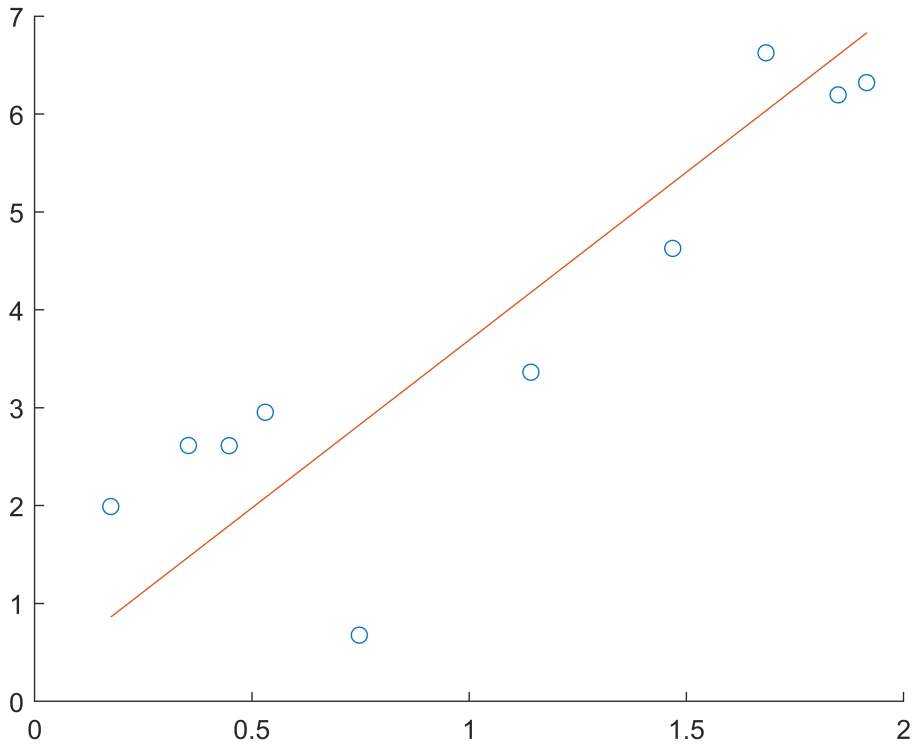



```
d_LS = double(subs(D_LS, m, m_value)) / (1+m_value^2)
```

```
d_LS = 1.0134
```

```
% TLS
M = [(x-x_mean) (y-y_mean)];
[P, D] = eig(M'*M);
[D_sort,index] = sort(diag(D),'descend');
```

```
D_sort = D_sort(index);
P_sort = P(:,index);
y_TLS = -P_sort(1,2)*(x-x_mean)/P_sort(2,2) + y_mean;
scatter(x,y), hold on
plot(x, y_TLS), hold off
```




```
d_TLS = norm(M*P_sort(:,2))^2
```

```
d_TLS = 0.8225
```