在此处键入公式。

Problem 3

Part1:

(a)

I got the meaning of the data from the readme.txt

x1: intercept term
x2: number of cylinders
x3: displacement
x4: horsepower
x5: weight
x6: acceleration
x7: model year

from the program I got

```
Wls =

   23.4282
   -0.6668
    0.9898
    0.0974
   -5.9046
    0.3229
    2.7529
```

In that case,

The number of cylinders and weight of a car draw a negative correlation to the result; i.e. with the number of cylinders go higher and weight got higher, the miles per gallon will go down. Other indexes draw positive correlation to the result, i.e. if they goes larger, the result of miles per gallon go higher.

Part1:

(b)

The result of Mean and standard deviation of MAE are listed below:

```
MAE_Mean =

    8.5719


MAE_StDeviation =
    1.3733
```

Part2:

(a)

Here I attached 2 examples of the result of the program

P = 1:

```
RMSE_Mean =

   10.4892


RMSE_StDeviation =

    1.3803
```

P = 2:

```
RMSE_Mean =

   10.6616


RMSE_StDeviation =

    0.3453
```

P = 3:

```
RMSE_Mean =

   10.5077


RMSE_StDeviation =

    1.1251
```

P = 4:

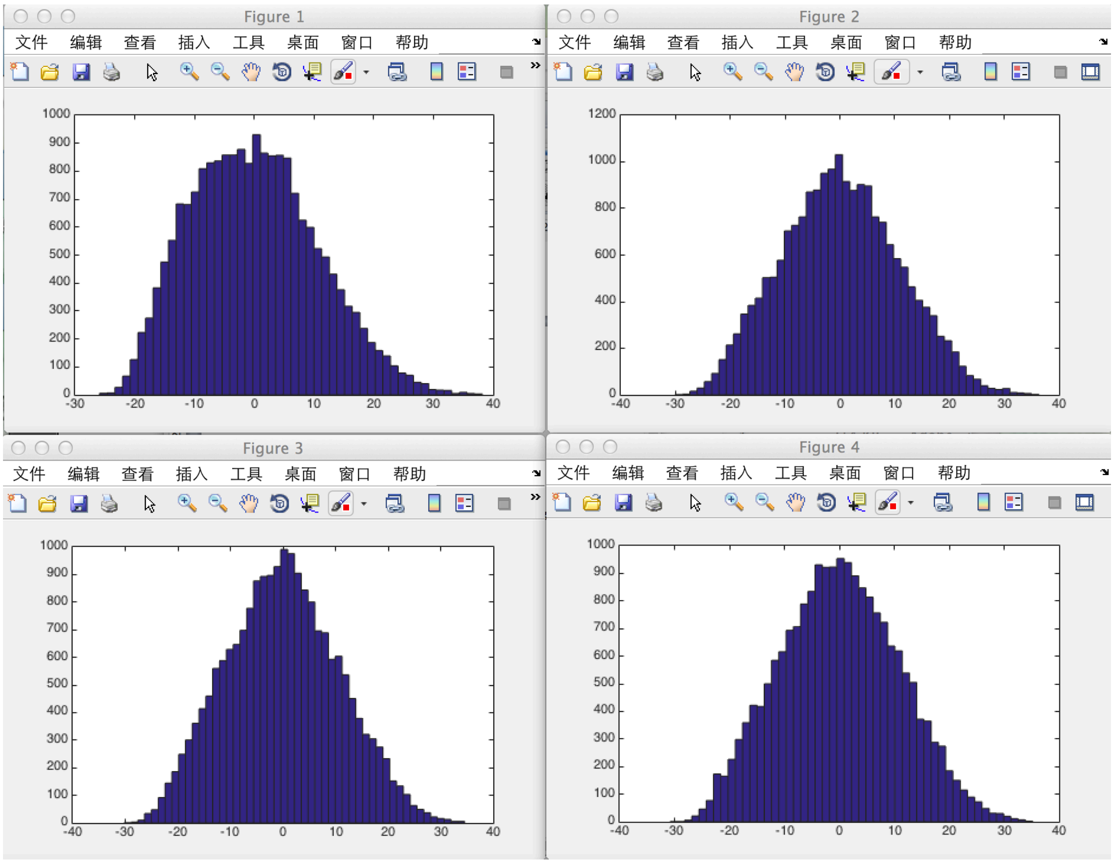
```
RMSE_Mean =

   10.6404


P = 1:

RMSE_Mean =

   10.4082


RMSE_StDeviation =

    1.3495

P = 2:

RMSE_Mean =

   10.6810


RMSE_StDeviation =

    1.6637

P = 3:

RMSE_Mean =

   10.6607


RMSE_StDeviation =

    1.5144

P = 4:

RMSE_Mean =

   10.6652
```

```
RMSE_StDeviation =

    0.5380



RMSE_StDeviation =

    0.4450
```

From the examples above and my other observations , for p=1,2,3,4 the mean of RMSE is roughly the same. However, the standard deviation varies significantly with respect to p. We can see from the examples, when p=3 or p=4, the standard deviations are significantly smaller than the others. With some more examples when I run programs, p=3 is better than p=4 on most occation. So p=3 maybe the best choice.

(b)

The histograms are listed below, figure k is for p=k;



(c)

The Maximum Likelihood obey the following equation

$$\text{ML:} \quad \arg\max_{w} \ -\frac{1}{2\sigma^2}\|y - Xw\|^2$$

So with this criterion, and the deriving method of in lecture, here is the result of the programming results :

```
P = 1:

mean =

    0.0906


variance =

  108.1178

P = 2:

mean =

   -0.1221


variance =

  113.9527

P = 3:

mean =

   -0.0444


variance =

  114.7182

P = 4:
```

```
mean =

    0.0768


variance =

  114.8160
```

For log likelihood function, the mean and variance still fit the former Guassian distribution. So the expected means and variance remains the same.

From the result above, we can see that when p=3, the means and variance is the smallest, so it correspond the conclusion we have in (a), i.e. p=3 may be the best choice.

Source Code Attachment

Problem 3, part 1, (a)

```
originX=importdata('X.txt');
originY=importdata('Y.txt');
% X1=transpose(originX)
% X2=inv(originX)
% randint(372,1, [1 ,392])
% originY(1,:)
All(:,1)=originY;
All(:,2:8)=originX;

nRows=size(All,1);
randRows=randperm(nRows);%# generate random ordering of row indices
AllTrain=All(randRows(1:372),:);
AllTest=All(randRows(373:end),:);
YTrain=AllTrain(:,1);
XTrain=AllTrain(:,2:8);
YTest=AllTest(:,1);
```

```
XTest=AllTrain(:,2:8);

Wls=inv(transpose(XTrain)*XTrain)*transpose(XTrain)*YTrain
```

Problem 3, part 1, (b)

```
originX=importdata('X.txt');
originY=importdata('Y.txt');
% X1=transpose(originX)
% X2=inv(originX)
% randint(372,1, [1 ,392])
% originY(1,:)
All(:,1)=originY;
All(:,2:8)=originX;
MAE=1:1000;

for i=1:1000
    nRows=size(All,1);
    randRows=randperm(nRows);%# generate random ordering of row indices
    AllTrain=All(randRows(1:372),:);
    AllTest=All(randRows(373:end),:);
    YTrain=AllTrain(:,1);
    XTrain=AllTrain(:,2:8);
    YTest=AllTest(:,1);
    XTest=AllTrain(:,2:8);
    Wls=inv(transpose(XTrain)*XTrain)*transpose(XTrain)*YTrain;
    sum1=0;
    for j=1:20
        sum1=sum1+abs(YTest(j,:)-XTest(j,:)*Wls);
    end
    MAE(1,i)=sum1/20;
end

sum2=0;
for k=1:1000
    sum2=sum2+MAE(1,k);
end
MAE_Mean=sum2/1000;
```

Problem 3, part 2, (a)

```
originX=importdata('X.txt');
```

```
originY=importdata('Y.txt');
% X1=transpose(originX)
% X2=inv(originX)
% randint(372,1, [1 ,392])
% originY(1,:)
originX2=originX.*originX;
originX3=originX2.*originX;
originX4=originX3.*originX;
All(:,1)=originY;
All(:,2:8)=originX;
All(:,9:15)=originX2;
All(:,16:22)=originX3;
All(:,23:29)=originX4;
RMSE=1:1000;

for p=1:4

disp(['P = ',num2str(p),':'])
for i=1:1000
   nRows=size(All,1);
   randRows=randperm(nRows);
   AllTrain=All(randRows(1:372),:);
   AllTest=All(randRows(373:end),:);
   YTrain=AllTrain(:,1);
   XTrain=AllTrain(:,2:7*p+1);
   YTest=AllTest(:,1);
   XTest=AllTrain(:,2:7*p+1);
   Wls=pinv(transpose(XTrain)*XTrain)*transpose(XTrain)*YTrain;
   sum1=0;
   for j=1:20
      sum1=sum1+(YTest(j,:)-XTest(j,:)*Wls)^2;
   end
   RMSE(1,i)=sqrt(sum1/20);
end

sum2=0;
for k=1:1000
   sum2=sum2+RMSE(1,k);
end
RMSE_Mean=sum2/1000;

sum3=0;
for l=1:1000
   sum3=sum3+(RMSE(1,i)-RMSE_Mean)^2;
```

```
end
RMSE_StDeviation=sqrt(sum3/1000);

RMSE_Mean
RMSE_StDeviation
end
```

Problem 3, part 2, (b)

```
originX=importdata('X.txt');
originY=importdata('Y.txt');
% X1=transpose(originX)
% X2=inv(originX)
% randint(372,1, [1 ,392])
% originY(1,:)
originX2=originX.*originX;
originX3=originX2.*originX;
originX4=originX3.*originX;
All(:,1)=originY;
All(:,2:8)=originX;
All(:,9:15)=originX2;
All(:,16:22)=originX3;
All(:,23:29)=originX4;
Error=20:1000;

for p=1:4

disp(['P = ',num2str(p),':'])
for i=1:1000
    nRows=size(All,1);
    randRows=randperm(nRows);
    AllTrain=All(randRows(1:372),:);
    AllTest=All(randRows(373:end),:);
    YTrain=AllTrain(:,1);
    XTrain=AllTrain(:,2:7*p+1);
    YTest=AllTest(:,1);
    XTest=AllTrain(:,2:7*p+1);
    Wls=pinv(transpose(XTrain)*XTrain)*transpose(XTrain)*YTrain;
    for j=1:20
        Error(j,i)=YTest(j,:)-XTest(j,:)*Wls;
    end
end
```

```
figure; hist(Error(:),50);

end
```

Problem 3, part 2, (c)

```
originX=importdata('X.txt');
originY=importdata('Y.txt');
% X1=transpose(originX)
% X2=inv(originX)
% randint(372,1, [1 ,392])
% originY(1,:)
originX2=originX.*originX;
originX3=originX2.*originX;
originX4=originX3.*originX;
All(:,1)=originY;
All(:,2:8)=originX;
All(:,9:15)=originX2;
All(:,16:22)=originX3;
All(:,23:29)=originX4;
Error=20:1000;

for p=1:4

disp(['P = ',num2str(p),':'])
for i=1:1000
   nRows=size(All,1);
   randRows=randperm(nRows);
   AllTrain=All(randRows(1:372),:);
   AllTest=All(randRows(373:end),:);
   YTrain=AllTrain(:,1);
   XTrain=AllTrain(:,2:7*p+1);
   YTest=AllTest(:,1);
   XTest=AllTrain(:,2:7*p+1);
   Wls=pinv(transpose(XTrain)*XTrain)*transpose(XTrain)*YTrain;
   for j=1:20
      Error(j,i)=YTest(j,:)-XTest(j,:)*Wls;
   end
end

% figure; hist(Error(:),50);
sum1=0;
```

```
sum2=0;

for j=1:1000
    for i=1:20
        sum1=sum1+Error(i,j);
    end
end
mean=sum1/20000

for j=1:1000
    for i=1:20
        sum1=sum1+(Error(i,j)-mean)^2;
    end
end
variance=sum1/20000

end
```