%Project
%Qiuying Li
%05/05/2016

%Result

% Summary of problem 1

%first 30 vaues represent omega:
%Columns 1 through 10

%-4.1056 -0.1242 -3.3016 -1.6735 2.4501 -4.8508 -0.3210 3.7928 -1.2820 0.2479

%Columns 11 through 20

%4.1126 -0.8100 -0.3688 3.8027 0.5084 -0.0471 -0.1307 4.2478 0.7133 -7.1841

%Columns 21 through 30

%6.5976 5.0381 4.1452 8.0556 -1.6099 0.1599 -0.3038 0.6853 0.6642 6.4163

%31st value represents gamma:-3.5688

%The object value is 0.0459

%Problem 2

% Report the number of misclassi ed points on the tuning set:
% When mu equals to : 5.000000e-05
%In the 100 cases of tuning set, misclassified number is   3
%The testing set error is 3.665006e+00
%When mu equals to : 1.000000e-04
%In the 100 cases of tuning set, misclassified number is   3
%The testing set error is 5.322415e+00
%When mu equals to : 1.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 6.458237e+00
%When mu equals to : 2.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 7.614409e+00
%When mu equals to : 2.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 8.139580e+00
%When mu equals to : 3.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 8.419897e+00
%When mu equals to : 3.500000e-04

%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 8.719700e+00
%When mu equals to : 4.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 9.334193e+00
%When mu equals to : 4.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 9.890028e+00
%When mu equals to : 5.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 1.028280e+01

%summary of problem 2
%Compare all the results from 10 mu, we can see that when u becomes larger,
%the misclassified number decrese. In addition, the testing error becomes
%smaller.
%compare all the 10 mu, 5e-4 has the best performance.
%The testing error for mu = 5e-4 eauqls to 10.2828 (1.028280e+01)
%The misclassified number when mu = 5e-4 equals to 2

%Problem 3
atts 2 3: misclass 100
atts 2 4: misclass 13
atts 2 5: misclass   8
atts 2 6: misclass   8
atts 2 7: misclass   7
atts 2 8: misclass   6
atts 2 9: misclass   6
atts 2 10: misclass   6
atts 2 11: misclass   6
atts 2 12: misclass   6
atts 2 13: misclass   6
atts 2 14: misclass   6
atts 2 15: misclass   6
atts 2 16: misclass   5
atts 2 17: misclass   5
atts 2 18: misclass   5
atts 2 19: misclass   5
atts 2 20: misclass   5
atts 2 21: misclass   5
atts 2 22: misclass   5
atts 2 23: misclass   2
atts 2 24: misclass   2
atts 2 25: misclass   2
atts 2 26: misclass   2
atts 2 27: misclass   2
atts 2 28: misclass   2
atts 2 29: misclass   2
atts 2 30: misclass   2

```
atts  3  4: misclass   2
atts  3  5: misclass   2
atts  3  6: misclass   2
atts  3  7: misclass   2
atts  3  8: misclass   2
atts  3  9: misclass   2
atts  3 10: misclass   2
atts  3 11: misclass   2
atts  3 12: misclass   2
atts  3 13: misclass   2
atts  3 14: misclass   2
atts  3 15: misclass   2
atts  3 16: misclass   2
atts  3 17: misclass   2
atts  3 18: misclass   2
atts  3 19: misclass   2
atts  3 20: misclass   2
atts  3 21: misclass   2
atts  3 22: misclass   2
atts  3 23: misclass   2
atts  3 24: misclass   2
atts  3 25: misclass   2
atts  3 26: misclass   2
atts  3 27: misclass   2
atts  3 28: misclass   2
atts  3 29: misclass   2
atts  3 30: misclass   2
atts  4  5: misclass   2
atts  4  6: misclass   2
atts  4  7: misclass   2
atts  4  8: misclass   2
atts  4  9: misclass   2
atts  4 10: misclass   2
atts  4 11: misclass   2
atts  4 12: misclass   2
atts  4 13: misclass   2
atts  4 14: misclass   2
atts  4 15: misclass   2
atts  4 16: misclass   2
atts  4 17: misclass   2
atts  4 18: misclass   2
atts  4 19: misclass   2
atts  4 20: misclass   2
atts  4 21: misclass   2
atts  4 22: misclass   2
atts  4 23: misclass   2
atts  4 24: misclass   2
atts  4 25: misclass   2
atts  4 26: misclass   2
```

atts 4 27: misclass 2
atts 4 28: misclass 2
atts 4 29: misclass 2
atts 4 30: misclass 2
atts 5 6: misclass 2
atts 5 7: misclass 2
atts 5 8: misclass 2
atts 5 9: misclass 2
atts 5 10: misclass 2
atts 5 11: misclass 2
atts 5 12: misclass 2
atts 5 13: misclass 2
atts 5 14: misclass 2
atts 5 15: misclass 2
atts 5 16: misclass 2
atts 5 17: misclass 2
atts 5 18: misclass 2
atts 5 19: misclass 2
atts 5 20: misclass 2
atts 5 21: misclass 2
atts 5 22: misclass 2
atts 5 23: misclass 2
atts 5 24: misclass 2
atts 5 25: misclass 2
atts 5 26: misclass 2
atts 5 27: misclass 2
atts 5 28: misclass 2
atts 5 29: misclass 2
atts 5 30: misclass 2
atts 6 7: misclass 2
atts 6 8: misclass 2
atts 6 9: misclass 2
atts 6 10: misclass 2
atts 6 11: misclass 2
atts 6 12: misclass 2
atts 6 13: misclass 2
atts 6 14: misclass 2
atts 6 15: misclass 2
atts 6 16: misclass 2
atts 6 17: misclass 2
atts 6 18: misclass 2
atts 6 19: misclass 2
atts 6 20: misclass 2
atts 6 21: misclass 2
atts 6 22: misclass 2
atts 6 23: misclass 2
atts 6 24: misclass 2
atts 6 25: misclass 2
atts 6 26: misclass 2

atts 6 27: misclass 2
atts 6 28: misclass 2
atts 6 29: misclass 2
atts 6 30: misclass 2
atts 7 8: misclass 2
atts 7 9: misclass 2
atts 7 10: misclass 2
atts 7 11: misclass 2
atts 7 12: misclass 2
atts 7 13: misclass 2
atts 7 14: misclass 2
atts 7 15: misclass 2
atts 7 16: misclass 2
atts 7 17: misclass 2
atts 7 18: misclass 2
atts 7 19: misclass 2
atts 7 20: misclass 2
atts 7 21: misclass 2
atts 7 22: misclass 2
atts 7 23: misclass 2
atts 7 24: misclass 2
atts 7 25: misclass 2
atts 7 26: misclass 2
atts 7 27: misclass 2
atts 7 28: misclass 2
atts 7 29: misclass 2
atts 7 30: misclass 2
atts 8 9: misclass 2
atts 8 10: misclass 2
atts 8 11: misclass 2
atts 8 12: misclass 2
atts 8 13: misclass 2
atts 8 14: misclass 2
atts 8 15: misclass 2
atts 8 16: misclass 2
atts 8 17: misclass 2
atts 8 18: misclass 2
atts 8 19: misclass 2
atts 8 20: misclass 2
atts 8 21: misclass 2
atts 8 22: misclass 2
atts 8 23: misclass 2
atts 8 24: misclass 2
atts 8 25: misclass 2
atts 8 26: misclass 2
atts 8 27: misclass 2
atts 8 28: misclass 2
atts 8 29: misclass 2
atts 8 30: misclass 2

atts 9 10: misclass   2
atts 9 11: misclass   2
atts 9 12: misclass   2
atts 9 13: misclass   2
atts 9 14: misclass   2
atts 9 15: misclass   2
atts 9 16: misclass   2
atts 9 17: misclass   2
atts 9 18: misclass   2
atts 9 19: misclass   2
atts 9 20: misclass   2
atts 9 21: misclass   2
atts 9 22: misclass   2
atts 9 23: misclass   2
atts 9 24: misclass   2
atts 9 25: misclass   2
atts 9 26: misclass   2
atts 9 27: misclass   2
atts 9 28: misclass   2
atts 9 29: misclass   2
atts 9 30: misclass   2
atts 10 11: misclass   2
atts 10 12: misclass   2
atts 10 13: misclass   2
atts 10 14: misclass   2
atts 10 15: misclass   2
atts 10 16: misclass   2
atts 10 17: misclass   2
atts 10 18: misclass   2
atts 10 19: misclass   2
atts 10 20: misclass   2
atts 10 21: misclass   2
atts 10 22: misclass   2
atts 10 23: misclass   2
atts 10 24: misclass   2
atts 10 25: misclass   2
atts 10 26: misclass   2
atts 10 27: misclass   2
atts 10 28: misclass   2
atts 10 29: misclass   2
atts 10 30: misclass   2
atts 11 12: misclass   2
atts 11 13: misclass   2
atts 11 14: misclass   2
atts 11 15: misclass   2
atts 11 16: misclass   2
atts 11 17: misclass   2
atts 11 18: misclass   2
atts 11 19: misclass   2

atts 11 20: misclass 2
atts 11 21: misclass 2
atts 11 22: misclass 2
atts 11 23: misclass 2
atts 11 24: misclass 2
atts 11 25: misclass 2
atts 11 26: misclass 2
atts 11 27: misclass 2
atts 11 28: misclass 2
atts 11 29: misclass 2
atts 11 30: misclass 2
atts 12 13: misclass 2
atts 12 14: misclass 2
atts 12 15: misclass 2
atts 12 16: misclass 1
atts 12 17: misclass 1
atts 12 18: misclass 1
atts 12 19: misclass 1
atts 12 20: misclass 1
atts 12 21: misclass 1
atts 12 22: misclass 1
atts 12 23: misclass 1
atts 12 24: misclass 1
atts 12 25: misclass 1
atts 12 26: misclass 1
atts 12 27: misclass 1
atts 12 28: misclass 1
atts 12 29: misclass 1
atts 12 30: misclass 1
atts 13 14: misclass 1
atts 13 15: misclass 1
atts 13 16: misclass 1
atts 13 17: misclass 1
atts 13 18: misclass 1
atts 13 19: misclass 1
atts 13 20: misclass 1
atts 13 21: misclass 1
atts 13 22: misclass 1
atts 13 23: misclass 1
atts 13 24: misclass 1
atts 13 25: misclass 1
atts 13 26: misclass 1
atts 13 27: misclass 1
atts 13 28: misclass 1
atts 13 29: misclass 1
atts 13 30: misclass 1
atts 14 15: misclass 1
atts 14 16: misclass 1
atts 14 17: misclass 1

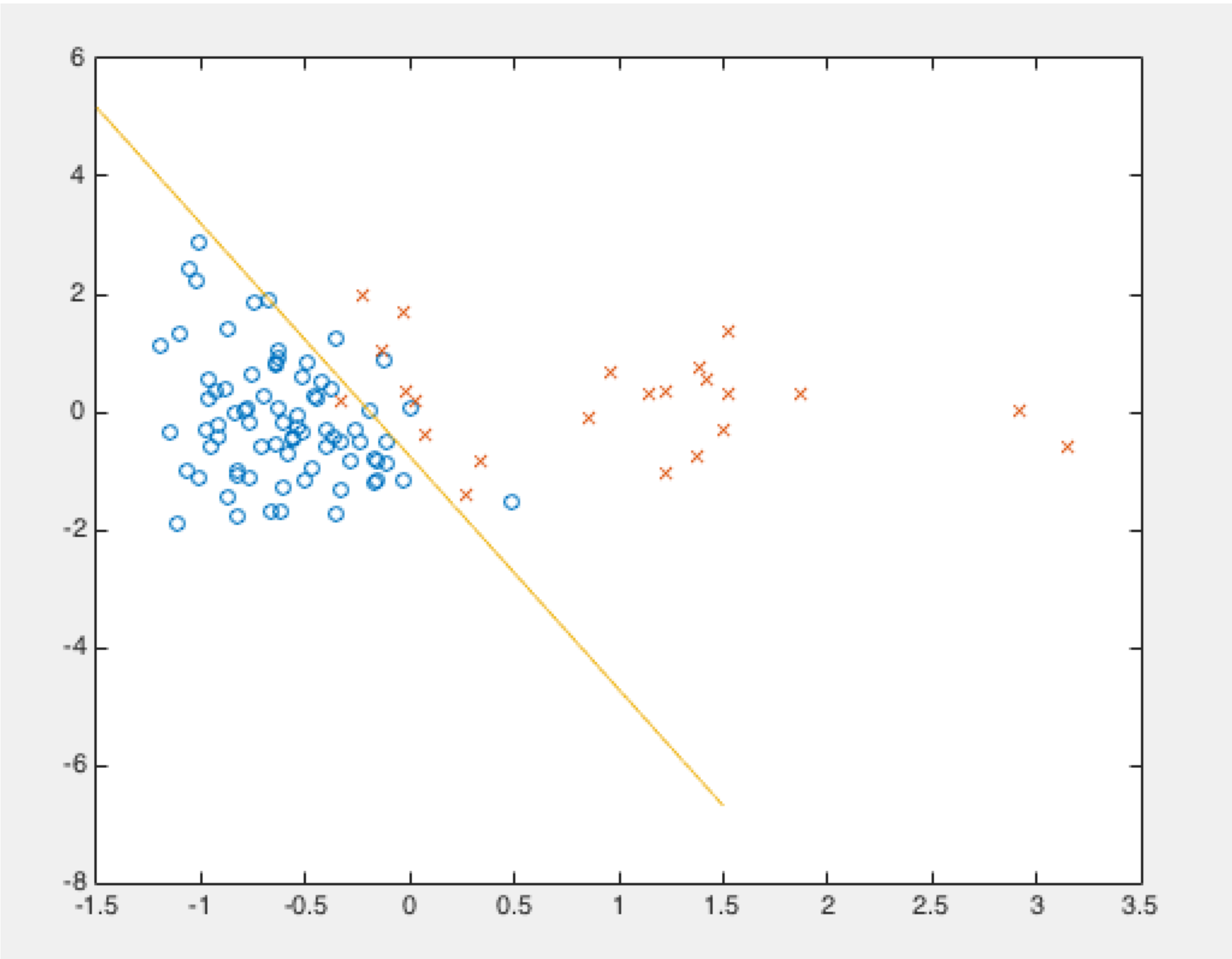
atts 14 18: misclass 1
atts 14 19: misclass 1
atts 14 20: misclass 1
atts 14 21: misclass 1
atts 14 22: misclass 1
atts 14 23: misclass 1
atts 14 24: misclass 1
atts 14 25: misclass 1
atts 14 26: misclass 1
atts 14 27: misclass 1
atts 14 28: misclass 1
atts 14 29: misclass 1
atts 14 30: misclass 1
atts 15 16: misclass 1
atts 15 17: misclass 1
atts 15 18: misclass 1
atts 15 19: misclass 1
atts 15 20: misclass 1
atts 15 21: misclass 1
atts 15 22: misclass 1
atts 15 23: misclass 1
atts 15 24: misclass 1
atts 15 25: misclass 1
atts 15 26: misclass 1
atts 15 27: misclass 1
atts 15 28: misclass 1
atts 15 29: misclass 1
atts 15 30: misclass 1
atts 16 17: misclass 1
atts 16 18: misclass 1
atts 16 19: misclass 1
atts 16 20: misclass 1
atts 16 21: misclass 1
atts 16 22: misclass 1
atts 16 23: misclass 1
atts 16 24: misclass 1
atts 16 25: misclass 1
atts 16 26: misclass 1
atts 16 27: misclass 1
atts 16 28: misclass 1
atts 16 29: misclass 1
atts 16 30: misclass 1
atts 17 18: misclass 1
atts 17 19: misclass 1
atts 17 20: misclass 1
atts 17 21: misclass 1
atts 17 22: misclass 1
atts 17 23: misclass 1
atts 17 24: misclass 1

atts 17 25: misclass 1
atts 17 26: misclass 1
atts 17 27: misclass 1
atts 17 28: misclass 1
atts 17 29: misclass 1
atts 17 30: misclass 1
atts 18 19: misclass 1
atts 18 20: misclass 1
atts 18 21: misclass 1
atts 18 22: misclass 1
atts 18 23: misclass 1
atts 18 24: misclass 1
atts 18 25: misclass 1
atts 18 26: misclass 1
atts 18 27: misclass 1
atts 18 28: misclass 1
atts 18 29: misclass 1
atts 18 30: misclass 1
atts 19 20: misclass 1
atts 19 21: misclass 1
atts 19 22: misclass 1
atts 19 23: misclass 1
atts 19 24: misclass 1
atts 19 25: misclass 1
atts 19 26: misclass 1
atts 19 27: misclass 1
atts 19 28: misclass 1
atts 19 29: misclass 1
atts 19 30: misclass 1
atts 20 21: misclass 1
atts 20 22: misclass 1
atts 20 23: misclass 1
atts 20 24: misclass 1
atts 20 25: misclass 1
atts 20 26: misclass 1
atts 20 27: misclass 1
atts 20 28: misclass 1
atts 20 29: misclass 1
atts 20 30: misclass 1
atts 21 22: misclass 1
atts 21 23: misclass 1
atts 21 24: misclass 1
atts 21 25: misclass 1
atts 21 26: misclass 1
atts 21 27: misclass 1
atts 21 28: misclass 1
atts 21 29: misclass 1
atts 21 30: misclass 1
atts 22 23: misclass 1

```
atts 22 24: misclass   1
atts 22 25: misclass   1
atts 22 26: misclass   1
atts 22 27: misclass   1
atts 22 28: misclass   1
atts 22 29: misclass   1
atts 22 30: misclass   1
atts 23 24: misclass   1
atts 23 25: misclass   1
atts 23 26: misclass   1
atts 23 27: misclass   1
atts 23 28: misclass   1
atts 23 29: misclass   1
atts 23 30: misclass   1
atts 24 25: misclass   1
atts 24 26: misclass   1
atts 24 27: misclass   1
atts 24 28: misclass   1
atts 24 29: misclass   1
atts 24 30: misclass   1
atts 25 26: misclass   1
atts 25 27: misclass   1
atts 25 28: misclass   1
atts 25 29: misclass   1
atts 25 30: misclass   1
atts 26 27: misclass   1
atts 26 28: misclass   1
atts 26 29: misclass   1
atts 26 30: misclass   1
atts 27 28: misclass   1
atts 27 29: misclass   1
atts 27 30: misclass   1
atts 28 29: misclass   1
atts 28 30: misclass   1
atts 29 30: misclass   1
atts 25 26 has the minimum misclass   1
```

```
%Summary of Problem 3
%Based on the all the caces of misclassification and test error
%I think 25,26 is the best.
```

% Problem 4
% Use the best performing answer from Part 3 above;
% find and print out the number of misclassied points on the testing set;
% plot all the testing set points on a two dimensional figure



%Summary of problem 4
%According to the description. 'o' stands for the benign points;
%'x'stands for the malignant points.
%Based on the polt, I find there are 4 'o' on the 'x' part, 1 'x' on the 'o' part;
%Moreover, there are 2 'o' on the line, and another 'o' is very close to
%the line.
%Thus, the number of misclassi ed points agrees with the plot and comment.

Code and comment

```
%problem 1
%Formulate  the  problem  as  a  quadratic  program.
%Solve  the  problem using the matrices M and B as training set using
%first 369 cases of of the wdbc.data
%Make sure you print out Omega and Gamma and the minimum value of the QP

[train,tune,test] = getdata('wdbc.data',30);
label = train(:,1);
mu =0.0001;
M = train(find(label=='M'),2:31);
B = train(find(label =='B'),2:31);

% set up the two matrices M and B, and set mu = 0.001.
% Matrix M stands for the malignant FNA.
% Matrix B stands for the benign FNA.

size_m = size(M,1);
size_b = size(B,1);
b_m = ones(size_m,1);
b_n = ones(size_b,1);
Q = eye(30)*mu;
Q = [Q, zeros(30,370)];
Q = [Q; zeros(370,400)];
c = [b_m'/size_m, b_n'/size_b];
c=[zeros(size(c,1),31) c];
H = [zeros(1,400)];
g = [0];
b = [-b_m',-b_n']';
A = [-M, b_m, diag(-b_m), zeros(size_m,size_b);B,-b_n,zeros(size_b,size_m),diag(-b_n)];
lb = [-inf(31, 1); zeros(369,1)];
ub = [inf(400,1)];
[x,obj]=cplexqp(Q,c',A,b,H,g,lb,ub);

% set up for the cplexqp function.

w=x(1:30)
r=x(31)


% Summary of problem 1

%first 30 vaues represent omega:
%Columns 1 through 10

%-4.1056 -0.1242 -3.3016 -1.6735  2.4501 -4.8508 -0.3210 3.7928 -1.2820 0.2479

%Columns 11 through 20
```

%4.1126 -0.8100 -0.3688 3.8027 0.5084 -0.0471 -0.1307 4.2478 0.7133 -7.1841

%Columns 21 through 30

%6.5976 5.0381 4.1452 8.0556 -1.6099 0.1599 -0.3038 0.6853 0.6642 6.4163

%31th value represents gamma:-3.5688

%The object value is 0.0459

% Problem 2

% Test the separating plane on the 100 cases of the tuning set.
% Report the number of misclassi ed points on the tuning set.
% What is the effect of mu?
% What is the best value of mu from this set?
% What is the testing set error
% What is the number of misclassifed points for this choice of mu?

```
clear;
for mu=[5e-5,1e-4,1.5e-4,2e-4,2.5e-4,3e-4,3.5e-4,4e-4,4.5e-4,5e-4];
[train,tune,test] = getdata('wdbc.data',30);
label = train(:,1);
M = train(find(label=='M'),2:31);
B = train(find(label =='B'),2:31);
m = size(M,1);
k = size(B,1);
b_m = ones(m,1);
b_n = ones(k,1);
Q = eye(30)*mu;
Q = [Q, zeros(30,370)];
Q = [Q; zeros(370,400)];
c = [b_m'/m, b_n'/k];
c=[zeros(size(c,1),31) c];
H = [zeros(1,400)];
g = [0];
b = [-b_m',-b_n']';
A = [-M, b_m, diag(-b_m), zeros(m,k);B,-b_n,zeros(k,m),diag(-b_n)];
lb = [-inf(31, 1); zeros(369,1)];
ub = [inf(400,1)];
[x,obj]=cplexqp(Q,c',A,b,H,g,lb,ub);
w=x(1:30);
r=x(31);
% set up for the cplexqp function,and get the reletive omega and gamma for
% each mu.
```

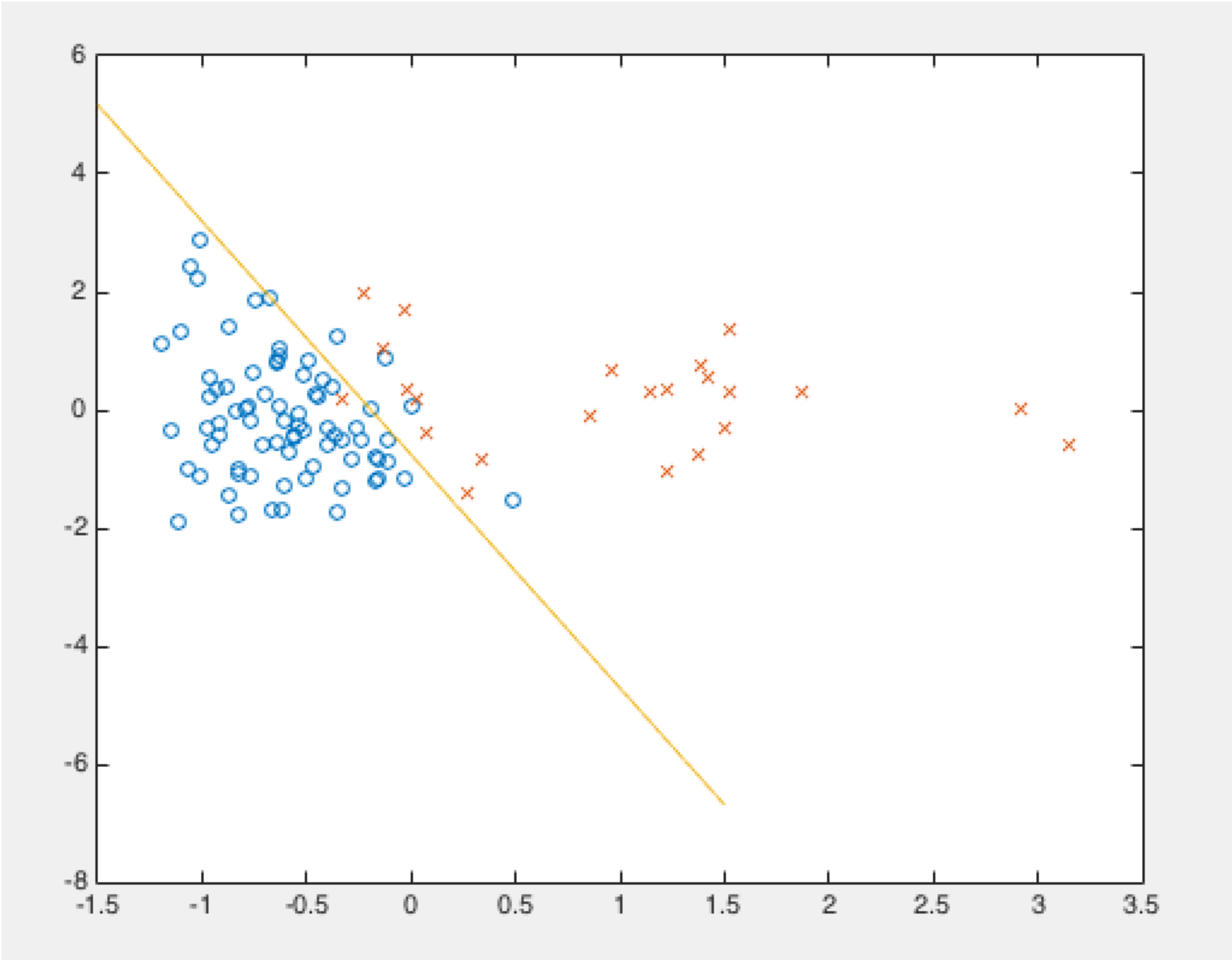
```
%Since there are 10 mu in total, so I created a loop to calculate 10 times
%for different mu

v=x(32:31+164);
t=x(32+164:369+31);
test_error =sum(v)+sum(t);

% According to the problem descrbtion, the sum of the distance from each
%point to the plane is the testing error.

mis_counter=0;
total_error=0;
error_M=0;
error_B=0;

% set up for the loop, which aims to determine the misclassfied points of
% Matrix B and Matrix M

for i = 1:100
    if w'*tune(i,2:31)'-r > 0
      if tune(i,1) == 77
      elseif tune(i,1)==66
            mis_counter =mis_counter+1;
            error_M = error_M-w'*tune(i,2:31)'+r;
          end
    else
           if w'*tune(i,2:31)'-r < 0
             if tune(i:1)==66
             elseif tune(i:1)==77
               mis_counter=mis_counter+1;
               error_B=error_B-w'*tune(i,2:31)'-r;
             end
           end
          total_error=error_M+ error_B;
      end




% Based on the description, f(x) = w'x-r. This is a function that seperates
%to the extent possible malignant points from benign ones.
%If f(x) > 0, then it is malignant.
%If f(x) <= 0. then it is Benign.
%Moreover, since 66 and 77 are two ways to differenciate the M and B.
%Thus, everytime when f(x) > 0 and tune data = 66;or when f(x) <=0, and
%tune data equals to 77. The misclassified number plus one.
```

```
end

fprintf('When mu equals to : %3d\n', mu)
fprintf('In the 100 cases of tuning set, misclassified number is %3d\n',mis_counter )
fprintf('The testing set error is %3d\n', test_error)
end

% Report the number of misclassi ed points on the tuning set:
% When mu equals to : 5.000000e-05
%In the 100 cases of tuning set, misclassified number is   3
%The testing set error is 3.665006e+00
%When mu equals to : 1.000000e-04
%In the 100 cases of tuning set, misclassified number is   3
%The testing set error is 5.322415e+00
%When mu equals to : 1.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 6.458237e+00
%When mu equals to : 2.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 7.614409e+00
%When mu equals to : 2.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 8.139580e+00
%When mu equals to : 3.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 8.419897e+00
%When mu equals to : 3.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 8.719700e+00
%When mu equals to : 4.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 9.334193e+00
%When mu equals to : 4.500000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 9.890028e+00
%When mu equals to : 5.000000e-04
%In the 100 cases of tuning set, misclassified number is   2
%The testing set error is 1.028280e+01

%summary of problem 2
%Compare all the results from 10 mu, we can see that when u becomes larger,
%the misclassified number decrese. In addition, the testing error becomes
%smaller.
%compare all the 10 mu, 5e-4 has the best performance.
%The testing error for mu = 5e-4 eauqls to 10.2828
%The misclassified number when mu = 5e-4 equals to 2
```

```
%Problem 3
%Determine which pair of attributes is most e ective in
%determining a correct diagnosis as follows.
%For each plane use the tuning set with the corresponding pair of
%attributes to determine the number of misclassifed cases.

%In the problem 2, we get the result that 5e-4 has the best performance.
%Thus I used mu = 5e-4 in the problem 3.

clear;
[train,tune,test] = getdata('wdbc.data',30);
label = train(:,1);
mu=5e-4;
misclassified=100;
for i=2:29
    j=i+1;
    while j<31;

M = train(find(label=='M'),[i,j]);
B = train(find(label =='B'),[i,j]);
size_m = size(M,1);
size_b = size(B,1);
b_m = ones(size_m,1);
b_n = ones(size_b,1);
Q = eye(2)*mu;
Q = [Q, zeros(2,370)];
Q = [Q; zeros(370,372)];
c = [b_m'/size_m, b_n'/size_b];
c=[zeros(size(c,1),3) c];
Aeq = [ ];
beq = [ ];
b = [-b_m',-b_n']';
A = [-M, b_m, diag(-b_m), zeros(size_m,size_b);B,-b_n,zeros(size_b,size_m),diag(-b_n)];
lb = [-inf(3, 1); zeros(369,1)];
ub = [inf(372,1)];
[x,obj]=cplexqp(Q,c',A,b,Aeq,beq,lb,ub);
w=x(1:2);
r=x(3);
%set up for the cplexqp function when mu = 5e-4.
%Omega is the first two numbers, and the Gamma is the third number.
y=x(4:3+164);
z=x(4+164:3+369);

mis_counter3=0;

  for i3 = 1:100
      if w'*tune(i3,[i,j])'-r > 0
          if tune(i3,1)==66
                  mis_counter3 =mis_counter3+1;
```

```
        end
      end

                if w'*tune(i3,[i,j])'-r < 0
                    if tune(i3:1)==77
                        mis_counter3=mis_counter3+1;

                    end
                 end

  end


% Based on the description, f(x) = w'x-r. This is a function that seperates
%to the extent possible malignant points from benign ones.
%If f(x) > 0, then it is malignant.
%If f(x) <= 0. then it is Benign.
%Moreover, since 66 and 77 are two ways to differenciate the M and B.
%Thus, everytime when f(x) > 0 and tune data = 66;or when f(x) <=0, and
%tune data equals to 77. The misclassified number plus one.



fprintf('atts %2d %2d: misclass %3d\n',i,j, misclassified);
  if mis_counter3<=misclassified
      ans1=i;
      ans2=j;
      misclassified=mis_counter3;
  end

  %I need to make sure that all my misclassified points smaller than 100.
  %If the this is correct, then replace the misclassified points with
  %miscounter number

 j=j+1;




    end
 end

 fprintf('atts %2d %2d has the minimum misclass %3d\n',ans1,ans2, misclassified);

 %Summary of Problem 3
 %Based on the all the caces of misclassification and test error
 %I think 25,26 is the best.
```

```
% Problem 4
% Use the best performing answer from Part 3 above;
% find and print out the number of misclassied points on the testing set;
% plot all the testing set points on a two dimensional figure



clear;
[train,tune,test] = getdata('wdbc.data',30);
label = train(:,1);
mu=5e-4;
M = train(find(label=='M'),[25,26]);
B = train(find(label =='B'),[25,26]);
m = size(M,1);
k = size(B,1);
b_m = ones(m,1);
b_n = ones(k,1);
Q = eye(2)*mu;
Q = [Q, zeros(2,370)];
Q = [Q; zeros(370,372)];
c = [b_m'/m, b_n'/k];
c=[zeros(size(c,1),3) c];
H = [ ];
g = [ ];
b = [-b_m',-b_n']';
A = [-M, b_m, diag(-b_m), zeros(m,k);B,-b_n,zeros(k,m),diag(-b_n)];
lb = [-inf(3, 1); zeros(369,1)];
ub = [inf(372,1)];
[x,obj]=cplexqp(Q,c',A,b,H,g,lb,ub);

%set up the cplexqp function

w=x(1:2);
r=x(3);
y=x(4:3+164);
z=x(4+164:3+369);


 miscounter4=0;
  for i4 = 1:100
        if w'*test(i4,[25,26])'-r > 0
            if tune(i4,1)==66
                        miscounter4 =miscounter4+1;

            end
        end

                    if w'*test(i4,[25,26])'-r < 0
                        if tune(i4:1)==77
```




```
                miscounter4=miscounter4+1;

            end
        end

 end
 %Because the best situation I find in the problem 3 is 25&26.
% Based on the description, f(x) = w'x-r. This is a function that seperates
%to the extent possible malignant points from benign ones.
%If f(x) > 0, then it is malignant.
%If f(x) <= 0. then it is Benign.
%Moreover, since 66 and 77 are two ways to differenciate the M and B.
%Thus, everytime when f(x) > 0 and tune data = 66;or when f(x) <=0, and
%tune data equals to 77. The misclassified number plus one.

 fprintf('atts %2d %2d: misclass %3d\n',25,26,  miscounter4);


B=test(test(:,1)==66,[25,26]);
M=test(test(:,1)==77,[25,26]);


 plot(B(:,1),B(:,2),'o',M(:,1),M(:,2),'x');
 hold on
```

```
xaxis=[-1.5:0.0001:1.5];
yaxis=[r-x(1)*xaxis]/x(2);
plot(xaxis,yaxis)
% creating plot, setting the x-axis and y-axis

%Summary of problem 4
%According to the description. 'o' stands for the benign points;
%'x'stands for the malignant points.
%Based on the polt, I find there are 4 'o' on the 'x' part, 1 'x' on the 'o' part;
%Moreover, there are 2 'o' on the line, and another 'o' is very close to
%the line.
%Thus, the number of misclassi ed points agrees with the plot and comment.
```