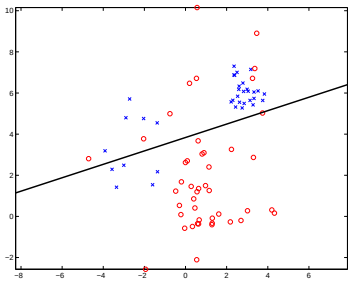
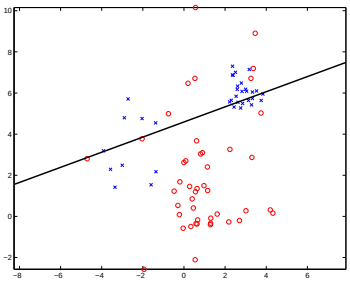
Problem Set 3
Solutions

```
% ps3 script
load -ascii class2d.ascii
x = class2d(:,1:(end-1));
y = class2d(:,end);

figure;
logreg(x,y);
figure;
perceptron(x,y);


function w = logreg(X,y)

d = size(X,2);
X = [X ones(size(X,1),1)];
w = zeros(size(X,2),1);
while(true)
    oldw = w;
    if (d==2)
        hold off;
        plot(X(y>0.5,1),X(y>0.5,2),'bx');
        hold on;
        plot(X(y<0.5,1),X(y<0.5,2),'ro');
        h = drawline(w(1:2),w(3));
        if (h>=0)
            set(h,'Color',[0 0 0]);
            set(h,'LineWidth',2);
        end;
        pause(1);
    end;
    p = exp(X*w);
    p = p./(1+p);
    U = diag(p.*(1-p));
    z = X*w + U\(y-p);
    w = (X'*U*X)\(X'*U*z);
    if (sum((oldw-w).*(oldw-w))<1e-6)
        break;
    end;
end;
```

```
function w = perceptron(X,y)

d = size(X,2);
n = size(X,1);
X = [X ones(size(X,1),1)];
w = rand(size(X,2),1)-0.5;
y = 2*y-1;
eta = 0.1;
nn = 1;
while(true)
    oldw = w;
    rperm = randperm(n);
    X = X(rperm,:);
    y = y(rperm);
    for i=1:n
        if (d==2)
            hold off;
            plot(X(y>0.5,1),X(y>0.5,2),'bx');
            hold on;
            plot(X(y<0.5,1),X(y<0.5,2),'ro');
            h = drawline(w(1:2),w(3));
            if (h>=0)
                set(h,'Color',[0 0 0]);
                set(h,'LineWidth',2);
            end;
            pause(0.001);
        end;
        yy = sign(X(i,:)*w);
        if (sign(yy)~=sign(y(i)))
            w = w + eta*X(i,:)'*y(i);
        end;
    end;
    if (sum((oldw-w).*(oldw-w))<1e-6)
        break;
    end;
    nn =nn+1 % to display # of iterations so far
    eta = eta*0.8;
end;
```



Solution from logistic regression



Solution from perceptron learning algorithm

It takes approximately 30 iterations (1 iteration is once through the dataset) for the perceptron algorithm to converge. It takes 5 iterations for logistic regression to converge. Logistic regression gives a better answer (in terms of fit on the training set). The perceptron learning algorithm has the advantage that it can iterate through the data quickly. Logistic regression must look at all of the data points (at least with iteratively reweighted least squares) before updating the weight vector. This might be a problem with many data points.

A larger problem is that IRLS requires solving a system of D linear equations. If the number of dimensions is large, this could be prohibative. However, we could resort to gradient descent which wouldn't have this problem.