Sample K-means Session

In this training session, we load a matrix of data into Matlab, we perform the $k-$means clustering algorithm, then we'll plot the data, coloring by cluster.

Download the "toy" data set `ToyClustering.mat`. This will load a matrix A that we'll use below (you can clear B, C for now).

You'll notice that we use two inputs to kmeans, and three outputs. The three outputs are idx (giving the index of the cluster for each data point, so it will be "number of points" by 1. The matrix C will hold the centers (it will be $k\times$dimension. and disterr gives us the three distortion errors, one for each cluster.

```
load ToyClustering;  %Use Matrix A for HW problem 1
clear B C

[idx,C,disterr]=kmeans(A,3);


% From here down are the plotting routines
%  The first set of commands will return the indices
%  of the data in cluster 1, 2, and 3 respectively.
%  It will put those values into idx1, idx2, idx3.

idx1=find(idx==1);
idx2=find(idx==2);
idx3=find(idx==3);

% Next we plot the three clusters.  The "hold on"
%  command allows us to plot over older data.

plot(A(idx1,1),A(idx1,2),'r^');
hold on
plot(A(idx2,1),A(idx2,2),'b*');
plot(A(idx3,1),A(idx3,2),'ko');
hold off
```