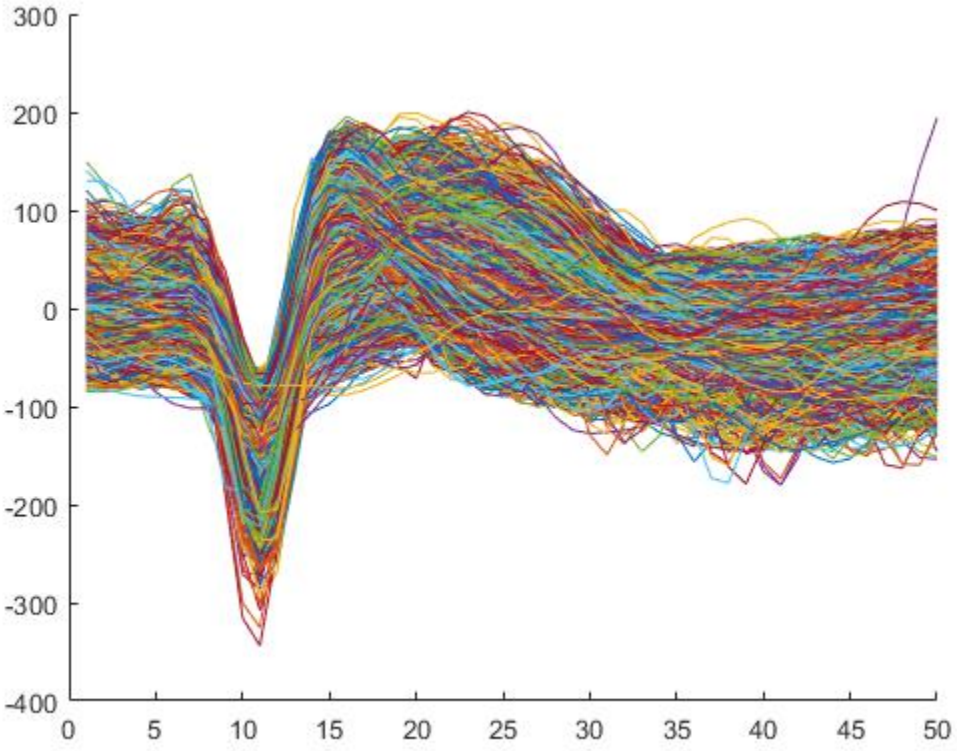
Mini Project 2


```
load('Spike_Events.mat');
load('templates.mat');
figure;
hold on;
for i = 1: size(SpikeEvents,2)
    plot(SpikeEvents(:,i));
end
hold off;
SpikeEventsdemean = bsxfun(@minus,SpikeEvents',mean(SpikeEvents'));
```




```
[coeff,score,latent] = pca(SpikeEvents','NumComponents',2);
% covarianceMat = cov(SpikeEventsdemean);
% [eigvec,eigvals] = eig(covarianceMat);
% figure;
% hold on;
% for i = 1:2
% plot(coeff(:,i));
% end
% hold off;
```

```
% legend('1','2');
% figure;
% biplot(coeff(:,1:2),'scores',score(:,1:2));
```

figure; hold on; scatter(score(:,1),score(:,2),'r.'); temp = SpikeEventsdemean*coeff; scatter(temp(:,1),temp(:,2),'b.'); hold off;

```
clear indeces;
numclusters = 5;
[idx,C,sumd,D] = kmeans(score(:,1:2),numclusters,'MaxIter',1000);
matrix = [score(:,1:2),idx];
figure;
hold on;
    for i = 1:numclusters
        indeces(1:length(find(matrix(:,3) == i)),i) = find(matrix(:,3) == i);

        scatter(matrix(nonzeros(indeces(:,i)),1),...
            matrix(nonzeros(indeces(:,i)),2),'.');


        variances(i) = sum(pdist2(score(nonzeros(indeces(:,i)),:),C(i,:)).^2)...
            /length(nonzeros(indeces(:,i)));
        covariances(i) = sum((score(nonzeros(indeces(:,i)),1) - C(i,1)).*...
            (score(nonzeros(indeces(:,i)),2) - C(i,2)))...
            /length(nonzeros(indeces(:,i)));
    end
legend('1','2','3','4','5');
scatter(score2(:,1),score2(:,2),'ko','DisplayName','Attempt Centroid');
scatter(C(:,1),C(:,2),'kx','DisplayName','Ground Truth Centroid');
hold off;
```

variances =

1×5 single row vector

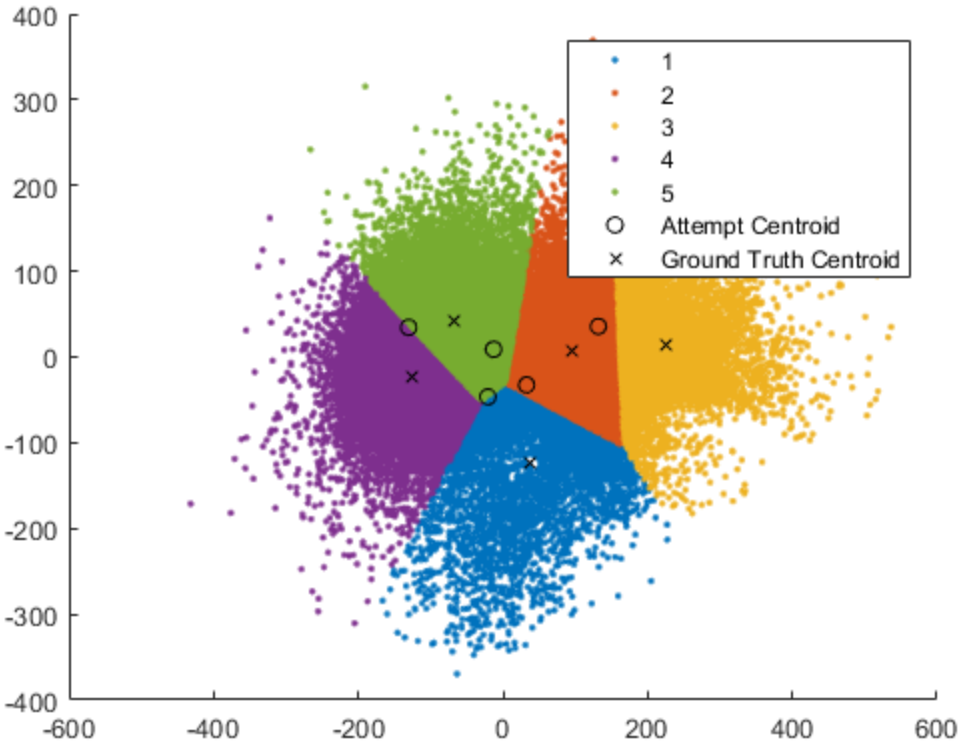
1.0e+03 *

8.4691 3.5450 6.4448 3.4662 3.5445

covariances =

1×5 single row vector

592.1907 -193.9375 49.6907 -256.4259 -402.5603



Mean Squared Error

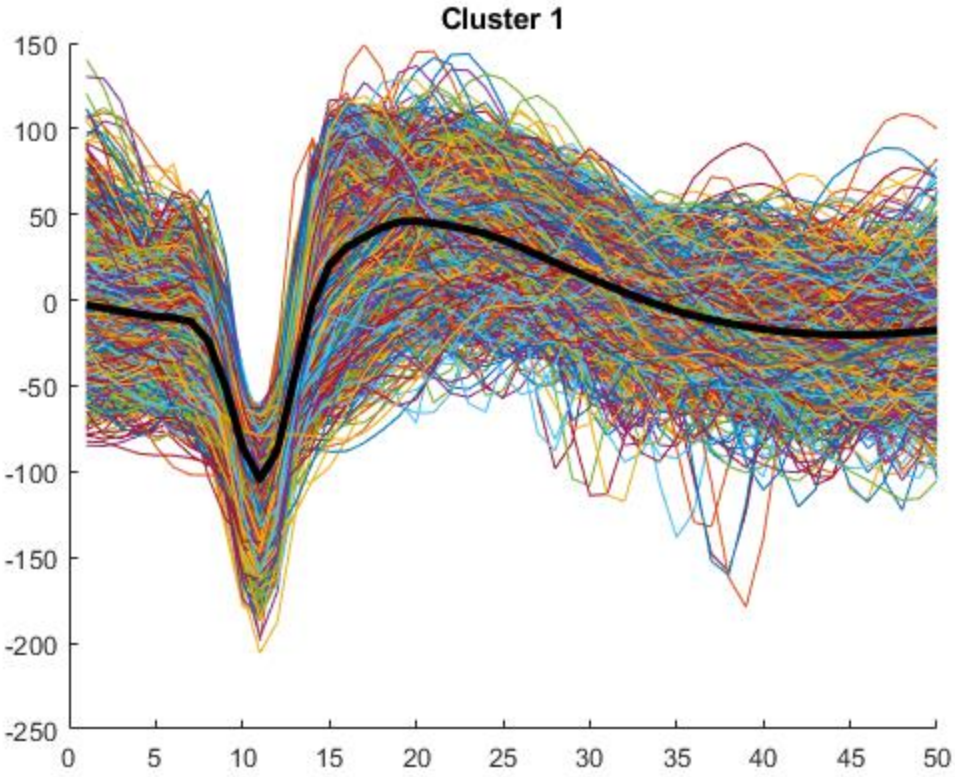
```
groundtruth = sortrows(score2,1);
attempt = sortrows(C,1);
MSE = sum(diag(pdist2(attempt,groundtruth)).^2);
```

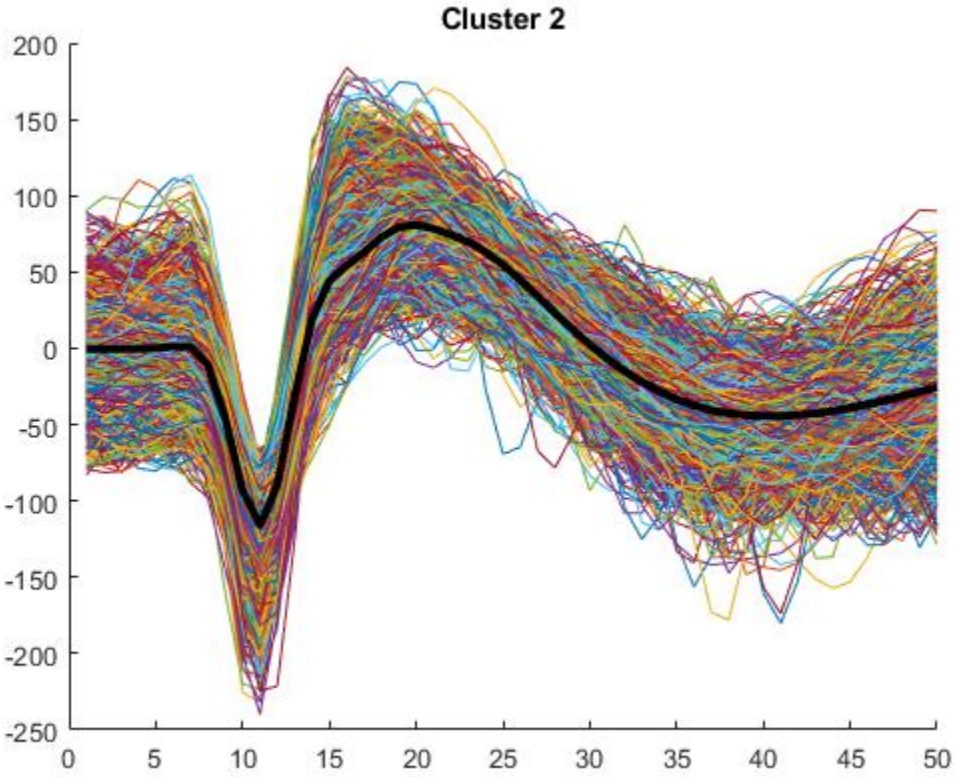
MSE =
4.8438e+04

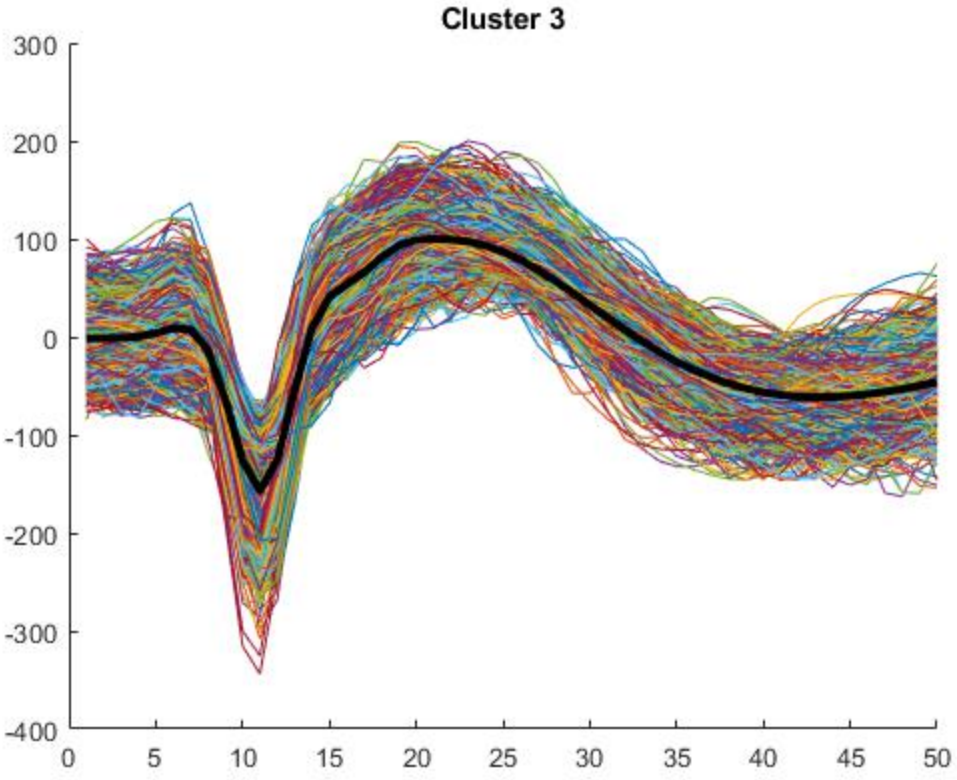
Mean squared error is pretty high, however ground truth labels don't seem to match clusters very well either.

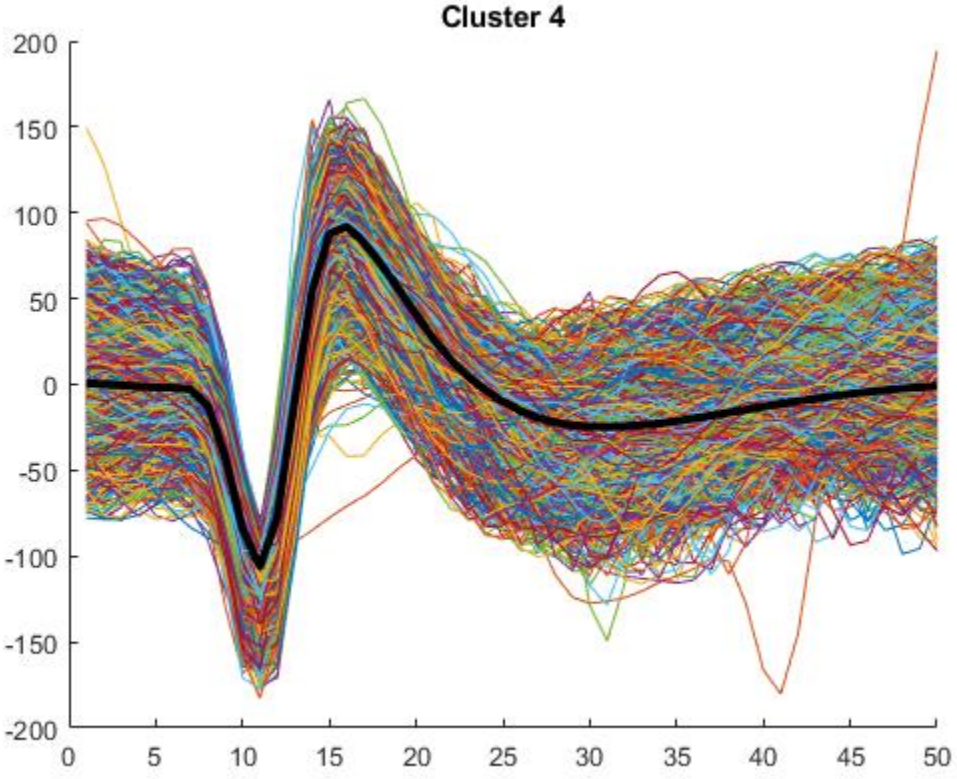
Create Mean Plots

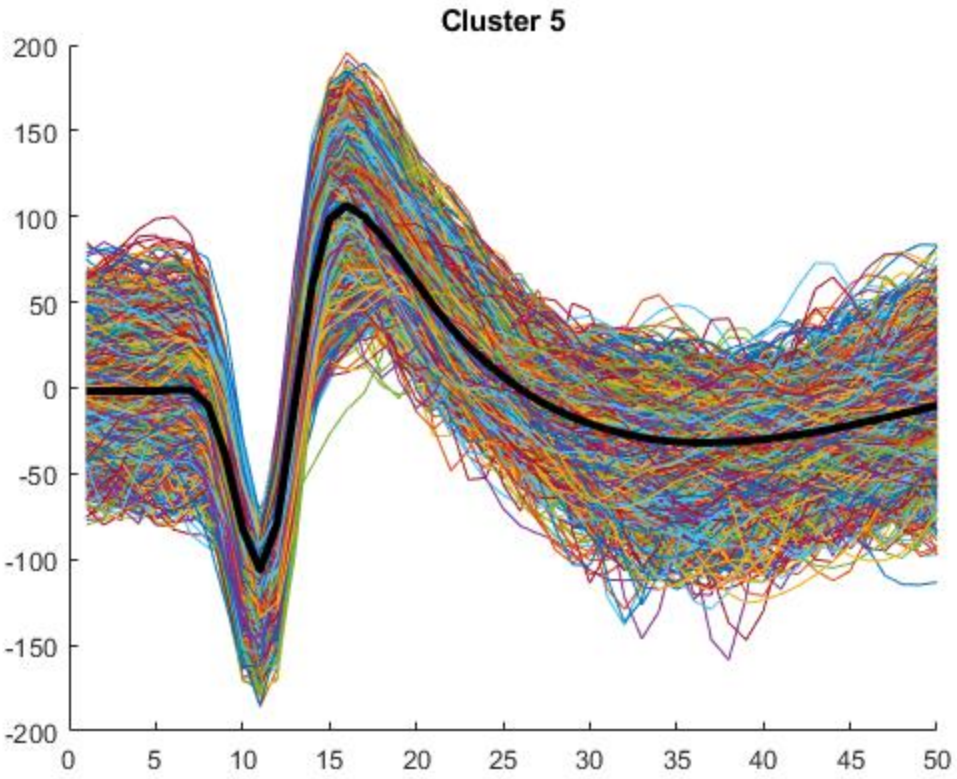
```
for i = 1:numclusters
    figure;
    hold on;
    plot(SpikeEvents(:,nonzeros(indeces(:,i))));
    plot(mean(SpikeEvents(:,nonzeros(indeces(:,i))),2),'k-','LineWidth',3);
    title(['Cluster ' num2str(i)]);
    hold off;
end
```


Cluster 1
150
100
50
0
-50
-100
-150
-200
-250
0
5
10
15
20
25
30
35
40
45
50


Cluster 2
200
150
100
50
0
-50
-100
-150
-200
-250
0
5
10
15
20
25
30
35
40
45
50


Cluster 3
300
200
100
0
-100
-200
-300
-400
0
5
10
15
20
25
30
35
40
45
50


Cluster 4
200
150
100
50
0
-50
-100
-150
-200
0
5
10
15
20
25
30
35
40
45
50



Published with MATLAB® R2017b

The optimal decision boundary does match the variances and the covariances well, the variances and the covariances are quite large, contributing to the decision boundaries squeezing up against each other to form straight lines.