P5q2

% Upload the MNIST dataset

% Work out the covariance matrix, eigen-value, eigen-vector

load mnist\_train.mat

Cov = cov(train\_X);

[V,D] = eig(Cov); % D : eigen-value V: normalised eigen-vector

% Choose out the largest 2 eigen-value

W = V(:,(size(V,2)-1):size(V,2));

z = (train\_X-mean(train\_X))\* W;

%[x1;x2;...]\*[v1,v2] => [x1v1,x1v2 ; x2v1,x2v2; ...] -> new data after dimension reduction

z\_new = [z,train\_labels];

figure;

for i = 1:max(z\_new(:,3))

t = z\_new(z\_new(:,3)==i,:);

scatter(t(:,1),t(:,2))

hold on

end

eigen\_value = sort(diag(D),'descend');

eigen\_value\_ratio = eigen\_value./sum(eigen\_value);

account\_for\_ratio = sum(eigen\_value\_ratio(1:2,:));

index = repmat(1:1:length(eigen\_value),1)';

ratio\_append = [index,eigen\_value\_ratio];

figure;

plot(ratio\_append(:,1),eigen\_value);

figure;

plot(ratio\_append(:,1),cumsum(eigen\_value\_ratio));