Pre-class Exercise

• The Mahalanobis distance can be used to classify an unknown feature vector.
  – Assume you know the class centers (centroids) $z_i$, and their covariances $C_i$
  – We find the class that has the smallest distance from its center to the point in feature space

• The Mahalanobis distance of feature vector $x$ to the $i^{th}$ class is
  \[ d_i = \sqrt{(x - z_i)^T C_i^{-1} (x - z_i)} \]

• where $C_i$ is the covariance matrix of the feature vectors in the $i^{th}$ class
clear all
close all

% Load iris data. This loads in variables "meas" and "species".
load fisheriris

% Keep only the first two classes, which are 1..50 and 51..100.
X1 = meas(1:50,1:2);        % Keep only 1st two measurement dimensions
X2 = meas(51:100,1:2);      % Keep only 1st two measurement dimensions
plot(X1(:,1), X1(:,2), 'ro', X2(:,1), X2(:,2), 'go');

% Compute the mean of each class.
m1 = mean(X1);
m2 = mean(X2);
hold on
plot(m1(1), m1(2), 'r+');
plot(m2(1), m2(2), 'g+');

% Compute the covariance of each class.
C1 = cov(X1);
C2 = cov(X2);
disp('C1:'), disp(C1);
disp('C2:'), disp(C2);

% This bit of code displays the covariance as an ellipse.
th=0:2*pi/30:2*pi;
ptsCircle = [cos(th) sin(th)]'; % Create points along a circle
n = size(ptsCircle,2);
p1 = chol(C1)'*ptsCircle + repmat(m1',1,n);
p2 = chol(C2)'*ptsCircle + repmat(m2',1,n);
plot(p1(1,:), p1(2,:), 'LineWidth', 2.0, 'Color', 'r');
plot(p2(1,:), p2(2,:), 'LineWidth', 2.0, 'Color', 'g');
• Let $X_{\text{test}} = [5.3, 3.1]$
• What is the Mahalanobis distance to class 1 (red) and class 2 (green)?