Unsupervised Learning
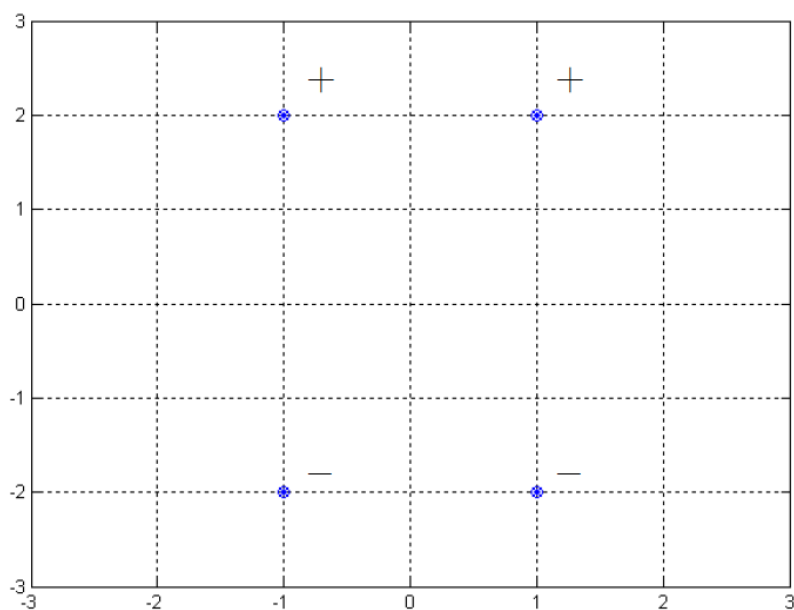
The two following classes are given:

$$x_1 = (-1, 2)^T, x_2 = (1, 2)^T, x_3 = (-1, -2)^T, x_4 = (1, 2)^T$$
$$y_1 = +1, y_2 = +1, y_3 = -1, y_4 = -1$$
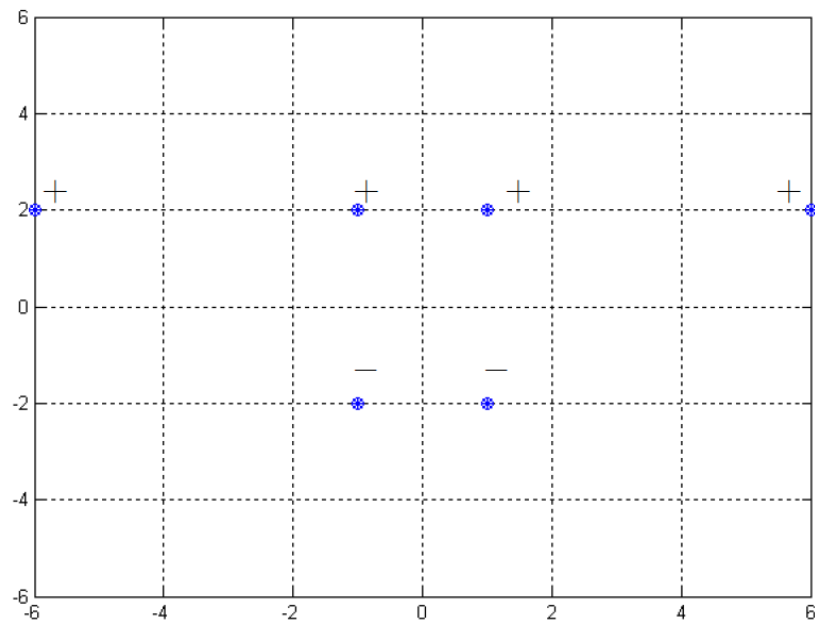

a. What is the first principal axis (u_1) of the given dataset? Center the dataset as needed and calculate the empirical covariance matrix C. Prove your answer regarding u_1.

b. Suggest a linear classifier that relies only on the projections of the dataset onto u_1 that you found and scores perfect accuracy.

Now the two following examples are added:

$$x_5 = (6, 2)^T, x_6 = (-6, 2)^T$$
$$y_5 = +1, y_6 = +1$$

c. What is u_1 of the new given dataset? Center the dataset as needed and calculate the empirical covariance matrix C. Prove your answer regarding u_1.

d. Draw the projections of the dataset onto u_1 that you found. Is it possible to classify perfectly the dataset using the projections and a linear classifier? or with a non-linear classifier?