In The Name of God, The Compassionate, The Merciful

Name: ....
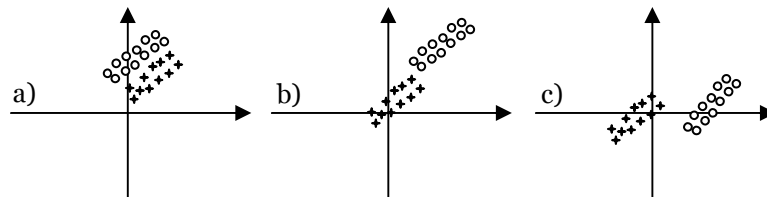
Student ID#: ....

Statistical Pattern Recognition (CE-725)
Department of Computer Engineering
Quiz #3 solution (Features) - Spring 2011

1.a (3 points) Consider the following three datasets:

For each dataset, which of three methods PCA, LDA and feature selection do you suggest for reducing dimensionality?

For (a) LDA is preferred, since the direction of maximum variance is not the same as direction of maximum separation.

For (b) PCA is preferred since the direction of maximum variance and direction of maximum separation are the same.

For (c), feature selection is best, since all samples can be separated by using only x (the horizontal feature)

1.b (7 points) Which of the following would be a good objective function to use instead of Fisher's one and which of them not? Give 1 sentence explanations.

$$J(v) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2} + \frac{(\mu_1 - \mu_2)^2}{\sigma_2^2}$$

Good one: maximizing J(v) maximizes the means distances and minimizes the within-class variabilities.

$$J(v) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 / \sigma_2^2}$$

Bad one: maximizing J(v) maximizes the class 2's within-class variability.

$$J(v) = \frac{\sigma_1^2 * \sigma_2^2}{(\mu_1 - \mu_2)^2}$$

Good one: minimizing J(v) maximizes the means distances and minimizes the within-class variabilities.