Machine Learning I: Statistical Learning Theory
Homework 2, Due Febuary 24

Problem 1 - Linear Discriminant Analysis: Consider the categorical learning problem consisting of a data set with two labels:

Label 1:

X_1	3.81	0.23	3.05	0.68	2.67
X_2	-0.55	3.37	3.53	1.84	2.74

Label 2:

X_1	-2.04	-0.72	-2.46	-3.51	-2.05
X_2	-1.25	-3.35	-1.31	0.13	-2.82

a) For each label above, the data follow a multivariate normal distribution $\mathcal{N}(\mu_i, \mathbf{\Sigma})$, where the covariance $\mathbf{\Sigma}$ is the same for both label 1 and for label 2. Fit a pair of Guassian discriminant functions to the labels by computing the covariances, means, and proportions of datapoints as discussed in the Linear Discriminant Analysis section of Lecture 5. You may use a computer, but you should not use an LDA solver. You should report the values for μ_i and $\mathbf{\Sigma}$.

b) Give the formula for the line forming the discretion boundary.

Problem 2 - Cubic Splines: *(Problem 5.1 in ESLII)* Lets consider data of the shape (X, y) with $X, y \in \mathbb{R}$. Consider fitting of piecewise-cubic polynomial splines to the data with continuous first and second derivatives at the two point ξ_1 and ξ_2.
a) As in Lecture 7, derive the equations relating the polynomials at each boundary point.

b) Show that

$$h_1(X) = 1\,, \qquad h_3(X) = X^2\,, \qquad h_5(X) = (X-\xi_1)_+^3\,,$$
$$h_2(X) = X\,, \qquad h_4(X) = X^3\,, \qquad h_6(X) = (X-\xi_2)_+^3\,.$$

is a basis for the cubic splines.

Problem 3 - Gradient Decent: The data below was in the second part of the datafile is drawn from the sum of two sine functions:

$$f(X) = \beta_0 + \beta_1 \sin(\alpha_1 x - \mu_1) + \beta_2 \sin(\alpha_2 x - \mu_2)\,.$$

Use gradient decent or SGD to fit the data below to the function by minimizing the RSS loss

$$RSS = \sum_{i=1}^{N} (y_i - f(X))^2\,.$$

Turn in any associated computations, your learning rate, and the parameters.