Machine Learning. Spring 2013. Homework 2.

Due: Tuesday 14/May. 2013.

Question 1. Regression

Suppose you have two variables x_1, x_2 and we want to perform linear regression using the function:

$$f(x_1, x_2) = w_0 + w_1 x_1 + w_2 x_2 + w_3 x_1 x_2 + w_4 x_1^2 + w_5 x_2^2 \quad (1)$$

What is the least squares formulation for this problem? Give a solutions for the w_i using a sample set $X = \{x_1^t, x_2^t, r^t\}$. What is the variance of this estimate?

Question 2. AdaBoost.

The AdaBoost learning algorithm takes an input dataset $\{(x_i, y_i) : i = 1, ..., m\}$. Describe the algorithm. What is a weak classifier? What is a strong classifier? How does AdaBoost select and weight weak classifiers? What criterion does AdaBoost minimize? How does this relate to the error function?

Apply AdaBoost to the one-dimensional problem where the data lies on the x-axis. There is one positive example at $x = 0$ and two negative examples at $x = \pm 1$. There are three weak classifiers are $h_1(x) = 1,\ x > 1/2, \&\ h_1(x) = -1,\ x < 1/2$, $h_2(x) = 1,\ x > -1/2, \&\ h_2(x) = -1,\ x < -1/2$, and $h_3(x) = 1,\ \forall x$. Show that this data can be classified correctly by a strong classifier which uses only three weak classifiers. Calculate the first two iterations of AdaBoost for this problem. Are they sufficient to classify the data correctly?

Question 3.

Consider a binary classification problem where the data is one-dimensional. There are two negative ($y = -1$) examples at $x = \pm 1$ and one positive ($y = 1$) example at $x = 0$.

Show that you cannot classify this data perfectly by linear separation (i.e. by a decision rule $sign(ax + b)$ for some a, b).

Now formulate this problem with slack variables $\{z_i\}$. The classifier with the largest margin is obtained by solving the primal problem:

$L_P = (1/2)a^2 + \gamma \sum_{i=1}^{3} z_i - \sum_{i=1}^{3} \alpha_i\{y_i(ax_i + b) - (1 - z_i)\} - \sum_{i=1}^{3} \mu_i z_i$.

where γ is a constant, the $\{\alpha_i, \mu_i\}$ are Lagrange multipliers (constrained to be non-negative) and the $\{z_i\}$ are non-negative.

Minimize L_p by searching for the minimum of $(1/2)a^2 + \gamma(z_1 + z_2 + z_3)$ subject to the constraints $y_i(ax_i + b) - (1 - z_i) \geq 0,\ \forall\ i \in \{1, 2, 3\}$ with $z_i \geq 0,\ \forall\ i \in \{1, 2, 3\}$. (Hint: exploit structure of the problem to guess where the decision boundary should be). What are the support vectors?

Question 4. Primal Dual Quadratic Optimization

The primal problem is formulated as follows:

$$L_p(\vec{a}, b, \{z_i\}; \{\alpha_i, \mu_i\}) = (1/2)|\vec{a}|^2 + \gamma \sum_{i=1}^{m} z_i - \sum_{i=1}^{m} \alpha_i\{\omega_i(\vec{a} \cdot \vec{x}_i + b) - (1 - z_i)\} - \sum_{i=1}^{m} \mu_i z_i. \tag{2}$$

Explain the meaning of all the terms and variables in this equation. What constraints do the variables satisfy? Calculate the form of the solution $\vec{a}$ by minimizing L_p with respect to $\vec{a}, b, \{z_i\}$. What are the support vectors?

Obtain the dual formulation by eliminating $\vec{a}, b, \{z_i\}$ from L_p. Describe a strategy for solving the primal problem.