HW 2B – ENGR 4399: Machine Learning

Instructor: Dr. Gonzalo D. Parra

Problem 1 *Show that the ridge regression objective function $f(\beta) = \|\mathbf{Y} - \mathbb{X}\beta\|^2 + \lambda\|\beta\|^2$ is convex in β. Find the optimal solution and provide justification for each step. Here*

$$\|\beta\| = \|\beta\|_2 = \sqrt{\beta_1^2 + \beta_2^2 + \cdots + \beta_p^2}.$$

You should get the solution

$$\beta_{ridge} = \left(\mathbb{X}^T\mathbb{X} + \lambda I_p\right)^{-1}\mathbb{X}^T\mathbb{Y}.$$

Now, assume that the columns of the data matrix $\mathbb{X}$ are orthonormal, i.e.,

$$\mathbb{X}^T\mathbb{X} = I_p.$$

*Show that under this assumption, the ridge solution **shrinks** the OLS solutions towards the all zero vector.*

$$\beta_{ridge} = \frac{\mathbb{X}^T\mathbb{Y}}{1+\lambda} = \frac{\beta_{ols}}{1+\lambda}.$$

Problem 2 *Show that the Elastic Net regression function, for $\lambda_1 > 0$ and $\lambda_2 > 0$,*

$$f(\beta) = \|\mathbb{Y} - \mathbb{X}\beta\|^2 + \lambda_1\|\beta\|_2^2 + 2\lambda_2\|\beta\|_1,$$

is convex in β. Here

$$\|\beta\|_1 = \sqrt{|\beta_1| + |\beta_2| + \cdots + |\beta_p|}.$$

Problem 3 *Define the BIC regression function*

$$f(\beta) = \|\mathbb{Y} - \mathbb{X}\beta\|^2 + \lambda\|\beta\|_0,$$

*where $\|\beta\|_0 = \sum_{i=1}^p \mathbb{I}_{\{\beta_i \neq 0\}}$. Show that the function $f(\beta)$ is **not convex**.*

Problem 4 *Define the logistic regression function*

$$f(\beta) = -\sum_{i=1}^n \left(Y_i\beta^T X_i + \log\frac{1}{1 + e^{\beta^T X_i}}\right).$$

Here $(X_1, Y_1), \ldots (X_n, Y_n)$ is our training data. Show that the function above is convex in β.

Problem 5 *Define the support vector machine objective function*

$$L(\beta, \beta_0, \lambda) = \frac{1}{2}\|\beta\|^2 + \sum_{i=1}^n \lambda_i \left(1 - Y_i(X_i^T\beta + \beta_0)\right).$$

Show that the function is convex in β and β_0.