Exercise Sheet #6: Trees and Forests and Supports etc.

Mhd Jawad Al Rahwanji, Wojciech Ptaś
7038980, 7042843

January 31, 2023

Problem 3 – Linear and Support Vector Regression

(a) First, we define the *Ridge Regression* objective function as

$$\min_{\beta}\left\{\lambda||\beta||_2^2+\sum_{i=1}^{N}(y_i-\beta_0-\beta^\top x_i)^2\right\} \tag{1.1}$$

Now, regarding the objective in question, we may divide it by C and plug in λ knowing that they are inversely proportional

$$\min_{\beta_0,\beta,\xi_1\ldots,\xi_N}\left\{\lambda||\beta||_2^2+\sum_{i=1}^{N}\xi_i^2+\tilde{\xi}_i^2\right\} \tag{1.2}$$

$$\textit{where} \qquad \lambda\approx\frac{1}{2C}$$

Next, we have the following inequality

$$-\tilde{\xi}_i\leq y_i-\beta_0-\beta^\top x_i\leq\xi_i$$

so we square the whole thing to get

$$(y_i-\beta_0-\beta^\top x_i)^2\leq\max(\xi_i^2,\tilde{\xi}_i^2)$$

which can also be rewritten as

$$(y_i-\beta_0-\beta^\top x_i)^2\leq\xi_i^2+\tilde{\xi}_i^2 \tag{1.3}$$

Lastly, we plug the lower bound from (1.3) into (1.2) since we intend to find – now – β that minimizes it to get

$$\min_{\beta}\left\{\lambda||\beta||_2^2+\sum_{i=1}^{N}(y_i-\beta_0-\beta^\top x_i)^2\right\} \tag{1.4}$$

thus, (1.4) ≡ (1.1), quod erat demonstrandum.

(b) We know that the *Support Vector Machine* optimization problem can be rewritten as

$$\min_{\beta_0,\beta,\ldots,\beta_p} \left\{ \sum_{i=1}^{N} \max[0, 1 - y_i f(x_i)]_+ + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}$$

The *SVM* equivalent for the loss function L using the same rationale from (a) takes the form of a hinge function

$$\sum_{i=1}^{N} \max[(\xi_i + \varepsilon)^2, (-\tilde{\xi}_i - \varepsilon)^2]_+$$

(c) The optimization objective involving the *2N* additional $\xi_i, \tilde{\xi}_i$ may be preferred for *SVRs* because it resembles a polynomial kernel which would better fit non-linear data if that were the case at hand.