TP2: Linear Hard -band SVR and Linear soft -band SVR

Objective:

The goal of this lab is to understand how Support Vector Machines (SVMs) works for regression and to implement Sequential Minimal Optimization (SMO) to train the model. We will implement Linear Hard -band SVR , Linear soft -band SVR and SMO using NumPy and Matplotlib for visualization.

Linear Hard -band SVR and Linear soft -band SVR

We give the mathematical models of the dual forms of Hard margin Linear Hard -band (SVR) and Linear soft -band (SVR) respectively:

Linear Hard -band SVR:

- Input: training set : $\{(x_i, y_i\}_{i=1}^n$ where $x_i \in \mathbb{R}^d, y_i \in \mathbb{R}, \varepsilon_{\text{inf}} > 0$
- Choose an approriate parameter $\varepsilon > 0$ and $\varepsilon \geq \varepsilon_{\text{inf}}$
- Contruct and solve the optimzation problem DLH $-\varepsilon-$ band $-SVR$ obtaining $\lambda^* = (\lambda_1^*, \ldots, \lambda_n^*)$ and $\mu^* = (\mu_1^*, \ldots, \mu_n^*)$

$$\text{DLH } -\varepsilon\text{-band } -SVR \begin{cases} \max \frac{1}{2} \sum_{i,j=1}^n (\mu_i - \lambda_i)(\mu_j - \lambda_j)\left(x_i{}^T x_i\right) - \varepsilon \sum_{i=1}^n (\mu_i + \lambda_i) + \sum_{i=1}^n y_i (\mu_i - \lambda_i) \\ \text{s.à } \sum_{i=1}^n (\mu_i - \lambda_i) = 0 \\ \lambda_i \geq 0, \mu_i \geq 0, \quad i = 1, \ldots, n \end{cases}$$

if $\lambda^* \neq 0$ and $\mu^* \neq 0$, the solution to the problem, (w^*, b^*), can be obtained in the following way

- $w^* = \sum_{i=1}^n (\mu_i - \lambda_i) x_i$
- for any nonzero component λ_j^* of $\lambda^*, b^* = y_j - (w^*)^T x_j + \varepsilon$
- Or for any nonzero component μ_j^* of $\mu^*, b^* = y_j - (w^*)^T x_j - \varepsilon$

$$h(x) = (w^*)^T x + b^*$$

Linear Soft-band SVR:

- Input: training set : $\{(x_i, y_i\}_{i=1}^n$ where $x_i \in \mathbb{R}^d, y_i \in \mathbb{R}, \varepsilon_{\text{inf}} \geq 0$
- Choose an approriate parameter $\varepsilon > \varepsilon_{\text{inf}}$ and penalty parameter $C > 0$
- Contruct and solve the optimzation problem DLS $-\varepsilon$-band $-SVR$ obtaining $\lambda^* = (\lambda_1^*, \ldots, \lambda_n^*)$ and $\mu^* = (\mu_1^*, \ldots, \mu_n^*)$

$$\text{DLS } -\varepsilon\text{-band } -SVR \begin{cases} \min \frac{1}{2} \sum_{i,j=1}^n (\mu_i - \lambda_i)(\mu_j - \lambda_j)\left(x_i^T x_i\right) - \varepsilon \sum_{i=1}^n (\mu_i + \lambda_i) + \sum_{i=1}^n y_i (\mu_i - \lambda_i) \\ \text{s.à } \quad \sum_{i=1}^n (\mu_i - \lambda_i) = 0 \\ C \geq \lambda_i \geq 0, C \geq \mu_i \geq 0, \quad i = 1, \ldots, n \end{cases}$$

if $\lambda^* \neq 0$ and $\mu^* \neq 0$, the solution to the problem, (w^*, b^*), can be obtained in the following way

- $w^* = \sum_{i=1}^{n} (\mu_i - \lambda_i) x_i$
- For any component $\lambda_j^* \in]0, C[$ of $\lambda^*, b^* = y_j - (w^*)^T x_j + \varepsilon$
- Or for any component $\mu_j^* \in]0, C[$ of $\mu^*, b^* = y_j - (w^*)^T x_j - \varepsilon \; h(x) = (w^*)^T x + b^*$

Method:

1. Choose two simple datasets of regression, that are not purely linear, one for simple regression (one x and one y) and another for Multiple linear regression ($x1, x2$ and y)
2. Perform exploratory analysis of the data and preprocess the data if necessary.
3. Split data into train and test
4. Implement the SMO optimization algorithm.
5. Implement Linear Hard -band (SVR) and Linear soft -band (SVR).
6. Train the two models, and compare their error mean squared errors (MSE), what can we conclude?
7. Visualize the regression results(the approximation hperplane in $2 \mathrm{D}$ or $3 \mathrm{D}$ depending on the type of dataset)