scikit-Learn

1.2.2. Regression

永田 純平

Support Vector Regression(SVR)

- SVMの回帰問題に拡張したもの(SVR,NuSVR)

- 汎化誤差の上限を最小化するように学習を行うため、高い汎化能力をもつ。

- カーネルを適用することで非線形関数近似に拡張することができる。

回帰に対する損失関数

- 二乗誤差損失関数

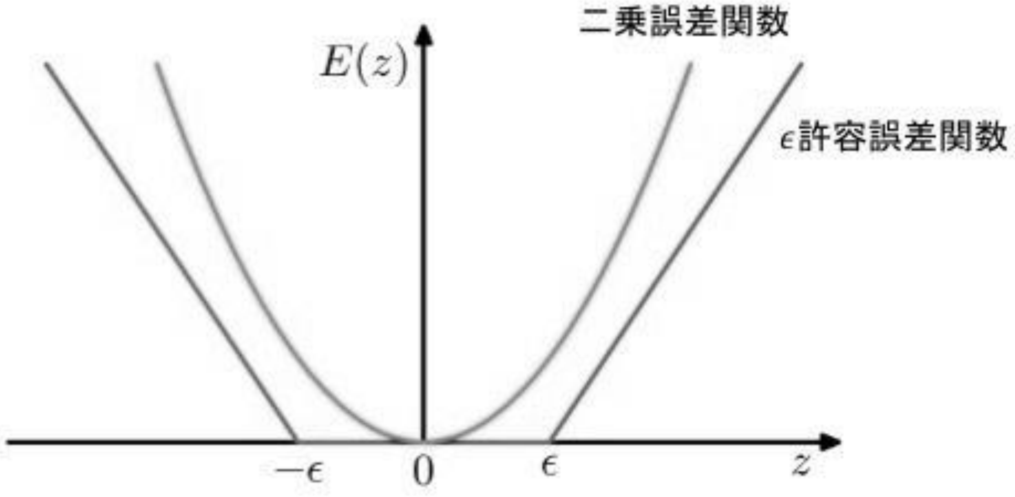
$$\frac{1}{2}\sum_{n=1}^{N}\{f(x_n)-y_n\}^2+\frac{1}{2}||w||^2$$

- ε許容誤差関数を用いた損失関数

$$C\sum_{n=1}^{N}E_{\varepsilon}(f(x_n)-y_n)+\frac{1}{2}||w||^2$$

※ ε許容誤差関数

$$E_\varepsilon(f(x_n) - y_n) = \begin{cases} 0 & |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & (otherwise) \end{cases}$$



スラック変数の導入

$f(x_n) - \varepsilon \leq y_n \leq f(x_n) + \varepsilon$より

$$\begin{cases} \xi_n \geq 0, \hat{\xi}_n \geq 0 \\ y_n - f(x_n) - \varepsilon \leq \xi_n \\ f(x_n) - y_n - \varepsilon \leq \hat{\xi}_n \end{cases}$$

※実測地と予測値の誤差がεより大きいと損失として値をとる。

$$\xi_n = \begin{cases} 0 & y_n \leq f(x_n) + \varepsilon \\ \xi_n & y_n > f(x_n) + \varepsilon \end{cases}, \hat{\xi}_n = \begin{cases} 0 & y_n \leq f(x_n) - \varepsilon \\ \hat{\xi}_n & y_n > f(x_n) - \varepsilon \end{cases}$$

$$C \sum_{n=1}^{N} (\xi_n + \hat{\xi}_n) + \frac{1}{2} ||w||^2$$

$$L = C\sum_{n=1}^{N}(\xi_n+\hat{\xi}_n) + \frac{1}{2}||w||^2 - \sum_{n=1}^{N}\left(\mu_n\xi_n + \hat{\mu}_n\hat{\xi}_n\right)$$

$$-\sum_{n=1}^{N} a_n\,(\varepsilon + \xi_n + f(x_n) - y_n) - \sum_{n=1}^{N}\hat{a}_n(\varepsilon + \hat{\xi}_n + y_n - f(x_n))$$

※$\mu_n, \hat{\mu}_n, a_n\,, \hat{a}_n$はそれぞれの制約条件に対応するラグランジュ乗数

$f(x) = w \cdot x - b$を代入し、$w, b, \xi_n, \hat{\xi}_n$をそれぞれ微分する

$$\frac{\partial L}{\partial w} = w - \sum_{n=1}^{N}(a_n - \hat{a}_n)x_n = 0 \rightarrow w = \sum_{n=1}^{N}(a_n - \hat{a}_n)x_n$$

$$\frac{\partial L}{\partial b} = \sum_{n=1}^{N}(a_n - \hat{a}_n) = 0\,,$$

$$\frac{\partial L}{\partial \xi_n} = C - a_n - \mu_n = 0 \rightarrow a_n + \mu_n = C$$

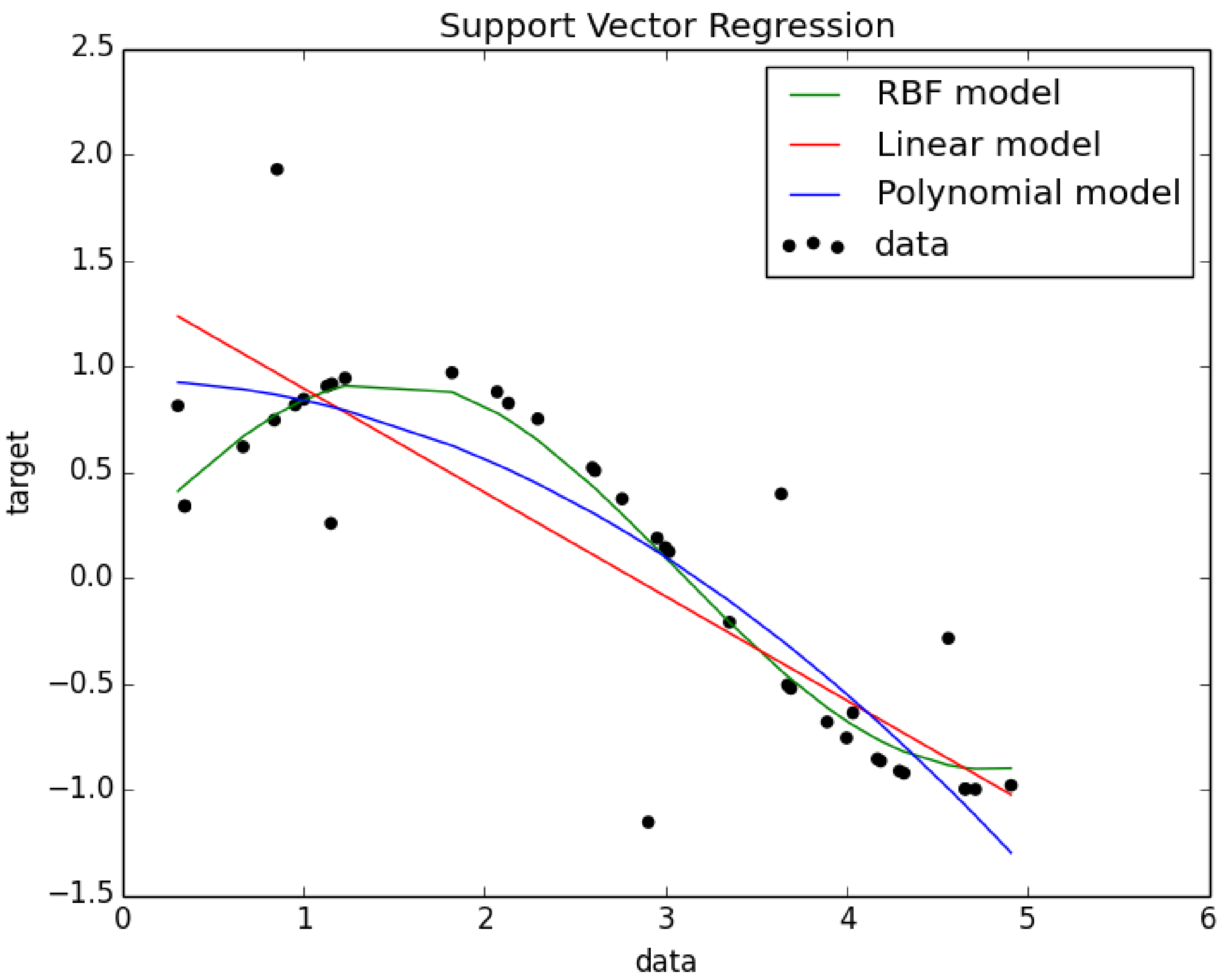
$$\frac{\partial L}{\partial \hat{\mu}_n} = C - \hat{a}_n - \hat{\mu}_n = 0 \rightarrow \hat{a}_n + \hat{\mu}_n = C$$

$$L(a_n, \hat{a}_n) = -\frac{1}{2}\sum_{n,m=1}^{N}(a_n - \hat{a}_n)(a_m - \hat{a}_m)(x_n \cdot x_m)$$

$$-\varepsilon\sum_{n=1}^{N}(a_n + \hat{a}_n) + \sum_{n=1}^{N}(a_n + \hat{a}_n)y_n$$

回帰モデル

$$f(x) = \sum_{n=1}^{N}(a_n - \hat{a}_n)(x \cdot x_n) - b$$


Support Vector Regression
RBF model
Linear model
Polynomial model
data
target
data
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
0
1
2
3
4
5
6