University of Pittsburgh

ECE 2195: Special Topics – Computers Machine Learning

Support Vector Regression

Mai Abdelhakim, PhD

Assistant Professor of ECE

Swanson School of Engineering

University of Pittsburgh

maia@pitt.edu

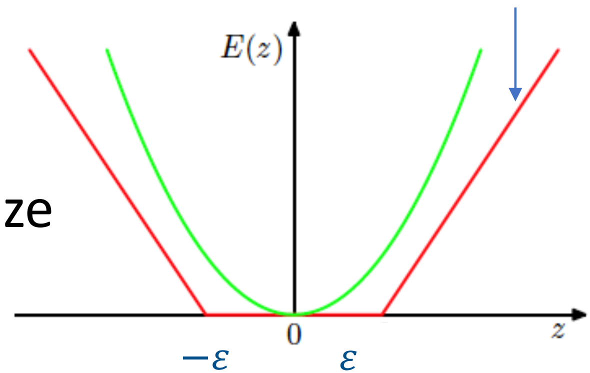
Recall Linear Regression with Regularization

- Objective function of linear regression:
  $\underbrace{\frac{1}{2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}_{\text{error RSS}} + \frac{\lambda}{2}||w||^2$ Ridge regularization
- Let's replace the loss $\frac{1}{2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ with ε-insensitive error function E(z)
  - Consider error only if they are more than ε
- The objective now becomes to minimize

$$C\sum_{i=1}^{n} E(y_i - \hat{y}_i) + \frac{1}{2}||w||^2$$

C is the inverse regularization parameter

ε-insensitive error



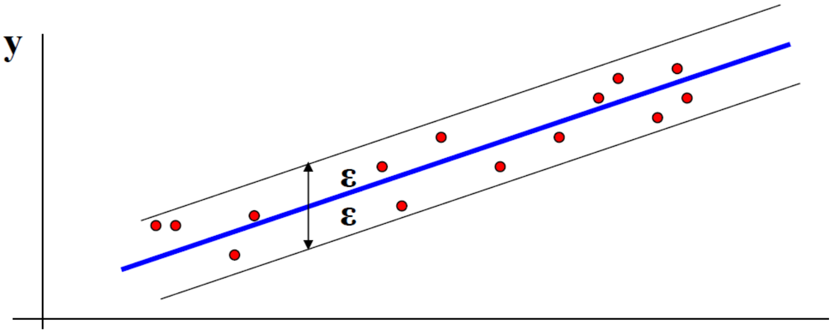
Regression using SVC

$$f(\mathbf{x}) = \mathbf{w}^{\mathrm{T}}\mathbf{x} + b = \langle \mathbf{w}, \mathbf{x} \rangle + b$$

- Linear function for regression
- If all points assumed to be with ε neighborhood
  - With ε-insensitive error function
- The objective will be

$$\text{minimize} \quad \frac{1}{2}\|w\|^2$$

$$\text{subject to} \quad \begin{cases} \mathrm{y}_i - \langle w_i, x_i \rangle - b \le \varepsilon \\ \langle w_i, x_i \rangle + b - \mathrm{y}_i \le \varepsilon \end{cases}$$

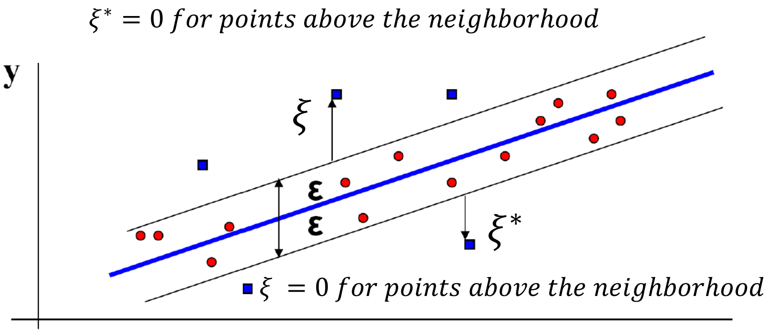


Get it more general

- Not all data could satisfy neighborhood (some have more error)
- Allow some errors but add constraint to limit the error
- Objective function

$$\text{minimize} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)$$

$$\text{subject to} \quad \begin{cases} y_i - \langle w_i, x_i \rangle - b \le \varepsilon + \xi_i \\ \langle w_i, x_i \rangle + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases}$$



$$< w, x >= w^T x$$

- Solve using Lagrange as before

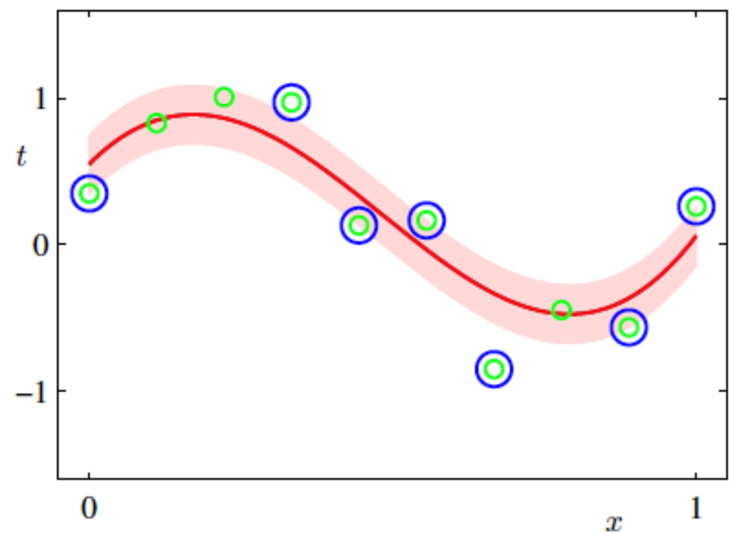
Solution

Dot product

- Use Lagrange, we get

$$f(x) = \sum_{i=1}^{n} (C_i) < x_i, x > +b$$

  - where C_i depends on Lagrange multipliers



example with Radial Basis Kernel

- We can apply Kernel as before
  - Replace dot product with kernel to get Get non linear function

$$f(x) = \sum_{i=1}^{n} (C_i) K(x_i, x) + b$$