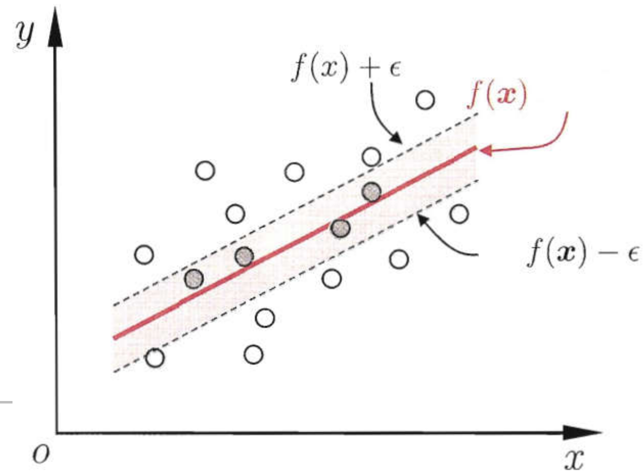
2.11 支持向量回归 （SVR）

CSDN学院

2017年11月

▶ 支持向量回归

- 在之前的线性回归模型中，只有当y与$f(\mathbf{x})$完全相等时，我们才认为损失为0（L2损失、L1损失、Huber损失）。
- 在支持向量回归中，我们能容忍y与$f(\mathbf{x})$之间有ϵ的偏差，即当y与$f(\mathbf{x})$之间的差异大于ϵ时才计算损失（ ε insensitive loss ）。
  - 相当于以为中心，构造一个宽度为2 ϵ 的间隔带。若样本落入此间隔带，认为预测正确。

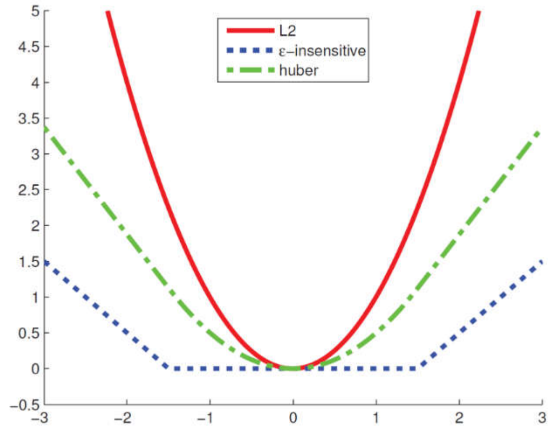


SVM for 回归

- ε insensitive loss：

$$L_\varepsilon(y,\hat{y})=\begin{cases}0 & if\ |y-\hat{y}|<\varepsilon \\ |y-\hat{y}| & otherwise\end{cases}$$

  – 误差较小时不惩罚



- 目标函数为

$$J(\mathbf{w},\lambda)=\sum_{i=1}^{N}L_\varepsilon(y_i,\hat{y}_i)+\lambda\|\mathbf{w}\|^2$$

- 亦可写成

$$J(\mathbf{w},C)=\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{N}L_\varepsilon(y_i,\hat{y}_i)$$

  – 为凸函数，但不可微

3

▶ SVM for 回归(cond.)

不止于代码

- 实际应用时，再加入松弛变量，用于表示每个点允许在ε管道外的程度

$$y_i \le f(\mathbf{x}_i) + \varepsilon + \xi_i^+$$

$$y_i \ge f(\mathbf{x}_i) - \varepsilon - \xi_i^-$$

- 则目标函数变为

$$J(\mathbf{w}, C) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\left(\xi_i^+ + \xi_i^-\right)$$

- 约束为 $\xi_i^+ \ge 0, \xi_i^- \ge 0$

▶ SVM for 回归(cond.)

- 目标函数变为

$$J(\mathbf{w}, C) = \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}\left(\xi_i^+ + \xi_i^-\right)$$

- 可以证明最优解为：$\hat{\mathbf{w}} = \sum_{i=1}^{N}\alpha_i \mathbf{x}_i$

- 预测为：$\hat{y}(\mathbf{x}) = \hat{w}_0 + \hat{\mathbf{w}}^T\mathbf{x} = \hat{w}_0 + \sum_{i=1}^{N}\alpha_i \mathbf{x}_i^T \mathbf{x}$
- 利用kernel trick 核化线性SVM：即用核函数 $k(\mathbf{x}_i, \mathbf{x})$代替点积 $\mathbf{x}_i^T\mathbf{x}$ ：

$$\hat{y}(\mathbf{x}) = \hat{w}_0 + \sum_{i=1}^{N}\alpha_i k\left(\mathbf{x}_i, \mathbf{x}\right)$$

THANK YOU