Passive Aggressive Algorithms

肖 麗穎

Introduction of Algorithms

- yt∈{+1,−1},(xt,yt)
- wt ∈R^n,(w·x)➔ |w·x|➔degree of confidence in this prediction
- hinge-loss function $\ell(\mathbf{w};(\mathbf{x},y)) = \begin{cases} 0 & y(\mathbf{w}\cdot\mathbf{x}) \geq 1 \\ 1 - y(\mathbf{w}\cdot\mathbf{x}) & \text{otherwise} \end{cases}$
- Predition of y $y_{t+1} = \arg\max_{w} (w^t \bullet \phi(x_t, y_t))$
- w🡐 $w_{t+1} = \arg\min_{w} \frac{1}{2}\|w - w_t\|^2 + C\xi$

Introduction of Algorithms

INPUT: aggressiveness parameter $C > 0$

INITIALIZE: $\mathbf{w}_1 = (0, \ldots, 0)$

For $t = 1, 2, \ldots$

- receive instance: $\mathbf{x}_t \in \mathbb{R}^n$
- predict: $\hat{y}_t = \text{sign}(\mathbf{w}_t \cdot \mathbf{x}_t)$
- receive correct label: $y_t \in \{-1, +1\}$
- suffer loss: $\ell_t = \max\{0\ ,\ 1 - y_t(\mathbf{w}_t \cdot \mathbf{x}_t)\}$
- update:

    1. set:

$$\tau_t = \frac{\ell_t}{\|\mathbf{x}_t\|^2} \qquad \text{(PA)}$$

$$\tau_t = \min\left\{C\ ,\ \frac{\ell_t}{\|\mathbf{x}_t\|^2}\right\} \qquad \text{(PA-I)}$$

$$\tau_t = \frac{\ell_t}{\|\mathbf{x}_t\|^2 + \frac{1}{2C}} \qquad \text{(PA-II)}$$

    2. update: $\mathbf{w}_{t+1} = \mathbf{w}_t + \tau_t y_t \mathbf{x}_t$

Figure 1: Three variants of the Passive-Aggressive algorithm for binary classification.

Fuction

- For classification, Passive Aggressive Classifier can be used with loss='hinge' (PA-I) or loss='squared_hinge' (PA-II). For regression, Passive Aggressive Regressor can be used with loss='epsilon_insensitive' (PA-I) or loss='squared_epsilon_insensitive' (PA-II).