Линейная классификация

Лекция 12

Бинарная классификация

Добавим в наши данные единичный признак, тогда:

$$a(x) = sign(\langle w, x \rangle) = sign(\sum_{j=0}^{d} w_j x^j)$$

Обучение линейного классификатора

$$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[a(x_i) \neq y_i]$$

$$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[sign(\langle w, x \rangle) \neq y_i]$$

$$Q(a, X) = \frac{1}{l}\sum_{i=1}^{l}[\underbrace{y_i \langle w, x \rangle}_{M} < 0]$$

Отступ

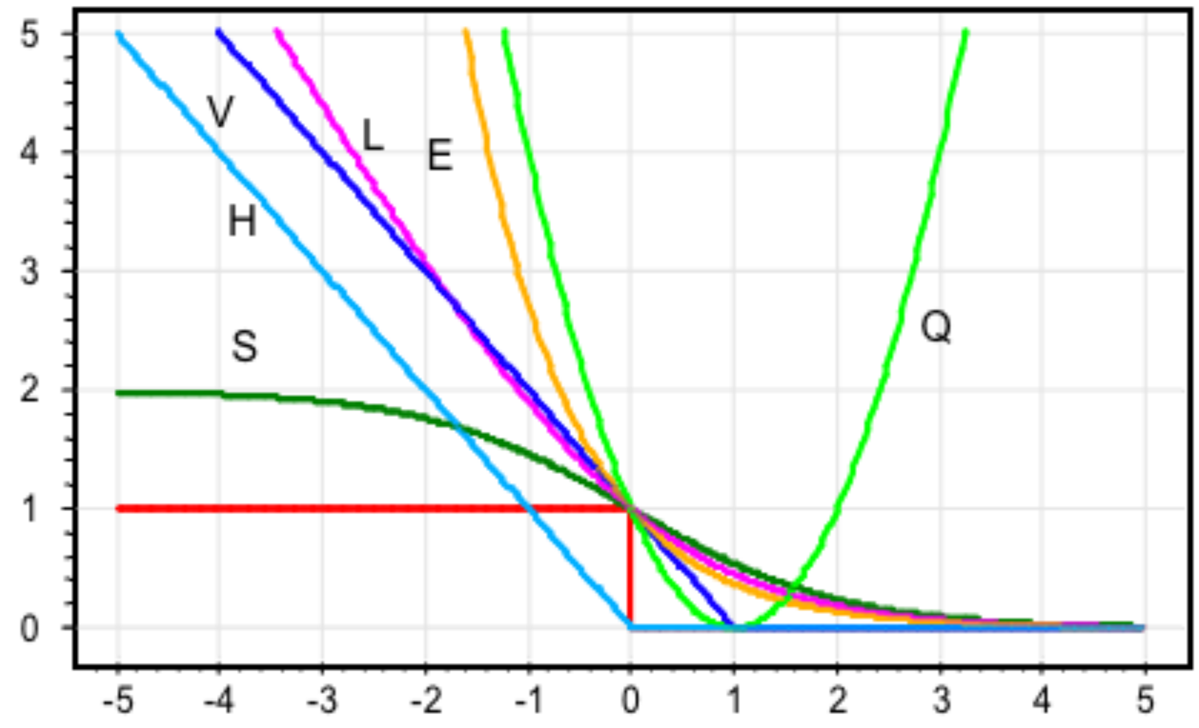
$\tilde{L}(M) = \log(1 + e^{-M})$ - логистическая функция потерь

$\tilde{L}(M) = \max(0, 1 \ - M)$ – кусочно-линейная функция потерь

$\tilde{L}(M) = max(0, -M)$ - экспоненциальная функция потерь

$\tilde{L}(M) = e^{-M}$ - экспоненциальная функция потерь

$\tilde{L}(M) = \frac{2}{1+e^{M}}$ - сигмоидная функция потерь



Обучение линейного классификатора

$$\tilde{Q}(w, X) = \frac{1}{l}\sum_{i=1}^{l} \log(1 + \exp(-y_i\langle w, x\rangle)) \to \min_{w}$$

$$\nabla_w \tilde{Q}(w, X) = -\frac{1}{l}\sum_{i=1}^{l} \frac{y_i x_i}{\log(1 + \exp(y_i\langle w, x\rangle))}$$

Логистическая регрессия

А что если мы хотим научиться предсказывать вероятности?

$$p_i = \frac{1}{1 + \exp(-\langle w, x \rangle)}$$

Метод максимального правдоподобия

$$p(y|X, w) = \prod_i p(y_i|x_i, w)$$

$$p(y|X, w) = \prod_i p_i^{y_i} (1 - p_i)^{1 - y_i}$$

Обучение

$$L(y, X, w) = -\sum_i y_i \log(p_i) + (1 \ - y_i)\log(1 \ - p_i)$$

$$\nabla_w L(y, X, w) = -\sum_i x_i (y_i - p_i)$$