Loss Functions

Tuesday, April 2, 2024 6:22 PM

Non-linear classification problem: Minimizing the problem:

$$\min_{x \in \mathbb{R}^n} \Big(c\,\underbrace{\rho}_{\text{greek } \rho}(x) + \frac{1}{\underbrace{p}_{\text{latin } p}} \sum_{i=1}^{p} \ell(x : z_i, y_i) \Big)$$

$c > 0$: regularization parameter

x : input vector

$\rho(x)$: regularizing function, strongly convex

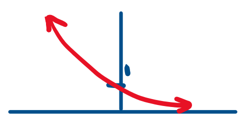
$\ell(x : z_i, y_i)$: loss function, usually convex, sometimes non-differentiable

$z_i \in \mathbb{R}^n$ feature vector (an instance of x) (a data point/sample)

$y_i \in \{+1, -1\}$ true label of sample

CONVEX LOSS FUNCTIONS :

- Logistic Regression Loss : STRONG

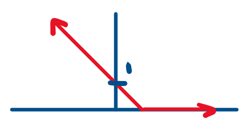


$$\ell(x : z, y) = \log\big(1 + \exp(-y\langle x, z\rangle)\big)$$
$$= \log\big(1 + \exp(-y(x^T z))\big)$$

$y \in \{-1, 1\}$ data point's label

↳ Logistic Regression (Linear) classifier Problem.

- Hinge Loss :

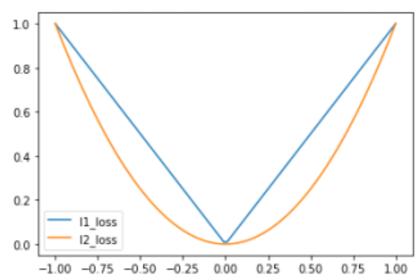


$$\ell(x : z, y) = \max\{0, 1 - y(x^T z)\}$$

↳ Maximum margin (Linear) classifier problems

- L1 Loss (Absolute Loss)

$$\ell(x : z, y) = \sum_{i=1}^{n} |z_i - y_i|$$



- L2 loss (Quadratic Loss)

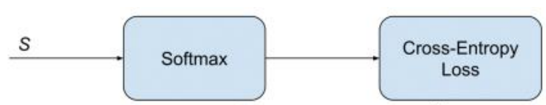
$$\ell(x : z, y) = \sum_{i=1}^{n} (z_i - y_i)^2$$

- Softmax Loss (Softmax Function → Cross Entropy Loss Function)

(Multi-Class y)

$$\ell(x : z, y) = \sum_{c=1}^{M} z_c \log(y_c)$$

M : # of classes



$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}} \qquad CE = -\sum_i^C t_i log(f(s)_i)$$