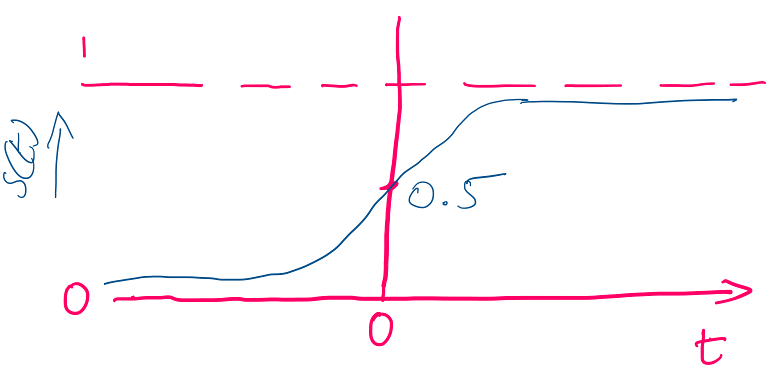
Logistic Regression

Saturday, 8 March 2025 9:14 PM

Purpose : Estimate the probability that a point belongs to one of two classes $y \in \{0,1\}$

Logistic function (Sigmoid)

$$s(t) = \frac{1}{1+e^{-t}} = \frac{e^t}{1+e^t}$$



For $t \gg 0$, $s(t) = 1$
$t \ll 0$, $s(t) = 0$

Logistic Regression

For a point a_i^T and its label $y \in \{0,1\}$

$$P(y=1 \mid a_i^T) = s(a_i^T x)$$
$$= \frac{1}{1+e^{-w^T x}}$$

The output $s(a_i^T x)$ is interpreted as a probability

$a_i^T x \gg 0 \Rightarrow s(a_i^T x) \gg 0.5$

$\Rightarrow P(y=1 \mid a_i^T) \approx 1$

$\therefore a_i^T$ is classified as class 1

$a_i^T x \ll 0 \Rightarrow s(a_i^T x) \ll 0.5$

$\Rightarrow P(y=1 \mid a_i^T) \approx 0$

$\therefore a_i^T$ is classified as 0.

Cost function / Loss

For data $\{(a_1^T, y_1), (a_2^T, y_2), \ldots, (a_n^T, y_n)\}$

estimate

$$P(y_i = 1 \mid a_i^T) = \frac{1}{1+e^{-a_i^T x}} = \pi_i$$

$$loss(x) = -\sum_{i=1}^{n} \left[y_i \cdot \ln \pi_i + (1-y_i) \ln(1-\pi_i) \right]$$

Cost Function Interpretation

Suppose there is only one datapoint say (a_i^T, y_i)

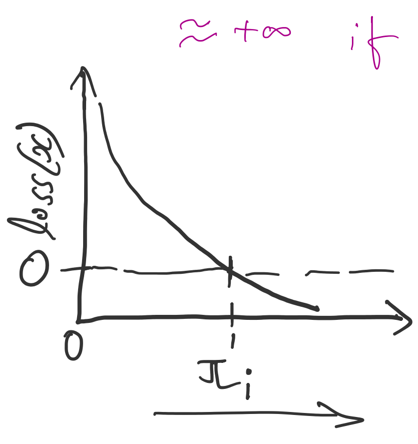
$$\therefore loss(x) = \begin{cases} -\ln \pi_i & \text{if } y_i = 1 \\ -\ln(1-\pi_i) & \text{if } y_i = 0 \end{cases}$$

Case $y_i = 1$

$loss(x) = -\ln \pi_i$

This is ≈ 0 if $y_i = 1$ and $\pi_i \approx 1$

$\approx +\infty$ if $y_i = 1$ and $\pi_i \approx 0$

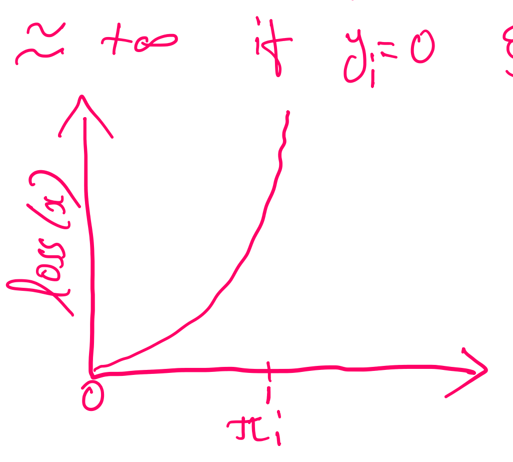


Case $y_i = 0$

$loss(x) = -\ln(1-\pi_i)$

≈ 0 if $y_i = 0$ & $\pi_i = 0$

$\approx +\infty$ if $y_i = 0$ & $\pi_i = 1$



So, training means give data Matrix A & label vector y
find $x \in R^d$
s.t. $loss(x)$ is minimized

$$x^* = \underset{x \in R^d}{\operatorname{argmin}} \; loss(x)$$

x^* is our model

For a new point a_{n+1}^T

Let $P = \frac{1}{1+e^{-a_{n+1}^T x^*}}$

if $P \geq 0.5$, class $= 1$
else class $= 0$

One vs all for Multiclass classification

→ Since prediction still depends on $a^T x$ it is a generalized linear model