Lecture Notes - Logistic Regression 02/01/2021

- HW #1 due today
- HW #2 posted later today

[From the last class]

$$P(Y=1 \mid X) = \frac{1}{1 + \exp\left(\log \frac{P(Y=0)}{P(Y=1)} + \sum_{j=1}^{d} \log \frac{P(x_j \mid Y=0)}{P(x_j \mid Y=1)} \right)}$$

$$\Rightarrow \log \frac{P(Y=1)}{P(Y=0)} + \sum_{j=1}^{d} \log \frac{P(x_j \mid Y=1)}{P(x_j \mid Y=0)} \propto \underset{\mathbb{R}^d}{w}^T x + \underset{\mathbb{R}}{w_0}$$

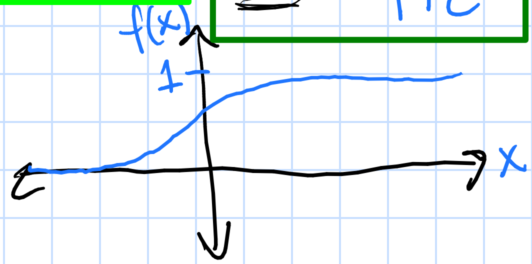
Side Note

$y \in \pm 1$ class labels $g(x) = w^T x + b$

Then

$$f_w(x) = \boxed{P(Y=1 \mid X)} = \frac{1}{1 + \exp(\underbrace{-w^T x - w_0}_{\text{regression}})} \quad \text{or} \quad \frac{1}{1 + e^{-w^T x}}$$

Logistic Function: $\boxed{f(x) = \frac{1}{1 + e^{-x}}}$



We need to find w and w_0!

$(x_1, 1), (x_2, 0)$

$f_w(x_1) \approx 1 \qquad 1 - f_w(x_2) \approx 1$

$$\log \frac{P(Y=1|X)}{P(Y=0|X)} = W^T X$$

The odds in favor of the event $Y=1$ is $\frac{p}{1-p}$ where p is the probability of the event.

LR assumes that the log odds is a linear function of X.

- $W^T X$ should have large negative values for $Y=0$
- $W^T X$ should have large positive values for $Y=1$

$\boxed{\text{Find } w \text{ and } w_0!}$

How do we find w?

① $g(x) =$

② $\log P(Y|X)$

With $f_w(x)$ we have a non-convex opt task,

$$f_w(x) = \frac{1}{1+e^{-w^T x}}, \quad J(w) = \frac{1}{2}\sum_{i=1}^{n}\left(f_w(x_i) - y_i\right)^2$$

This is a non convax optiomization problem
we can't just take the derivative.

$y_i \in \{0, 1\}$

$f_w(x) \in [0, 1]$

Let us find the parameters, w (w_0), that maximize the likelihood of the data.

$$L(w) = \prod_{i=1}^{n} p(y_i \mid x_i ; w)$$

$$\boxed{f_w(x_i)}$$

Now we want to maximize the likelihood fcn

$$w^* = \arg\max_{w \in \mathbb{R}^d} \prod_{i=1}^{n} p(y_i \mid x_i ; w)$$

$$= \arg\max \log \left\{ \prod_{i=1}^{n} p(y_i \mid x_i ; w) \right\}$$

$$= \arg\max \sum_{i=1}^{n} \log \boxed{p(y_i \mid x_i ; w)} \quad (1)$$

$y_i \in \{0, 1\}$;

we are working $P(Y=1 \mid X)$

(2) $$= \arg\max \sum_{i=1}^{n} \left[y_i \log\left(p(y_i \mid x_i ; w)\right) + (1-y_i) \log\left(1 - p(y_i \mid x_i ; w)\right) \right]$$

$$= \arg\min - \sum_{i=1}^{n} y_i \log\left(f_w(x_i)\right) + (1-y_i) \log\left(1 - f_w(x_i)\right)$$

$$= \arg\min - \sum_{i=1}^{n} \left[y_i \log \hat{y}_i + (1-y_i) \log(1 - \hat{y}_i) \right]$$

* Convex

Has no closed form solution.