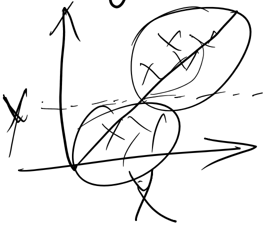
Logistic Regression

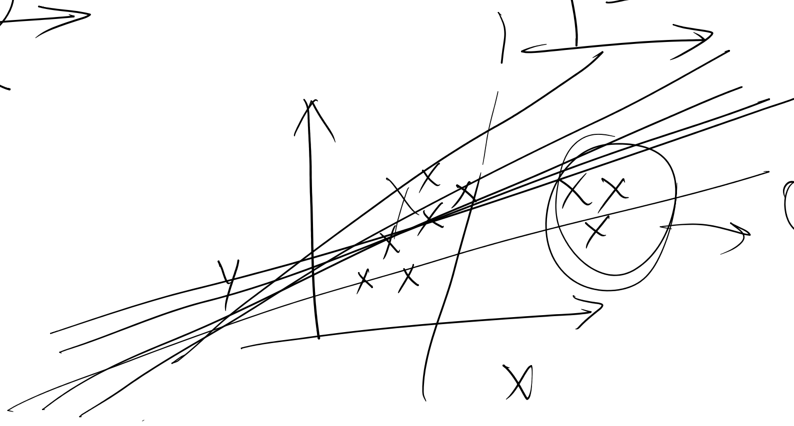
26 March 2023 19:14

Why not LR for classification task?



LR plot

Prob → X → P(0.5)
Y → P(0.2)
Z → P(0.1)



Outliers

0
60
70
101 —
200 —

? Multiclass
↳ LR will not solve?

LOGISTIC REGRESSION
↓
But it is actually CLASSIFICATION Algorithm

→ supervised.
→ log odds prop.

$$\text{odds} = \frac{P(\text{event})}{1 - P(\text{event})}$$

odds → ↑
↓

* conditional prob =

$$P(y=1/X) = P($$

class

Reqt → We need a prob model curve where predictor domain $X \in R_d$, $p(X) \in [0,1]$

↓
Logistic Regression

Sigmoid f^n : $p(X) = \quad 1$

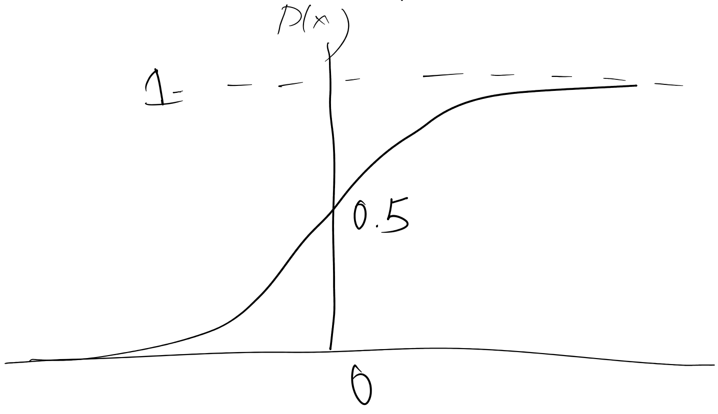
Logistic Regression

$\boxed{\text{Sigmoid } f^n}$: $p(X) = \dfrac{1}{1 + e^{-\beta X}}$

F^n in more admissible form

$\left[\log \left[\dfrac{p(X)}{1 - p(X)} \right] \right] = \beta X$

→ logit v/s X ⇒ linear



~~P(~~

Parameter Estimation

→ Estimate $\hat{\beta}$ vector

Gradient Descent
Least Sq. Methods

Linear

Max. Likelihood

Logistic.

Consider N samples w/ labels 0 & 1

* Label 1 : Estimate $\hat{\beta}$ s.t. that $p(x) \approx 1$
  Label 0 : $(1 - p(X)) \approx 1$

Presenting Mathematically,

* For every label 1 data pt. we want to estimate $\hat{\beta}$ such that product of all cond^nal prob of class 1 ≈ 1.

$\prod_{s \text{ in } y_i = 1} p(x_i)$ $\qquad\qquad$ $\prod_{s \text{ in } y_i = 0} [1 - p(x_i)]$

Maximize,

$L(\beta) = \prod_{s \text{ in } y_i = 1} p(x_i) \times \prod_{s \text{ in } y_i = 0} [1 - p(x_i)]$ } needs to be optimised ⇓ c/a likelihood

$L(\beta) = \prod_{s} p(x_i)^{y_i} \times [1 - p(x_i)]^{1 - y_i}$ [combining products]

$$L(\beta) = \prod_{s} p(x_i)^{y_i} \times [1 - p(x_i)]^{1-y_i}$$ [combining products]

$$l(\beta) = \sum_{i=1}^{n} y_i \log p(x_i) + (1-y_i) \log [1 - p(x_i)]$$

p(x) in exponent form

$$l(\beta) = \sum_{i=1}^{n} y_i \log\left(\frac{1}{1+e^{-\beta x_i}}\right) + (1-y_i)\log\left[\frac{e^{-\beta x_i}}{1+e^{-\beta x_i}}\right]$$

grouping coeff of y_i

$$= \sum_{i=1}^{n} y_i \left[\log\left(\frac{1}{1+e^{-\beta x_i}}\right) - \log\left(\frac{e^{-\beta x_i}}{1+e^{-\beta x_i}}\right)\right]$$

$$+ \log\left(\frac{e^{-\beta x_i}}{1+e^{-\beta x_i}}\right)$$

Simplification

$$l(\beta) = \sum_{i=1}^{n} y_i \left[\log e^{\beta x_i}\right] + \log\left(\frac{e^{-\beta x_i}}{1+e^{-\beta x_i}} \cdot \frac{e^{\beta x_i}}{e^{\beta x_i}}\right)$$

$$= \sum_{i=1}^{n} y_i \beta x_i + \log\left(\frac{1}{1+e^{\beta x_i}}\right)$$

$$\boxed{l(\beta) = \sum_{i=1}^{n} y_i \beta x_i + \log\left(\frac{1}{1+e^{\beta x_i}}\right)}$$

→ find β to maximize this f^n.

$$l(\beta) = \sum_{i=1}^{n} y_i \beta x_i - \log\left(1+e^{\beta x_i}\right)$$

→ transcedental eq^n

We apply numerical method

Newton Raphson method to

maximize f^n

$$\boxed{p(X) = \frac{1}{1+e^{-\beta X}}}$$