COL774
Machine Learning
Sep 3, 2021

Unless mandated by institute o.w.

Logistics:

① Audit Criteria:-

(Equivalent of B-) OR
or better

(At least 30% of the score in exams (Minor+Major)
AND At least 30% of the score in assignment
AND At least 40% of the overall score)

II :- We will release sol. to past papers] Do Not release them publicly.

III :- Forum:- Any logistical issues - → Being on teams / ↳ COL774 | DC list / ↳ others.

→ TA:- assignment :-

IV :- We will have make-up of Sunday 8am - 9am] 3rd Make-up class

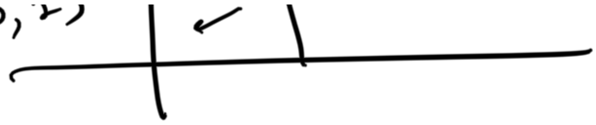
⇒ Last class:-

Logistic Regression: + separator (linear)

$\{x^{(i)}, y^{(i)}\}_{i=1}^{m}$

$y^{(i)} | x^{(i)}; \theta \sim \text{Bernoulli}(\phi^{(i)})$,

$\{0,1\}$



$$\phi^{(i)} = \underbrace{\left[\frac{1}{1+e^{-\theta^T x^{(i)}}}\right]}_{\in[0,1]} g(\theta^T x^{(i)})$$

$\Rightarrow$

$$\underbrace{LL(\theta)}_{\text{concave}} = \log \prod_{i=1}^{m} p(y^{(i)}|x^{(i)};\theta) = \sum_{i=1}^{m} \log p(y^{(i)}|x^{(i)};\theta)$$

$\hookrightarrow$ concave — HW / Exam

$$\left.\begin{array}{l} y^{(i)}=1 \\ y^{(i)}=0 \end{array}\right\}$$

$$\arg\max_{\theta} LL(\theta)$$

$$\hookrightarrow \frac{1}{m}\nabla_\theta LL(\theta) = \frac{1}{m}\sum_{i=1}^{m}\Big[y^{(i)} - \underbrace{\frac{1}{1+e^{-\theta^T x^{(i)}}}}_{p(y^{(i)}=1|x^{(i)};\theta)}\Big]x^{(i)}$$

Define $h_\theta(x) = \underbrace{p(y=1|x;\theta)}$

$$= \frac{1}{1+e^{-\theta^T x}} \leftarrow \text{linear reg: } \theta^T x^{(i)} \quad \hookrightarrow \text{Logistic}$$

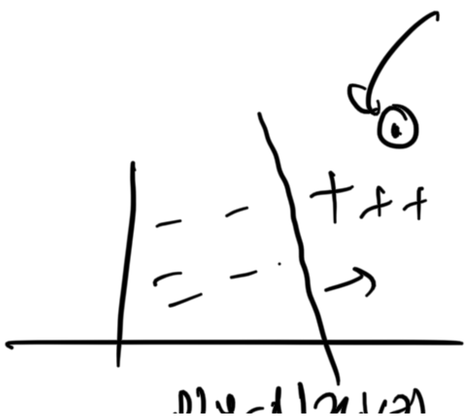
do {
$\theta^{(t+1)} \leftarrow \theta^{(t)} + \gamma \cdot \nabla_\theta LL(\theta)$
$t \leftarrow t+1$;
} while (! converged) $\rightarrow$

$$\downarrow \nabla_\theta LL(\theta) = \sum_{i=1}^{m}\underbrace{\left(y^{(i)} - h_\theta(x^{(i)})\right)}x^{(i)}\frac{1}{m}$$

$$\hat{y} := 1\{h_\theta(x) \geq 0.5\}$$

Linear separator.

$$p(y=1|x;\theta) = \frac{1}{1+e^{-\theta^T x}}$$



At decision boundary ① $p(y=1|x;\theta) = 0.5$

$\Rightarrow \frac{1}{2} = \frac{1}{1+e^{-\theta^T x}}$

$\Rightarrow 1+e^{-\theta^T x} = 2$

$\Rightarrow e^{-\theta^T x} = 1$

$\Rightarrow -\theta^T x = \log 1 \Rightarrow \boxed{\theta^T x = 0}$

Generalized Linear Models (GLM):-

↳ Linear Regression
↳ Logistic Regression

$\{x^{(i)}, y^{(i)}\}_{i=1}^{m}$

$y^{(i)} | x^{(i)}; \theta \sim$ Dist.

$P(y;\eta)$ → function [θ & x] natural parameter

$y;\eta \sim$ Exponential family Distribution

$\underbrace{\hphantom{Exponential family Distribution}}_{\text{Class}}$

① ↳ $P(y;\eta) = b(y) \exp(\eta y - a(\eta))$

→ $E[y|x;\theta]$

$b(y)$ ⇓ some fn. of y

→ H.W:- Prove: ① Bernoulli (ϕ) ② Normal Dist. ($\mu, \sigma^2 = 1$) $\in$ Exp. family

$y^{(i)}, x$ $\{x^{(i)}, y^{(i)}\}_{i=1}^{m}$

$LL(\eta^{(i)}) = \log \prod_{i=1}^{m} (P(y^{(i)};\eta))$

$(\theta, x^{(i)})$

$$= \sum_{i=1}^{m} \log p(y^{(i)}; \eta)$$

$$= \sum_{i=1}^{m} \left[\log\left[b(y^{(i)})\right] + \left[\eta^{(i)} y^{(i)} - a(\eta^{(i)})\right]\right]$$

$$\nabla_{\eta^{(i)}} LL(\eta^{(i)}) = \sum_{i=1}^{m} 0 + \left[y^{(i)} - a'(\eta^{(i)})\right]$$

(II) $\eta = \theta^T x$:- Linear dependence between x & y.

$\eta^{(i)} = \theta^T x^{(i)}$

$\nabla_{\eta^{(i)}} LL(\eta^{(i)})$

$$\nabla_\theta LL(\theta) = \sum_{i=1}^{m} \left[y^{(i)} - a'(\eta^{(i)})\right] \cdot \nabla_\theta \underbrace{\left[\eta^{(i)}\right]}_{\theta^T x^{(i)}}$$

$$= \sum_{i=1}^{m} \left[y^{(i)} - a'(\eta^{(i)})\right] x^{(i)}$$

(III) $h_\theta(x) = E[y|x;\theta]$ $\quad \eta = \theta^T x$

↳ $\theta^T x$

↳ $\frac{1}{1+e^{-\theta^T x}}$

$$E[y|x;\theta] = g(\eta) \equiv g(\theta^T x)$$

$$g^{-1}\{E[y|x;\theta]\} = \theta^T x \quad :-$$

Linearity assumption.

↓ Link function

↳ — Linear Regression :→

g = identity

$$E[y|x;\theta] = \theta^T x$$

Logistic Regression :-

$$\underbrace{E[y|x;\theta]}_{\phi} = \frac{1}{1+e^{-\theta^T x}}$$

$$\phi = \frac{1}{1+e^{-\theta^T x}}$$

$$\Rightarrow 1+e^{-\theta^T x} = \frac{1}{\phi}$$

$$e^{-[\theta^T x]} = \frac{1}{\phi} - 1 = \frac{1-\phi}{\phi}$$

$$\Rightarrow -\theta^T x = \log\left(\frac{1-\phi}{\phi}\right)$$

$$\boxed{\log\left(\frac{\phi}{1-\phi}\right) = \theta^T x}$$ Logistic reg.

$\xrightarrow{g^{-1}}$ log odds / logit

ϕ, μ: Expected value

$$\boxed{\mu = \theta^T x}$$

Poisson - Regression $y|x;\theta \sim \text{Poisson}(\lambda)$

GDA :- Generative Models] classification
(Gaussian Discriminant Analysis) → classification prob

| { $y^{(i)} \sim \text{Bernoulli}(\phi)$

Random

Logistic Regression

$y^{(i)} | x^{(i)} ; \theta \sim \text{Bernoulli}(\phi^{(i)})$

$p(y^{(i)} | x^{(i)} ; \theta)$

$\prod_{i=1}^{m} p(y^{(i)} | x^{(i)} ; \theta)$

$LL(\theta)$

$(x^{(i)} | y^{(i)} ; \theta) \sim \text{Normal}(\mu, \Sigma)$

$(x^{(i)}, y^{(i)}) \sim$

$p(x^{(i)}, y^{(i)} ; \theta) \leftarrow$

$= \sim p(y^{(i)} ; \theta)$

$\hookrightarrow p(x^{(i)} | y^{(i)} ; \theta)$

$\prod_{i=1}^{m} p(x^{(i)}, y^{(i)} ; \theta)$

$LL(\theta) \uparrow$