Poisson regression .

08.05.2018

~~Let λ be average rate of occuring an event E. Poisson random variable is defined as at any given "instance" the value of occurance of event E.~~

Poisson distribution is the discrete probability distribution of the #. of events occuring in a given time period, given the average #. of times the event occurs over that period.

Range of poisson Random variable $R = \{0, 1, 2, \ldots\}$

Probability mass function (Pmf) :

$$P(y) = \frac{e^{-\lambda} \lambda^{y}}{y!} \quad \forall y \in R$$

Given a data set, we assume target variable "y" is distributed as Poisson distribution

$$y \mid x ; \theta \sim \text{Poisson}(\lambda)$$

Show poisson distribution is a member of Exponential family.

$$P(y|x;\theta) = \frac{e^{-\lambda} \lambda^{y}}{y!} \quad -(1)$$

It poisson distribution is in Exponential family iff its pmf can be written as

$$P(y|x;\eta) = b(\eta) \exp(\eta^{T} T(y) - a(\eta)) - (2)$$

from (1), $P(y|x;\theta) = \frac{1}{y!} \exp(\log(e^{-\lambda} \lambda^{y}))$

Log here is base "e"

$$= \frac{1}{y!} \exp(-\lambda + y \log(\lambda))$$

Comparing with eqn. (2)

$\boxed{b(y) = \frac{1}{y!}}$, $\boxed{T(y) = y}$, $\boxed{\eta = \log(\lambda)}$, $\boxed{a(\eta) = \lambda}$

$\lambda = e^{\eta}$ $\qquad$ $a(\eta) = e^{\eta}$

Building generalized linear model-(Constructing hypothesis for m...

$y \mid x ; \theta \sim$ Exponential family (λ) (Poisson (λ))

expected Value, $h_\theta(x) = E[T(y) \mid x] = E[y \mid x] = \lambda = e^{\eta}$ — ③

$\nearrow$ hypothesis

$\eta = \theta^T x$ — ④, from eqn. ③ and ④ $\boxed{h_\theta(x) = e^{\theta^T x}}$

what is the cost function ? — log likelihood function

& $h_\theta(x) = \lambda = e^{\theta^T x}$

$y \mid x ; \theta \sim$ Poisson (λ)

Pmf: $p(y \mid x ; \theta) = \dfrac{e^{-\lambda} \lambda^{y}}{y!}$

likelihood, $\boxed{L(\theta) = \prod_{j=1}^{m} \dfrac{e^{-\lambda} \lambda^{y^{(j)}}}{y^{(j)}!}}$ $m = \#$ of training examp...

$$= \prod_{j=1}^{m} \frac{e^{-e^{\theta^T x^{(j)}}} \, e^{\theta^T x^{(j)} y^{(j)}}}{y^{(j)}!}$$

$$l(\theta) = \log L(\theta) = \frac{1}{m} \sum_{j=1}^{m} \log e^{-e^{\theta^T x^{(j)}}} + \log e^{\theta^T x^{(j)} y^{(j)}} - \log(y^{(j)}!)$$

$$= \frac{1}{m} \sum_{j=1}^{m} -e^{\theta^T x^{(j)}} + \theta^T x^{(j)} y^{(j)} - \log(y^{(j)}!)$$

Since $l(\theta)$ is ~~convex~~ concave function ; $-l(\theta)$ will be convex

maximizing $l(\theta)$ = minimizing $-l(\theta)$

negative log likelihood, $l(\theta) = -\dfrac{1}{m} \sum_{j=1}^{m} -e^{\theta^T x^{(j)}} + \theta^T x^{(j)} y^{(j)} - \log y^{(j)}!$

$$l(\theta) = \frac{1}{m}\sum_{j=1}^{m} e^{\theta^T x^{(j)}} + \log y^{(j)}! - \theta^T x^{(j)} y^{(j)}$$

Derivation ~~maximizing~~ <u>minimizing</u> negative log likelihood ?

~~minimization $l(\theta)$~~ $\rightarrow$ $J(\theta)$ using gradient descent rule

$$J(\theta) = l(\theta)$$

$\leftarrow$ negative log likelihood

$$\frac{\partial}{\partial\theta_i} J(\theta) = \frac{1}{m}\sum_{j=1}^{m} \frac{\partial}{\partial\theta_i}\left(e^{\theta^T x^{(j)}} + \log y^{(j)}! - \theta^T x^{(j)} y^{(j)}\right)$$

$$= \frac{1}{m}\sum_{j=1}^{m} e^{\theta^T x^{(j)}} \frac{\partial(\theta^T x^{(j)})}{\partial\theta_i} - \frac{\partial}{\partial\theta_i} \theta^T x^{(j)} y^{(j)})$$

$$= \frac{1}{m}\sum_{j=1}^{m} (e^{\theta^T x^{(j)}} x_i^{(j)} - x_i^{(j)} y^{(j)})$$

$$= \frac{1}{m}\sum_{j=1}^{m} \left(e^{\theta^T x^{(j)}} - y^{(j)}\right) x_i^{j}$$

How to update parameters ?

Repeat until convergence {

$$\theta_i := \theta_i - \frac{\alpha}{m}\sum_{j=1}^{m}\left(e^{\theta^T x^{(j)}} - y^{(j)}\right) x_i^{(j)}$$

$\forall i \in \{0, \ldots, n\}$

}

n = #. of features +1

$\theta_0 x_0 + \theta_1 x_1$ $f\theta_n x_n$

$x_0 = 1$