Lesson Plan 4/19

Thursday, April 19, 2018 3:23 PM

Admin

- midterm on Tuesday

HDSC

- The ML algorithms used in self driving cars https://www.kdnuggets.com/2017/06/machine-learning-algorithms-used-self-driving-cars.html

Review

Outcome $y: \{0, 1\}$

predictors $x_1, \dots, x_p$

Logistic Regression (log odds)

$$\log \frac{P(Y=1 \mid X=x)}{\quad} = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

$P(Y=0|X=x)$

odds

Q1) $p(Y=1|X=x) = 25\%$

odds ? $\dfrac{1/4}{3/4} = \dfrac{1}{3}$

Q2) $\log \dfrac{P(Y=1|X_1, X_2)}{P(Y=0|X_1, X_2)} = -6 + 0.05X_1 + 1X_2$

$= -6 + 0.05(30) + 3.5$

$$= -2.5 + 1.5$$

$$\log\left\{\frac{P(y=1|x)}{P(y=0|x)}\right\} = -1$$

$$f(\beta; x) = \beta_0 + \beta_1 x + \dots + \beta_p x_p$$

$$P(y=1 | X) = \frac{e^{f(\beta; x)}}{1 + e^{f(\beta; x)}}$$

$$\frac{e^{-1}}{1+e^{-1}} = \frac{1/e}{1+1/e} = \frac{e \cdot 1/e}{e \cdot (1 + 1/e)} = \frac{1}{1+e}$$

Q3) Given $p(y=1 | ; x) = .5$ what is the value of x_1?

Last time was

1) Gradient descent to solve linear regression

Model $f(\beta; x) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$

Loss function

$$L(\beta) = \frac{1}{2} \sum_{i=1}^{n} \left(y_i - f(\beta; x_i)\right)^2$$

$$\vec{\beta}^{k+1} = \vec{\beta}^{k} + \alpha \sum_{i=1}^{n} \left(y_i - f(\beta; \vec{x}_i)\right) \vec{x}_i$$

residual

$= -\nabla_\beta L(\beta)$

Step size

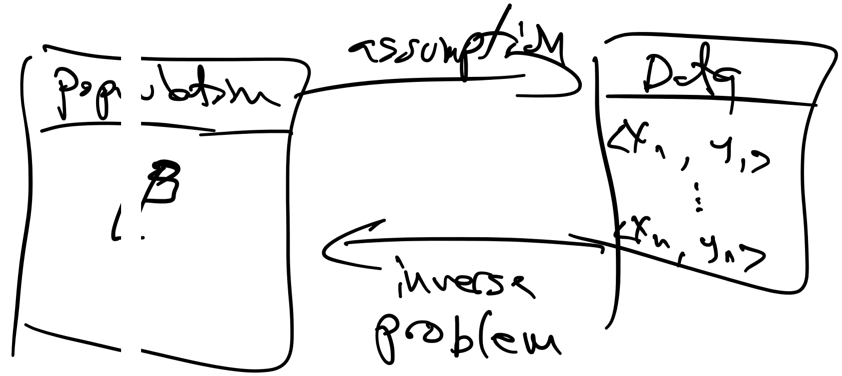
$$\beta^{k+1} = \beta^{k} - \alpha \boxed{\nabla_\beta L(\beta)}$$

$$\begin{bmatrix} \beta_0^{k+1} \\ \beta_1^{k+1} \\ \vdots \\ \beta_p^{k+1} \end{bmatrix} = \begin{bmatrix} \beta_0^{k} \\ \beta_1^{k} \\ \vdots \\ \beta_p^{k} \end{bmatrix} - \alpha \begin{bmatrix} \frac{\partial L(\beta)}{\partial \beta_0} \\ \frac{\partial L(\beta)}{\partial \beta_1} \\ \vdots \\ \frac{\partial L(\beta)}{\partial \beta_p} \end{bmatrix}$$

What is the loss function for logistic regression??

outcome $Y \in \{0, 1\}$

Predictors $X_1, \ldots, X_p$



Assumption:

$$Y_i \sim \text{Bernoulli}(p(\beta; X_i))$$

independent

What is the inverse problem?

- Find β that maximizes probability of observed data under

$$p(Y_i = y_i \mid X = x_i) = p(\beta; x_i)^{y_i} \left(1 - p(\beta; x_i)\right)^{(1-y_i)}$$

$$\mathcal{L}(\beta) = \prod_{i=1}^{n} p(\beta; x_i)^{y_i} \left(1 - p(\beta; x_i)\right)^{(1-y_i)}$$

likelihood

$\Downarrow$ loss function

negative log likelihood: loss function

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

1. Assume a probability model (Bernoulli)
2. Derive data probability (likelihood)
3. Use negative log likelihood as loss function
4. Define update rule

For logistic regression:

$$\beta^{k+1} = \beta^{k} + \alpha \sum_{i=1}^{n} \underbrace{\left(y_i - p(\beta; x_i)\right)}_{\text{"residual"}} \vec{x}_i$$

$$- \alpha \nabla_\beta L(\beta)$$