SMGT 432: Bradley-Terry Lecture

Announcements:

1. Rice Soccer home opener vs. UH tomorrow night!
2. Today is a lecture (points opportunity!)
3. Assignment #1 is due next Friday

Questions:

1. What distinguishes the Bradley-Terry model from other regression models?
2. What are one strength and one weakness of the Bradley-Terry model?
3. How can you avoid overfitting when estimating a Bradley-Terry model?

$\boxed{\text{Story}}$ first time trying to estimate this model as a grad student

$\boxed{\text{Pause}}$ Has anyone heard of the Bradley-Terry model?

Bradley-Terry Model

$$Y \sim \text{Normal}(\eta, \sigma^2)$$

$$\eta = \alpha + \beta_{HOME} - \beta_{AWAY}$$

$$\vec{y} = \begin{bmatrix} 23 \\ -4 \\ \dots \\ 1 \\ 9 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & -1 \\ 1 & -1 & 0 & \dots & 1 & 0 \\ & & & \dots & & \\ 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Rasch Model

$$Y \sim \text{Normal}(\eta, \sigma^2)$$

$$\eta = \alpha + \beta_{OFFENSE} + \gamma_{DEFENSE}$$

$$\vec{y} = \begin{bmatrix} 96 \\ 87 \\ \dots \\ 101 \\ 91 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ 1 & 0 & 1 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & & \dots & & & & \dots & \\ 1 & 0 & 0 & \dots & 1 & 0 & 1 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$\boxed{\text{Emphasize}}$ What makes BT unique is that the design matrix X is ~~sparse~~ sparse and always 0/1/-1.

$\boxed{\text{Pause}}$ Confess love for BT model

How does it work?

Maximum Likelihood

$$\mathcal{L}(\alpha, \beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(y_i-\eta_i)^2}{2\sigma^2}}$$

$$\ell(\alpha, \beta) \equiv \log(\mathcal{L}(\alpha, \beta)) = -\frac{n}{2}\log(2\pi\sigma^2) + \sum_{i=1}^{n} \frac{-(y_i-\eta_i)^2}{2\sigma^2}$$

$$\hat{\alpha}, \hat{\beta} = \arg\max_{\alpha, \beta} \sum_{i=1}^{n} -(y_i-\eta_i)^2$$

$$= \arg\min_{\alpha, \beta} \sum_{i=1}^{n} \cdot$$

Ex 1: $\alpha = 3$ $\beta_{Rice} = 4$ $\beta_{UH} = 0$

Rice beats UH by 7.

$$\ell(\alpha, \beta) = -(7-(3+4-0))^2 = 0$$

Rice beats UH by 10.

$$\ell(\alpha, \beta) = -(10-(3+4-0))^2 = -9$$

Generalizations

Generalized Linear Models (GLMs):

$$Y \sim \text{Bernoulli}\left(\frac{e^{\eta}}{1+e^{\eta}}\right)$$

$$Y \sim \text{Poisson}(e^{\eta})$$

Ex. 2: (H) AFC Richmond 3 - 2

$$\mathcal{L}(\eta) = P(Y=3) = \frac{\lambda^3 e^{-\lambda}}{3!} \qquad \lambda = e^{\eta}$$

$$\ell(\eta) = \log\left(\frac{\lambda^3 e^{-\lambda}}{3!}\right) = 3\log\lambda - \lambda - \log 3!$$

$$\frac{\partial}{\partial\lambda}\ell(\eta) = 3/\lambda - 1$$

$$\frac{\partial}{\partial\lambda}\ell(\eta) = 0 \Rightarrow \lambda = 3 \qquad \Rightarrow \eta = \log 3 = 1.1$$

$$\eta = \alpha + \beta_{Richmond} + \gamma_{WH}$$

$$1.1 = 0.37 + \beta_{Richmond} + 0.35$$

$$\Rightarrow \beta_{Richmond} = 0.38$$

Pause This means we're estimating Richmond to be the strongest offense in the league. How do we feel about this?

Preventing Overfitting

1. Random/Mixed-Effect Model (Traditional Statistics)
2. Bayesian Regression (Bayesian Statistics)
3. Regularization (Statistical Machine Learning)