REGRESSION MODELS FOR ORDINAL DATA: A CONCISE SUMMARY

ABHAY SHANKAR K: CS21BTECH11001 AND KARTHEEK SRIRAM TAMMANA: CS21BTECH11028

1. Paper Summary

- The paper proposes two models for ordinal data, namely the proportional odds model and the proportional hazards model.
- The proportional odds model is a generalisation of the logistic regression model for ordinal data. Here, the odds of the response variable $Y \leq j$ are given by

$$\kappa_j = \kappa_j \exp(-\boldsymbol{\beta^T}\mathbf{x})$$

.

- The proportional hazards model considers a hazard function $\lambda(t)$, which expresses the probability of failure at time t, of the form

$$\lambda(t) = \lambda_0(t) \exp\left(-\boldsymbol{\beta^T}\mathbf{x}\right)$$

.

- The paper proposes a generalised empirical logit transform for the two models.
- The paper also discusses
  - The properties of the two models, proposing a few alternative link functions.
  - Invariances of the models under reversal of the ordering.
  - Asymptotic properties of the two models.
  - Parameter estimation for both models.
  - Application of the models to real data.

2. Parameter Estimation

Revising the notation from the paper, we have the probabilities of the k ordered categories of the response variable Y given by $\{\pi_1, \dots, \pi_k\}$, as a function of the covariant vector $\mathbf{x}$, and their cumulative probabilities given by $\gamma_j = \sum_{i=1}^{j} \pi_j$. The cumulative odds are thus $\boldsymbol{\kappa}_j = \frac{\gamma_j}{1-\gamma_j}$.

We then have the likelihood function $\boldsymbol{\kappa}_j = \kappa_j \exp\left(\boldsymbol{\beta^T}\mathbf{x}\right)$, which we can reframe as

$$\begin{aligned} \frac{\gamma_j}{1-\gamma_j} &= \exp\left(\theta_j - \boldsymbol{\beta^T}\mathbf{x}\right) \\ \implies \gamma_{j+1} &= \frac{1}{1+\exp\left(\theta_j - \boldsymbol{\beta^T}\mathbf{x}\right)} \end{aligned} \tag{1}$$

where we set $\theta_0 = 0$. We also define $R_j = \sum_{i=1}^{j} n_i$.

Date: October 8, 2023.

We have the likelihood function:

$$
\begin{aligned}
p(\mathbf{y}|\mathbf{x},\boldsymbol{\beta}) &= \prod_{j=1}^{k} \pi_j^{n_j} \\
&= \pi_1^{n_1} \prod_{j=1}^{k-1} (\gamma_{j+1} - \gamma_j)^{n_{j+1}} \\
&= \pi_1^{n_1} \prod_{j=1}^{k} (\gamma_{j+1} - \gamma_j)^{R_{j+1} - R_j} \\
&= \prod_{n=1}^{k} \left(\frac{\gamma_j}{\gamma_{j+1}}\right)^{R_j} \left(1 - \frac{\gamma_j}{\gamma_{j+1}}\right)^{R_{j+1} - R_j}
\end{aligned}
\tag{2}
$$

Taking the logarithm, we have

$$
\begin{aligned}
-\ln p(\mathbf{y}|\mathbf{x},\boldsymbol{\beta}) &= \sum_{j=1}^{k} \left[R_j \ln\left(\frac{1+\exp\left(\theta_{j-1} - \boldsymbol{\beta^T}\mathbf{x}\right)}{1+\exp\left(\theta_j - \boldsymbol{\beta^T}\mathbf{x}\right)}\right) - (R_{j+1} - R_j)\left(\ln\left(\frac{e^{\theta_j} - e^{\theta_{j-1}}}{1+\exp\left(\theta_j - \boldsymbol{\beta^T}\mathbf{x}\right)}\right) - \boldsymbol{\beta^T}\mathbf{x}\right)\right] \\
&= \sum_{j=1}^{k} \left[R_j \ln\left(\frac{1+\boldsymbol{\kappa}_{j-1}}{1+\boldsymbol{\kappa}_j}\right) - (R_{j+1} - R_j) \ln\left(\frac{\boldsymbol{\kappa}_j - \boldsymbol{\kappa}_{j-1}}{1+\boldsymbol{\kappa}_j}\right)\right]
\end{aligned}
\tag{3}
$$

Differentiating and equating to 0 yields

This does not have a closed form solution. We can use gradient descent to find the optimal weights $\boldsymbol{\beta}$ or intervals $\boldsymbol{\theta} = \begin{pmatrix}\theta_1 & \dots & \theta_{k-1}\end{pmatrix}^T$.

3. Code

Refer to `wine.ipynb`.