gren: Group-regularized logistic elastic net regression

Magnus Münch[1,2]
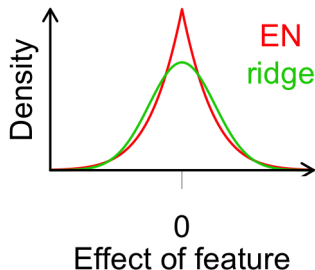
[1]Dep. of Epidemiology & Data Science, Amsterdam University Medical Centers

Our group: www.bigstatistics.nl

Model

$$y_i | \boldsymbol{\beta} \sim \text{Binom}(m_i, \text{expit}(\boldsymbol{x}_i^{\mathsf{T}} \boldsymbol{\beta}))$$

$$\beta_j \overset{ind}{\sim} \text{EN}(\alpha, \lambda, \lambda'_{g(j)})$$

Density

EN
ridge

0
Effect of feature

Predictions through posterior mode:

$$\hat{\boldsymbol{\beta}} = \text{argmax}_\beta \mathcal{L}(\mathbf{y}; \boldsymbol{\beta}) - \alpha\lambda \sum_{g=1}^{G} \sqrt{\lambda'_g} ||\boldsymbol{\beta}_g||_1 - (1-\alpha)\lambda \sum_{g=1}^{G} \lambda'_g ||\boldsymbol{\beta}_g||_2^2$$

Estimation

Estimation in three steps:

1. Cross validation for λ

2. Expectation*-maximisation for $\boldsymbol{\lambda}'$:

$$\boldsymbol{\lambda}'^{(k+1)} = \underset{\boldsymbol{\lambda}'}{\operatorname{argmax}} \, \mathbb{E}_{\boldsymbol{\beta}|\mathbf{y}}[\mathcal{L}(\mathbf{y};\boldsymbol{\beta})\pi_{\boldsymbol{\lambda}'}(\boldsymbol{\beta})|\boldsymbol{\lambda}'^{(k)}]$$

3. Coordinate descent for $\hat{\boldsymbol{\beta}}$ (`glmnet`)

---

*Expectation approximated with variational Bayes

Software

R-package `gren`
- Installation:
  `install.packages("gren")`

- Depends on `Rcpp` and `glmnet`

- `gren`: Estimation of λ, λ', and β

- `predict.gren`: Predictions for new data

- `cv.gren`: Cross-validation of predictive performance

Data

Two datasets:

1. Colorectal cancer with microRNAs
   - ▶ Predict treatment response
   - ▶ Comparison of tumour and normal tissue microRNAs in previous experiment available

2. Cervical cancer disease with microRNAs
   - ▶ Classify high grade cervical lesions
   - ▶ Conservation status of microRNAs across different species available

Course info

Course website:
`https://magnusmunch.github.io/co-data_learning/`