

gren: Group-regularized logistic elastic net regression

Magnus Münch^{1,2}

¹Dep of Epidemiology and Biostatistics, VU University medical center

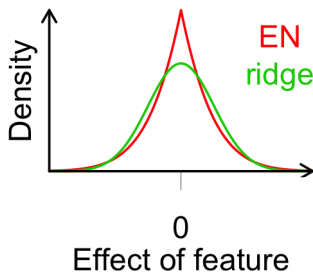
²Mathematical Institute, Leiden University, NL

Our group: www.bigstatistics.nl

Model

$$y_i | \beta \sim \mathcal{B}(m_i, \text{expit}(\mathbf{x}_i^T \beta))$$

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



Predictions through posterior mode:

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

Estimation

Estimation in three steps:

1. Cross validation for λ
2. Expectation*-maximisation for λ' :

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

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

*Expectation approximated with variational Bayes

Software

R-package gren (developmental version)

- Installation:

```
library(devtools)  
install_github("magnusmunch/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/

Acknowledgements:

Cervical cancer data: Saskia Wilting, Barbara Snoek

Colorectal cancer data: Maarten Neerincx, Henk Verheul

gren: Mark van de Wiel, Carel Peeters, Aad van der Vaart