Using gemmR

Jeffrey S. Chrabaszcz and Joe W. Tidwell

September 7, 2014

Motivation

Often in the social sciences, we ask locational questions:

- Do people who train on certain computer tasks have *higher* cognitive ability? [4]

- Are there *more* murders per capita in more honor-focused cultures? [3]

- Does the native language *change* acquisition of definite articles in a second language? [1]

These questions make no mention of the specific distance between relative groups and instead focus on the order of outcome magnitudes. While the statistics applied to these questions are usually variants of the general linear model, there is no reason to impose the assumption of linearity on the reality underlying these tests. One alternative is to apply the general monotone model (`GeMM`) as proposed by **(author?)** [2].

`GeMM` uses a search and scale procedure to: first, find the optimal relative weights for a set of predictors and, second, scale these weights to minimize the order-constrained squared error. This first, computationally-intensive step is accomplished by using a genetic algorithm to maximize Kendall's τ between an observed outcome and a weighted set of predictors. Use of τ in this case assures relative weights that maximally reflect the monotone relationship between the outcome and model predictions. We then regress the original outcome onto the relative-weighted model predictions to compute an intercept and scaling factor that minimizes squared error conditioned on this ordinal constraint.

Fitting a `gemm` model

We implement `GeMM` with the `gemmR` package, which uses `Rcpp` to speed up repeated calculation of Kendall's τ for use in the genetic search process. As `GeMM` serves as a functional replacement for the linear model, a similar syntax is used to fit a `GeMM` model.

Helper functions

The `gemmR` package includes a number of S3 methods and a few novel functions to help extract information from `gemm` objects.

References

[1] Anna Chrabaszcz and Nan Jiang. The role of the native language in the use of the english nongeneric definite article by l2 learners: A cross-linguistic comparison. *Second Language Research*, 30(3):351–379, 2014.

[2] Michael R Dougherty and Rick P Thomas. Robust decision making in a nonlinear world. *Psychological review*, 119(2):321, 2012.

[3] Michael R Dougherty, Rick P Thomas, Ryan P Brown, Jeffrey S Chrabaszcz, and Joe W Tidwell. An introduction to the general monotone model with application to two problematic datasets. Sociological Methods.

[4] Joe W Tidwell, Michael R Dougherty, Jeffrey R Chrabaszcz, Rick P Thomas, and Jorge L Mendoza. What counts as evidence for working memory training? problems with correlated gains and dichotomization. *Psychonomic bulletin & review*, 21(3):620–628, 2014.