

Getting More from Generalized Linear Mixed Models in R

Sydney Benson and Christina Knudson, PhD

University of St. Thomas, St. Paul, Minnesota

This research was made possible in part by the UST Center for Applied Mathematics

Linear Modeling Techniques

Linear model assumptions:

- Independent responses
- Normally distributed responses
- Responses have equal variances

Generalized linear models assume independent, but not normally distributed responses. Occurs when:

- Log odds of your favorite sports team winning a game (binomial)
- Log mean number of students per class at your university (Poisson or negative binomial)

Linear mixed models assume normally distributed, but not independent responses. Occurs when:

- Repeated measurements on individuals
- Measurements on siblings, parents, relatives

Generalized Linear Mixed Models

Assumes neither independence, nor normally distributed responses. Uses fixed and random effects to account for the differences caused by a changing variable and individual or group fluctuations, respectively.

Parallel Computing in R

1. Decide the number of cores to use²
2. Make² and register³ cluster
3. Import necessary packages, variables and functions²
4. Split up calculations among cores^{4, 5}
5. Output results from each core
6. Close cluster²



Parallel Computing for glmm

The parallel computing process is completed in two separate steps:

1. Value of the log-likelihood approximation and gradient vector
2. Hessian matrix

Salamander Example

Do salamanders prefer to mate with others from the same location?

- Correlated responses
- Non-normal responses
- Independent, random effects

```
> sal <- glmm(Mate ~ 0 + Cross,
  random = list( ~ 0 + Female, ~ 0 + Male ),
  varcomps.names = c( "F" , "M" ),
  data = salamander, m = 10^4,
  family.glmm = bernoulli.glmm)
```

Fixed Effects:

	Estimate	Std. Error	z value	Pr(> z)
CrossR/R	1.4629	0.2720	5.378	7.53e-08 ***
CrossR/W	0.3781	0.2527	1.496	0.134612
CrossW/R	-1.7398	0.3157	-5.512	3.55e-08 ***
CrossW/W	1.0345	0.2683	3.857	0.000115 ***

Cross	RR	WW	RW	WR
Probability of mating	0.812	0.798	0.584	0.149

Pre-parallelization (1 core): 43.33 minutes

Post-parallelization (3 cores): 36.6 minutes

Selected References

1. Knudson C. (2015). *glmm: Generalized Linear Mixed Models via Monte Carlo Likelihood Approximation*. R package version 1.0.2, URL <http://CRAN.R-project.org/package=glmm>.
2. Ripley B., Tierney L., Urbanek S. (2017). *Package 'parallel'* R package version 3.3.1, URL <http://stat.ethz.ch/R-manual/R-devel/library/parallel/doc/parallel.pdf>.
3. Microsoft Corporation and Steve Weston (2017). *doParallel: Foreach Parallel Adaptor for the 'parallel' Package*. R package version 1.0.11, URL <https://CRAN.R-project.org/package=doParallel>.
4. Microsoft and Steve Weston (2017). *foreach: Provides Foreach Looping Construct for R*. R package version 1.4.4, URL <https://CRAN.R-project.org/package=foreach>.
5. Steve Weston and Hadley Wickham (2014). *itertools: Iterator Tools*. R package version 0.1.3, URL <https://CRAN.R-project.org/package=itertools>.

For a full list of references, visit

https://github.com/bensonsyd/GLMM/blob/master/Univ_St_Thomas/R_CAM_Paper.pdf.