

July 22, 2019

The results below are generated from an R script.

```
// BaselineHMNL.stan
//Adapted from From: https://discourse.mc-stan.org/t/speeding-up-a-hierarchical-multinomial-logit-model/
// Justin Yap's final 2x faster model
// See yegBaselineHMNL.stan
// Code from YAP is changed as noted. Unmodified code is in YapSpeededHMNLReference.stan
// Feeding data. Should work if we provide data as per // https://mc-stan.org/rstan/reference/stan.html
//     When an element is of type list, it is supposed to make it easier to pass data
//     for those declared in Stan code such as "vector[J] y1[I]" and "matrix[J,K] y2[I]".
//     Using the latter as an example, we can use a list for y2 if the list has "I" elements,
//     each of which is an array (matrix) of dimension "J*K". However, it is not possible
//     to pass a list for data declared such as "vector[K] y3[I,J]"; the only way for it is to
//     use an array with dimension "I*J*K".
// J Yapp's original Data section
// from : https://discourse.mc-stan.org/t/speeding-up-a-hierarchical-multinomial-logit-model/1538/5
// data {
//   int<lower=2> C; // Number of alternatives (choices) in each scenario
//   int<lower=1> K; // Number of // variables in design matrix
//   int<lower=1> R; // Number of respondents
//   int<lower=1> S; // Number of scenarios per respondent
//   int<lower=1,upper=C> YB[R, S]; // best choices
//   int<lower=1,upper=C> YW[R, S]; // worst choices
//   matrix[C, K] X[R, S]; // matrix of attributes for each obs
// }

// variable changes Yapp -> this file
// C -> #nRespCat -- # of alternatives -> responseCategories ;
// R -> nSj respondent -- # of respondents- subjects/participants/persons
// K is number of columns in design matrix (same as Yapp )
// S is not used.
// Added
// nT -> total number of observation trials. Should be equal to S x R in Yapp's model
// for fully balanced model (every respondent (participant) gets exactly same set of Scenarios)
// where each observation trial is a single trial for a single participant.
// we have to identify the Subject from an additional data variable
data {
  int<lower=2> nRespCat; // Number of response category alternatives (choices) in each scenario
  int<lower=1> K; // Number of columns in each design matrix list X[t]
  int<lower=1> nSj; // Number of subjects (respondants/ agents/ persons/perceivers)
  int <lower=1> nT; // total number of observation trials (grand trials) //// int<lower=1> S; // Number
  int<lower=1,upper=nRespCat> Y[nT]; // YB[nSj, S]; // best choices
  // int<lower=1,upper=nRespCat> YW[nSj, S]; // worst choices
  matrix[nRespCat, K] X[nT]; // was matrix[nRespCat, K] X[nSj, S]; // matrix of attributes for each obs
  int<lower=1, upper=nSj> SjID[nT]; // Added tmn. Serial identifier for participant on each observation
```

```

}

parameters {
  vector[K] Beta[nSj];
  vector[K - 1] Theta_raw;
  cholesky_factor_corr[K] L_Omega;
  vector<lower=0, upper=pi()/2>[K] L_sigma_unif;
}

transformed parameters {
  vector<lower=0>[K] L_sigma;
  matrix[K, K] L_Sigma;
  vector[nRespCat] XB[nT]; // was vector[nRespCat] XB[nT] XB[nSj, S];
  vector[K] Theta;

  for (k in 1:K) {
    L_sigma[k] = 2.5 * tan(L_sigma_unif[k]);
  }

  L_Sigma = diag_pre_multiply(L_sigma, L_Omega);

  Theta[1] = 0;
  for (k in 1:(K-1)) {
    Theta[k + 1] = Theta_raw[k];
  }

  for (t in 1:nT) { // was for (r in 1:nSj) {
    // was: for (s in 1:S) {
      XB[t] = X[t] * Beta[SjID[t]] ; // was :XB[r,s] = X[r,s] * Beta[r];
      // was: }
    }
  }
}

model {
  //priors
  Theta_raw ~ normal(0, 10);
  L_Omega ~ lkj_corr_cholesky(4);

  //likelihood
  Beta ~ multi_normal_cholesky(Theta, L_Sigma);
  for (t in 1:nT) { // was for (r in 1:nSj) {
    Y[t] ~ categorical_logit(XB[t]);
    // was YB[r,s] ~ categorical_logit(XB[r,s]);
    // was: for (s in 1:S) {
      // was: YW[r,s] ~ categorical_logit(-XB[r,s]);
      // was: }
    }
  }
}

// DONE LIST
// 1. Remove YW (worst choice) references lines 11 and 56
// 2. Change YB references to just Y (best/only choice)

```

```
// Replace [nSj,S] dimensioning of X and Y and XB with nT
// where nT is total number of scenarios for all subjects.
// For balanced complete design nT would equal nSj * X
//
// Provide a respondent index PID[nT] that selects current Beta[r] as Beta[PID[n]]

// provide an PID[nT] vector of indices for respondent.
// XB

## Error: <text>:1:1: unexpected '/'
## 1: /
## ~
```

The R session information (including the OS info, R version and all packages used):

```
sessionInfo()

## R version 3.5.2 (2018-12-20)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Mojave 10.14.5
##
## Matrix products: default
## BLAS: /System/Library/Frameworks/Accelerate.framework/Versions/A/Frameworks/vecLib.framework/Versions/
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_CA.UTF-8/en_CA.UTF-8/en_CA.UTF-8/C/en_CA.UTF-8/en_CA.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] usethis_1.5.0
##
## loaded via a namespace (and not attached):
## [1] Rcpp_1.0.1      highr_0.8       compiler_3.5.2  pillar_1.3.1
## [5] plyr_1.8.4      prettyunits_1.0.2 tools_3.5.2     packrat_0.5.0
## [9] pkgbuild_1.0.3  evaluate_0.13   tibble_2.1.1    gtable_0.3.0
## [13] pkgconfig_2.0.2 rlang_0.3.4     cli_1.1.0       rstudioapi_0.10
## [17] parallel_3.5.2 xfun_0.6        loo_2.1.0       gridExtra_2.3
## [21] stringr_1.4.0   knitr_1.22      dplyr_0.8.0.1   fs_1.2.7
## [25] stats4_3.5.2    grid_3.5.2      tidyselect_0.2.5 glue_1.3.1
## [29] inline_0.3.15   R6_2.4.0        processx_3.3.0  rstan_2.18.2
## [33] ggplot2_3.1.1   callr_3.2.0     purrr_0.3.2     magrittr_1.5
## [37] scales_1.0.0    ps_1.3.0        StanHeaders_2.18.1 matrixStats_0.54.0
## [41] assertthat_0.2.1 colorspace_1.4-1 tinytex_0.12     stringi_1.4.3
## [45] lazyeval_0.2.2  munsell_0.5.0   crayon_1.3.4

Sys.time()

## [1] "2019-07-22 14:06:26 MDT"
```