Small Simulation Study on Kronecker FLASH in All Null Setting

David Gerard

2016-03-29

Abstract

I run some simulations to see how well FLASH does in an all-null setting

Matrix FLASH

First, I generate conditions for matrix flash.

```
library(flashr)
library(tensr)
set.seed(638)
p <- c(10, 10)
n <- length(p)

lower_cov <- 0.25
upper_cov <- 4

cov_half_list <- list()
for (mode_index in 1:n) {
    cov_half_list[[mode_index]] <- sqrt(seq(lower_cov, upper_cov, length = p[mode_index]))
}
```

The data are a matrix with 2 rows and 10, 10 columns. The columns and rows have standard deviations (0.5, 0.8, 1, 1.2, 1.4, 1.5, 1.7, 1.8, 1.9, 2).

```
itermax <- 100
all_null <- rep(NA, length = itermax)
cor_of_sig <- matrix(NA, ncol = n, nrow = itermax)

for(iter_index in 1:itermax) {
    E <- array(rnorm(prod(p)), dim = p)
    Y <- tensr::atrans(E, lapply(cov_half_list, diag))
    tout <- tflash_kron(Y)

    all_null[iter_index] <- all(abs(tout$post_mean[[1]]) < 10 ^ -6)

    for(mode_index in 1:n) {
        cor_of_sig[iter_index, mode_index] <-
            cor(1 / tout$sigma_est[[mode_index]], cov_half_list[[mode_index]] ^ 2)
    }

    ## cat("Prop Correct:", mean(all_null, na.rm = TRUE), "\n")
    ## cat("     Sig Cor:", colMeans(cor_of_sig, na.rm = TRUE), "\n\n")
}
```

The proportion of the iterations where the mean estimate was a matrix of zeros was 0.8. The mean correlations between the true variances and the estimated variances was (0.72, 0.68).

More General Sims

We apply the exact same conditions as in the previous scenario except under different dimensions.

```r
rm(list = ls())
library(flashr)
library(tensr)
library(xtable)
set.seed(783)
plist <- list(c(50, 50),
              c(10, 10, 10),
              c(10, 10, 10, 10))


all_null_list <- list()
cor_list <- list()

for(pindex in 1:length(plist)) {
  p <- plist[[pindex]]
  n <- length(p)

  lower_cov <- 0.25
  upper_cov <- 4

  cov_half_list <- list()
  for (mode_index in 1:n) {
      cov_half_list[[mode_index]] <- sqrt(seq(lower_cov, upper_cov, length = p[mode_index]))
  }

  itermax <- 100
  all_null <- rep(NA, length = itermax)
  cor_of_sig <- matrix(NA, ncol = n, nrow = itermax)

  for(iter_index in 1:itermax) {
      E <- array(rnorm(prod(p)), dim = p)
      Y <- tensr::atrans(E, lapply(cov_half_list, diag))
      tout <- tflash_kron(Y)

      all_null[iter_index] <- all(abs(tout$post_mean[[1]]) < 10 ^ -6)

      for(mode_index in 1:n) {
          cor_of_sig[iter_index, mode_index] <-
              cor(1 / tout$sigma_est[[mode_index]], cov_half_list[[mode_index]] ^ 2)
      }
  }
  cor_list[[pindex]] <- cor_of_sig
  all_null_list[[pindex]] <- all_null
}
mean_vec <- sapply(all_null_list, mean)
```

```r
library(xtable)
cat_num <- function(x) {
  paste0("(", paste(x, collapse = ", "), ")")
}

xncol <- max(sapply(cor_list, ncol))
xnrow <- length(plist)
cor_table <- matrix(NA, ncol = xncol, nrow = xnrow)
for(pindex in 1:length(plist)) {
  vals <- colMeans(cor_list[[pindex]], na.rm = TRUE)
  cor_table[pindex, 1:length(vals)] <- vals
}
colnames(cor_table) <- paste0("Sigma ", 1:xncol)
rownames(cor_table) <- lapply(plist, cat_num)

print(xtable(cor_table, caption = "Average correlation between
             estimated variances and true variances.", label = "tab:cor.var"))
```

% latex table generated in R 3.2.4 by xtable 1.8-2 package % Tue Mar 29 09:16:59 2016

	Sigma 1	Sigma 2	Sigma 3	Sigma 4
(50, 50)	0.92	0.88		
(10, 10, 10)	0.97	0.97	0.97	
(10, 10, 10, 10)	1.00	1.00	1.00	1.00

Table 1: Average correlation between estimated variances and true variances.

The proportion of the iterations where the mean estimate was a tensor of zeros was (0.89, 0.93, 1). The mean correlations between the true variances and the estimated variances can be found in Table 1.