Fine Mapping Benchmark

https://github.com/gaow/mvarbvs/tree/master/dsc

May 9, 2018

Benchmark status

Data-set

Genotype

- A GTEx sample region (*FMO2*) of size $N = 698$, $P = 7492$
- A GUEVADIS sample region of size $N = 343$, $P = 1001$
- Parameters to "trim" P, eg from 7492 also to, say, 1001

Phenotype

- The original GTEx Throid and Lung expression for the GTEx sample region
  - should have around 3 eQTLs
- A simple simulated GUEVADIS expression data from DAP-g paper

Simulation themes

Univariate

- Simple point mass + `rnorm()` simulation, as in DAP-g paper
- Point mass + mixture of normal, as in ASH paper
  - spiky, near-normal, flat-top, skew, big-normal, bimodal

Multivariate

- Column-wise stacking of univariate simulations
- Point mass + mixture of multivariate normal, as in MASH paper
  - All "canonical" prior covariances
  - Have to provide grid

About LD

There are mechansim to

- Plot save LD heatmap for input data
- Put signals to the most “LD-convoluted” blocks
- Ensure signals are from independent LD blocks
- Permute & break LD structure?

Fine-mapping methods

From Stephens Lab

- `varbvs`
- `susie`
- `M&M ASH`

From the field

- `DAP-g`
- `FINEMAP`
- `CAVIAR`

Each with multiple module "flavors" (parameters)

Single-replicate diagnosis

- When avaiable, compare scattered plots of $\tilde{\beta}$ or $\hat{\beta}$ vs β plots
- Show PIP and log10BF
- ...
- Customized diagnostic plots for SSE methods: `susie` and `M&M`.

Still working on unifying output from methods and annotate with eg LD info.

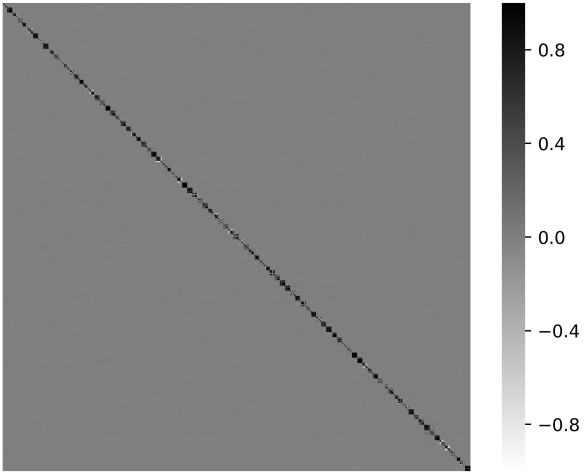
Cross-replicate evaluation

Next to-do:

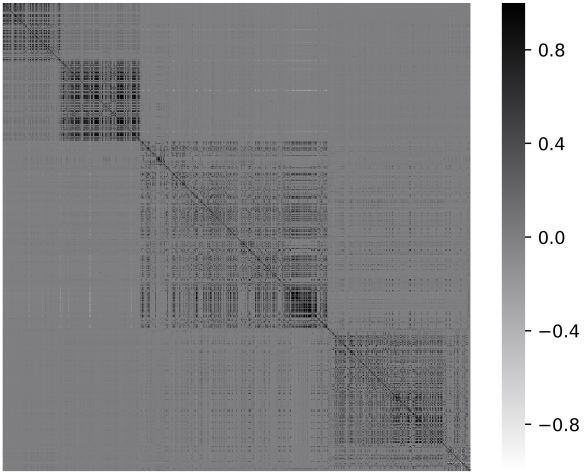
- Power vs false positive: ROC
- Point signal level: PIP / lfsr
- Set signal level: cluster PIP

LD situation

GUEVADIS sample



Trimmed GTEx sample (to the size of GUEVADIS sample)



Simple GUEVADIS simulation

Simple GUEVADIS simulation

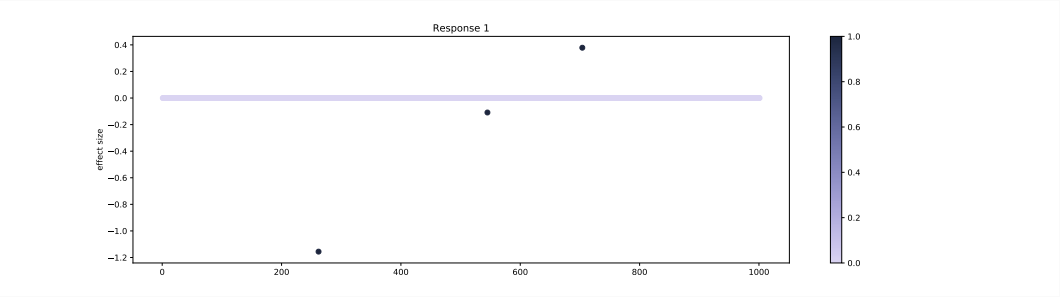


Figure 1: Simple GUEVADIS simulation

varbvs

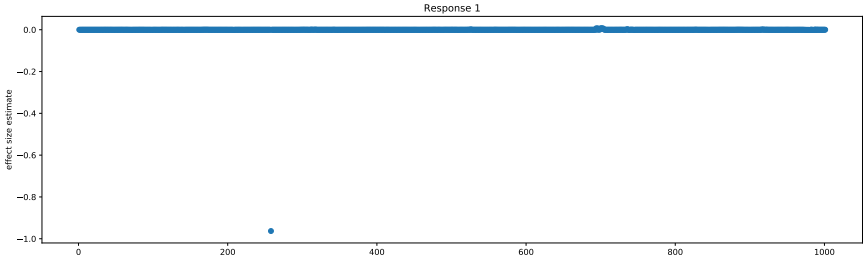


Figure 2: varbvs

susie

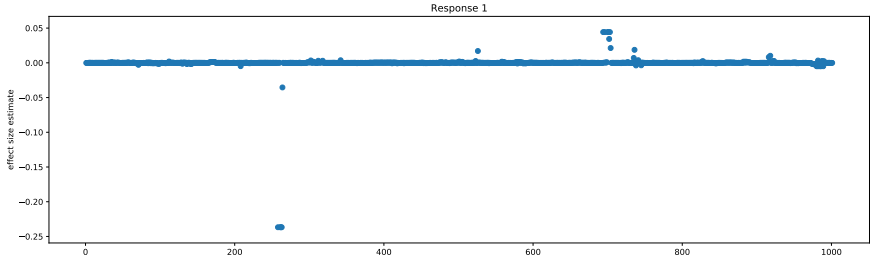
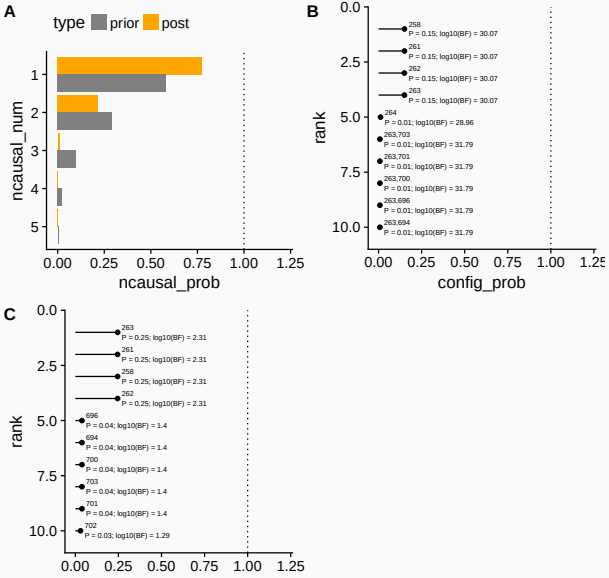
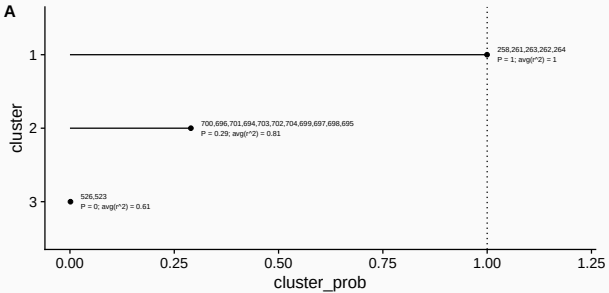


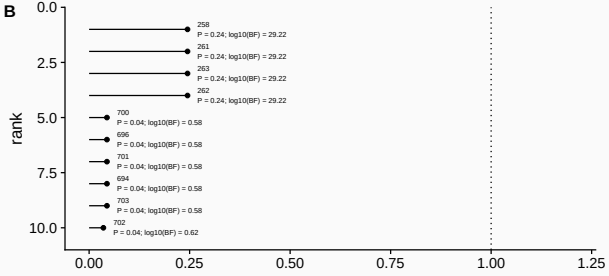
Figure 3: `susie`

FINEMAP

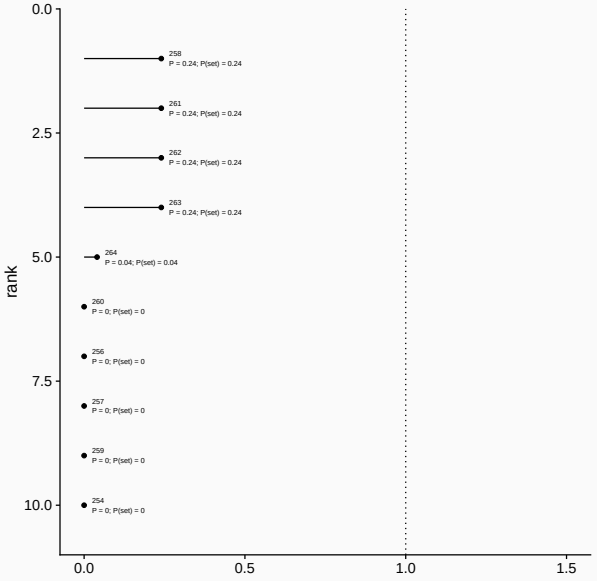

A
type
prior
post
ncausal_num
1
2
3
4
5
0.00
0.25
0.50
0.75
1.00
1.25
ncausal_prob
B
rank
0.0
2.5
5.0
7.5
10.0
258
P = 0.15; log10(BF) = 30.07
261
P = 0.15; log10(BF) = 30.07
262
P = 0.15; log10(BF) = 30.07
263
P = 0.15; log10(BF) = 30.07
264
P = 0.01; log10(BF) = 28.96
263,703
P = 0.01; log10(BF) = 31.79
263,701
P = 0.01; log10(BF) = 31.79
263,700
P = 0.01; log10(BF) = 31.79
263,696
P = 0.01; log10(BF) = 31.79
263,694
P = 0.01; log10(BF) = 31.79
config_prob
C
rank
263
P = 0.25; log10(BF) = 2.31
261
P = 0.25; log10(BF) = 2.31
258
P = 0.25; log10(BF) = 2.31
262
P = 0.25; log10(BF) = 2.31
696
P = 0.04; log10(BF) = 1.4
694
P = 0.04; log10(BF) = 1.4
700
P = 0.04; log10(BF) = 1.4
703
P = 0.04; log10(BF) = 1.4
701
P = 0.04; log10(BF) = 1.4
702
P = 0.03; log10(BF) = 1.29

DAP


A
cluster
1
2
3
258,261,263,262,264
P = 1; avg(r^2) = 1
700,696,701,694,703,702,704,699,697,698,695
P = 0.29; avg(r^2) = 0.81
526,523
P = 0; avg(r^2) = 0.61
0.00
0.25
0.50
0.75
1.00
1.25
cluster_prob


B
rank
0.0
2.5
5.0
7.5
10.0
258
P = 0.24; log10(BF) = 29.22
261
P = 0.24; log10(BF) = 29.22
263
P = 0.24; log10(BF) = 29.22
262
P = 0.24; log10(BF) = 29.22
700
P = 0.04; log10(BF) = 0.58
696
P = 0.04; log10(BF) = 0.58
701
P = 0.04; log10(BF) = 0.58
694
P = 0.04; log10(BF) = 0.58
703
P = 0.04; log10(BF) = 0.58
702
P = 0.04; log10(BF) = 0.62
0.00
0.25
0.50
0.75
1.00
1.25

CAVIAR


0.0
2.5
5.0
7.5
10.0
rank
258
P = 0.24; P(set) = 0.24
261
P = 0.24; P(set) = 0.24
262
P = 0.24; P(set) = 0.24
263
P = 0.24; P(set) = 0.24
264
P = 0.04; P(set) = 0.04
260
P = 0; P(set) = 0
256
P = 0; P(set) = 0
257
P = 0; P(set) = 0
259
P = 0; P(set) = 0
254
P = 0; P(set) = 0
0.0
0.5
1.0
1.5

GTEx with MASH “simple het” covariance

Trimmed GTEx, response 1

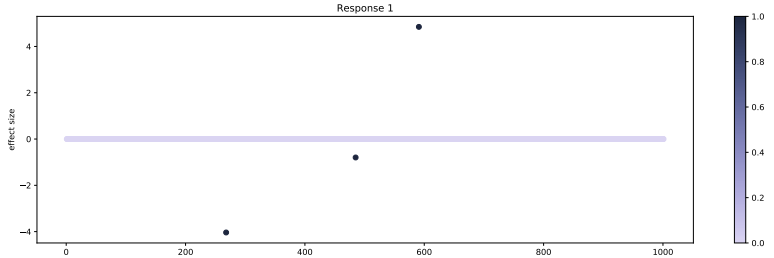


Figure 4: Simulated GTEx tissue 1

Trimmed GTEx, response 2

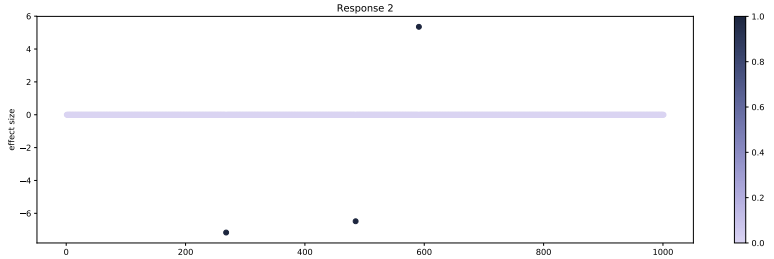


Figure 5: Simulated GTEx tissue 2

varbvs, response 1

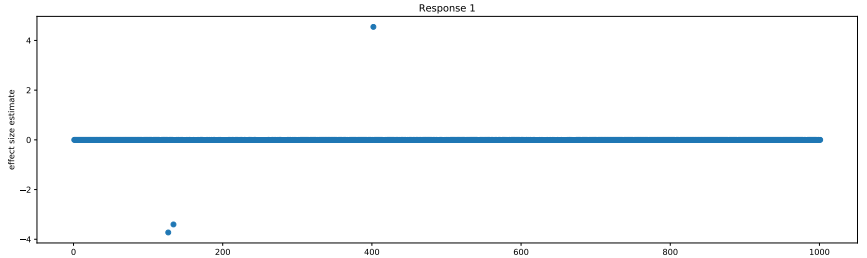


Figure 6: `varbvs`

varbvs, response 2

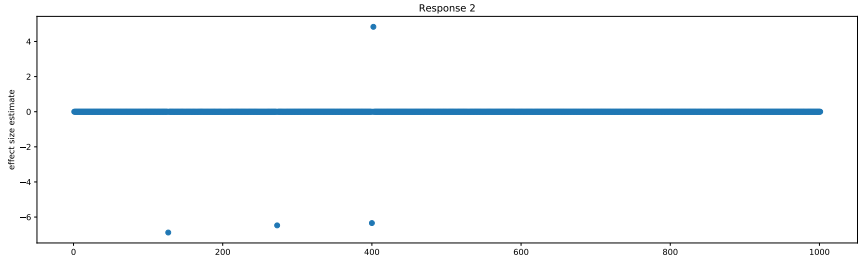


Figure 7: `varbvs`

susie, response 1

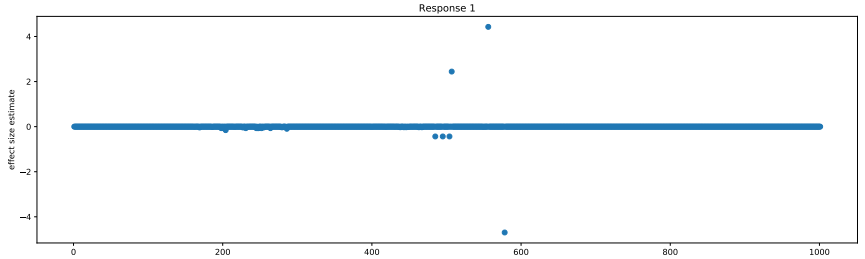


Figure 8: `susie`

susie, response 2

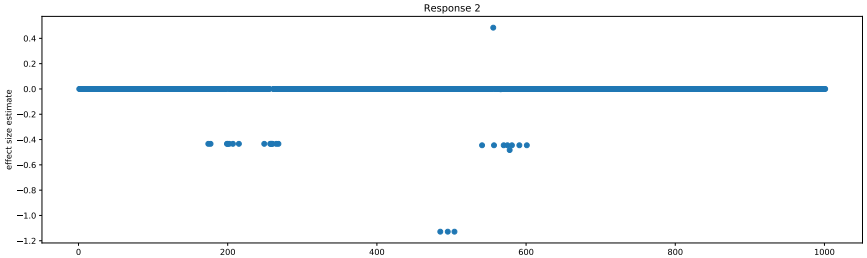


Figure 9: susie

M&M, response 1

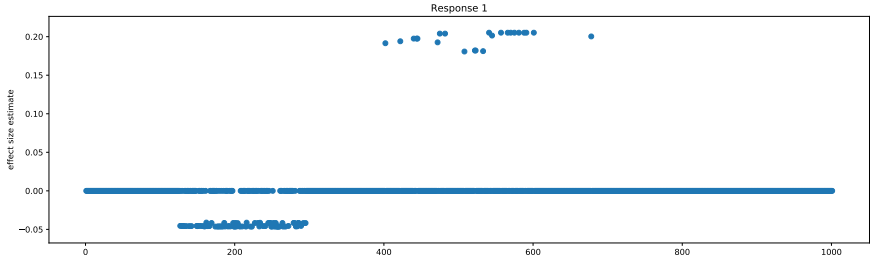


Figure 10: M&M

M&M, response 2

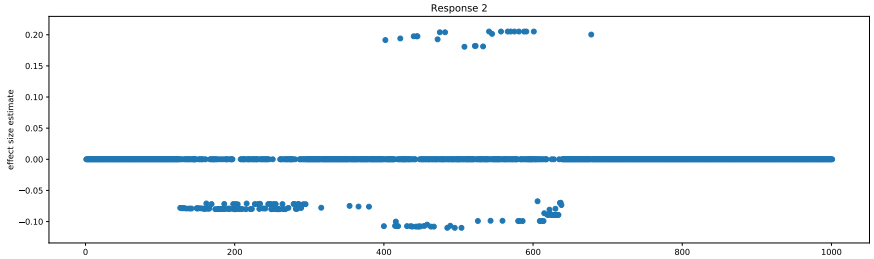


Figure 11: M&M