Estimate Null Correlation in MASH

Yuxin Zou

December 23, 2020

1 Background

The MASH model is

$$\hat{\boldsymbol{b}}_j|\boldsymbol{b}_j, \hat{\boldsymbol{S}}_j \sim N_R(\boldsymbol{b}_j, \hat{\boldsymbol{S}}_j\boldsymbol{V}\hat{\boldsymbol{S}}_j), \tag{1.1}$$

$$\boldsymbol{b}_j|\boldsymbol{\pi} \sim \sum_{k=1}^{K}\sum_{l=1}^{L}\pi_{kl}N_R(\boldsymbol{0}, \omega_l\boldsymbol{U}_k). \tag{1.2}$$

Let P = KL, $\Sigma_p = \omega_l\boldsymbol{U}_k$, we can re-write (1.2) as

$$\boldsymbol{b}_j|\boldsymbol{\pi} \sim \sum_{p=1}^{P}\pi_p N_R(\boldsymbol{0}, \Sigma_p) \tag{1.3}$$

The goal is estimating $\boldsymbol{V}$ and $\boldsymbol{\pi}$ by maximum likelihood.

$$p(\hat{\boldsymbol{B}}) = \prod_{j=1}^{J}p(\hat{\boldsymbol{b}}_j) = \prod_{j=1}^{J}\sum_{p=1}^{P}\pi_p N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j\boldsymbol{V}\hat{\boldsymbol{S}}_j + \Sigma_p) \tag{1.4}$$

Specifically, we estimate them by coordinate ascend. Given $\boldsymbol{V}$, we estimate $\boldsymbol{\pi}$ by solving a concave problem. Given $\boldsymbol{\pi}$, we want to estimate $\boldsymbol{V}$ by maximum likelihood.

2 The Fake Method

The method described below is not the exact correct EM updates. There is a problem in M step and I described it in the last part of this section. The exact one is in Section 3.

Given $\boldsymbol{\pi}$, it is hard to estimate $\boldsymbol{V}$ by maximizing log of (1.4), so we use EM algorithm. Given $\pi_{(t)}$, we augment each $\hat{\boldsymbol{b}}_j$ with corresponding $\boldsymbol{b}_j$ and its mixture membership index $\gamma_j \in [1, \cdots, p]$,

$p(\gamma_j = p) = \boldsymbol{\pi}_{(t)p}$. The complete likelihood is

$$p(\hat{\boldsymbol{B}}, \boldsymbol{B}, \boldsymbol{\gamma}) = \prod_{j=1}^{J} \prod_{p=1}^{P} \left[\boldsymbol{\pi}_{(t)p} N_R(\boldsymbol{b}_j; \boldsymbol{0}, \Sigma_p) N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{b}_j, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j) \right]^{\mathbb{I}(\gamma_j = p)} \tag{2.1}$$

2.1 E step

To get the objective function in E step, we need to compute the posterior distribution for $\boldsymbol{\gamma}$, $\boldsymbol{B}$ given $\hat{\boldsymbol{B}}$,

$$p(\gamma_j = p, \boldsymbol{b}_j | \hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) = \frac{p(\gamma_j = p, \boldsymbol{b}_j, \hat{\boldsymbol{b}}_j | \boldsymbol{V}, \boldsymbol{\pi})}{p(\hat{\boldsymbol{b}}_j | \boldsymbol{V}, \boldsymbol{\pi})} = \frac{p(\hat{\boldsymbol{b}}_j | \boldsymbol{b}_j, \boldsymbol{V}) p(\boldsymbol{b}_j | \gamma_j = p) p(\gamma_j = p)}{p(\hat{\boldsymbol{b}}_j | \boldsymbol{\pi})} \tag{2.2}$$

$$= \frac{\pi_p N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{b}_j, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j) N_R(\boldsymbol{b}_j; \boldsymbol{0}, \Sigma_p)}{\sum_{p'} \pi_{p'} N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j + \Sigma_{p'})} \tag{2.3}$$

$$= \frac{\pi_p N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j + \Sigma_p)}{\sum_{p'} \pi_{p'} N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j + \Sigma_{p'})} \frac{N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{b}_j, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j) N_R(\boldsymbol{b}_j; \boldsymbol{0}, \Sigma_p)}{N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j + \Sigma_p)}. \tag{2.4}$$

Let

$$\tilde{\pi}_{jp} = P(\gamma_j = p | \hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) = \frac{\pi_p N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j + \Sigma_p)}{\sum_{p'} \pi_{p'} N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{0}, \hat{\boldsymbol{S}}_j \boldsymbol{V} \hat{\boldsymbol{S}}_j + \Sigma_{p'})}, \tag{2.5}$$

the posterior for $\boldsymbol{\gamma}$, $\boldsymbol{B}$ given $\hat{\boldsymbol{B}}$ is

$$p(\gamma_j = p, \boldsymbol{b}_j | \hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) = \tilde{\pi}_{jp} P(\boldsymbol{b}_j | \gamma_j = p, \hat{\boldsymbol{b}}_j, \boldsymbol{V}) \tag{2.6}$$

The posterior of $\boldsymbol{b}_j$ given $\gamma_j = p$ is

$$\boldsymbol{b}_j | \hat{\boldsymbol{b}}_j, \boldsymbol{V}, \gamma_j = p \sim N(\tilde{\boldsymbol{\mu}}_{jp}, \tilde{\Sigma}_{jp}) \tag{2.7}$$

$$\tilde{\Sigma}_{jp} = \Sigma_p (I + \hat{\boldsymbol{S}}_j^{-1} \boldsymbol{V}^{-1} \hat{\boldsymbol{S}}_j^{-1} \Sigma_p)^{-1} \tag{2.8}$$

$$\tilde{\boldsymbol{\mu}}_{jp} = \tilde{\Sigma}_{jp} \hat{\boldsymbol{S}}_j^{-1} \boldsymbol{V}^{-1} \hat{\boldsymbol{S}}_j^{-1} \hat{\boldsymbol{b}}_j \tag{2.9}$$

Integrating over γ_j, the posterior of $\boldsymbol{b}_j$ is

$$\boldsymbol{b}_j | \hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi} \sim \sum_{p=1}^{P} \tilde{\pi}_{jp} N(\tilde{\boldsymbol{\mu}}_{jp}, \tilde{\Sigma}_{jp}), \tag{2.10}$$

with the following first and second moments:

$$\tilde{\boldsymbol{\mu}}_j = \mathbb{E}(\boldsymbol{b}_j|\hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) = \sum_{p=1}^{P} \tilde{\pi}_{jp}\tilde{\boldsymbol{\mu}}_{jp} \tag{2.11}$$

$$\mathbb{E}(\boldsymbol{b}_j\boldsymbol{b}_j^T|\hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) = \sum_{p=1}^{P} \tilde{\pi}_{jp}(\tilde{\Sigma}_{jp} + \tilde{\boldsymbol{\mu}}_{jp}\tilde{\boldsymbol{\mu}}_{jp}^T) \tag{2.12}$$

$$Q_j \equiv \mathbb{E}((\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)(\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)^T|\hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) = \hat{\boldsymbol{b}}_j\hat{\boldsymbol{b}}_j^T - \hat{\boldsymbol{b}}_j\tilde{\boldsymbol{\mu}}_j^T - \tilde{\boldsymbol{\mu}}_j\hat{\boldsymbol{b}}_j^T + \mathbb{E}(\boldsymbol{b}_j\boldsymbol{b}_j^T|\hat{\boldsymbol{b}}_j, \boldsymbol{V}, \boldsymbol{\pi}) \tag{2.13}$$

We replace $\boldsymbol{V}$ and $\boldsymbol{\pi}$ with estimates from the previous step, $\boldsymbol{V}_{(t)}, \boldsymbol{\pi}_{(t)}$. Let $q(\cdot) = \mathbb{E}_{\boldsymbol{b}_j|\hat{\boldsymbol{B}}, \boldsymbol{V}_{(t)}, \boldsymbol{\pi}_{(t)}}(\cdot)$.

Taking log of (2.1),

$$\log p(\hat{\boldsymbol{B}}, \boldsymbol{B}, \boldsymbol{\gamma}) = \sum_{j=1}^{J}\sum_{p=1}^{P} \mathbb{I}(\gamma_j = p)\left[\log \boldsymbol{\pi}_{(t)p} + \log N_R(\boldsymbol{b}_j; \boldsymbol{0}, \Sigma_p) + \log N_R(\hat{\boldsymbol{b}}_j; \boldsymbol{b}_j, \hat{\boldsymbol{S}}_j\boldsymbol{V}\hat{\boldsymbol{S}}_j)\right] \tag{2.14}$$

$$= \sum_{j=1}^{J}\sum_{p=1}^{P} \mathbb{I}(\gamma_j = p)\left[\log \boldsymbol{\pi}_{(t)p} - \frac{1}{2}\log|\boldsymbol{V}| - \frac{1}{2}(\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)^T\hat{\boldsymbol{S}}_j^{-1}\boldsymbol{V}^{-1}\hat{\boldsymbol{S}}_j^{-1}(\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)\right] + C \tag{2.15}$$

where C is a constant that does not depend on $\boldsymbol{V}$.

Taking expectations of (2.15), we have

$$
\begin{aligned}
&\mathbb{E}_{\boldsymbol{B},\boldsymbol{\gamma}|\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} \log p(\hat{\boldsymbol{B}}, \boldsymbol{B}, \boldsymbol{\gamma}) && (2.16)\\
&= \mathbb{E}_{\boldsymbol{\gamma}|\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} \left[\mathbb{E}_{\boldsymbol{B}|\boldsymbol{\gamma},\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} \left[\log p(\hat{\boldsymbol{B}}, \boldsymbol{B}, \boldsymbol{\gamma})\right]\right] && (2.17)\\
&= \mathbb{E}_{\boldsymbol{\gamma}|\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} \Bigg[\sum_{j=1}^{J}\sum_{p=1}^{P} \mathbb{I}(\gamma_j = p) \log \boldsymbol{\pi}_{(t)p} - \frac{1}{2} J \log |\boldsymbol{V}| - \\
&\qquad \frac{1}{2}\sum_{j=1}^{J}\sum_{p=1}^{P} \mathbb{I}(\gamma_j = p)\mathbb{E}_{\boldsymbol{b}_j|\gamma_j=p,\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} (\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)^T \hat{\boldsymbol{S}}_j^{-1}\boldsymbol{V}^{-1}\hat{\boldsymbol{S}}_j^{-1}(\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)\Bigg] && (2.18)\\
&= \sum_{j=1}^{J}\sum_{p=1}^{P} \tilde{\pi}_{jp} \log \boldsymbol{\pi}_{(t)p} - \frac{1}{2} J \log |\boldsymbol{V}| - \\
&\qquad \frac{1}{2}\sum_{j=1}^{J}\sum_{p=1}^{P} \tilde{\pi}_{jp}\mathbb{E}_{\boldsymbol{b}_j|\gamma_j=p,\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} (\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)^T \hat{\boldsymbol{S}}_j^{-1}\boldsymbol{V}^{-1}\hat{\boldsymbol{S}}_j^{-1}(\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j) && (2.19)\\
&= \sum_{j=1}^{J}\sum_{p=1}^{P} \tilde{\pi}_{jp} \log \boldsymbol{\pi}_{(t)p} - \frac{1}{2} J \log |\boldsymbol{V}| - \frac{1}{2}\sum_{j=1}^{J} \mathbb{E}_{\boldsymbol{b}_j|\hat{\boldsymbol{B}},\boldsymbol{V}_{(t)},\boldsymbol{\pi}_{(t)}} (\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j)^T \hat{\boldsymbol{S}}_j^{-1}\boldsymbol{V}^{-1}\hat{\boldsymbol{S}}_j^{-1}(\hat{\boldsymbol{b}}_j - \boldsymbol{b}_j) && (2.20)\\
&= \sum_{j=1}^{J}\sum_{p=1}^{P} \tilde{\pi}_{jp} \log \boldsymbol{\pi}_{(t)p} - \frac{1}{2} J \log |\boldsymbol{V}| - \frac{1}{2}\sum_{j=1}^{J} tr(\boldsymbol{V}^{-1}\hat{\boldsymbol{S}}_j^{-1} Q_j \hat{\boldsymbol{S}}_j^{-1}) && (2.21)
\end{aligned}
$$

Note that (2.19) to (2.20) is true because $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y))$.

2.2 Fake M step

We want to maximize (2.21) over $\boldsymbol{V}$. There is a constraint on $\boldsymbol{V}$, the diagonal of $\boldsymbol{V}$ must be 1 since it is a correlation matrix. Let $\boldsymbol{V} = \boldsymbol{DCD}$, $\boldsymbol{C}$ is the covariance matrix, $\boldsymbol{D} = diag(1/\sqrt{\boldsymbol{C}_{jj}})$. We maximize (2.21) over $\boldsymbol{C}$ and ignore the diagonal matrix $\boldsymbol{D}$ which depends on $\boldsymbol{C}$. The estimated $\boldsymbol{V}$ is the corresponding correlation matrix for $\boldsymbol{C}$. Since we maximize (2.21) with respect to $\boldsymbol{C}$, not $\boldsymbol{V}$, the log likelihood may drop. We call this fake M step. Although it is not a true EM algorithm, we use it in mashr. Because the log likelihood increases a lot in the first few iterations, we perform the algorithm with a few iterations.

The objective function with respect to $\boldsymbol{C}$ is

$$
\begin{aligned}
f(\boldsymbol{C}) &= \sum_{j=1}^{J} -\frac{1}{2}\log|\boldsymbol{V}| - \frac{1}{2} tr(\boldsymbol{V}^{-1}\hat{\boldsymbol{S}}_j^{-1} Q_j \hat{\boldsymbol{S}}_j^{-1}) && (2.22)\\
&= \sum_{j=1}^{J} -\frac{1}{2}\log|\boldsymbol{DCD}| - \frac{1}{2} tr(\boldsymbol{C}^{-1}\boldsymbol{D}^{-1}\hat{\boldsymbol{S}}_j^{-1} Q_j \hat{\boldsymbol{S}}_j^{-1}\boldsymbol{D}^{-1}). && (2.23)
\end{aligned}
$$

Taking derivative with respect to $\boldsymbol{C}$,

$$f(\boldsymbol{C})' = \sum_{j=1}^{J} -\frac{1}{2}\boldsymbol{C}^{-1} + \frac{1}{2}\boldsymbol{C}^{-1}\boldsymbol{D}^{-1}\hat{\boldsymbol{S}}_j^{-1}Q_j\hat{\boldsymbol{S}}_j^{-1}\boldsymbol{D}^{-1}\boldsymbol{C}^{-1} = 0 \tag{2.24}$$

$$\boldsymbol{C} = \frac{1}{J}\boldsymbol{D}^{-1}\left[\sum_{j=1}^{J}\hat{\boldsymbol{S}}_j^{-1}Q_j\hat{\boldsymbol{S}}_j^{-1}\right]\boldsymbol{D}^{-1}. \tag{2.25}$$

We can update $\boldsymbol{C}$ and $\boldsymbol{V}$ as

$$\hat{\boldsymbol{C}}_{(t+1)} = \hat{\boldsymbol{D}}_{(t)}^{-1}\frac{1}{J}\left[\sum_{j=1}^{J}\hat{\boldsymbol{S}}_j^{-1}Q_j\hat{\boldsymbol{S}}_j^{-1}\right]\hat{\boldsymbol{D}}_{(t)}^{-1} \tag{2.26}$$

$$\hat{\boldsymbol{D}}_{(t+1)} = diag(1/\sqrt{\hat{\boldsymbol{C}}_{(t+1)jj}}) \tag{2.27}$$

$$\hat{\boldsymbol{V}}_{(t+1)} = \hat{\boldsymbol{D}}_{(t+1)}\hat{\boldsymbol{C}}_{(t+1)}\hat{\boldsymbol{D}}_{(t+1)} \tag{2.28}$$

The resulting $\hat{\boldsymbol{V}}_{(t+1)}$ is equivalent as

$$\hat{\boldsymbol{C}}_{(t+1)} = \frac{1}{J}\left[\sum_{j=1}^{J}\hat{\boldsymbol{S}}_j^{-1}Q_j\hat{\boldsymbol{S}}_j^{-1}\right] \tag{2.29}$$

$$\hat{\boldsymbol{D}}_{(t+1)} = diag(1/\sqrt{\hat{\boldsymbol{C}}_{(t+1)jj}}) \tag{2.30}$$

$$\hat{\boldsymbol{V}}_{(t+1)} = \hat{\boldsymbol{D}}_{(t+1)}\hat{\boldsymbol{C}}_{(t+1)}\hat{\boldsymbol{D}}_{(t+1)} \tag{2.31}$$

We notice that updating $\hat{\boldsymbol{V}}$ requires the posterior of $\boldsymbol{b}_j$, which is obtained by mash model. We perform coordinate ascend. Given $\boldsymbol{V}$, we estimate $\boldsymbol{\pi}$ by solving a concave problem. Given $\boldsymbol{\pi}$, we perform one step of EM.

The algorithm is

Algorithm 1 Estimate Null Correlation Fake

Require: mash data, covariance matrices $\boldsymbol{U}$s, initial value of $\boldsymbol{V}$

```
1:  Given V, estimate π                                   ▷ concave problem
2:  repeat
3:      Given π, estimate V:
4:          E step: compute the posterior distribution of b
5:          Update C ← 2.29
6:          Convert C to V ← 2.31
7:      Given V, estimate π
8:      Compute loglikelihood
9:  until logliklihood does not increase
10: return V
```

3 Exact updates

Given $\boldsymbol{\pi}$, we need to find the $\boldsymbol{V}$ maximize the objective function, (1.4) or (2.21). The objective function is not a concave function for $\boldsymbol{V}$, so it is unclear how to do the optimization with the constrain that the diagonal elements of $\boldsymbol{V}$ are 1.

Pinheiro and Bates (1996) show different parametrizations for variance-covariance matrices that leaves the estimation problem unconstrained. We use the spherical parametrization, which is based on the Cholesky decomposition, $\boldsymbol{V} = \boldsymbol{L}^T\boldsymbol{L}$, $\boldsymbol{L}$ is an upper triangular matrix. Let $\boldsymbol{L}_i$ be the ith column of $\boldsymbol{L}$, and $\boldsymbol{l}_i$ be the spherical coordinates of the first i elements of $\boldsymbol{L}_i$. Since the diagonal elements of $\boldsymbol{V}$ are one, $\boldsymbol{l}_{i,1} = 1$ for all $i = 1, \cdots, R$.

$$\boldsymbol{L}_{i,1} = \cos(\boldsymbol{l}_{i,2}) \tag{3.1}$$
$$\boldsymbol{L}_{i,2} = \sin(\boldsymbol{l}_{i,2})\cos(\boldsymbol{l}_{i,3}) \tag{3.2}$$
$$\cdots \tag{3.3}$$
$$\boldsymbol{L}_{i,i-1} = \sin(\boldsymbol{l}_{i,2})\cdots\cos(\boldsymbol{l}_{i,i}) \tag{3.4}$$
$$\boldsymbol{L}_{i,i} = \sin(\boldsymbol{l}_{i,2})\cdots\sin(\boldsymbol{l}_{i,i}) \tag{3.5}$$

To have unconstrained estimation, we define $\boldsymbol{\theta}$ as follows:

$$\theta_{(i-2)(i-1)/2+(j-1)} = \log\left(\frac{\boldsymbol{l}_{i,j}}{\pi - \boldsymbol{l}_{i,j}}\right) \quad i = 2, \cdots, R,\ j = 2, \cdots, i \tag{3.6}$$

Therefore, we find $\boldsymbol{\theta}$ that maximize the objective function, not $\boldsymbol{V}$, then convert the estimated $\boldsymbol{\theta}$ to $\boldsymbol{V}$.

Algorithm 2 Estimate Null Correlation MLE

Require: mash data, covariance matrices $\boldsymbol{U}$s, initial value of $\boldsymbol{V}$
1: Given $\boldsymbol{V}$, estimate $\boldsymbol{\pi}$ ▷ concave problem
2: **repeat**
3: Parameterize $\boldsymbol{V}$ using $\boldsymbol{\theta}$
4: Given $\boldsymbol{\pi}$, estimate $\boldsymbol{\theta}$ that maximize (1.4) or (2.21) ▷ `optim` or `nlminb`
5: Convert $\boldsymbol{\theta}$ to $\boldsymbol{V}$
6: Given $\boldsymbol{V}$, estimate $\boldsymbol{\pi}$ ▷ concave problem
7: Compute objective function
8: **until** objective function does not increase
9: **return** $\boldsymbol{V}$

4 Speed Considerations

Although the exact updates maximize the objective function, it is very slow to find a length $R(R-1)/2$ vector $\boldsymbol{\theta}$ that achieves the maximum. So we use the Fake method in Section 2 in mashr package.

The fake method also needs some time to converge, because the mash model is fitted at each iteration. There are several things we can do to reduce the running time. First of all, we can use a good initial value for $\boldsymbol{V}$. We set it as the empirical correlation matrix of the z scores for those effects that have (absolute) z score < 2 in all conditions. Moreover, we can set the number of iterations to a small number (i.e. 3). Since there is a large improvement in the log likelihood within the first few iterations, running the algorithm with small number of iterations provides estimates of $\boldsymbol{V}$ that is better then the initial value. Finally, we can estimate $\boldsymbol{V}$ using a random subset of $\hat{\boldsymbol{b}}_j$, $j = 1, \cdots, J$.

The simulation results are in https://zouyuxin.github.io/mash_application/EstimateCorIndex.html.

References

Pinheiro, J. C. and D. M. Bates (1996). Unconstrained parametrizations for variance-covariance matrices. Technical report.