Materials and Methods

Let $\boldsymbol{b_j}$ represents the genetic effect of SNP-gene pair j across $R = 44$ tissues.

We assume the following mixture prior for the R dimensional vector of true effects,

$$\boldsymbol{b_j}|\boldsymbol{\pi}, \mathbf{U}, \boldsymbol{\omega} \sim \sum_{\mathbf{k,l}} \pi_{\mathbf{k,l}}\ N_{\mathbf{R}}(\mathbf{x}; \mathbf{0}, \omega_{\mathbf{l}}\mathbf{U_k}) \tag{1}$$

Where $N_R(.; \mathbf{0}, \omega_l U_k)$ denotes the density of a normal distribution with mean $\mathbf{0}$ and variance $\omega_l U_k$.

Each component of the mixture distribution is characterized by these prior covariance matrices, U_k which capture the pattern of effects across tissues. Critically, this prior distribution is the same for all J - hence the hierarchical incorporation of shared information.

0.1 Covariance Matrices

For a given ω_l, we specify 4 'types' of RxR prior covariance matrices $U_{k,l}$.

1. $U_{k=1,l} = \omega_l\ \mathbf{I}_R$
2. $U_{k=2,l} = \omega_l \mathrm{X}_z$ The (naively) estimated tissue covariance matrix as estimated from the column-centered J $\times R$ matrix of Z statistics, Z_{center}: $\frac{1}{J}\ {Z_{center}}^t\ Z_{center}$
3. $U_{k=3,l} = \omega_l\ \frac{1}{J}\ V_{1\ldots p}\ d^2_{1\ldots p}\ V^t_{1..p}$ is the rank p eigenvector approximation of the tissue covariance matrices, i.e., the sum of the first p eigenvector approximations, where ${}_{1\ldots p}$ represent the eigenvectors of the covariance matrix of tissues and ${}_{1\ldots p}$ are the first p eigenvalues.
4. $U_{k=4:4+Q-1,l} = \frac{1}{J}((\Lambda\mathbf{F})^t\Lambda\mathbf{F})_q$ corresponding to the q_{th} sparse factor representation of the tissue covariance matrix
5. $U_{k=4+Q,l} = \frac{1}{J}\ (\Lambda\mathbf{F})^t\Lambda\mathbf{F}$ is the sparse factor representation of the tissue covariance matrix, estimated using all q factors.
6. $U_{k=5+Q:R+4+Q,l} = \frac{1}{J}\ ([100..]'[100...])$
7. $U_{k=R+5+Q,l} = \frac{1}{J}\ ([111...]'[111...])$
8. [1000...] or [111...] represent configurations such that given membership,$\boldsymbol{b_j}$ arise from the same prior variance.

0.2 Deconvolution

To retrieve a 'denoised' or 'deconvoluted' estimate of the non-single rank dimensional reduction matrices, we then perform deconvolution after initializing the EM algorithm with the matrices specified in (2), (3) and (5). The final results of this iterative procedure preserves the rank of

the initialization matrix, and allows us to use the 'true' effect at each component component $\boldsymbol{b_j}$ as missing data in deconvoluting the prior covariance matrices. In brief, this algorithm works by treating not only the component identity but also the true effect $\boldsymbol{b_j}$ as unobserved data, and maximizing the likelihood over the expectation of the complete data likelihood, considering the values $\boldsymbol{b_j}$ as extra missing data (in addition to the indicator variables q_{ij}) (Bovy et al, 2014). This allows us to write down the 'full data' log likelihood as follows:

$$\phi = \sum_J \sum_K q_{jk} ln \alpha_k N(\hat{\boldsymbol{b_j}} | 0, U_k + V_j)$$
$$\phi = \sum_J \sum_K q_{jk} ln \alpha_k N(\boldsymbol{b_j} | 0, U_k)$$

(2)

Where α_k represents π_k and q_{jk} is the latent identifier variable.

0.3 Likelihood

By maximum likelihood in each tissue separately, we can easily obtain the observed estimates of the standardized genotype effect sizes, $\hat{\boldsymbol{b}}_j$, and their observed squared standard errors recorded on the diagonal of an $R \times R$ matrix noted $\hat{V}_j = \mathbb{V}(\hat{\boldsymbol{b}}_j)$. We assume that the matrix of standard errors of $\hat{\boldsymbol{b}}_j$, V_j as approximated by $\hat{V}_j$ is diagonal and that $\hat{V}_j$ is an accurate point estimate for the standard error and that these standard errors are independent between tissues.

If we now view $\hat{\boldsymbol{b}}_j$ and $\hat{V}_j$ as *observed data*, we can write a new "likelihood" using only the sufficient statistics, $\hat{\boldsymbol{b}}_j$ and $\hat{V}_j$:

$$\hat{\boldsymbol{b}}_j | \boldsymbol{b}_j \sim \mathcal{N}_R(\mathbf{x}; \mathbf{b_j}, \hat{\mathbf{V}}_\mathbf{j}) \tag{3}$$

0.4 Posterior Quantities

We aim to report posterior quantities for a given gene-snp pair **j**. We know that for a single multivariate *Normal* the posterior on $\boldsymbol{b}|U$ is simply:

$$\boldsymbol{b}|\hat{\boldsymbol{b}} \sim N_R(\tilde{\boldsymbol{\mu}}, \tilde{U})$$

where:

- $\tilde{\boldsymbol{\mu}} = \tilde{U}(\hat{V}^{-1}\hat{\boldsymbol{b}})$
- $\tilde{U} = (U^{-1} + \hat{V}^{-1})^{-1}$.

Let us concatenate the list of all KxL combinations of prior covariance matrices U_k and their scaling parameters ω_l into a KxL list and assign this length K for simplicity of notation.

Now each U_k imparts information about both *scale* and *direction* . Furthermore, a mixture-multivariate normal prior and a normal likelihood yields a mixture multivariate posterior, where the final posterior distribution is simply a weighted combination of multivariate normal distributions, where for each gene-snp pair **j** is now characterized by it's posterior mean $\tilde{\boldsymbol{\mu}}_{jk}$ and covariance $\tilde{U}_{jk} = (U_k^{-1} + \hat{V}_j^{-1})^{-1}$.

$$\boldsymbol{b}_j|\hat{\boldsymbol{b}}_j, \hat{V}_j, \hat{\boldsymbol{\pi}} \sim \sum_k^K \tilde{\pi}_{jk} N_R(\mathbf{x}; \tilde{\boldsymbol{\mu}}_{\mathbf{jk}}, \tilde{\mathbf{U}}_{\mathbf{jk}}) \tag{5}$$

Where again, $N_R(.; \mathbf{0}, \omega_l U_k)$ denotes the density of a normal distribution with mean $\tilde{\boldsymbol{\mu}}_k$ and variance $\tilde{U}_k$ and the posterior mixture weight $\tilde{\pi}_k$ is simply

$$\tilde{\pi}_{jk} = \frac{p(\hat{\boldsymbol{b}}_j|\hat{V}_j, z_j = k)\hat{\pi}_k}{\sum_{k=1}^K p(\hat{\boldsymbol{b}}_j|\hat{V}_j, z_j = k)\hat{\pi}_k} \tag{6}$$

Where $z_j = k$ is the latent variable indicator of the component identity and each $\hat{\pi}_k$ represents the Maximum Likelihood Estimate of the prior mixture weights assigned to each component.

0.5 Reported Quantities

For every gene-snp pair 'j', we aim to report the effect size as the posterior mean, defined as:

$$E(\boldsymbol{b}_j|\hat{\boldsymbol{b}}_j, \hat{V}_j, \hat{\boldsymbol{\pi}}) = \sum_k^K \tilde{\pi}_k \tilde{\boldsymbol{\mu}}_k \tag{7}$$

And the local false sign rate, or posterior probability of incorrectly identifying the sign of the effect for a given tissues 'r' as :

$$P(b_{jr}) = 1 - max[\sum_k p(b_{j,r} > 0|\hat{\boldsymbol{b}}_j, \hat{V}_j, z_j = k)\tilde{\pi}_k, \sum_k p(b_{j,r} < 0|\hat{\boldsymbol{b}}_j, \hat{V}_j, z_j = k)\tilde{\pi}_{jk}] \tag{8}$$