FaST-LMM & epistasis: Variance on the coefficients

Ava Khamseh[1]

[1] *MRC Human Genetics Units, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh EH9 3FD, United Kingdom.*

(Dated: 25/04/2021)

This document details the derivation of the fixed effect coefficients variance in a Linear Mixed Model. For consistency, we use the notation outlined in the supplementary material of the original FaST-LMM paper.

I. VARIANCE ON THE ESTIMATED $\hat{\beta}$ COEFFICIENTS

Let ϵ be normally distributed as $\mathcal{N}\left(0;\sigma_g^2\mathbf{K}+\sigma_e^2\mathbf{I}\right)$, so that

$$\mathbf{y}_{n\times 1}=\mathbf{X}_{n\times d}\boldsymbol{\beta}_{d\times 1}+\boldsymbol{\epsilon}_{n\times 1}. \tag{1}$$

Also, for notational convenience, let $\mathcal{Z}$ denote:

$$\mathcal{Z}_{d\times n}=\left(\left(\mathbf{U}^T\mathbf{X}\right)^T\right)_{d\times n}\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}_{n\times n}\left(\mathbf{U}^T\right)_{n\times n}. \tag{2}$$

Then,

$$\hat{\boldsymbol{\beta}}_{d\times 1}=\Big[\left(\mathbf{U}^T\mathbf{X}\right)^T\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}\left(\mathbf{U}^T\mathbf{X}\right)\Big]^{-1}_{d\times d}\Big[\left(\mathbf{U}^T\mathbf{X}\right)^T\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}\left(\mathbf{U}^T\mathbf{y}\right)\Big]_{d\times 1} \tag{3}$$

can be written in a concise form:

$$\begin{aligned}\hat{\boldsymbol{\beta}}=&\left[\mathcal{Z}\mathbf{X}\right]^{-1}\left[\mathcal{Z}\mathbf{y}\right]=\left[\mathcal{Z}\mathbf{X}\right]^{-1}\left[\mathcal{Z}(\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\epsilon})\right]\\=&\boldsymbol{\beta}+\left(\mathcal{Z}\mathbf{X}\right)^{-1}\left(\mathcal{Z}\boldsymbol{\epsilon}\right).\end{aligned} \tag{4}$$

Hence it can be clearly observed that when $\mathcal{Z}$ and $\mathbf{X}$ are not stochastic and using $\mathbb{E}[\epsilon]=0$, we have $\mathbb{E}[\hat{\boldsymbol{\beta}}]=\boldsymbol{\beta}$, as required. In other words, if the ground truth data is generated from the LMM in Eq. 1, then the estimator $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ is unbiased. The variance on each of the coefficients in $\hat{\boldsymbol{\beta}}$ can now be derived as the diagonal elements of the following equation:

$$\begin{aligned}\mathrm{Cov}[\hat{\boldsymbol{\beta}}]=&\mathbb{E}\left[\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)^T\right]=\mathbb{E}\left[\left(\left(\mathcal{Z}\mathbf{X}\right)^{-1}\mathcal{Z}\boldsymbol{\epsilon}\right)\left(\left(\mathcal{Z}\mathbf{X}\right)^{-1}\mathcal{Z}\boldsymbol{\epsilon}\right)^T\right]\\=&\left(\mathcal{Z}\mathbf{X}\right)^{-1}\underbrace{\mathcal{Z}\ \mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T]\ \mathcal{Z}^T}_{\text{expand}}\left(\left(\mathcal{Z}\mathbf{X}\right)^{-1}\right)^T\\=&\sigma_g^2\left(\mathcal{Z}\mathbf{X}\right)^{-1}\left(\mathbf{U}^T\mathbf{X}\right)^T\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}\mathbf{U}^T\mathbf{U}\left(\mathbf{S}+\delta\mathbb{1}\right)\mathbf{U}^T\mathbf{U}\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}\mathbf{U}^T\mathbf{X}\left(\left(\mathcal{Z}\mathbf{X}\right)^{-1}\right)^T\\=&\sigma_g^2\left(\mathcal{Z}\mathbf{X}\right)^{-1}\underbrace{\left(\mathbf{U}^T\mathbf{X}\right)^T\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}\mathbf{U}^T}_{\mathcal{Z}}\mathbf{X}\left(\left(\mathcal{Z}\mathbf{X}\right)^{-1}\right)^T\\=&\sigma_g^2\left(\left(\mathcal{Z}\mathbf{X}\right)^{-1}\right)^T=\sigma_g^2\left(\mathcal{Z}\mathbf{X}\right)^{-1}.\end{aligned} \tag{5}$$

In the last equality, we have used the relation $(A^{-1})^T=(A^T)^{-1}$ and the fact that the matrix $\mathcal{Z}\mathbf{X}$ is symmetric, *i.e.* , $(\mathcal{Z}\mathbf{X})^T=\mathcal{Z}\mathbf{X}$. Therefore, the variance on the estimate $\hat{\boldsymbol{\beta}}_j$ of $\boldsymbol{\beta}_j$ is given by the (j,j)-th diagonal entry of the covariance matrix $\mathrm{Cov}[\hat{\boldsymbol{\beta}}]$, *i.e.* ,

$$\boxed{\mathrm{Var}[\hat{\boldsymbol{\beta}}_j]=\sigma_g^2\Big[\left(\left(\mathbf{U}^T\mathbf{X}\right)^T\left(\mathbf{S}+\delta\mathbb{1}\right)^{-1}\left(\mathbf{U}^T\mathbf{X}\right)\right)^{-1}\Big]_{j,j}} \tag{6}$$

for each of the $j=1,2,\ldots,d$ coefficients.

Remark:
As a sanity check, we can compare Eq. 6 with the variance on coefficients in a simple OLS, where $\mathbf{y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon}\sim\mathcal{N}(0,\sigma^2)$:

$$\mathrm{Cov}[\hat{\boldsymbol{\beta}}]_{\mathrm{OLS}}=\sigma^2(\mathbf{X}^T\mathbf{X})^{-1}\ ,\ \hat{\sigma}^2=\frac{\hat{\boldsymbol{\varepsilon}}^T\hat{\boldsymbol{\varepsilon}}}{n}. \tag{7}$$

Indeed, we observe that the form of Eq. 6 is similar to Eq. 7, apart from the transformation via U. Notice also, that σ_g^2 together with $(\mathbf{S}+\delta\mathbb{1})$ play the role of σ^2 in Eq. 7, *i.e.* , the covariance of the noise term.