Web-based Supplementary Materials for

GENE-TRAIT SIMILARITY REGRESSION FOR

MULTIMARKER-BASED ASSOCIATION ANALYSIS

by

Tzeng JY, Zhang D, Chang SM, Thomas DC, Davidian M

Web Appendix A. Derivation of the distribution of U_b

Equation (4) of the paper expresses the score statistic as $U_b = \frac{1}{2}\left(Y - \widehat{\mu}^0\right)^T V^{-1}\Omega^{-1}S_0\Omega^{-1}V^{-1}\left(Y - \widehat{\mu}^0\right)$. To derive its distribution, we first apply Taylor expansion on $\widehat{\mu}^0$ and get that under the null hypothesis,

$$\widehat{\mu}^0 = \delta(X\widehat{\gamma}) \;\approx\; \mu^0 + DX(\widehat{\gamma} - \gamma),$$

where $D = diag\left\{\partial\delta\left(\eta_i\right)/\partial\eta_i\right\}$ and $\eta_i = X_i^T\gamma$. By the score function $U_\gamma = X^T DV^{-1}\left(Y - \mu^0\right)$ and that $\mathrm{E}(\partial U_\gamma/\partial\gamma) = -X^T DV^{-1}DX$ under H_0,

$$(\widehat{\gamma} - \gamma) \;\approx\; (X^T DV^{-1}DX)^{-1}X^T DV^{-1}(Y - \mu^0).$$

Therefore

$$\begin{aligned} Y - \widehat{\mu}^0 &\approx Y - \left\{\mu^0 + DX\left(X^T DV^{-1}DX\right)^{-1}X^T DV^{-1}\left(Y - \mu^0\right)\right\} \\ &= \left\{I - DX\left(X^T DV^{-1}DX\right)^{-1}X^T DV^{-1}\right\}\left(Y - \mu^0\right). \end{aligned}$$

Multiplying both sides by matrix V^{-1}, we have $V^{-1}\left(Y - \widehat{\mu}\right) \approx Q\left(Y - \mu^0\right)$ with $Q = V^{-1} - V^{-1}DX\left(X^T DV^{-1}DX\right)^{-1}X^T DV^{-1}$, the projection matrix under H_0. Therefore,

$$\begin{aligned} U_b &\approx \frac{1}{2}\left(Y - \mu^0\right)^T Q\Omega^{-1}S_0\Omega^{-1}Q\left(Y - \mu^0\right) \\ &= \frac{1}{2}\widetilde{Y}^T C\widetilde{Y}, \end{aligned}$$

where $\widetilde{Y} = V^{-\frac{1}{2}} (Y - \mu^0)$ and is the standardized Y under H_0, i.e., its ith element is equal to $(Y_i - \mu_i^0) / \sqrt{\mathrm{v}_i^0}$. Matrix $C = V^{\frac{1}{2}} Q \Omega^{-1} S_0 \Omega^{-1} Q V^{\frac{1}{2}} / 2$.

Let $\Lambda = diag(\lambda_i)$ with $\lambda_1 \geq \cdots \geq \lambda_n$ the ordered eigenvalues of matrix C, and let E be an $n \times n$ matrix consisting of vector e_i, the corresponding orthonormal eigenvectors of λ_i. We then have

$$U_b \approx \widetilde{Y}^T C \widetilde{Y} = \widetilde{Y}^T E \Lambda E^T \widetilde{Y} = Z^T \Lambda Z,$$

where $Z = diag\{Z_i\}$ with $Z_i \equiv e_i^T \widetilde{Y}$. Provided that each e_i is not dominated by a few elements, Z_i will be approximately independently standard normal random variables. Therefore under H_0, the distribution of U_b is approximately the same as that of $\sum_{i=1}^n \lambda_i \chi^2_{1,i}$, the weighted chi-squared distribution.

Web Appendix B. Variance of Y in Section 3

$$\text{cov}(Y_i, Y_j \mid X, H) = \text{E}_\beta \{\text{cov}(Y_i, Y_j \mid \beta, X, H) \mid X, H\} + \text{cov}_\beta \{\mu_i(\beta), \mu_j(\beta) \mid X, H\}.$$

The first is zero by conditional independence. Here we tentatively rewrite μ_i as $\mu_i(\beta)$ to denote that μ_i is a function of β. By the Taylor expansion of $\mu_i(\beta)$ around $\text{E}\beta = 0$,

$$\begin{aligned} \text{cov}_\beta(\mu_i, \mu_j \mid X, H) &\approx \text{cov}_\beta \left\{\mu_i(0) + \mu_i'(0)^T \beta,\ \mu_j(0) + \mu_i'(0)^T \beta \mid X, H\right\} \\ &= \mu_i'(0)^T \text{var}(\beta)\, \mu_j'(0) \\ &= \left\{g'\left(\mu_i^0\right) g'\left(\mu_j^0\right)\right\}^{-1} \times \tau H_i^T R_\beta H_j, \end{aligned}$$

as $\mu_i'(\beta) = \partial \mu_i(\beta) / \partial \beta = \{g'(\mu_i)\}^{-1} H_i$.