

**Supplementary Materials for *Computationally efficient, exact
covariate-adjusted multivariate methods for genetic analysis leveraging
summary statistics from large biobanks***

Theorem 1 (Equation 8). *Let $\mathbf{y}_1 = \mathbf{X}\beta_1 + \varepsilon_1$, $\mathbf{y}_2 = \mathbf{X}\beta_2 + \varepsilon_2, \dots$, $\mathbf{y}_m = \mathbf{X}\beta_m + \varepsilon_m$. Then, for $c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \dots + c_m\mathbf{y}_m = \mathbf{X}\beta$,*

$$\hat{\beta} = c_1\hat{\beta}_1 + c_2\hat{\beta}_2 + \dots + c_m\hat{\beta}_m.$$

Proof. Define $\mathbf{y} = c_1\mathbf{y}_1 + c_2\mathbf{y}_2 + \dots + c_m\mathbf{y}_m$. Then,

$$\begin{aligned} c_1\mathbf{X}'\mathbf{y}_1 + c_2\mathbf{X}'\mathbf{y}_2 + \dots + c_m\mathbf{X}'\mathbf{y}_m &= c_1 \begin{bmatrix} \sum_{i=1}^n y_{1i} \\ \sum_{i=1}^n x_{1i}y_{1i} \\ \vdots \\ \sum_{i=1}^n x_{pi}y_{1i} \end{bmatrix} + c_2 \begin{bmatrix} \sum_{i=1}^n y_{2i} \\ \sum_{i=1}^n x_{1i}y_{2i} \\ \vdots \\ \sum_{i=1}^n x_{pi}y_{2i} \end{bmatrix} + \dots + c_m \begin{bmatrix} \sum_{i=1}^n y_{mi} \\ \sum_{i=1}^n x_{1i}y_{mi} \\ \vdots \\ \sum_{i=1}^n x_{pi}y_{mi} \end{bmatrix} \\ &= \begin{bmatrix} c_1 \sum_{i=1}^n y_{1i} + c_2 \sum_{i=1}^n y_{2i} + \dots + c_m \sum_{i=1}^n y_{mi} \\ c_1 \sum_{i=1}^n x_{1i}y_{1i} + c_2 \sum_{i=1}^n x_{1i}y_{2i} + \dots + c_m \sum_{i=1}^n x_{1i}y_{mi} \\ \vdots \\ c_1 \sum_{i=1}^n x_{pi}y_{1i} + c_2 \sum_{i=1}^n x_{pi}y_{2i} + \dots + c_m \sum_{i=1}^n x_{pi}y_{mi} \end{bmatrix} \\ &= \mathbf{X}'\mathbf{y}. \end{aligned}$$

Therefore,

$$\begin{aligned} \beta &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\ &= (\mathbf{X}'\mathbf{X})^{-1}(c_1\mathbf{X}'\mathbf{y}_1 + c_2\mathbf{X}'\mathbf{y}_2 + \dots + c_m\mathbf{X}'\mathbf{y}_m) \\ &= c_1(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_1 + c_2(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_2 + \dots + c_m(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}_m \\ &= c_1\hat{\beta}_1 + c_2\hat{\beta}_2 + \dots + c_m\hat{\beta}_m, \end{aligned}$$

as desired. □