Supplementary Note

Equivalence of M_a definitions

M_a can be defined in 3 equivalent ways. First, it can be defined in terms of fourth moments of the effect size distribution:

$$M_a = \frac{3M}{\kappa_a}, \quad \kappa_a = \frac{3E(\alpha^2\beta^2) - 2E(\beta^4)}{E(\alpha\beta)^2}, \tag{1}$$

where β is the causal effect size distribution (i.e., $Y = X\beta + \epsilon$), and $\alpha = R\beta$.

Second, it can be defined in terms of the *average unit of heritability*. Suppose that $\beta \sim N(\mathbf{0}, \Sigma)$, where Σ is a diagonal matrix with entries $\sigma_1^2, ..., \sigma_M^2$. The average unit of heritability is defined as:

$$E_{h^2}(\alpha^2) = \frac{1}{h^2}\sum_i \sigma_i^2 E(\alpha_i^2 | R, \Sigma), \tag{2}$$

where $h^2 = Tr(\Sigma)$. $E_{h^2}(\alpha^2)$ is proportional to κ_a:

$$\begin{aligned}
3E(\alpha^2\beta^2) =& \frac{3}{M}\sum_i E(\beta_i^2\alpha_i^2|\Sigma) \\
=& \frac{3}{M}\sum_{i,j} r_{ij}^2 E(\beta_i^2\beta_j^2|\Sigma) \\
=& \frac{3}{M}\sum_i [E(\beta_i^4|\Sigma) + \sum_{j\neq i} r_{ij}^2 E(\beta_i^2|\Sigma)E(\beta_j^2|\Sigma)] \\
=& \frac{3}{M}\sum_i [2\sigma_i^4 + \sigma_i^4 + \sum_{j\neq i} r_{ij}^2\sigma_i^2\sigma_j^2] \\
=& 2E(\beta^4) + \frac{3}{M}\sum_i \sigma_i^2 E(\alpha_i^2|\Sigma),
\end{aligned} \tag{3}$$

where we have used the fact that $E(\beta_i^4|\Sigma) = 3\sigma_i^4$. Rearranging,

$$\frac{3E(\alpha^2\beta^2) - 2E(\beta^4)}{E(\alpha\beta)^2} = \frac{\frac{3}{M}\sum_i \sigma_i^2 E(\alpha_i^2|\Sigma)}{E(\alpha\beta)^2} = \frac{3}{E(\alpha\beta)} E_{h^2}(\alpha^2), \tag{4}$$

where $h^2 = ME(\alpha\beta)$, and we have a second definition of M_a:

$$M_a = \frac{h^2}{E_{h^2}(\alpha^2)}. \tag{5}$$

Third, define $S = \Sigma^{1/2}R\Sigma^{1/2}$. Then:

$$\begin{aligned}
Tr(S^2) &= \sum_{i,j} r_{ij}^2\sigma_i^2\sigma_j^2 \\
&= \sum_i \sigma_i^2 \sum_j r_{ij}^2\sigma_j^2 \\
&= h^2 E_{h^2}(\alpha).
\end{aligned} \tag{6}$$

Thus, we obtain another equivalent definition:

$$M_a = \frac{Tr(S)^2}{Tr(S^2)}, \tag{7}$$

where $Tr(S) = h^2$. This definition is slightly more general than definition (5), since it does not require that Σ is a diagonal matrix. (Note that in the diagonal case, $Tr(S) = Tr(\Sigma)$; more generally, these definitions are different, corresponding to the difference between $E(\beta^2)$ and $E(\alpha\beta)$.) We remark that this definition of M_a is symmetric with respect to the dual fixed objects in the model, R and Σ (i.e. $M_a(R,\Sigma) = M_a(\Sigma, R)$).

Derivation of moment condition

We assume that:

$$cov(\alpha_i^2, \beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) \approx r_{ij}^2 cov(\alpha_j^2, \boldsymbol{\beta}_j^2 | \ell_i^{(2)}, \ell_i^{(4)}). \tag{8}$$

We use this approximation as follows. First, we split up $E(\alpha_i^4)$:

$$E(\alpha_i^4 | \ell_i^{(2)}, \ell_i^{(4)}) = E(\alpha_i^2 [\sum_j r_{ij}^2 \beta_j^2 + \sum_{k \neq j} r_{ik} r_{ij} \beta_k \beta_j] | \ell_i^{(2)}, \ell_i^{(4)}).$$

Next, we use the fact that

$$E(\alpha_i^2 [\sum_j r_{ij}^2 \beta_j^2]) = E([\sum_j r_{ij}^2 \beta_j^2]^2)$$

and that

$$E(\alpha_i^2 [\sum_{k \neq j} r_{ik} r_{ij} \beta_k \beta_j]) = E(2 \sum_{k \neq j} r_{ik}^2 r_{ij}^2 \beta_k^2 \beta_j^2)$$

to obtain:

$$\begin{aligned} E(\alpha_i^4 | \ell_i^{(2)}, \ell_i^{(4)}) &= E([\sum_j r_{ij}^2 \beta_j^2]^2 + 2 \sum_{j \neq k} r_{ij}^2 r_{ik}^2 \beta_j^2 \beta_k^2 | \ell_i^{(2)}, \ell_i^{(4)}) \\ &= 3 \sum_j r_{ij}^2 [E(\alpha_i^2 \beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) - \frac{2}{3} r_{ij}^2 E(\boldsymbol{\beta}_j^4 | \ell_i^{(2)}, \ell_i^{(4)})]. \end{aligned} \tag{9}$$

Now, we are ready to use (8) to break down $E(\alpha_i^2 \beta_j^2) = cov(\alpha_i^2, \beta_j^2) + E(\alpha_i^2) E(\beta_j^2)$:

$$\begin{aligned} E(\alpha_i^4 | \ell_i^{(2)}, \ell_i^{(4)}) &= 3 \sum_j r_{ij}^2 [cov(\alpha_i^2, \beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) + E(\alpha_i^2 | \ell_i^{(2)}, \ell_i^{(4)}) E(\beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) - \frac{2}{3} r_{ij}^2 E(\boldsymbol{\beta}_j^4 | \ell_i^{(2)}, \ell_i^{(4)})] \\ &\approx 3 \sum_j r_{ij}^2 [r_{ij}^2 (E(\alpha_j^2 \beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) - E(\alpha_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) E(\beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)})) \\ &\quad + E(\alpha_i^2 | \ell_i^{(2)}, \ell_i^{(4)}) E(\beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) - \frac{2}{3} r_{ij}^2 E(\boldsymbol{\beta}_j^4 | \ell_i^{(2)}, \ell_i^{(4)})] \\ &= 3E(\alpha_i^2 | \ell_i^{(2)}, \ell_i^{(4)})^2 + \sum_j r_{ij}^4 [3E(\alpha_j^2 \beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) - 3E(\alpha_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) E(\beta_j^2 | \ell_i^{(2)}, \ell_i^{(4)}) - 2E(\boldsymbol{\beta}_j^4 | \ell_i^{(2)}, \ell_i^{(4)})]. \end{aligned} \tag{10}$$

Similar to LD score regression, we assume that SNPs in LD with regression SNPs (i.e. SNPs j which are in LD with SNP i) are representative of a larger population of SNPs (e.g. all common SNPs), allowing us to replace $E(\cdot_j | \ell_i^{(2)}, \ell_i^{(4)})$ with $E(\cdot)$:

$$E(\alpha_i^4 | \ell_i^{(2)}, \ell_i^{(4)}) = 3E(\alpha_i^2 | \ell_i^{(2)}, \ell_i^{(4)})^2 + (3E(\alpha^2 \beta^2) - 2E(\beta^4) - 3E(\alpha^2) E(\beta^2)) \ell_i^{(4)}. \tag{11}$$

We restate this equation for a randomly chosen SNP (rather than for a particular SNP i):

$$E(\alpha^4 | \ell^{(2)}, \ell^{(4)}) = 3E(\alpha^2 | \ell^{(2)})^2 + \ell^{(4)} K, \tag{12}$$

where

$$K = 3E(\beta^2)[E_{h^2}(\alpha^2) - E(\alpha^2)]. \tag{13}$$

Polygenic prediction accuracy

If Σ is given, then it is clear what the optimal risk prediction scheme is. Given an estimate $\hat{\boldsymbol{\alpha}}$ of α, the expected phenotypic value of an individual with genotype X is:

$$E(\mathbf{X}\boldsymbol{\beta}|\hat{\boldsymbol{\alpha}},\Sigma,X) = XE(\boldsymbol{\alpha}|\hat{\boldsymbol{\alpha}},\Sigma). \tag{14}$$

The prediction r^2 is:

$$\begin{aligned} r^2 =& E((X\boldsymbol{\beta})(XE(\boldsymbol{\beta}|\hat{\boldsymbol{\alpha}},\Sigma)|\Sigma)) \\ =& E(\boldsymbol{\beta}^T RE(\boldsymbol{\beta}|\hat{\boldsymbol{\alpha}},\Sigma)) \\ =& E(\boldsymbol{\beta}^T E(\boldsymbol{\alpha}|\hat{\boldsymbol{\alpha}},\Sigma)) \end{aligned} \tag{15}$$

If $\hat{\boldsymbol{\alpha}} \sim N(\boldsymbol{\alpha}, \frac{1}{N}R)$, then:

$$E(\boldsymbol{\alpha}|\hat{\boldsymbol{\alpha}},\Sigma) = \frac{\hat{\boldsymbol{\alpha}}E(\hat{\boldsymbol{\alpha}}^2|\Sigma)}{1/N + E(\hat{\boldsymbol{\alpha}}^2|\Sigma)}, \tag{16}$$

and taking an expectation over SNPs,

$$\begin{aligned} r^2 =& E(\alpha\beta\frac{E(\alpha^2|\Sigma)}{1/N + E(\alpha^2|\Sigma)}|\Sigma) \\ =& h^2 E_{h^2}(\frac{\alpha^2}{1/N + \alpha^2}). \end{aligned} \tag{17}$$

When N is large, r^2 converges to h^2; when N is small, r^2 is approximately $Nh^2E_{h^2}(\alpha^2)$.