Additional file 1: Approximation of
\[ \text{diag} \left( Z' A^{-1}_{gg} Z \right) \] and
\[ \text{diag} \left( Z' A^{gn} (A^{nn})^{-1} A^{ng} Z \right) \]

Let \( Z \) be a matrix that contains the SNP genotypes (coded as 0 for one homozygous genotype, 1 for the heterozygous genotype, or 2 for the alternate homozygous genotype) centered by their observed means. The observed mean of the \( i \)-th SNP is equal to twice the allele frequency \( p_i \) (for the allele coded as 1) of that SNP.

Let \( A^{-1} = \begin{bmatrix} A^{nn} & A^{ng} \\ A^{gn} & A^{gg} \end{bmatrix} \) be the inverse of the pedigree relationship matrix partitioned between \( n_n \) ungenotyped animals (denoted as \( n \)) and \( n_g \) genotyped animals (denoted as \( g \)).

Below we derive formula to approximate \( \text{diag} \left( Z' A^{-1}_{gg} Z \right) \) and \( \text{diag} \left( Z' A^{gn} (A^{nn})^{-1} A^{ng} Z \right) \) under some assumptions. Finally, we compare the approximations to the exact values based on the genotype dataset used in this manuscript.

**Approximation of \( \text{diag} \left( Z' A^{-1}_{gg} Z \right) \)**

Let’s assume that there is no inbreeding and that all genotyped animals are unrelated (i.e. \( A^{-1}_{gg} = I \)). If follows that:

\[ \text{diag} \left( Z' A^{-1}_{gg} Z \right) = \text{diag} \left( Z' Z \right) \]

The \( i \)-th diagonal element of \( \text{diag} \left( Z' A^{-1}_{gg} Z \right) \) can therefore be approximated as

\[ z'_i A^{-1}_{gg} z_i = z'_i z_i \approx 2 p_i (1 - p_i) n_g \tag{1} \]

where \( z_i \) is the \( i \)-th column of \( Z \), and \( E (z'_i z_i) = 2 p_i (1 - p_i) n_g \) under the assumption of Hardy-Weinberg equilibrium [33].

To relax the assumption of no inbreeding and no relationships among genotyped animals, let’s assume an average pedigree-based inbreeding coefficient \( \bar{f} \) for the \( n_g \) genotyped animals, yielding an expected value for the diagonal elements of \( A_{gg} \) is equal to \( 1 + \bar{f} \). If we assume random mating and ignore that the rate of inbreeding across generations within the data is likely different from
0, then the expected value for the off-diagonal elements of $A_{gg}$ is equal to $2\bar{f}$, because the inbreeding coefficient is equal to the kinship, and thus half the additive genetic relationship, between the parents [34, 35].

Therefore, we can approximate $A_{gg}$ as:

$$A_{gg} \approx I \left( 1 - \bar{f} \right) + 2\bar{f}2'$$

where $2'$ is a column-vector with values equal to $\sqrt{2}$.

Then, using the Woodbury matrix identity, it follows that:

$$A_{gg}^{-1} \approx \frac{1}{1 - \bar{f}} - \frac{1}{1 - \bar{f}}2\left( \bar{f}^{-1} + 2'12^{-1} \right)2'1 \frac{1}{1 - \bar{f}}$$

$$\approx \frac{1}{1 - \bar{f}} \left( I - 2 \frac{\bar{f}^{-1} + \frac{2n_g}{1 - \bar{f}}}{2'} \right)2'1 \frac{1}{1 - \bar{f}}$$

$$\approx \frac{1}{1 - \bar{f}} \left( I - 2 \frac{\bar{f}}{2n_g} \right)2'1 \frac{1}{1 - \bar{f}}$$

$$\approx \frac{1}{1 - \bar{f}} \left( I - 11' \frac{1}{n_g} \right)$$

(2)

with $\left( \bar{f}^{-1} + \frac{2n_g}{1 - \bar{f}} \right)^{-1} \approx \frac{1 - \bar{f}}{2n_g}$ for large $n_g$.

Using Eq. 2, the $i$-th diagonal element of the product $Z'A_{gg}^{-1}Z$ can be approximated as:

$$z'_i A_{gg}^{-1} z_i \approx z'_i \left( \frac{1}{1 - \bar{f}} \left( I - 11' \frac{1}{n_g} \right) \right) z_i$$

$$\approx \frac{1}{(1 - \bar{f})} \left( z'_i z_i - \frac{1}{n_g} z'_i11'z_i \right)$$

$$\approx \frac{1}{(1 - \bar{f})} (z'_i z_i)$$

$$\approx \frac{2}{(1 - \bar{f})} p_i (1 - p_i) n_g$$

(3)

where $z'_i1 = 0$.

It is worth noting that Eq. 3 is equal to Eq. 1 when $\bar{f} = 0$. In our implementation, we assumed no inbreeding and implemented the following approximation:

$$z'_i A_{gg}^{-1} z_i \approx 2p_i (1 - p_i) n_g$$

**Approximation of $\text{diag} \left( Z' \left( A^{gn} (A^{nn})^{-1} A^{ng} \right) Z \right)$**

The product $Z' \left( A^{gn} (A^{nn})^{-1} A^{ng} \right) Z$ can be written as:

$$Z' A^{gn} (A^{nn})^{-1} A^{ng} Z = Z' A^{gg} Z - Z' A_{gg}^{-1} Z$$
Assuming an average pedigree inbreeding coefficient \( \bar{f} \), no missing parents, no mating among genotyped animals, no parent-offspring relationship between genotyped animals, it follows that the matrix \( A_{gg} \) is a diagonal matrix, and that the \( j \)-th diagonal element of \( A_{gg} \) is equal to, following Henderson’s rules [22]:

\[
\frac{2 + 0.5 \bar{n}_{\text{offspring}}}{1 - \bar{f}} = \frac{0.5 (4 + \bar{n}_{\text{offspring}})}{1 - \bar{f}}
\]

where \( \bar{n}_{\text{offspring}} \) is the average number of offspring of a genotyped animal.

The \( i \)-th diagonal element of the product \( Z^T A_{gg} Z \) can therefore be approximated as:

\[
z_i^T A_{gg} z_i \approx z_i^T (0.5 (4 + \bar{n}_{\text{offspring}}) \frac{1}{1 - \bar{f}}
\]

\[
= 2 p_i (1 - p_i) n_g (\bar{n}_{\text{offspring}} + 4) \frac{1}{1 - \bar{f}}
\]

\[
= \frac{1}{1 - \bar{f}} p_i (1 - p_i) n_g (\bar{n}_{\text{offspring}} + 4)
\]

(4)

Finally, from Eq. (3) and (4), it follows that \( i \)-th diagonal element of the product \( Z^T A_{gn} (A_{nn})^{-1} A_{ng} Z \) can be approximated as:

\[
z_i^T A_{gn} (A_{nn})^{-1} A_{ng} z_i = z_i^T A_{gg} z_i - z_i^T A_{gg}^{-1} A_{ng} z_i
\]

\[
\approx \frac{1}{1 - \bar{f}} p_i (1 - p_i) n_g (\bar{n}_{\text{offspring}} + 4) - \frac{2}{1 - \bar{f}} p_i (1 - p_i) n_g
\]

\[
\approx \frac{1}{1 - \bar{f}} p_i (1 - p_i) (\bar{n}_{\text{offspring}} n_g + 2 n_g)
\]

\[
\approx \frac{1}{1 - \bar{f}} p_i (1 - p_i) (n_{\text{offspring}} + 2 n_g)
\]

(5)

where \( n_{\text{offspring}} \) is the total number of offspring of all genotyped animals.

In our implementation, we assumed no inbreeding and implemented the following approximation:

\[
z_i^T A_{gn} (A_{nn})^{-1} A_{ng} z_i \approx p_i (1 - p_i) (n_{\text{offspring}} + 2 n_g).
\]

Illustration

Figures 1 and 2 illustrate the approximations of \( z_i^T A_{gg}^{-1} z_i \) by \( 2 p_i (1 - p_i) n_g \) and of \( z_i^T A_{gn} (A_{nn})^{-1} A_{ng} z_i \) by \( p_i (1 - p_i) (n_{\text{offspring}} + 2 n_g) \), respectively, based on the 260,591 genotypes including 37,995 segregating SNPs and used in this manuscript. The pedigree included 29,885,286 animals, and the total number of offspring for all genotyped animals was equal to 18,799,162. While it was not considered in this illustration and in our implementation, the average pedigree inbreeding coefficient for all genotyped animals was equal to 0.05.
Figure 1: Calculated and approximated values of \( \text{diag}(Z'A_{gg}^{-1}Z) \)
Figure 2: Calculated and approximated values of $\text{diag}\left(Z'A^n (A^n)^{-1} A^{ng} Z\right)$