File S1

Supporting Methods

1. Equivalence between penalized and Bayesian regressions

We show the equivalence first for the **Ridge Regression** (RR) and then for LASSO. The same steps can be used to derive the Bayesian equivalents of other methods, such as Bridge regression. The solution to the optimization problem of the RR is

$$\left(\hat{\mu},\hat{\boldsymbol{\beta}}\right) \underset{\arg\min}{=} \left\{ \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\}.$$

Multiplying the objective function by -1/2 and switching from minimization to maximization preserves the solution, therefore:

$$\left(\hat{\mu},\hat{\boldsymbol{\beta}}\right) \underset{\arg\max}{=} \left\{ -\frac{1}{2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 - \frac{1}{2}\lambda \sum_{j=1}^{p} \beta_j^2 \right\}.$$

Further, dividing the objective function by any positive constant preserves the solution; therefore for any $\sigma^2 > 0$ we have,

$$\left(\hat{\mu},\hat{\boldsymbol{\beta}}\right) \underset{\arg\max}{=} \left\{ -\frac{1}{2\sigma^2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 - \frac{\lambda}{2\sigma^2} \sum_{j=1}^{p} \beta_j^2 \right\}.$$

Moreover, for any positive value of σ_β^2 such that $\lambda = \sigma^2\sigma_\beta^{-2}$ we have:

$$\left(\hat{\mu},\hat{\boldsymbol{\beta}}\right) \underset{\arg\max}{=} \left\{ -\frac{1}{2\sigma^2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 - \frac{1}{2\sigma_\beta^{-2}} \sum_{j=1}^{p} \beta_j^2 \right\}.$$

Finally, applying any monotonic transformation to the objective function also preserves the solution, therefore:

$$
\begin{aligned}
\left(\hat{\mu},\hat{\boldsymbol{\beta}}\right) &\underset{\arg\max}{=} \left\{ -\frac{1}{2\sigma^2}\sum_i \left(y_i-\mu-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2 -\frac{1}{2\sigma_\beta^{-2}}\sum_{j=1}^{p}\beta_j^2 \right\}. \\
&\underset{\arg\max}{=} \left\{ \exp\left[-\frac{1}{2\sigma^2}\sum_i \left(y_i-\mu-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2 -\frac{1}{2\sigma_\beta^{-2}}\sum_{j=1}^{p}\beta_j^2 \right] \right\} \\
&\underset{\arg\max}{=} \left\{ \exp\left[-\frac{1}{2\sigma^2}\sum_i \left(y_i-\mu-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2 \right] \exp\left[-\frac{1}{2\sigma_\beta^{-2}}\sum_{j=1}^{p}\beta_j^2 \right] \right\} \\
&\underset{\arg\max}{=} \left\{ \prod_{i=1}^{n} \exp\left(-\frac{\left(y_i-\mu-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2}{2\sigma^2} \right) \right\} \left\{ \prod_{j=1}^{p} \exp\left(-\frac{\beta_j^2}{2\sigma_\beta^{-2}} \right) \right\}.
\end{aligned}
$$

The first term in the above expression, is proportional to a Gaussian likelihood for data (y_i) with mean $\mu+\sum_{j=1}^{p}x_{ij}\beta_j$ and residual variance σ^2. And the second term is proportional to a Gaussian prior for marker effects with mean equal to zero and variance σ_β^2. Specifically, the solution to RR optimization problem is equivalent to the posterior mode of the following Bayesian model:

$$
\begin{aligned}
p\left(\mathbf{y},\boldsymbol{\beta},\mu \middle| \sigma^2,\sigma_\beta^2\right) &\propto p\left(\mathbf{y}\middle|\boldsymbol{\beta},\mu,\sigma^2,\sigma_\beta^2\right)p\left(\boldsymbol{\beta}\middle|\sigma_\beta^2\right) \\
&\propto \prod_{\mathrm{i}=1}^{\mathrm{n}} N\left(y_i \middle| \mu+\sum_{j=1}^{p}x_{ij}\beta_j,\sigma^2\right)\prod_{\mathrm{j}=1}^{\mathrm{p}} N\left(\beta_j^2 \middle| 0,\sigma_\beta^2\right).
\end{aligned}
$$

A similar reasoning can be used to show the equivalence for the LASSO and in general for Bridge regression. For the LASSO, we replace the penalty $\sum_{j=1}^{p}\beta_j^2$ with $\sum_{j=1}^{p}\left|\beta_j\right|$; therefore

$$
\begin{aligned}
\left(\hat{\mu}, \hat{\boldsymbol{\beta}}\right) &\underset{\arg\min}{=} \left\{ \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 + \lambda \sum_{j=1}^{p} \left|\beta_j\right| \right\} \\
&\underset{\arg\max}{=} \left\{ -\frac{1}{2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 - \frac{\lambda}{2} \sum_{j=1}^{p} \left|\beta_j\right| \right\} \\
&\underset{\arg\max}{=} \left\{ -\frac{1}{2\sigma^2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 - \frac{\lambda}{2\sigma^2} \sum_{j=1}^{p} \left|\beta_j\right| \right\}, \text{ for any } \sigma^2 > 0 \\
&\underset{\arg\max}{=} \left\{ -\frac{1}{2\sigma^2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 - \tilde{\lambda} \sum_{j=1}^{p} \left|\beta_j\right| \right\}, \text{ for any } \tilde{\lambda} > 0, \text{ such that } \tilde{\lambda} = \frac{\lambda}{2\sigma^2} \\
&\underset{\arg\max}{=} \left\{ \exp\left\{ -\frac{1}{2\sigma^2} \sum_i \left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2 \right\} \exp\left\{ -\tilde{\lambda} \sum_{j=1}^{p} \left|\beta_j\right| \right\} \right\} \\
&\underset{\arg\max}{=} \left\{ \prod_{i=1}^{n} \exp\left\{ -\frac{\left(y_i - \mu - \sum_{j=1}^{p} x_{ij}\beta_j\right)^2}{2\sigma^2} \right\} \prod_{j=1}^{p} \exp\left\{ -\tilde{\lambda} \left|\beta_j\right| \right\} \right\}
\end{aligned}
$$

As with the RR, the first term is proportional to the following Gaussian likelihood. The second term is proportional to the product of *p* IID Double-Exponential, or Laplace, priors densities for marker effects.

2. On the joint density of genetic values of genotyped and un-genotyped individuals

In this section we consider the problem of deriving the joint density of genetic values when some individuals (set 1) were not genotyped and others (set 2) were genotyped. We show that the joint density of the genetic values of these two sets of individuals, denoted as $\mathbf{g}_1$ and $\mathbf{g}_2$, respectively, in the RR-BLUP model is a mixture of multivariate normal densities.

When all individuals are genotyped. Following standard assumptions, the marginal distribution of genomic values in RR-BLUP is as follows:

$$p\begin{pmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{pmatrix| \mathbf{X}_1, \mathbf{X}_2, \sigma_u^2 \Big) = MVN\left(\begin{matrix} \mathbf{0} \\ \mathbf{0} \end{matrix}, \begin{bmatrix} \mathbf{X}_1\mathbf{X}_1' & \mathbf{X}_1\mathbf{X}_2' \\ \mathbf{X}_2\mathbf{X}_1' & \mathbf{X}_2\mathbf{X}_2' \end{bmatrix} \sigma_u^2 \right), \qquad [1]$$

where, $\mathbf{X}_1$ and $\mathbf{X}_2$ are matrices of marker genotypes and σ_u^2 is a variance parameter and MVN denotes a multivariate normal density.

When some individuals are not genotyped. Consider the case where only individuals in set 2 are genotyped. In this case, we need to derive the joint density of genetic values given $\mathbf{X}_2$, pedigree relationships (denoted as P) and σ_u^2 that is, $p(\mathbf{g}_1, \mathbf{g}_2 | \mathbf{X}_2, P, \sigma_u^2)$. To derive this density we first augment the probability model by introducing $\mathbf{X}_1$, and subsequently integrating it out:

$$\begin{aligned} p(\mathbf{g}_1, \mathbf{g}_2 | \mathbf{X}_2, P, \sigma_u^2) &= \int p(\mathbf{g}_1, \mathbf{g}_2, \mathbf{X}_1 | \mathbf{X}_2, P, \sigma_u^2) \partial \mathbf{X}_1 \\ &= \int p(\mathbf{g}_1, \mathbf{g}_2 | \mathbf{X}_1, \mathbf{X}_2, \sigma_u^2) p(\mathbf{X}_1 | \mathbf{X}_2, P) \partial \mathbf{X}_1 \\ &= \int MVN\left(\begin{matrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{matrix} \middle| \begin{matrix} \mathbf{0} \\ \mathbf{0} \end{matrix}, \begin{bmatrix} \mathbf{X}_1\mathbf{X}_1' & \mathbf{X}_1\mathbf{X}_2' \\ \mathbf{X}_2\mathbf{X}_1' & \mathbf{X}_2\mathbf{X}_2' \end{bmatrix} \sigma_u^2 \right) p(\mathbf{X}_1 | \mathbf{X}_2, P) \partial \mathbf{X}_1 \qquad [2] \end{aligned}$$

The first density on the right-hand side is simply the MVN density of expression [1]. The second component of the right-hand side, $p(\mathbf{X}_1 | \mathbf{X}_2, P)$ gives the probability density function of the unknown-genotypes given the observed genotypes and the pedigree. This is the density we would use, for instance, in pedigree-based imputation algorithms. For every realization of $\mathbf{X}_1$ we have a peculiar MVN with a particular co-variance structure (see right-hand side of expression [2]). Therefore, we conclude that the joint density of genetic values is a mixture of scaled-multivariate normal densities.

Existing proposals for joint analysis of genotyped and un-genotyped individuals (e.g., CHRISTENSEN and LUND 2010; AGUILAR *et al.* 2010) assume that the joint density of genetic values of these two groups of individuals is MVN. In light of the above-results, these methods should be considered linear approximation to a non-linear problem.