Overview

This document describes the models that can be fit using various options to the R functions `blasso` and `blasso.vs`. See Hans (2009) and Hans (2010) for more details about the models. The likelihood for all models is the normal linear regression model given by

\[
y \mid \beta, \sigma^2 \sim N(X\beta, \sigma^2 I_n),
\]

where \( \beta \) is a \( p \times 1 \) vector of regression coefficients. Please note that both functions mean center both \( y \) and the columns of \( X \), and both functions standardize the columns of \( X \) so that each predictor has unit sample variance. No intercept is included in the model due to the mean centering.

Function `blasso`

The function `blasso` can only be used when \( p \leq n \). For the \( p > n \) case, see the function `blasso.vs`. Several models that can be fit by `blasso` are described below by identifying the prior distribution and the corresponding call to `blasso` that is used to obtain \( M \) samples from the posterior. In all models below, \(-\infty < \beta_j < \infty, \sigma^2 > 0 \) and \( \tau > 0 \).

Model 1

Prior distribution:

\[
p(\beta \mid \tau) = \left( \frac{\tau}{2} \right)^p \exp(-\tau \|\beta\|_1),
\]

where \( \|\beta\|_1 \) is the \( L_1 \)-norm of \( \beta \). The parameters \( \tau \) and \( \sigma^2 \) are considered known and fixed at values Tau and Sig2. Given a vector of starting values for \( \beta \), `beta.start`, samples from the posterior distribution \( p(\beta \mid y, \sigma^2, \tau) \) are obtained with the call:

\[
blasso(y, X, iters=M, beta=beta.start, sig2=Sig2, tau=Tau, beta.prior="classic", fixsig=TRUE, fixtau=TRUE)
\]

Model 2

Prior distribution:

\[
p(\beta \mid \sigma^2, \tau) = \left( \frac{\tau}{2\sigma} \right)^p \exp(-\tau \sigma^{-1} \|\beta\|_1).
\]
The parameters $\tau$ and $\sigma^2$ are considered known and fixed at values $\text{Tau}$ and $\text{Sig2}$. Given a vector of starting values for $\beta$, \texttt{beta.start}$, samples from the posterior distribution $p(\beta \mid y, \sigma^2, \tau)$ are obtained with the call:

\begin{verbatim}
blasso(y, X, iters=M, beta=rep(0,dim(X)[2]), sig2=Sig2, tau=Tau, beta.prior="scaled", fixsig=TRUE, fixtau=TRUE)
\end{verbatim}

\textbf{Model 3}

Prior distribution:

\begin{align*}
   p(\beta \mid \sigma^2, \tau) &= \left(\frac{\tau}{2\sigma}\right)^p \exp(-\tau\sigma^{-1}||\beta||_1), \\
   p(\sigma^2) &= \frac{b^a}{\Gamma(a)}(\sigma^2)^{-(a+1)}\exp(-b/\sigma^2).
\end{align*}

The parameter $\tau$ is considered known and fixed at value $\text{Tau}$, and the hyperparameters $a$ and $b$ are fixed at values $a$ and $b$. The values $a = b = 0$ result in the improper prior $p(\sigma^2) \propto \sigma^{-2}$. Given a vector of starting values for $\beta$, \texttt{beta.start}$, and a starting value for $\sigma^2$, \texttt{sigma2.start}$, samples from the posterior distribution $p(\beta, \sigma^2 \mid y, \tau)$ are obtained with the call:

\begin{verbatim}
blasso(y, X, iters=M, beta=beta.start, sig2=sig2.start, tau=Tau, beta.prior="scaled", sig2prior=c(a,b), fixtau=TRUE)
\end{verbatim}

If prior (1) is desired in place of (2), replace \texttt{beta.prior="scaled"} with \texttt{beta.prior="classic"}.

\textbf{Model 4}

Prior distribution:

\begin{align*}
   p(\beta \mid \sigma^2, \tau) &= \left(\frac{\tau}{2\sigma}\right)^p \exp(-\tau\sigma^{-1}||\beta||_1), \\
   p(\sigma^2) &= \frac{b^a}{\Gamma(a)}(\sigma^2)^{-(a+1)}\exp(-b/\sigma^2), \\
   p(\tau) &= \frac{s^r}{\Gamma(r)}\tau^{r-1}\exp(-s\tau).
\end{align*}

The parameter $\sigma^2$ is considered known and fixed at value $\text{Sig2}$, and the hyperparameters $r$ and $s$ are fixed at values $r$ and $s$. Given a vector of starting values for $\beta$, \texttt{beta.start}$, and a starting value for $\tau$, \texttt{tau.start}$, samples from the posterior distribution $p(\beta, \tau \mid y, \sigma^2)$ are obtained with the call:

\begin{verbatim}
blasso(y, X, iters=M, beta=beta.start, sig2=Sig2, tau=tau.start, beta.prior="scaled", fixsig=TRUE, tauprior=c(r,s))
\end{verbatim}

If prior (1) is desired in place of (3), replace \texttt{beta.prior="scaled"} with \texttt{beta.prior="classic"}.

\textbf{Model 5}

Prior distribution:

\begin{align*}
   p(\beta \mid \sigma^2, \tau) &= \left(\frac{\tau}{2\sigma}\right)^p \exp(-\tau\sigma^{-1}||\beta||_1), \\
   p(\sigma^2) &= \frac{b^a}{\Gamma(a)}(\sigma^2)^{-(a+1)}\exp(-b/\sigma^2), \\
   p(\tau) &= \frac{s^r}{\Gamma(r)}\tau^{r-1}\exp(-s\tau).
\end{align*}
The hyperparameters $a$, $b$, $r$ and $s$ are fixed at values $a$, $b$, $r$ and $s$. Given a vector of starting values for $\beta$, `beta.start`, a starting value for $\sigma^2$, `sigma2.start`, and starting value for $\tau$, `tau.start`, samples from the posterior distribution $p(\beta, \sigma^2, \tau \mid y)$ are obtained with the call:

```r
lasso(y, X, iters=M, beta=beta.start, sig2=sig2.start, tau=tau.start, 
beta.prior="scaled", sig2prior=c(a,b), tauprior=c(r,s))
```

If prior (1) is desired in place of (4), replace `beta.prior="scaled"` with `beta.prior="classic"`.

**Function `lasso.vs`**

The function `lasso.vs` implements a variable selection Gibbs sampler for the Bayesian lasso regression model. Several models are described below by identifying the prior distribution and the corresponding call to `lasso.vs` that is used to obtain $M$ samples from the posterior. In all models below, $-\infty < \beta_j < \infty$, $\sigma^2 > 0$, $\tau > 0$ and $0 < \phi < 1$.

**Model 6**

Prior distribution:

$$p(\beta \mid \tau, \phi) = \prod_{j=1}^{p} \left\{ (1 - \phi) \delta_0(\beta_j) + \phi \left( \frac{\tau}{2} \right) \exp\left(-\tau |\beta_j|\right) \right\},$$

where $\delta_0(\beta_j)$ is a unit mass at zero. The parameters $\sigma^2$, $\tau$ and $\phi$ are considered known and fixed at values Sig2, Tau and Phi. Given a vector of starting values for $\beta$, `beta.start`, samples from the posterior distribution $p(\beta \mid y, \sigma^2, \tau, \phi)$ are obtained with the call:

```r
lasso.vs(y, X, iters=M, beta=beta.start, sig2=Sig2, tau=Tau, phi=Phi, 
beta.prior="classic", fixsig=TRUE, fixtau=TRUE, fixphi=TRUE)
```

**Model 7**

Prior distribution:

$$p(\beta \mid \sigma^2, \tau, \phi) = \prod_{j=1}^{p} \left\{ (1 - \phi) \delta_0(\beta_j) + \phi \left( \frac{\tau}{2\sigma} \right) \exp\left(-\tau \sigma^{-1} |\beta_j|\right) \right\},$$

The parameters $\sigma^2$, $\tau$ and $\phi$ are considered known and fixed at values Sig2, Tau and Phi. Given a vector of starting values for $\beta$, `beta.start`, samples from the posterior distribution $p(\beta \mid y, \sigma^2, \tau, \phi)$ are obtained with the call:

```r
lasso.vs(y, X, iters=M, beta=beta.start, sig2=Sig2, tau=Tau, phi=Phi, 
beta.prior="scaled", fixsig=TRUE, fixtau=TRUE, fixphi=TRUE)
```
Model 8

Prior distribution:

\[ p(\beta \mid \sigma^2, \tau, \phi) = \prod_{j=1}^{p} \left\{ (1 - \phi)\delta_0(\beta_j) + \phi \left( \frac{\tau}{2\sigma} \right) \exp(-\tau\sigma^{-1}|\beta_j|) \right\} \]  

\[ p(\sigma^2) = \frac{b^a}{\Gamma(a)} \sigma^{2a} \exp(-b/\sigma^2) \]  

\[ p(\tau) = \frac{s^r}{\Gamma(r)} \tau^{r-1} \exp(-s\tau) \]  

\[ p(\phi) = \frac{\Gamma(g+h)}{\Gamma(g)\Gamma(h)} \phi^{g-1}(1-\phi)^{h-1}, \]

The hyperparameters \( a, b, r, s, g \) and \( h \) are fixed at values \( a, b, r, s, g \) and \( h \). Given a vector of starting values for \( \beta \), \texttt{beta.start}, a starting value for \( \sigma^2 \), \texttt{sig2.start}, a starting value for \( \tau \), \texttt{tau.start}, and a starting value for \( \phi \), \texttt{phi.start}, samples from the posterior distribution \( p(\beta, \sigma^2, \tau, \phi \mid y) \) are obtained with the call:

\[
\texttt{lasso.vs(y, X, iters=M, beta=beta.start, sig2=sig2.start, tau=tau.start,}
\phi=\texttt{phi.start, beta.prior="scaled", sig2prior=c(a,b), tauprior=c(r,s),}
\texttt{phiprior=c(g,h))}
\]

If prior (5) is desired in place of (6), replace \texttt{"beta.prior=\\texttt{\textquotedbl}scaled\texttt{\textquotedbl} with \texttt{"beta.prior="classic".}

Other Models

Any combination of the parameters \( \sigma^2 \), \( \tau \) and \( \phi \) can be fixed at specific values by setting \texttt{fixphi=TRUE} (or \texttt{fixsig} or \texttt{fixtau}) and removing the corresponding argument regarding the prior.

Acknowledgements

The research and software development were supported in part by U.S. National Science Foundation grant DMS-0706948. Any opinions, findings and conclusions, or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the NSF.

References
