Brief Technical Report to Accompany the `R` Package `blasso`

Bayesian Lasso Regression

Chris Hans
hans@stat.osu.edu
http://www.stat.osu.edu/~hans/

December 2009

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

$$y \mid \beta, \sigma^2 \sim \mathrm{N}(X\beta, \sigma^2 I_n),$$

where β 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:

$$p(\beta \mid \tau) = \left(\frac{\tau}{2}\right)^p \exp(-\tau \|\beta\|_1), \tag{1}$$

where $\|\beta\|_1$ is the L_1-norm of β. The parameters τ and σ^2 are considered known and fixed at values `Tau` and `Sig2`. Given a vector of starting values for β, `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 τ and σ^2 are considered known and fixed at values Tau and Sig2. Given a vector of starting values for β, 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=rep(0,dim(X)[2]), sig2=Sig2, tau=Tau,
       beta.prior="scaled", fixsig=TRUE, fixtau=TRUE)
```

Model 3

Prior distribution:

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

The parameter τ is considered known and fixed at value 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 β, beta.start, and a starting value for σ^2, sigma2.start, samples from the posterior distribution $p(\beta, \sigma^2 \mid y, \tau)$ are obtained with the call:

```
blasso(y, X, iters=M, beta=beta.start, sig2=sig2.start, tau=Tau,
       beta.prior="scaled", sig2prior=c(a,b), fixtau=TRUE)
```

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

Model 4

Prior distribution:

$$
\begin{aligned}
p(\beta \mid \sigma^2, \tau) &= \left(\frac{\tau}{2\sigma}\right)^p \exp(-\tau\sigma^{-1}\|\beta\|_1), \qquad (3)\\
p(\tau) &= \frac{s^r}{\Gamma(r)}\tau^{r-1}\exp(-s\tau).
\end{aligned}
$$

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

```
blasso(y, X, iters=M, beta=beta.start, sig2=Sig2, tau=tau.start,
       beta.prior="scaled", fixsig=TRUE, tauprior=c(r,s))
```

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

Model 5

Prior distribution:

$$
\begin{aligned}
p(\beta \mid \sigma^2, \tau) &= \left(\frac{\tau}{2\sigma}\right)^p \exp(-\tau\sigma^{-1}\|\beta\|_1), \qquad (4)\\
p(\sigma^2) &= \frac{b^a}{\Gamma(a)}(\sigma^2)^{-(a+1)}\exp(-b/\sigma^2),\\
p(\tau) &= \frac{s^r}{\Gamma(r)}\tau^{r-1}\exp(-s\tau).
\end{aligned}
$$

The hyperparameters a, b, r and s are fixed at values `a`, `b`, `r` and `s`. Given a vector of starting values for β, `beta.start`, a starting value for σ^2, `sigma2.start`, and starting value for τ, `tau.start`, samples from the posterior distribution $p(\beta, \sigma^2, \tau \mid y)$ are obtained with the call:

```
blasso(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 `blasso.vs`

The function `blasso.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 `blasso.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:

$$p(\beta \mid \tau, \phi) = \prod_{j=1}^{p} \left\{ (1-\phi)\delta_0(\beta_j) + \phi \left(\frac{\tau}{2}\right) \exp(-\tau|\beta_j|) \right\}, \tag{5}$$

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

```
blasso.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:

$$p(\beta \mid \sigma^2, \tau, \phi) = \prod_{j=1}^{p} \left\{ (1-\phi)\delta_0(\beta_j) + \phi \left(\frac{\tau}{2\sigma}\right) \exp(-\tau\sigma^{-1}|\beta_j|) \right\}. \tag{6}$$

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

```
blasso.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:

$$
\begin{aligned}
p(\beta \mid \sigma^2, \tau, \phi) &= \prod_{j=1}^{p} \left\{ (1-\phi)\delta_0(\beta_j) + \phi\left(\frac{\tau}{2\sigma}\right)\exp(-\tau\sigma^{-1}|\beta_j|) \right\} \qquad (7)\\
p(\sigma^2) &= \frac{b^a}{\Gamma(a)}(\sigma^2)^{-(a+1)}\exp(-b/\sigma^2)\\
p(\tau) &= \frac{s^r}{\Gamma(r)}\tau^{r-1}\exp(-s\tau)\\
p(\phi) &= \frac{\Gamma(g+h)}{\Gamma(g)\Gamma(h)}\phi^{g-1}(1-\phi)^{h-1},
\end{aligned}
$$

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 β, `beta.start`, a starting value for σ^2, `sig2.start`, a starting value for τ, `tau.start`, and a starting value for ϕ, `phi.start`, samples from the posterior distribution $p(\beta, \sigma^2, \tau, \phi \mid y)$ are obtained with the call:

```
blasso.vs(y, X, iters=M, beta=beta.start, sig2=sig2.start, tau=tau.start,
       phi=phi.start, beta.prior="scaled", sig2prior=c(a,b), tauprior=c(r,s),
       phiprior=c(g,h))
```

If prior (5) is desired in place of (6), replace `"beta.prior=scaled"` with `beta.prior="classic"`.

Other Models

Any combination of the parameters σ^2, τ and ϕ can be fixed at specific values by setting `fixphi=TRUE` (or `fixsig` or `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

Hans, C. (2009). Bayesian lasso regression. *Biometrika* **96**, 835–845.

Hans, C. (2010). Model uncertainty and variable selection in Bayesian lasso regression. *Statistics and Computing* **20**, 221–229.