Skew-Normal Regression Gibbs Sampler

Carter Allen

10/1/2018

Regression Model

$$Y_i = \mathbf{x}_i\beta + \psi z_i + \sigma\epsilon$$

Where $\mathbf{x}_i = [1, x_{1i}]$, $\beta^T = [\beta_0, \beta_1]$, $Z \sim N_+(0,1)$, and $\epsilon \sim N(0,1)$.

$$\delta = \frac{\alpha}{\sqrt{1+\alpha^2}}$$

$$\psi = \omega\delta$$

$$\sigma^2 = \frac{\omega^2}{1+\alpha^2}$$

$$\Rightarrow Y_i \sim SN(\mathbf{x}_i\beta, \omega^2, \alpha)$$

Priors

$$\beta \sim N_2(\beta_0, T_0)$$

$$1/\sigma^2 = \tau \sim Gamma(a,b)$$

Full Conditionals

$\underline{\beta^*|Y, \mathbf{X}^*, \tau}$: Let $\mathbf{X}^* = [\mathbf{X}|\mathbf{z}]$, and $\beta^* = [\beta_0, \beta_1, \psi]$.

$$\pi(\beta^*|Y, \mathbf{X}^*, \tau) \propto \pi(Y|\beta, \mathbf{X}^*, \tau)\pi(\beta)$$

$$\propto exp\{-\frac{1}{2}(Y - \mathbf{X}^*\beta)^T\tau I_n(Y - \mathbf{X}^*\beta)\}exp\{-\frac{1}{2}(\beta - \beta_0)^T T_0(\beta - \beta_0)\}$$

$$= exp\{-\frac{1}{2}[\tau Y^T Y - 2\tau\beta^T\mathbf{X}^{*T}Y + \tau\beta^T\mathbf{X}^{*T}\mathbf{X}^*\beta + \beta^T T_0\beta - 2\beta^T T_0\beta_0 + \beta_0^T T_0\beta_0]\}$$

$$\propto exp\{-\frac{1}{2}[\beta^T(\tau\mathbf{X}^{*T}\mathbf{X}^* + T_0)\beta - \beta^T(2\tau\mathbf{X}^*Y + 2T_0\beta_0)]\}$$

Let $V = \tau\mathbf{X}^{*T}\mathbf{X}^* + T_0$ and $M = V^{-1}(T_0\beta_0 + \tau\mathbf{X}^{*T}Y)$

$$\Rightarrow \beta^*|Y, \mathbf{X}^*, \tau \sim N_3(M, V)$$

$\underline{\tau|\beta^*, \mathbf{X}^*, Y}$:

$$\pi(\tau|\beta^*, \mathbf{X}^*, Y) \propto \pi(Y|\beta, \mathbf{X}^*, \tau)\pi(\tau)$$

$$\propto \tau^{n/2} exp\{-\frac{1}{2}\tau(Y - \mathbf{X}^*\beta^*)^T(Y - \mathbf{X}^*\beta^*)\}\tau^{a-1}exp\{-b\tau\}$$

$$= \tau^{n/2+\alpha-1} exp\{-\tau(\frac{1}{2}(Y - \mathbf{X}^*\beta^*)^T(Y - \mathbf{X}^*\beta^*) + b)\}$$

$$\Rightarrow \tau|\beta^*, \mathbf{X}^*, Y \sim Gamma(n/2 + \alpha, \frac{1}{2}(Y - \mathbf{X}^*\beta^*)^T(Y - \mathbf{X}^*\beta^*) + b)$$

$\underline{z_i|y_i, \mathbf{x}_i, \beta, \tau}$:

$$\pi(z_i|y_i, \mathbf{x}_i, \beta, \tau, \psi) \propto \pi(y_i|\mathbf{x}_i, z_i, \beta, \tau, \psi)\pi(z_i)$$

$$\propto exp\{-\frac{\tau}{2}((y_i - \mathbf{x}_i\beta) - \psi z_i)^2\}exp\{-\frac{z_i^2}{2}\}$$

$$= exp\{-\frac{\tau}{2}((y_i - \mathbf{x}_i\beta)^2 - 2\psi z_i(y_i - \mathbf{x}_i\beta) + \psi^2 z_i^2)\}exp\{-\frac{z_i^2}{2}\}$$

$$\propto exp\{-\frac{1}{2}[\tau\psi^2 z_i^2 + z_i^2 - 2\psi z_i\tau(y_i - \mathbf{x}_i\beta)]\}$$

$$= exp\{-\frac{1}{2}[z_i^2(\tau\psi^2 + 1) - 2z_i\psi\tau(y_i - \mathbf{x}_i\beta)]\}$$

$$= exp\{-\frac{1}{2}[a(z_i + d)^2 + e]\}$$

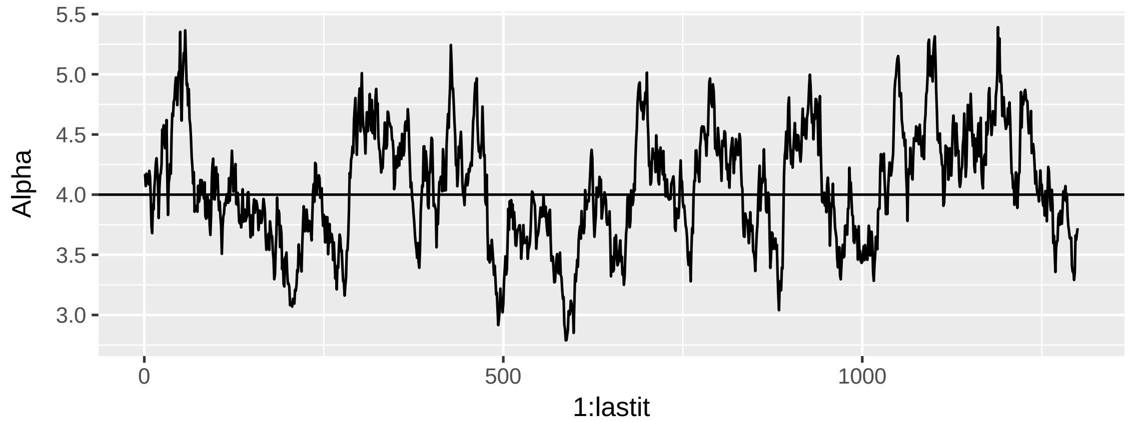
where $a = \tau\psi^2 + 1$, $d = \frac{b}{2a} = \frac{\psi\tau(y_i - \mathbf{x}_i\beta)}{(\tau\psi^2+1)}$, and e does not depend on z_i.

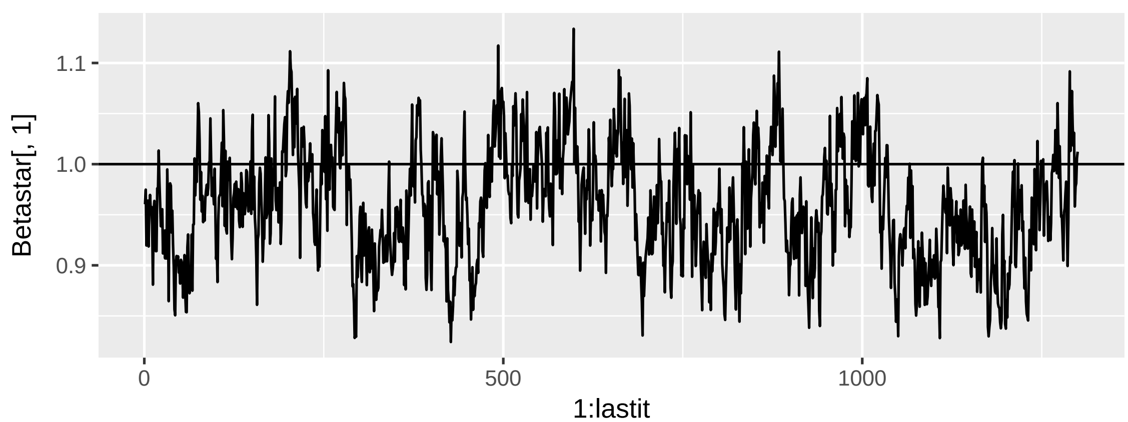
$$\Rightarrow z_i|y_i, \mathbf{x}_i, \beta, \tau \sim N_+(\frac{\psi\tau(y_i - \mathbf{x}_i\beta)}{\tau\psi^2 + 1}, \frac{1}{\tau\psi^2 + 1})$$

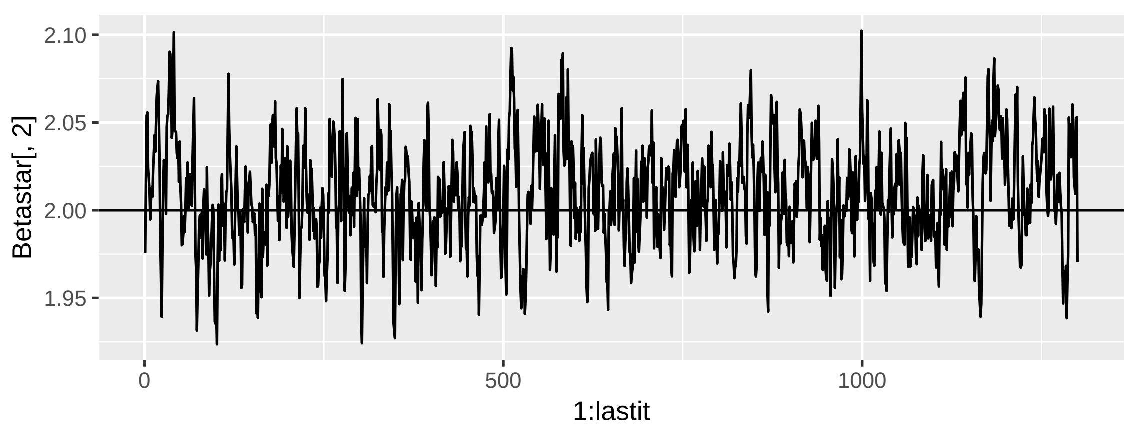
Gibbs Sampler

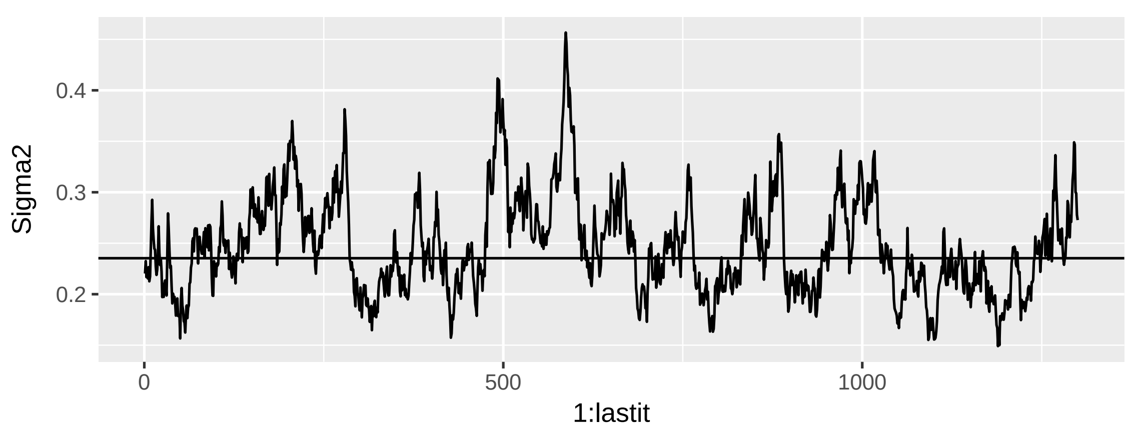
```
# Skew normal bayesian sampler
#
# Carter Allen
# Fall 2018

library(sn) # for rsn
library(truncnorm) # for rtruncnorm
library(mvtnorm) # for rmvtnorm
library(ggplot2)

# Generate Data
# set.seed(1801)
n <- 1000 # number of observations
x <- rnorm(n) # random normal predictors
X <- cbind(1,x) # design matrix w/ intercept
beta <- c(1,2) # true beta vector: beta0, beta1
xi <- X %*% beta # location
alpha <- 4 # skew
omega <- 2 # scale
delta <- alpha/sqrt(1+alpha^2) # coefficient
psi <- omega * delta # coefficient
y <- rsn(n = n, xi = xi, omega = omega, alpha = alpha) # simulated skew outcomes

# Priors
p <- ncol(X) + 1 # dimensions of X plus and column for Z
betastar0 <- rep(0,p) # prior mean for beta with psi
T0 <- diag(0.01,p) # prior precision for beta with psi
a <- b <- 0.001 # gamma hyper parameters for tau

# Initialize
tau <- 1 # initial precision
psi <- 0 # initial psi
betastar <- c(0,0,psi) # initial beta0, beta1, psi
z <- rep(0,n) # empty z storage

# Sample storage
nsim<-1500 # Number of MCMC Iterations
thin<-1 # Thinning interval
burn<-200   # Burnin
lastit<-(nsim-burn)/thin # Last stored value
Betastar<-matrix(0,lastit,p) # storage for each betastar
Sigma2<-rep(0,lastit) # storage for each 1/tau
Alpha <- rep(0,lastit) # storage for each alpha
Omega <- rep(0,lastit) # storage for each omega
Resid<-matrix(0,nsim,n)  # Store resids
Dy<-matrix(0,lastit,512) # Store density values for residual density plot
Qy<-matrix(0,lastit,100) # Store quantiles for QQ plot

# Gibbs
tmp<-proc.time()
for(i in 1:nsim)
```

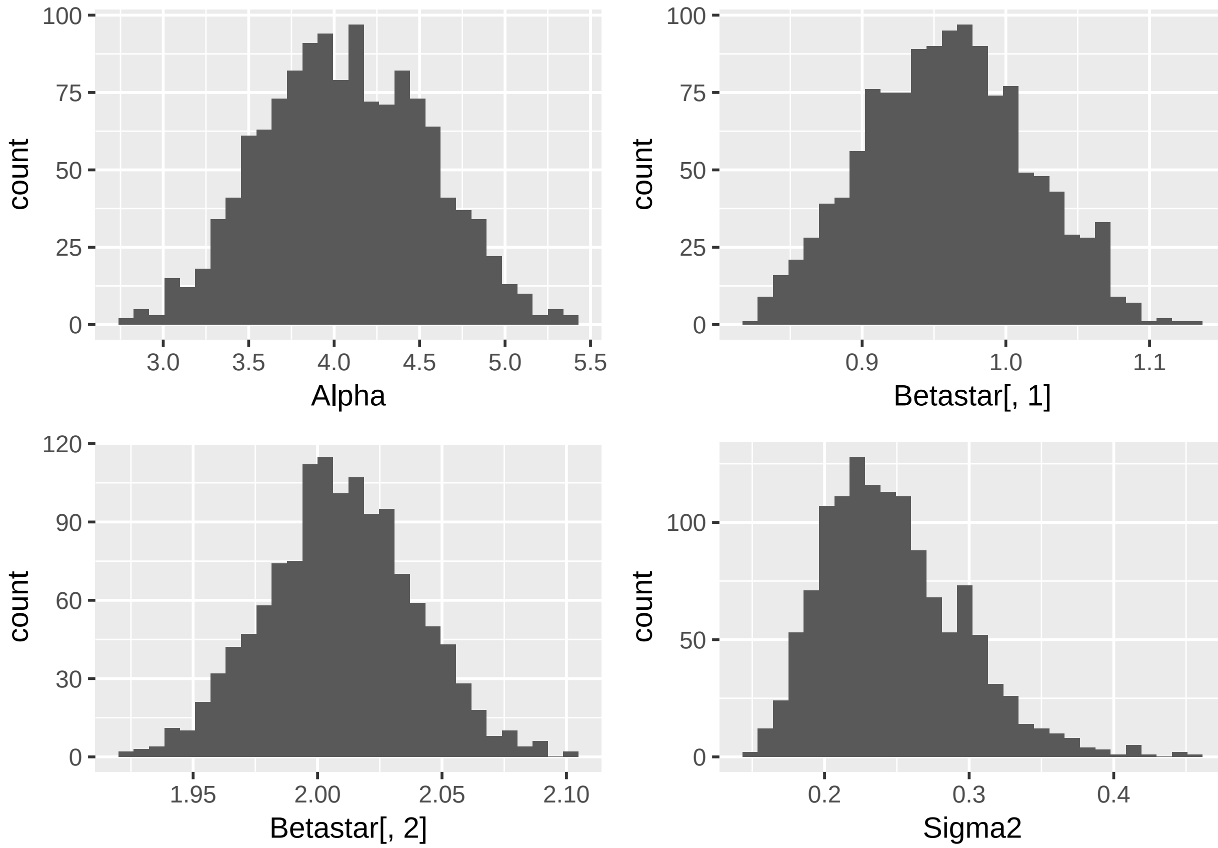
```r
{
  # update z
  for(j in 1:n)
  {
    v_j <- 1/(tau * psi^2 + 1) # full conditional variance of z
    m_j <- tau * v_j * betastar[3] * (y[j]-t(X[j,]) %*% betastar[-3]) # full conditional mean of z
    z[j] <- rtruncnorm(n = 1,a = 0,b = Inf,mean = m_j,sd = sqrt(v_j)) # update jth component of z
  }

  # Form Xstar, augmented design matrix of X and z
  Xstar <- cbind(X,z)

  # Update betastar
  v<-solve(T0+tau*crossprod(Xstar)) # full conditional var of betastar
  m<-v%*%(T0%*%betastar0+tau*crossprod(Xstar,y)) # full conditional mean of betastar
  betastar<-c(rmvnorm(1,m,v)) # draw new betastar

  # extract psi
  psi <- betastar[3]

  # Update tau
  tau <- rgamma(1,a+n/2,b+crossprod(y-Xstar%*%betastar)/2)

  # Store Results
  if (i> burn & i%%thin==0) {
    j<-(i-burn)/thin
    Betastar[j,]<-betastar
    Sigma2[j]<-1/tau
    Alpha[j] <- betastar[3]*sqrt(tau)
    Omega[j] <- sqrt(1/tau + betastar[3]^2)
    Resid[j,]<-resid<-y-Xstar%*%betastar                          # Raw Resid
    Dy[j,]<-density(resid/sd(resid),from=-5,to=5)$y  # Density of Standardized Resids
    Qy[j,]<-quantile(resid/sd(resid),probs=seq(.001,.999,length=100)) # Quantiles for QQ Plot
  }
}
beepr::beep(); run.time<-proc.time()-tmp

p1 <- ggplot() +
  geom_line(aes(x = 1:lastit, y = Alpha)) +
  geom_hline(yintercept = alpha)
p2 <- ggplot() +
  geom_line(aes(x = 1:lastit, y = Betastar[,1])) +
  geom_hline(yintercept = beta[1])
p3 <- ggplot() +
  geom_line(aes(x = 1:lastit, y = Betastar[,2])) +
  geom_hline(yintercept = beta[2])
p4 <- ggplot() +
  geom_line(aes(x = 1:lastit, y = Sigma2)) +
  geom_hline(yintercept = (omega^2)/(1+alpha^2))
```


5.5
5.0
4.5
4.0
3.5
3.0
Alpha
0
500
1000
1:lastit


1.1
1.0
0.9
Betastar[, 1]
0
500
1000
1:lastit


2.10
2.05
2.00
1.95
Betastar[, 2]
0
500
1000
1:lastit


0.4
0.3
0.2
Sigma2
0
500
1000
1:lastit



Below is the summary of the same model fit with Azzalini's `selm` function.

Parameter	Truth	MLE	Gibbs
beta_0	1	0.95	0.96
beta_1	2	2.007	2.009
sigma2	0.235	0.232	0.247
alpha	4	4.134	4.067