

/

2009 / 05 / 05

2009 / 03 / 15

Abstract

In this paper we consider parameter estimation in a linear regression setting with inequality linear constraints on the regression parameters. Most other research on this topic has typically been addressed from a Bayesian perspective. In this paper we apply Bayesian approach with Gibbs sampler to generate samples from the posterior distribution. However, these implementations can often exhibit poor mixing and slow convergence. This paper overcomes these limitations with a new implementation of the Gibbs sampler. In addition, this procedure allows for the number of constraints to exceed the parameter dimension and is able to cope with equality linear constraints.

Gibbs Sampling

Linear

$$Y = X\beta + \varepsilon \quad \dots\dots\dots (1)$$

k X

$$\varepsilon \sim N(0, \sigma^2 I) \quad \beta$$

$$\beta_i \geq 0$$

$$\beta_1 + \beta_2 + \dots + \beta_k = 1$$

β (1)

$$B\beta \leq b, \quad C\beta \leq c$$

c, b C, B

Liew (1976) Judge and Takayama (1966)

Inequality Constrained Least Squares β
 (ICLS) .Dantzig – Cottle (ICLS)

Binding

.(ICLS)

Geweke (1986)

. Lovell and Prescott (1970)

$$(1) \quad \beta$$

$$B\beta \leq b \quad \dots\dots\dots (2)$$

Geweke (1986)

Conventional

Indicator Function

Uninformative Distribution

Importance Sampling

Prior Distribution

Truncation Region

Gelfand et al. (1992)

Monte Carlo Markov Chain (MCMC)

Geman and Geman

θ D

$p(\theta|D)$

$p(\theta|D)$ (1984)

θ_i $\theta = (\theta_1, \theta_2, \dots, \theta_q)'$ θ

$p(\theta_i|D, \theta_j, i \neq j)$

$p(\theta_i|D, \theta_j, i \neq j)$

$\theta^{(t)}$ t $\theta^{(0)}$

Gelfand and t $p(\theta|D)$

Smith (1990)

O'Hagan (1994), Roberts (1996), Gilks and Roberts (1996)

θ

Geweke (1996)

(2)

Mixing

Rodriguez – Yam et al. (2002)

Bayesian Constrained Regression

.2

(1)

y (2)

: $B\beta \leq b$ β

$$L(B, \sigma^2, y) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{(y - X\beta)'(y - X\beta)}{2\sigma^2}\right\} \dots\dots\dots (3)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{(n-k)\hat{\sigma}^2}{2\sigma^2} - \frac{(\beta - \hat{\beta})' X' X (\beta - \hat{\beta})}{2\sigma^2}\right\} \dots\dots\dots (4)$$

σ^2, β $\hat{\sigma}^2, \hat{\beta}$

: (4) (3)

: $Y - X\beta$ $X\hat{\beta}$

$$Y - X\hat{\beta} + X\hat{\beta} - X\beta = Y - \hat{Y} + X\hat{\beta} - X\beta$$

$$= \varepsilon + X\hat{\beta} - X\beta \dots\dots\dots (5)$$

$$\hat{Y} = X\hat{\beta}, \quad \varepsilon = Y - \hat{Y}$$

$$\begin{aligned} (Y - X\beta)'(Y - X\beta) &= (\varepsilon + X\hat{\beta} - X\beta)'(\varepsilon + X\hat{\beta} - X\beta) \\ &= (\varepsilon + X(\hat{\beta} - \beta))'(\varepsilon + X(\hat{\beta} - \beta)) \\ &= \varepsilon'\varepsilon + \varepsilon'X(\hat{\beta} - \beta) + (\hat{\beta} - \beta)'X'\varepsilon + (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta) \dots\dots\dots (6) \end{aligned}$$

$$: \hat{\beta} = (X'X)^{-1}X'Y$$

$$\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y$$

$$\begin{aligned} \varepsilon = Y - \hat{Y} &= Y - X(X'X)^{-1}X'Y \\ &= (I - X(X'X)^{-1}X')Y = PY \dots\dots\dots (7) \end{aligned}$$

$$: P = I - X(X'X)^{-1}X'$$

1) $PX = 0$

2) $P = P'$

3) $P = P^2$

: (6)

$$\begin{aligned} \varepsilon'X(\hat{\beta} - \beta) &= (PY)'X(\hat{\beta} - \beta) \\ &= Y'PX(\hat{\beta} - \beta) \\ &= Y'.0.(\hat{\beta} - \beta) = 0 \end{aligned}$$

$$\begin{aligned} (\hat{\beta} - \beta)'X'\varepsilon &= (\hat{\beta} - \beta)'X'(PY) \\ &= (\hat{\beta} - \beta)'(Y'PX)' \\ &= (\hat{\beta} - \beta)'.0 = 0 \end{aligned}$$

(6)

$$\varepsilon' \varepsilon = \frac{\sum \sigma^2}{(n-k)} \cdot (n-k) = (n-k) \hat{\sigma}^2$$

Non – informative

: $\theta = (\beta, \sigma^2)$

$$p(\beta, \sigma^2) \propto 1/\sigma^2, \quad B\beta \leq b$$

θ Posterior Distribution

$$p(\beta, \sigma^2 | y)$$

: y

$$p(\beta, \sigma^2 | y) \propto L(\beta, \sigma^2; y) p(\beta, \sigma^2) \dots \dots \dots (8)$$

: (8) (4)

$$\beta | (\sigma^2, y) \sim N(\hat{\beta}, \sigma^2 (X'X^{-1})) \quad , \quad B\beta \leq B \dots \dots \dots (9)$$

: (8) (3)

$$S(\beta) \sigma^2 | (\beta; y) \sim \chi_n^2 \dots \dots \dots (10)$$

. n

$$\chi_n^2 \quad S(\beta) = (y - x\beta)' (y - x\beta)$$

$$\eta = A\beta \quad A(X'X)^{-1} = 1 \quad A$$

$$b \leq B\beta = BA^{-1}A\beta = BA^{-1}\eta \quad BA^{-1}\eta \leq b \quad \eta$$

$$\alpha = A\hat{\beta} \quad D = BA^{-1}$$

$$\eta | (\sigma^2, y) \sim N(\alpha, \sigma^2 I) \quad , \quad D\eta \leq b \dots \dots \dots (11)$$

$$\theta = (\eta, \sigma^2)$$

: (11) . η

$$\eta_i | (\eta, \sigma^2, y) \sim N(\alpha_j, \sigma^2) \dots \dots \dots (12)$$

: η_j

$$d_j \eta_j \leq b - D_{-j} \eta_j \dots \dots \dots (13)$$

: .3

Tudge and Takayama (1966) Telser (1963)

1943 1925 .

P_{ij}

j i

Telser (1963)

Telsar (1963)

$$y_{it} = \sum_{i=1}^3 y_{i,t-1} p_{ij} + u_{jt} \quad , \quad t = 1, 2, 3 \quad \dots\dots\dots(14)$$

$$u_{it} \quad t = 1, 2, \dots, T \quad t \quad j \quad y_{it}$$

p_{ij}

$$\sum_{i=1}^3 p_{ij} = 1 \quad \text{for all } i \quad \dots\dots\dots(15)$$

$$p_{ij} \geq 0 \quad \text{for all } i \text{ and } j \quad \dots\dots\dots(16)$$

:

(11)

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} W & 0 & 0 \\ 0 & W & 0 \\ 0 & 0 & W \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} + \begin{bmatrix} u \\ u_2 \\ u_3 \end{bmatrix} \quad , \quad \dots\dots\dots(17)$$

$3 \times T$

$$W \quad y_j = [y_{j2}, \dots, y_{jT}]$$

$$j \quad p_j \quad (14)$$

(14)

u_j

Hocking (1996)

(17)

y

$$p_{i3} = 1 - p_{i1} - p_{i2} \quad , \quad i = 1, 2, 3$$

:

$$y - W_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = [W_1 - W_3 \quad W_2 - W_3] \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + u \quad , \quad \dots\dots\dots(18)$$

$$p_{i1} + p_{i2} \leq 1 \quad , \quad i = 1, 2, 3 \quad \dots\dots\dots(19)$$

$$p_{ij} \geq 0 \quad , \quad i = 1, 2, 3 \quad , \quad j = 1, 2 \quad \dots\dots\dots(20)$$

(17)

u

Telser (1963)

Telser (1963)

	p_0	λ_1	λ_2	λ_3	p_{11}	p_{21}	p_{31}	p_{12}	p_{22}	p_{32}	p_{13}	p_{23}	p_{33}	w
e	$\begin{cases} 1 \\ 1 \\ 1 \end{cases}$				1	1		1	1	1	1	1	1	
$X^1 y^1$	$\begin{cases} 2.576 \\ 2.143 \\ 1.942 \end{cases}$	1	1	1	2.658	2.193	1.988							
$X^2 y^2$	$\begin{cases} 2.283 \\ 2.111 \\ 1.750 \end{cases}$	1	1	1				2.658	2.193	1.988				
$X^3 y^3$	$\begin{cases} 1.980 \\ 1.705 \\ 1.509 \end{cases}$	1	1	1				2.193	2.069	1.700	2.658	2.193	1.988	
								1.988	1.700	1.516	1.988	1.700	1.516	$-I$

(1)

Judge and Takayama (1966)

(1)

$$\text{var}(u) = \sigma^2 I$$

.Telsar (1963)

(1)

. Telsar (1963)

$$. u \sim N(0, \sigma^2 I)$$

$$(\sigma_1^2, \sigma_2^2, \sigma_3^2)$$

(18)

(20) (19)

$$A(X'X)^{-1}A = I$$

$$\eta = A\beta$$

.MATLAB

(12)

(1)

WinBUGS

5000

$$. \beta, \sigma^2, \alpha$$

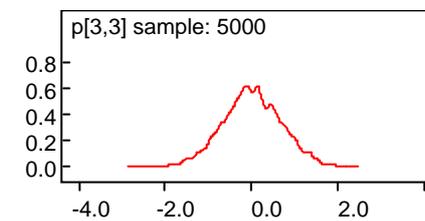
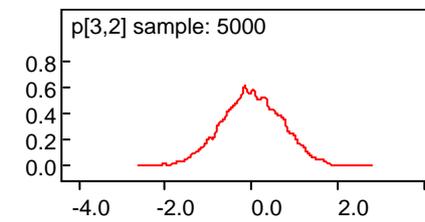
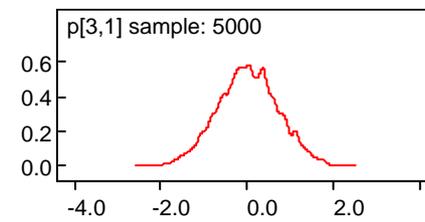
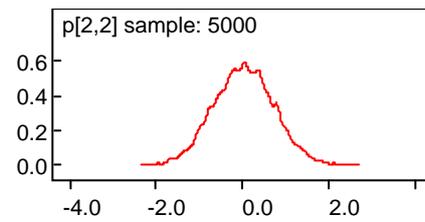
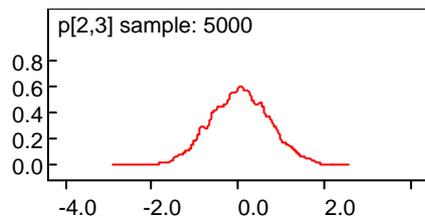
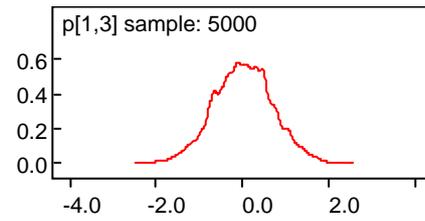
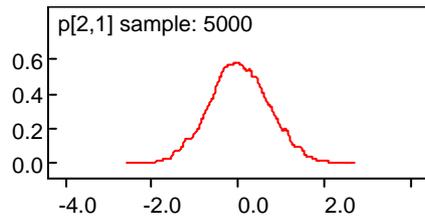
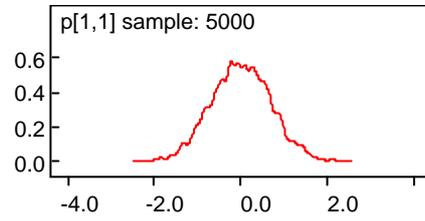
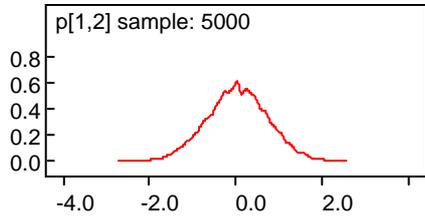
:

$$P = \begin{bmatrix} 0.163 & 0.092 & 0.745 \\ 0.389 & 0.074 & 0.537 \\ 0.219 & 0.513 & 0.268 \end{bmatrix}$$

2

$$p_{ij} \geq 0 \quad \forall i, j$$
$$\sum p_{ij} = 1$$

:



:

:

References

- 1) Gelfand, A. E. and Smith, A. F. M. "Sampling-based Approaches to Calculating Marginal Densities," *J. Amer. Statist. Assoc.*, 85, 398-409. (1990).
- 2) Gelfand, A. E., Smith, A. F. M. and Lee, T. M. "Bayesian Analysis of Constrained Parameters and Truncated Data Problems" *J. Amer. Statist. Assoc.*, 87, 523-532. (1992).
- 3) Geman, S. and Geman, D. "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images," *IEEE trans. pattern anal. mach. intell.*, 6, 721-741. (1984).
- 4) Geweke, J. "Exact Inference in the Inequality Constrained Normal Linear Regression Model," *J. Appl. Econ.*, 1, 127-141. (1986).
- 5) Geweke, J. "Bayesian Inference for Linear Models Subject to Linear Inequality Constraints," In: Zellner, A., Lee, J. S. (Eds.), *Modeling and Prediction: Honouring Seymour Geisser*. Springer, New York. (1996).

-
-
- 6) Gilks, W. R. and Roberts, G. O. "Strategies for Improving MCMC," In Markov Chain Monte Carlo in Practice (eds W. R. Gilks, S. Richardson and D. J. Spiegelhalter), Chapman & Hall, London, 89-114. (1996).
 - 7) Hocking, R. R. "Methods and Applications of Linear Models: Regression and the Analysis of Variance," Wiley, New York. (1996).
 - 8) Judge, G. C. and Takayama, T. "Inequality Restrictions In Regression Analysis," *J. Amer. Statist. Assoc.*, 61, 166-181. (1966).
 - 9) Liew, C. K. "Inequality Constrained Least-Squares Estimation," *J. Amer. Statist. Assoc.*, 71, 746-751. (1976).
 - 10) Lovell, M. C. and Prescott, E.. "Multiple Regression with inequality constraints: Pretesting Bias, Hypothesis Testing, and Efficiency." *J. Amer. Statist. Assoc.*, 65, 913-925. (1970).
 - 11) O'Hagan, A.. "Kendall's Advanced Theory of Statistics 2B: Bayesian Inference." London: Edward Arnold. (1994).
 - 12) Roberts, G. O. "Markov Chain Concepts Related to Sampling Algorithms." In Markov Chain Monte Carlo in Practice (eds W. R. Gilks, S. Richardson and D. J. Spiegelhalter), 45-57. London: Chapman & Hall. (1996).
 - 13) Rodriguez-Yam, G. A., Davis, R. A. and Scharf, L. L. "A Bayesian Model and Gibbs Sampler for Hyperspectral Imaging," Proc. 2002 IEEE Sensor Array and Multichannel Signal Processing Workshop, Washington, D.C. 105-109. (2002).
 - 14) Rodriguez-Yam, G. A. "Estimation for State-Space Models and Bayesian Regression Analysis with Parameter Constraints," Ph.D. Dissertation, Department of Statistics, Colorado State University, USA. (2003).
 - 15) Telser, L. G. "Least Squares Estimates of Transition Probabilities," in Christ, C. F., Friedman, M., Goodman, L. A., Griliches, Z., Harberger, A. C., Liviatan, N., Mincer, J. Mundlak, Y. Nerlove, M. Patinkin, D., Telser, L. G. and Theil, H. (Eds.), *Measurement in Economics: Studies in Mathematical Economics and Econometrics: In memory of Yehuda Grunfeld*. Stanford University Press, Stanford. (1963).