Bayesian Inference for Linear Models Subject to Linear Inequality Constraints

John F. Geweke
Research Department, Federal Reserve Bank of Minneapolis
and Department of Economics
University of Minnesota

Abstract

The normal linear model, with sign or other linear inequality constraints on its coefficients, arises very commonly in many scientific applications. Given inequality constraints Bayesian inference is much simpler than lassical inference, but standard Bayesian computational methods become impractical when the posterior probability of the inequality constraints (under a diffuse prior) is small. This paper shows how the Gibbs sampling algorithm can provide an alternative, attractive approach to inference subject to linear inequality constraints in this situation, and how the GHK probability simulator may be used to assess the posterior probability of the constraints.

1 Introduction

The normal linear regression model subject to linear inequality constraints for the coefficients arises commonly in applied econometrics as well as other scientific applications. Typically the motivating economic model restricts the signs of certain coefficients or of known linear combinations of coefficients. A well-known pedagogical example is provided by Pindyck and Rubinfeld (1981, p. 44) who take up the demand for student housing near the University of Michigan. Rent paid per person is a linear function of the number of rooms per person, with a positive coefficient, and the distance from campus, with a negative coefficient. (We return to this example below.)

The subsequent empirical work in this and other linear regression models subject to linear inequality constraints for the coefficients then focuses on two related but distinct questions. First, how plausible are the linear inequality restrictions delivered by the economic model? Second, conditional on these restrictions what is to be inferred about the coefficients of the regression model? These turn out to be nontrivial questions, and historically investigators have taken different, usually informal, approaches to these tasks. Difficulties for classical inference are discussed in Judge and Takayama (1966) and Lovell and Prescott (1970); for classical testing in Gourieroux, Holly, and Monfort (1982) and Wolak (1987).

J. C. Lee et al. (eds.), *Modelling and Prediction Honoring Seymour Geisser*

classical inference are discussed in Judge and Takayama (1966) and Lovell and Prescott (1970); for classical testing in Gourieroux, Holly, and Monfort (1982) and Wolak (1987).

This work takes up a Bayesian approach to the problem of linear regression with linear inequality constraints on the coefficients. Extending earlier analytical work by Davis (1978), Chamberlain and Leamer (1976), and Leamer and Chamberlain (1976), it uses fast numerical methods for the determination of posterior moments and probabilities, advancing the methods reported in Geweke (1986). But whereas Geweke (1986) takes up any inequality constraints on the coefficients, this paper limits attention to inequality constraints that are linear. The more specialized algorithms provide faster, more accurate numerical approximations to posterior moments than do the more general ones. In particular, when the posterior probability of linear inequality constraints is low or the number of coefficients is large the methods in Geweke (1986) may be slow to the point of impracticality. In contrast computation time in the approach taken here does not increase systematically with the inverse of the posterior probability of the inequality constraints, and increases only linearly with the number of coefficients.

In standard notation the normal linear regression model is

$$y_{T\times 1} = X_{T\times k}\beta_{k\times 1} + \varepsilon_{T\times 1}, \quad \varepsilon \sim N(0, \sigma^2 I_T) \tag{1}$$

where y is the vector of dependent variables, X is the matrix of explanatory variables (regressors), and ε is the vector of disturbances. There are T observations and k explanatory variables. In the unconstrained case a standard diffuse reference prior for the parameter vector (β', σ) is

$$f(\beta,\sigma) \propto \sigma^{-1}. \tag{2}$$

The inequality constraints are expressed

$$a_{k\times 1} \leq D_{k\times k}\beta \leq w_{k\times 1}. \tag{3}$$

In this expression the inequalities are to read line by line: $a_i \leq \sum_{j=1}^{k} d_{ij}\beta_j \leq w_i (i = 1, ..., k)$. The matrix D is composed of real numbers and is nonsingular. The vectors a and w are composed of extended real numbers, with -∞ and +∞ explicitly permitted, thus allowing single-sided inequality constraints. Since a constraint in (3) has no effect if $a_i = -\infty$ and $w_i = +\infty$, fewer than k linear inequality restrictions—perhaps only one—may be involved. Inequality constraints on more than k linear combinations are precluded by (3), a point to which the concluding section of the paper returns briefly. Extending the standard reference prior of the unconstrained model the prior distribution employed in this model is

$$\begin{aligned} f(\beta,\sigma) &\propto \sigma^{-1} \text{ if } \beta \in Q, \ f(\beta,\sigma) = 0 \text{ if } \beta \notin Q; \\ Q &\equiv \{\beta : a \leq D\beta \leq w\}. \end{aligned} \tag{4}$$