Economics Letters 126 (2015) 51–56

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Evaluating simulation-based approaches and multivariate quadrature on sparse grids in estimating multivariate binary probit models

Kibrom A. Abay

Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K, Denmark

HIGHLIGHTS

- I evaluate the performance of Sparse Grids Integration (SGI) and GHK simulator.
- I evaluate the performance of these estimators using Monte Carlo experiments.
- In lower dimension multivariate probit models SGI and GHK perform comparably.
- But as the dimension of integration or dependence among equations increases, the GHK outshines the SGI.

ARTICLE INFO

Article history:
Received 5 March 2014
Received in revised form
31 August 2014
Accepted 17 November 2014
Available online 24 November 2014

JEL classification:
C15
C25
C35

Keywords:
Multivariate probit model
Simulation approaches
GHK simulator
Multivariate quadrature-based approaches
Sparse grids integration

ABSTRACT

This paper evaluates the performance of a recently emerging multivariate quadrature-based Sparse Grids Integration (SGI) and the well-known Geweke–Hajivassiliou–Keane (GHK) simulator in estimating multivariate binary probit models. Monte Carlo exercises demonstrate that in lower dimension multivariate binary probit models, the multivariate quadrature-based SGI estimator with few quadrature points performs very well and comparable with the GHK simulator. But as the dimension of integration or dependence (error correlation) among equations increases, the GHK simulator outshines the SGI estimator. This indicates that for integration problems involving higher dimension multivariate probit models, and those with strong dependence among variables, the GHK simulator remains to be a more efficient approach.

1. Introduction

Estimating correlated multivariate limited dependent variable models involves evaluating intractable multidimensional integrals. There have been different approaches to approximate such multidimensional integrals in computing likelihood functions in these models; the most common being simulation and quadrature-based approaches. Within the simulation-based approaches the Geweke–Hajivassiliou–Keane (GHK) simulator, named after Geweke (1991), Hajivassiliou and McFadden (1998), and Keane (1994), is found to be the most efficient approach (see Geweke et al., 1994; Hajivassiliou et al., 1996).[1] Other competing alternatives to evaluate multidimensional integrals in estimating discrete choice models rely on multivariate quadrature-based approaches. Recently, Heiss and Winschel (2008) propose a multivariate quadrature based on sparse grids constructed using the Smolyak rule (Smolyak, 1963), and they demonstrate that it can serve as a powerful alternative to simulation approaches in evaluating multidimensional econometric integrals. This approach, Sparse Grids Integration (SGI), constructs nodes and weights for multivariate quadrature by combining univariate quadrature rules as in the product rule but it does by selecting fewer combinations in a more "clever" way, and hence does not suffer from the curse of dimensionality.

However, the relative performance of the SGI has not been evaluated in different econometric models involving multidimensional integration problems. In this paper, I evaluate the performance of the GHK simulator and the SGI estimator in estimating multivariate binary probit models. Towards this end, I estimate a fully-correlated multivariate binary probit model and compare the

E-mail address: Kibrom.Araya.Abay@econ.ku.dk.

[1] Empirically, the GHK simulator has been commonly used in estimating unordered multinomial probit models (see Train, 2009) and recently in an ordered response framework (for example, Abay et al., 2013; Bhat et al., 2010).

http://dx.doi.org/10.1016/j.econlet.2014.11.021

GHK simulator and the quadrature-based SGI approach in recovering the "true" parameters of interest as well as in terms of their computational cost. In comparing both integration approaches, I implement Monte Carlo experiments considering progressively increasing dimension of integration and alternative error correlation structures. Such an investigation helps to identify alternative and more efficient estimation approaches that can ease the complexity and computational cost of the existing techniques in estimating discrete choice models. Furthermore, the relative performance of the GHK simulator and the quadrature-based SGI estimator may vary depending on the dimension and dependence structure of the integration problem. Hence, evaluating the relative advantages of these alternative integration approaches in estimating multivariate probit models of different type and structure is crucial.

2. Model and estimation approaches

2.1. Multivariate binary probit model

Consider the following series of *M-variate* fully-correlated latent continuous (unobserved) outcomes of interest representing a sequence of M different binary outcome variables or a univariate binary outcome of interest at M different times:

$$y_{nm}{}^{*} = \beta' X_{nm} + \varepsilon_{nm} \tag{1}$$

where $n = 1 \dots N$ stands for individuals and $m = 1 \dots M$ stands for the number of binary outcome variables or choice situations of interest. This model can be used to model correlated package of binary outcome variables or a panel univariate model with unrestricted error covariance structure across time. But for convenience, this paper focuses on a cross-sectional multivariate probit model specification. As usual, the relation between these latent continuous outcomes and the observed sequence of binary outcomes can be expressed as:

$$y_{nm} = \begin{cases} 1 & \text{if } y_{nm}{}^{*} > 0 \\ 0 & \text{if } y_{nm}{}^{*} \leq 0. \end{cases}$$

In the above formulation, the regression parameters are fixed to be equal across equations for simplicity. The error terms of the series of equations, $\varepsilon_n = (\varepsilon_{n1}, \varepsilon_{n2} \dots \varepsilon_{nM})$, are identically and independently distributed across individuals but allowed to be fully-correlated across equations. To complete the model structure, one needs to be explicit about the distribution and structure of the *M-variate* error terms of the equations. In this paper, the correlated error terms of the equations are assumed to be distributed multivariate normal as $\varepsilon_n \sim MVN_M(0, \Sigma_M)$, with Σ_M assuming the following structure:

$$\mathbf{\Sigma}_M = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1M} \\ \rho_{21} & 1 & \rho_{23} & \cdots & \rho_{2M} \\ . & . & & .. & .. \\ . & . & & .. & .. \\ \rho_{M1} & \rho_{M2} & \rho_{M3} & \cdots & 1 \end{bmatrix}. \tag{2}$$

The diagonal elements in the above matrix are normalized to be 1 for identification purpose. Given this structure, we can construct the likelihood function for all possible response outcomes considered. For instance, the probability that an individual chooses "1" in all binary choice variables can be written as:

$$\begin{aligned} &\Pr(y_{n1} = 1, y_{n2} = 1, \dots, y_{nM} = 1) \\ &= L(\beta, \Sigma) = \int_{\ell_1=-\infty}^{\beta' X_{n1}} \int_{\ell_2=-\infty}^{\beta' X_{n2}} \cdots \int_{\ell M=-\infty}^{\beta' X_{nM}} \phi_M\,(\ell_1, \ell_2, \\ &\quad \ell_3, \dots \ell M | \sum M \Big)\, d\ell_1, d\ell_2 \dots d\ell_M \end{aligned} \tag{3}$$

where $\phi_M(\cdot)$ stands for an *M-variate* standard normal density function.

2.2. Estimation approaches

Evaluating the above joint probability for a sequence of events entails *M-dimensional* rectangular integration for each individual. This integrand can be approximated using simulation approaches, and within these approaches the GHK simulator is found to be the most efficient method to approximate multidimensional integrals of the form in Eq. (3). The GHK simulator works based on recursive conditioning of multivariate outcomes or sequence of outcomes. For instance, the joint probability that an individual chooses "1" in all binary choice variables considered can be expressed as a product of conditional probabilities as:

$$\begin{aligned} &\Pr(y_{n1} = 1, \dots, y_{nM} = 1) \\ &= \mathrm{pr}(y_{n1} = 1)\, \mathrm{pr}(y_{n2} = 1 | y_{n1} = 1) \dots \mathrm{pr}(y_{nM} \\ &= 1 | y_{nm} - 1 = 1, \dots, y_{n1} = 1). \end{aligned} \tag{4}$$

This joint probability for a sequence of events involves conditioning on prior events, which are all correlated due to the allowance of a correlated error structure. For convenience, we can express the correlated error structure as $\varepsilon_n = (\varepsilon_{n1}, \varepsilon_{n2} \dots \varepsilon_{nM})$, as $\varepsilon_n = C\eta_n$, where C stands for the lower triangular matrix of the Cholesky decomposition of the correlation matrix Σ_M, while the random terms η_n are now uncorrelated to each other and distributed as $\eta_n \sim MVN_M(0, I_M)$. Train (2009) provides a brief exposition on all steps in implementing the GHK simulator.

Another competing approach to evaluate multidimensional integrals of the type in Eq. (3) is an approximation based on multivariate quadrature. Multivariate quadrature based on simple tensor product of univariate quadrature rules involves exponentially increasing computational cost. Heiss and Winschel (2008) propose a multivariate quadrature on sparse grids constructed by the Smolyak rule, and they demonstrate its powerful performance in estimating mixed-logit models. The notion of the Smolyak rule assumes that some tensor product combinations are more important than others, and it provides a design to strategically select few tensor product combinations satisfying some criteria. Judd et al. (2013) best describe the Smolyak rule as a design that indicates which tensor product combinations should be selected for constructing multivariate (sparse grids) quadrature. Compared to the conventional tensor product rule way to expand univariate quadrature to multiple dimensions, constructing multivariate quadrature using the Smolyak rule reduces computational cost dramatically (see Heiss and Winschel, 2008; Judd et al., 2013).

To formally express the way the Smolyak rule selects the most important tensor product combinations, let i be a unidimensional accuracy level that defines a sequence of univariate quadrature rules (for each dimension) $Q_{i1}, Q_{i2} \dots Q_{iD}$ which generate a sequence of $R_{i1}, R_{i2} \dots R_{iD}$ univariate quadrature points and corresponding $w_{i1}, w_{i2} \dots w_{iD}$ weights. Furthermore, let k be the underlying accuracy level chosen for constructing a multivariate quadrature of dimension D.[2] Then, the Smolyak rule selects tensor products with different possible combinations of accuracy levels $\mathbf{i} = [i_1, i_2 \dots i_D]$ which satisfy that $\max(D, k+1) \leq |i| = \sum_{d=1}^{D} i_d \leq D + k$ (see Judd et al., 2013). With this criteria, and as described in Heiss and Winschel (2008), and Judd et al. (2013) the Smolyak rule combines univariate quadrature rules in such a way[3]:

$$\begin{aligned} Q_k^D = &\sum_{\mathbf{max}(D,k+1) \leq |i| \leq D+k} (-1)^{D+k-|i|} \binom{D-1}{D+k-|i|} \\ &\times (Q_{i1} \otimes Q_{i2} \dots \otimes Q_{iD}) \end{aligned} \tag{5}$$

[2] This underlying level of accuracy is the parameter that controls the number of tensor product combinations to be constructed. Higher level of this parameter implies larger number of tensor product combinations are chosen.

[3] For a more general treatment and examples, the reader is referred to Heiss and Winschel (2008), and Judd et al. (2013).