Computational Statistics and Data Analysis 53 (2008) 272–288

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Improving MCMC, using efficient importance sampling

Roman Liesenfeld [a,*], Jean-François Richard [b]

[a] Christian-Albrechts-Universität zu Kiel, Department of Economics, Ohlshausenstr. 40-60, 24118 Kiel, Germany

[b] University of Pittsburgh, Department of Economics, 4S01 Wesley W. Posvar Hall, Pittsburgh, PA 15260, USA

ARTICLE INFO

Article history:
Received 19 February 2008
Received in revised form 19 July 2008
Accepted 22 July 2008
Available online 26 July 2008

ABSTRACT

A generic Markov Chain Monte Carlo (MCMC) framework, based upon Efficient Importance Sampling (EIS) is developed, which can be used for the analysis of a wide range of econometric models involving integrals without analytical solution. EIS is a simple, generic and yet accurate Monte-Carlo integration procedure based on sampling densities which are global approximations to the integrand. By embedding EIS within MCMC procedures based on Metropolis–Hastings (MH) one can significantly improve their numerical properties, essentially by providing a fully automated selection of critical MCMC components, such as auxiliary sampling densities, normalizing constants and starting values. The potential of this integrated MCMC–EIS approach is illustrated with simple univariate integration problems, and with the Bayesian posterior analysis of stochastic volatility models and stationary autoregressive processes.

1. Introduction

Monte Carlo (MC) simulation methods are widely used to analyze a broad range of econometric models involving integrals, for which no analytical solution exists. Excellent surveys on MC based econometric methods are available (see, e.g., Geweke (1999), Gilks et al. (1996), Gourieroux and Monfort (1996) and Stern (1997)). Two methods dominate the field. Importance Sampling (IS) was introduced in econometrics by Kloek and van Dijk (1978) and widely used in the eighties. It became progressively supplanted by Markov Chain Monte Carlo (MCMC) in the nineties. Seminal papers are that of Gelfand and Smith (1990), proposing MCMC techniques for Bayesian computations, and that of Tanner and Wong (1987), who introduced data augmentation for the treatment of latent variables. The main argument against the use of IS is the potential non-existence of the MC sampling variance, which could lead to disastrous properties of IS estimates (see, e.g., Geweke (1989, 1999) and more recently Koopman and Shephard (unpublished)). In the present paper we will argue that the IS and MCMC methods are more closely related than is generally recognized and, in particular, that the argument of non-existing variance which contributed to the relative demise of IS also applies to Metropolis–Hastings (MH) procedures which represent, together with the Gibbs-sampling techniques, the most widely used MCMC algorithms.

The fundamental insight which motivates the present paper lies in the observation that the sampling properties of both IS and MCMC critically depend upon the adequacy of an auxiliary sampler m, meant to approximate (up to a proportionality constant) a density kernel φ which needs to be numerically integrated (by itself and/or to compute expectations of functions of interest). Efficient Importance Sampling (EIS), as proposed by Richard and Zhang (2007), provides a generic and essentially automated Least Squares (LS) procedure to construct such (near) optimal approximations within a preassigned parametric class $\mathcal{M}$ of auxiliary samplers. Whence we should be able to facilitate the design as well as improve the sampling properties of MCMC by relying upon auxiliary EIS regressions to construct m. Most importantly, once it is recognized how closely related

* Corresponding author. Tel.: +49 431 880 3810; fax: +49 431 880 7605.
E-mail address: liesenfeld@stat-econ.uni-kiel.de (R. Liesenfeld).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2008.07.028

EIS and MCMC are, the key issue of deciding which approach to use for a particular application becomes one of objective comparison of their respective ease of implementation and statistical adequacy, within the context of that application, not of subjective preference for one method or the other.

Typically, EIS can be expected to have comparative advantages at integrating out high-dimensional dynamic latent variables. The efficiency of EIS in such situations is illustrated by its application for the computation of the likelihood of various dynamic latent variable models (see, e.g., in Bauwens and Hautsch (2006), Bauwens and Galli (in press), Jung and Liesenfeld (2001) and Liesenfeld and Richard (2003a,b, 2008)). On the other hand, intricate Bayesian posterior densities are generally more amenable to MCMC integration.

A striking illustration of how such complementarity can be taken advantage of is provided in Section 7.1, where we conduct a multi-block EIS–MCMC Bayesian analysis of the standard stochastic volatility (SV) model. Specifically, we rely upon an original multi-block implementation of EIS to construct MCMC proposal densities for the latent variables conditional on the parameters, providing an efficient solution to the slow convergence of MCMC in the context of highly correlated latent variables (see, e.g., Carter and Kohn (1994) and Liu et al. (1994)). Moreover, even in applications where MCMC is unequivocally to be preferred, we can still consider embedding auxiliary EIS steps within the MCMC algorithm, in order to facilitate the construction of the MCMC sampler. In Section 7.2 we propose an innovative full Bayesian MCMC analysis of a stationary AR(p) autoregressive process under the non-linear parametrization associated with the roots of the process (a parametrization which is more easily interpretable and of more fundamental interest than the linear coefficients of the AR polynomial). Our MCMC posterior sampler relies upon an auxiliary EIS approximation of the likelihood function. Not only is it fully automated, but also proposal draws satisfy the stationarity conditions by construction. This produces a numerically efficient MCMC implementation illustrated by a 6th order pilot application with a pair of complex roots very close to the unit circle.

The rest of the paper is organized as follows. In Section 2, we review the principles of the MC integration procedures under consideration. Sections 3 and 4 contain a brief description of the EIS technique and MCMC approaches, respectively. In Section 5 we discuss key benefits offered by the combination of EIS and MH. This is illustrated in Sections 6 and 7 with numerical examples. In particular, we discuss two simple univariate integration problems (Sections 6.1 and 6.2), a Bayesian analysis of a SV model (Section 7.1), and a MCMC analysis of the roots of AR models (Section 7.2). Section 8 concludes.

2. MC integration

Consider a continuous random variable X with support $\Delta \subset \mathbb{R}^T$. We assume that its density function $f(x)$ is characterized by a density kernel $\varphi(x)$, whose integrating constant on Δ is unknown. That is to say

$$f(x) = c^{-1} \cdot \varphi(x), \quad \text{with } c = \int_\Delta \varphi(x)\mathrm{d}x. \tag{1}$$

Let $g(x)$ denote a φ-integrable function. Its expectation on f is given by

$$I_g = \mathrm{E}_f\left[g(X)\right] = \frac{\int_\Delta g(x)\varphi(x)\mathrm{d}x}{\int_\Delta \varphi(x)\mathrm{d}x}. \tag{2}$$

IS and MCMC (MH) are commonly used to evaluate such (ratios of) integrals. In order to motivate our paper, we first present both methods in a particular way which serves highlighting their intrinsically close relationship. Technical assumptions validating these methods are extensively discussed in the literature – see, e.g., Geweke (1989) and Robert and Casella (2004) – and are omitted here.

All MC integration techniques considered in the present paper share the characteristic that they rely upon draws $\left\{\tilde{x}_i; i = 1, \dots, S\right\}$ from a proposal density. They differ only by how they use these draws: IS uses i.i.d. draws with weights given by the ratios $\varphi(\tilde{x}_i)/m(\tilde{x}_i)$, where m denotes the proposal density; Similarly, Accept–Reject (AR) uses them with $\{0, 1\}$ weights determined by auxiliary uniform draws; MH allows for repetitions in addition to rejections, resulting in integer weights $\{0, 1, 2, \dots\}$. An objective comparison between these alternative procedures ought to focus on how the proposal draws are implicitly weighted in the construction of the corresponding MC estimates of I_g. For the purpose of such comparisons, it proves convenient to reinterpret IS, AR and MH estimates of I_g as randomized weighted sums of the function g evaluated at the proposal draws, say

$$\bar{I}_g = \frac{\sum_{i=1}^{S} g(\tilde{x}_i) \cdot \tilde{\rho}_i}{\sum_{i=1}^{S} \tilde{\rho}_i}, \tag{3}$$

where $\tilde{\rho}_i$ denotes the random weight assigned to the proposal draw $\tilde{x}_i$. Let μ denote the joint distribution of the proposal draws, and all additional uniform draws associated with the acceptance, rejection, and repetition steps of the particular MC procedure under consideration. Let m denote the marginal density of the proposal draw $\tilde{x}_i$. IS relies upon the weights

$$\tilde{\rho}_i \equiv \omega(\tilde{x}_i) = \frac{\varphi(\tilde{x}_i)}{m(\tilde{x}_i)}, \quad \text{with } \mathrm{E}_m\left[\omega(\tilde{x}_i)\right] = c. \tag{4}$$