ARTICLE IN PRESS

Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

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

Efficient importance sampling in mixture frameworks

Tore Selland Kleppe [a,*], Roman Liesenfeld [b]

[a] *Department of Mathematics, University of Bergen, Bergen, Norway*

[b] *Department of Economics, Christian-Albrechts-Universität, Kiel, Germany*

ARTICLE INFO

Article history:
Received 23 April 2012
Received in revised form 28 January 2013
Accepted 29 January 2013
Available online xxxx

Keywords:
Data augmentation
Dynamic latent variable model
Importance sampling
Marginalized likelihood
Mixture
Monte Carlo
Realized volatility

ABSTRACT

A flexible importance sampling procedure for the likelihood evaluation of dynamic latent variable models involving mixtures of distributions leading to possibly heavy tailed or multi-modal target densities is provided. The procedure is based upon the efficient importance sampling (EIS) approach and exploits the mixture structure of the model via data augmentation when constructing importance sampling distributions as mixtures of distributions. The proposed mixture EIS procedure is illustrated with ML estimation of a Student-*t* state space model for realized volatilities. MC simulations are used to characterize the sampling distribution of the ML estimator based upon the mixture EIS approach.

1. Introduction

In recent decades Monte Carlo (MC) procedures based upon Importance Sampling (IS) have been successfully applied for the analysis of econometric models involving multiple integrals for which no analytical solutions exist. Important applications of IS are the evaluation of Bayesian posterior expectations of functions of parameters of interest and that of likelihood functions in the presence of unobservable latent variables; see, e.g., Kloek and Van Dijk (1978), Geweke (1989), and Durbin and Koopman (1997).

It is well-known that the reliable and efficient use of IS requires that the IS auxiliary density closely mimics the target density kernel which needs to be integrated, and exhibits tails that do not decay more quickly than the tails of the target density (see, Geweke, 1989; Robert and Casella, 2004). This implies that IS implementations have to be tailored to the problem under consideration, which has proved to be a significant obstacle to routine applications of IS. This is especially true for applications with ill behaved, and therefore, difficult to approximate target densities. A survey of IS approaches is found, e.g., in Liesenfeld and Richard (2001).

Another critical issue is that most of the existing IS approaches do not appear to be applicable to highly multidimensional integration problems. Prominent exceptions are the high-dimensional IS methods proposed by Shephard and Pitt (1997), Durbin and Koopman (1997, 2000), and the Efficient Importance Sampling (EIS) procedure of Richard and Zhang (2007). Existing implementations of those methods rely on IS densities from the exponential family of distributions, which, in the case of EIS, considerably simplifies the implementation. While the IS approaches of Shephard and Pitt (1997) and Durbin and Koopman (1997, 2000) use Gaussian IS densities constructed from local Taylor-series approximations to the target density, the IS densities of the EIS approach of Richard and Zhang (2007) are based upon global approximations to the target obtained via a sequence of low-dimensional auxiliary least-squares regressions.

* Correspondence to: Universitetet i Bergen, Matematisk institutt, Postboks 7803, 5020 Bergen, Norway. Tel.: +47 95922364; fax: +47 55589672.
E-mail address: tore.kleppe@math.uib.no (T.S. Kleppe).

0167-9473/$ – see front matter
doi:10.1016/j.csda.2013.01.025

Please cite this article in press as: Kleppe, T.S., Liesenfeld, R., Efficient importance sampling in mixture frameworks. Computational Statistics and Data Analysis (2013), doi:10.1016/j.csda.2013.01.025

Those high-dimensional (E)IS approaches have been successfully applied to the computation of the likelihood for a broad range of dynamic latent variable (DLV) models, where the target densities are reasonably well behaved, such that they can be well approximated by IS densities from the exponential family of distributions (see, e.g., Sandmann and Koopman, 1998; Liesenfeld and Richard, 2003, 2010; Bauwens and Galli, 2009; Koopman et al., 2011; Kleppe and Skaug, 2012). However, for DLV models with pathological target densities, featuring e.g. multi-modality or fat tails, those high-dimensional procedures based upon IS densities from the exponential family might have severe convergence problems. In the context of DLV models, such ill-behaved target densities are often caused by finite or infinite mixture-of-distributions specifications assumed for some of the variables. An example, to be discussed further below are state space models with Student-t measurement errors. Hence, there exists a need for high-dimensional IS procedures based upon flexible IS densities beyond the exponential family of distributions.

In the present paper, we extend the (high-dimensional) EIS approach of Richard and Zhang (2007), by introducing mixtures of distributions as flexible classes of IS distributions allowing to approximate target densities which are possibly heavy-tailed or multi-modal. Our approach is particularly well adapted to the likelihood evaluations for DLV models involving variables characterized by a mixture of distributions, which can be exploited when constructing the IS densities via a simple data-augmentation step. Under appropriate simplifying conditions our proposed mixture EIS procedures rely, similarly to EIS implementations for IS densities from the exponential family, on a simple sequence of auxiliary least-squares regressions used to obtain close approximations to the integrand.

Alternative IS procedures using flexible mixtures of distributions as IS densities are the 'split'-Student IS approach of Geweke (1989), the 'defensive' mixture technique proposed by Hesterberg (1995) and the adaptive method of Ardia et al. (2009) using a mixture of Student-t distributions. While those approaches have been successfully applied to lower dimensional Bayesian integration problems, they do not appear to be applicable to very high-dimensional integrals, which needs to be approximated, e.g., for the likelihood evaluation of DLV models. A very promising high-dimensional IS sampling approach for the analysis of DLV models using flexible IS densities beyond the exponential family was recently proposed by McCausland (2012). His so-called HESSIAN method relies on IS densities from a parametric family of perturbed Gaussian densities. However, its current version is restricted to the application of DLV models with latent variables following a univariate Gaussian AR(1) process.

The rest of the paper is organized as follows. In Section 2, we briefly review the generic principles of EIS, and in Section 3 we introduce the mixture EIS approach. Section 3 also provides a simple low-dimensional example illustrating the benefits of the proposed mixture EIS. A sequential implementation of the mixture EIS approach for high-dimensional integration required for a likelihood analysis of realistic models is discussed in Section 4 and illustrated through the ML estimation of a Student-t state space model for realized volatilities. Section 5 concludes.

2. Efficient importance sampling (EIS)

2.1. General principle

Consider the problem of evaluating an integral of the form

$$I = \int \varphi(z) dz, \tag{1}$$

where $\varphi : \Delta \mapsto \mathbb{R}^+$ with $\Delta \subseteq \mathbb{R}^T$. Of special interest in this paper focusing on likelihood evaluations of DLV models is the case where the econometric model under consideration leads to an initial factorization of the integrand of the form

$$\varphi(z) = g(z) \cdot p(z), \tag{2}$$

where p is a probability density function for z referred to as the initial or natural (model based) sampler, and $g : \Delta \mapsto \mathbb{R}^+$ is a p-integrable function. The function g typically represents a conditional density of data given latent variable z as a function of z given the data.

IS integration consists of selecting an IS density, say $m(z)$, and rewriting Eq. (1) as

$$I = \int \left[\frac{\varphi(z)}{m(z)}\right] \cdot m(z) dz, \tag{3}$$

where $\varphi(z)/m(z)$ is the importance weight function. The corresponding MC–IS estimator of I is given by

$$\hat{I} = \frac{1}{M} \sum_{j=1}^{M} \frac{\varphi(z^{(j)})}{m(z^{(j)})}, \tag{4}$$

where $\{z^{(j)}\}_{j=1}^{M}$ denotes a set of M identically independently distributed draws from m.

The technical conditions, under which the IS estimator (4) converges almost surely to I and its variance is finite are discussed in Geweke (1989) and Robert and Casella (2004). A sufficient condition for the finiteness the variance of $\hat{I}$ is that $|\varphi(z)/m(z)|$ be bounded above on Δ.