Journal of Econometrics 201 (2017) 198–211

Contents lists available at ScienceDirect

Journal of Econometrics

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

Bayesian estimation of state space models using moment conditions

A. Ronald Gallant [a,*], Raffaella Giacomini [b], Giuseppe Ragusa [c]

[a] *Penn State University, USA*
[b] *University College London, United Kingdom*
[c] *Luiss University, Italy*

ARTICLE INFO

Article history:
Available online 1 September 2017

JEL classification:
C32
C36
E27

Keywords:
State space models
Bayesian estimation
Moment equations
Structural models
DSGE models
Particle filter

ABSTRACT

We consider Bayesian estimation of state space models when the measurement density is not available but estimating equations for the parameters of the measurement density are available from moment conditions. The most common applications are partial equilibrium models involving moment conditions that depend on dynamic latent variables (e.g., time–varying parameters, stochastic volatility) and dynamic general equilibrium models when moment equations from the first order conditions are available but computing an accurate approximation to the measurement density is difficult.

1. Introduction

We propose a method for conducting Bayesian inference regarding the parameters of a nonlinear structural model that has dynamic latent variables. By latent variables we mean all endogenous and exogenous variables in the model that are not observed.

The general approach to dealing with dynamic latent variables in econometrics is to resort to filtering techniques (e.g., the particle filter), which, in connection with Markov Chain Monte Carlo (MCMC) methods, deliver estimates of the structural parameters (see Andrieu et al., 2010). To implement a particle filter one needs to be able to: (1) draw from the transition density of the latent variables, which specifies the distribution of the latent variables conditional on their past history; and (2) evaluate the measurement density, which specifies the distribution of the observable variables conditional on the latent variables.

In this paper, we maintain the assumption that one can draw from the transition density of the latent variables but we assume that a measurement density is not available and/or it is difficult to approximate numerically. What is available is instead a set of moment conditions that provide estimating equations for the parameters of the measurement density. The most common applications in econometrics where this situation arises are (1) partial equilibrium models that involve moment conditions depending on dynamic latent variables (e.g., time-varying parameters, stochastic volatility); and (2) dynamic general equilibrium structural models when moment equations from the first order conditions are available but computing an accurate approximation to the measurement density is difficult. There are currently no econometric methods that apply to the first class of models, and for the second class of models our method can be considered as an alternative to existing approaches that does not rely on approximations or numerical solutions of the model.

The method of moments has a powerful appeal in economic research and researchers are increasingly keen to use prior information as a means to deal with data limitations. The method we propose here has potential to become a useful tool in applied economic research, because – as argued by Cochrane (2005) – most researchers find evidence based on method of moments more persuasive than evidence based on fully specified likelihoods. Our contribution is to show that combining method of moments and priors is viable theoretically and practically in economic models where the presence of dynamic latent variables makes it impossible to apply standard GMM estimation.

In fact, if one considers calibration to be Bayesian method of moments with extremely strong priors, then most of the science that matters in our daily lives uses Bayesian method of moments. In particular, climate models and macro models. The main exception is health, but this is mostly due to government regulation. Also, the exceptions one finds in macro are mostly due to the pressure of central banks. Our view is that if statistics is to become relevant

* Correspondence to: P.O. Box 659, Chapel Hill NC 27514, USA.
E-mail address: aronldg@gmail.com (A.R. Gallant).

http://dx.doi.org/10.1016/j.jeconom.2017.08.003
0304-4076/

to major policy decisions, then something along the lines of what we propose has to become viable.

We illustrate the usefulness of our method by applying it to the problem of estimating the latent endowment process in a Lucas (1978) economy given only knowledge of the agent's first order conditions and of the transition density of the latent process. The process we extract differs markedly from measured consumption and suggests the presence of stochastic volatility and jumps.

The central idea of the paper is to show that the moment conditions can be used to construct a "GMM representation" of the measurement density that one can substitute for the measurement density as an input into an otherwise standard filtering MCMC algorithm.

To illustrate, suppose we have a set of M moment conditions

$\mathcal{E}[g(y_{t+1}, x_{t+1}, \theta)] = 0$[1]

implied by a structural model. We observe a realization $y = \{y_1, \ldots, y_T\}$ from the stochastic process $\{\ldots, y_{t-1}, y_t, y_{t+1}, \ldots\}$ but we do not observe $\{\ldots, x_{t-1}, x_t, x_{t+1}, \ldots\}$ which is thus the latent process. What we know about the latent process is a parametric specification for its transition density. The objective is to obtain the posterior distribution of the structural parameter θ (comprised of the parameters of both the moment conditions and the transition density) and the posterior distribution of the latent process. Formally, the posterior is given by

$$p^o(\theta, x|y) \propto p^o(y|x, \theta)p^o(x|\theta)p^o(\theta)$$

where the measurement density $p^o(y|x, \theta)$ is unknown aside from the restrictions implicitly imposed by the moment conditions, the joint density of the latent variables $p^o(x|\theta)$ is pinned down by the transition density, and the prior $p^o(\theta)$ of the parameters is specified by the researcher. The contribution of this paper is twofold. We first show that the moment conditions induce a probability structure that allows us to replace the unknown transition density $p^o(y|x, \theta)$ with a known density $p^*(y|x, \theta)$. We then propose a numerical algorithm that uses the particle filter and a Metropolis algorithm to draw from the posterior $p^*(\theta, x|y) \propto p^*(y|x, \theta)p^o(x|\theta)p^o(\theta)$.

Regarding the first contribution, we build on and extend the results of Gallant and Hong (2007) and Gallant (2016a, b, c) to an environment with dynamic latent variables. The key insight is to show how to replace the probability space over $(\mathcal{Y} \times \mathcal{X} \times \Theta, \mathcal{C}^o, P^o)$ implied by the structural model and a prior for θ (where $\mathcal{Y} \times \mathcal{X}$ is the support of the observable and latent variables, Θ is the support of θ, and $\mathcal{C}^o$ is the collection of Borel subsets of $\mathcal{Y} \times \mathcal{X} \times \Theta$) by an alternative probability space $(\mathcal{Y} \times \mathcal{X} \times \Theta, \mathcal{C}^*, P^*)$. The alternative probability space is such that $\mathcal{C}^*$ is a subset of $\mathcal{C}^o$ and the density of P^* is the same as P^o except that the measurement density is replaced by a density function evaluated at the sample moment conditions g_T (scaled to have variance equal to the identity matrix, i.e., $p^*(y \,|\, x, \theta) = \psi([\Sigma(y, x, \theta)]^{-1/2} g_T(y, x, \theta)])$. We call this density function the "GMM representation" of the measurement density. Because we are concerned with subjective Bayesian inference, we assume that the density function ψ is specified by the user.[2] In practice, we suggest using the standard normal density, which is motivated by the asymptotic normality of the sample moments under the standard regularity assumptions. The key insight that allows us to substitute the unknown measurement density with its GMM representation is the fact that both probability measures assign the same probability to sets in $\mathcal{C}^*$. Naturally, because $\mathcal{C}^*$ is a subset of $\mathcal{C}^o$, some information is lost. Intuitively this is similar to the information loss that occurs when one divides the range of a continuous variable into intervals and uses a discrete distribution to assign probability to each interval. Both the continuous and discrete distributions assign the same probability to each interval but the discrete distribution cannot assign probability to subintervals. How much information is lost depends on how well one chooses moment conditions. An in-depth investigation of the effects of moment choice on inference is beyond the scope of this paper, but we provide some advice on choice strategy for some key economic applications. In many instances, as in the application of Section 6, discussion of the choice of moments is moot because the economics of the situation dictate the choice.

In the state-space literature to which we contribute, (cf. Flury and Shephard, 2011; Fernandez-Villaverde and Rubio-Ramirez, 2006) the assumption that one can draw from the transition density is standard. Our contribution is to be able to perform Bayesian inference without knowledge of the measurement density.

The importance of the first contribution is easy to overlook. What it does is establish the methodology as exact within the Bayesian paradigm given the information that the researcher chooses to use. Leaving aside specification error, inaccurate algorithms, etc. that plague all statistical methods, we are proposing exact Bayesian methods, not approximate Bayesian methods.

Regarding our second contribution, which builds on ideas from Beaumont (2003), Andrieu and Roberts (2009), Andrieu et al. (2010) and Flury and Shephard (2011), the computational strategy we propose consists of two steps: a conditional particle filter step that draws x given y, θ, and the previously drawn x and a Metropolis step that draws θ given y, x, and the previously drawn θ. The validity of the algorithm follows from the results of Andrieu et al. (2010) as it can be thought of as an adaptation of their particle Gibbs sampler when one has to resort to the GMM representation of the measurement density. The application of the algorithm results in an MCMC chain in (θ, x) and thus parameter estimates, standard deviations, and other characterizations of the posterior distribution can be computed from this chain in the standard way (Gamerman and Lopes, 2006).

The main attraction of the method we propose is that one does not have to solve the structural model. For partial equilibrium models this is crucial because, in general, there do not exist practicable alternatives.

We also expect that an important application for our results will be statistical inference regarding general equilibrium models in macroeconomic applications such as dynamic stochastic general equilibrium models (DSGE). For analytically intractable DSGE models there are alternatives to what we propose that rely on being able to solve the model numerically. For instance, one can use perturbation methods to approximate the model, use the approximation to obtain an analytical expression for the measurement density, and then use some method of numerical integration such as particle filtering to eliminate the latent variables along the lines proposed by Fernandez-Villaverde and Rubio-Ramirez (2006) and Flury and Shephard (2011). Alternatively, one can solve the model only to the point of being able to simulate it and then use the methods proposed by either Gallant and McCulloch (2009), who use an SNP (Gallant and Nychka, 1987) representation of the measurement density, or Gallant and Tauchen (2015), who use an EMM (Gallant and Tauchen, 1996) representation of the measurement density.

In the case of DSGE models, the main reason one might want to consider our alternative to the existing procedures is that one has misgivings about the quality of the numerical methods one has used to solve the structural model. For instance, perturbation

[1] Expectation for moment conditions is determined by context. If θ is regarded as exogenous and a likelihood $p(x, y \,|\, \theta)$ is well defined, then the meaning is $\iint g(y_{t+1}, x_{t+1}, \theta)\, p(x, y \,|\, \theta)\, dy\, dx = 0$. If θ is regarded as endogenous, then the meaning is $\iiint g(y_{t+1}, x_{t+1}, \theta)\, p(x, y, \theta)\, dy\, dx\, d\theta = 0$. Eq. (43) of the application in Section 6 is an instance of the latter case. The examples in Section 5 are instances of the former.

[2] While this article was in press, methods for determining Ψ from primitives were proposed in Gallant (2016d).