Likelihood-based Estimation of Dynamic Equilibrium Models in the Frequency Domain: A MCMC Approach PRELIMINARY

Christian Matthes
NYU *

May 19, 2009

Abstract

This paper describes a Markov Chain Monte Carlo algorithm that can be used to perform likelihood-based inference in the frequency domain for linear Gaussian state space models. Certain frequencies of the data and the time series of observables implied by the model can be omitted, leading to estimates based only on the likelihood for non-omitted frequencies. The algorithm thus allows a model-consistent investigation of issues such as seasonality and low frequency movements in economic time series. A Monte Carlo study is carried out using a benchmark 'New Keynesian' dynamic equilibrium model of Negro & Schorfheide (2004), which is log-linearized to fit into the framework presented here.

*email: cm1518@nyu.edu. I would like to thank Francisco Barillas, Bill Dupor, Sydney Ludvigson and especially Tim Cogley, Thomas Sargent and Tao Zha for helpful comments

1 Introduction

This paper presents an extension the Metropolis-Hastings algorithm commonly used in macroeconomics to estimate linear(ized) dynamic equilibrium models. In contrast to the existing literature, the approach presented here is based on a frequency domain representation of the likelihood function and thus lends itself easily to study the influence of certain frequencies of the data on parameter estimates. Thus, this algorithm can be seen as an alternative to common detrending procedures such as the HP-filter. Certain frequencies of the data can explicitly be removed from the likelihood, thus making clear what frequencies are still used to draw inferences about the model at hand. Also, one can compare estimates from this algorithm to estimates obtained using the standard time domain method and detrended data. This can serve as a check whether or not inference based on detrended data is valid. Both Bayesian and Classical (Maximum Likelihood) inference can be carried out within this framework. Furthermore, one could use this approach to detect possible model misspecification by comparing parameter estimates using different frequency bands. As an illustration, I perform a Monte Carlo study using the model of Negro & Schorfheide (2004) as a laboratory. In this Monte Carlo study I investigate the effect of removing low frequencies from the likelihood on parameter estimates when the model is correctly specified.

2 Related Literature

Bayesian estimation of dynamic equilibrium models using the Metropolis-Hastings algorithm has become a standard approach in macroeconomics over the past decade. One of the early contributions in this literature has been Smets & Wouters (2003). The papers in this tradition usually restrict themselves to analysis in the time domain, though. Previous studies who have estimated dynamic macro models in the frequency domain are Christiano & Vigfusson (2003) and Diebold, Ohanian & Berkowitz (1998).

These papers do not use a Metropolis-Hastings algorithm though, and restrict themselves to Maximum Likelihood estimation. Using a Metropolis-Hastings algorithm leads to more robust inference since the entire shape of the posterior is traced out.

Hansen & Sargent (1993) investigate asymptotical properties of the frequency domain Maximum Likelihood estimator of a dynamic equilibrium model when there is possible misspecification at seasonal frequencies. Cogley (2001) uses the same approach to study the effect of detrending the series of observables before estimation and finds that estimation of models that are misspecified at low frequencies with detrended data leads to unsatisfactory estimates.

The frequency domain approximation to the Gaussian likelihood function was originally developed by Peter Whittle and is commonly used in econometrics, in particular in the study of long memory time series models, see for example Hurvich, Moulines & Soulier (2005).

3 The State Space Model

The following two sections follow Hansen & Sargent (2005). We are interested in estimating the matrices A^0, C, G, D and R in a model of the following form:

$$x_{t+1} = A^0 x_t + C w_{t+1} \tag{1}$$

$$y_t = G x_t + v_t \tag{2}$$

$$v_t = D v_{t-1} + \eta_t \tag{3}$$

where $w_t \sim_{iid} N(0, I)$, $\eta_t \sim_{iid} N(0, R)$, x_t is the vector of possibly unobserved states and y_t the vector of time t observables. The model is restricted to belong to the class of models for which $\{y_t\}$ is asymptotically stationary. For notational simplicity, the elements of the matrices A^0, C, G, D and R are collected in a n dimensional vector z.
The spectrum of this model can be calculated via the following formula:

$$S_y(\omega) = G(I - A^0 e^{-i\omega})^{-1} C C' (I - A^{0\prime} e^{+i\omega})^{-1} G' + (I - D e^{-i\omega})^{-1} R (I - D' e^{+i\omega})^{-1} \tag{4}$$

with $\omega \in \Re \cap [-\pi, \pi]$. While the spectrum characterizes second moments of the model, we allow for non-zero expectation of y_t[1]:

$$E(y_t) = \mu \tag{5}$$

4 Frequency Domain Representation of a Gaussian Likelihood Function

Given a vector $\{y_t\}_{t=1}^T$ of observations [2] we can calculate the fourier transform of frequency ω $y(\omega)$ and the periodigram $J_y(\omega)$

$$y(\omega) = \sum_{t=1}^{T} y_t e^{-i\omega t} \tag{6}$$

$$J_y(\omega) = \frac{1}{T} y(\omega) \overline{y(\omega)}' \tag{7}$$

These formulas now allow calculating a frequency domain approximation to the log likelihood function of the state space system given by 1 to 3. This approximate log likelihood is given by:[3]

[1] for the empirical part we will look at a state space system for demeaned variables and calculate the mean only when computing the likelihood function

[2] T is assumed to be even

[3] For a derivation see Harvey (2003)

$$\log(L(z|\{y_t\}_{t=1}^{T})) = -\frac{1}{2}(T+Tp)\log 2\pi - \sum_{j=1}^{T/2+1} \log\{\det S_y(\omega_j)\} - \sum_{j=1}^{T/2+1} trace[S_y(\omega_j)^{-1} J_y(\omega_j)] \tag{8}$$
$$-\frac{T}{2} trace\{S_y(0)^{-1}[T^{-1}\sum_{t=1}^{T} y_t - \mu][T^{-1}\sum_{t=1}^{T} y_t - \mu]'\}$$
$$\omega_j = \frac{2\pi j}{T}$$
$$j = 1, ..., T$$

Let $v_j \in \{0,1\} \forall j$ be an indicator function for certain frequencies. Then the contribution to the log likelihood of all frequencies for which $v_j = 1$ has the following form

$$\log(L(z|\{y_t\}_{t=1}^{T}, \{v_j\}_{j=1}^{T/2+1})) = -\frac{1}{2}(T+Tp)\log 2\pi - \sum_{j=1}^{T/2+1} v_j(\log\{\det S_y(\omega_j)\})$$
$$-\sum_{j=1}^{T/2+1} v_j trace[S_y(\omega_j)^{-1} J_y(\omega_j)] - \frac{T}{2} trace\{S_y(0)^{-1}[T^{-1}\sum_{t=1}^{T} y_t - \mu][T^{-1}\sum_{t=1}^{T} y_t - \mu]'\} \tag{9}$$

Note that even though this approximation does not necessarily integrate to 1 this is not crucial for the algorithm presented in the next section since that algorithm uses ratios of posteriors which includes the ratio of two likelihoods, i.e. any normalizing constant that would need to be computed for the likelihood to integrate to 1 will cancel.

5 The Algorithm

This section first describes the algorithm on a general level. The actual implementation is discussed later.
We combine the information coming from the likelihood $L(z|\{y_t\}_{t=1}^{T}, \{v_j\}_{j=1}^{T/2+1})$ with prior information summarized in a prior density $p(z)$ to form the posterior $f(z)$[4]:

$$f(z) \propto L(z|\{y_t\}_{t=1}^{T}, \{v_j\}_{j=1}^{T/2+1})p(z) \tag{10}$$

To form an estimate of the parameter vector we need to calculate an integral of a function involving the posterior [5]. In practice this integral has to be evaluated using Monte Carlo integration, which requires draws from the posterior. It is in general not possible to directly draw from the posterior. Markov Chain Monte

[4]Throughout this section, the dependence of $f(z)$ and $L(z|\{y_t\}_{t=1}^{T}, \{v_j\}_{j=1}^{T/2+1})$ on data and the frequencies included in the calculation of the likelihood is suppressed

[5]The actual form of the integral depends on what estimate one wants to report, which in turn depends on the loss function a researcher chooses

Carlo (MCMC) algorithms circumvent this problem by generating draws from a Markov Chain whose stationary distribution is the posterior.

To start the algorithm, we need to fix the length of the chain M, a starting value z_0 and a proposal density $q(\cdot,\cdot)$ generating new candidate draws for the algorithm. q is allowed to depend on the previous accepted draw from the posterior, i.e. $q(a,b) = q(a|b)$. The algorithm also gives an approximate Maximum Likelihood estimate $z*$. If M is large enough and the prior puts non-zero mass on the ML estimate then $z*$ will give a good approximation of the ML estimator. If one is only interested in ML estimation, one can further pick a uniform prior with large variance. Then the posterior will approximately resemble the likelihood function.
The probability α of accepting a draw from the proposal density is defined as:

$$\alpha(z,w) = \min\left\{\frac{f(w)\,q(w,z)}{f(z)\,q(z,w)}, 1\right\} \tag{11}$$

The algorithm is then given by the following 5 steps:

1. Initialize the algorithm with z_0 and M. Evaluate the likelihood at z_0 and set $z_0 = z*$ and $L(z_0) = L*$
2. Set $j = 1$.
3. Generate z_j^* from $q\left(z_{j-1}, z_j^*\right)$ and u from a uniform distribution $\mathcal{U}[0,1]$.
4. If $u \leq \alpha\left(z_{j-1}, z_j^*\right)$ then $z_j = z_j^*$, if $u > \alpha\left(z_{j-1}, z_j^*\right)$ then $z_j = z_{j-1}$. Note that this step involves evaluating the posterior and as part of that the likelihood. If $L(z_j^*) > L*$ then set $z* = z_j^*$ and $L* = L(z_j^*)$
5. If $j \leq M$ then $j \rightsquigarrow j+1$ and go to 3.

$\{z_j\}_{j=1}^M$ can be used to carry out Monte Carlo integration with respect to the posterior. For the Monte Carlo study below, I will pick a random walk candidate density $q(w,z)$ such that

$$w \sim N(z,H) \tag{12}$$

and H is picked to be the negative inverse Hessian at the mode of $\log f(\cdot)$.

6 A Monte Carlo Study

6.1 The Model of Del Negro and Schorfheide

In this section I give a short summary of the model of Negro & Schorfheide (2004).

6.1.1 The Representative Agent

The representative agent maximizes the following objective function by choosing $\{C_t, M_t, h_t, B_t\}_{t=0}^{\infty}$:

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{(C_t/A_t)^{1-\tau} - 1}{1-\tau} + \chi \log \frac{M_t}{P_t} - h_t \right] \tag{13}$$

where C_t is real consumption in period t, A_t is the level of technology at time t, P_t is the price level, M_t nominal money holdings and h_t hours worked. The period t budget constraint is given by:

$$C_t + \frac{B_t}{P_t} + \frac{M_t}{P_t} + \frac{T_t}{P_t} = W_t h_t + \frac{M_{t-1}}{P_t} + R_{t-1} \frac{B_{t-1}}{P_t} + D_t \tag{14}$$

where R_{t-1} is the return on one period nominal government bonds, T_t is a lump sum transfer from the government and D_t are dividends paid by the unit mass of monopolistically competitive firms in the economy.

6.1.2 Firms

Firms face the following demand function:

$$P_t(j) = P_t \left(\frac{X_t(j)}{X_t} \right)^{-1/\nu} \tag{15}$$

Each firm has access to a linear production technology:

$$X_t(j) = A_t h_t(j) \tag{16}$$

The evolution of the technology level A_t induces a stochastic trend into the variables of the model:

$$\log(A_t) = \log \gamma + \log(A_{t-1}) + \widetilde{z}_t \tag{17}$$

$$\widetilde{z}_t = \rho_z \widetilde{z}_{t-1} + \epsilon_{z,t} \tag{18}$$

Firm specific variables are indexed by j. Firms choose $h_t(j)$ and $P_t(j)$ to maximize

$$E_0 \left[\sum_{t=0}^{\infty} Q_t D_t(j) \right] \tag{19}$$

where current period dividends are given by:

$$D_t(j) = \frac{P_t(j)}{P_t} X_t(j) - W_t h_t(j) - \phi/2 \left(\frac{P_t(j)}{P_{t-1}(j)} - \pi \right)^2 X_t(j) \tag{20}$$

Dividend streams are discounted using a discount factor Q_t which is equal to the marginal utility of consumption of the representative agent. In this economy firms face a quadratic adjustment cost if they want to change their price by more or less than the economy wide rate of inflation π. Hours worked of the representative agent h_t is calculated by applying a CES aggregator function to firm specific hours worked $h_t(j)$.

6.1.3 Policy

Monetary policy in this model is conducted via the following interest rate rule:

$$\frac{R_t}{R*} = \left(\frac{R_{t-1}}{R*}\right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi*}\right)^{\psi_1} \left(\frac{X_t}{X_t*}\right)^{\psi_2}\right]^{1-\rho_R} \exp(\epsilon_{R,t}) \tag{21}$$

where $R*$ is an interest rate target. $\pi*$ an inflation target and X_t* potential output at time t, which is normalized to A_t. Once this policy rule is log linearized, it bears close resemblance to a standard Taylor rule.
The government consumes a fraction ζ_t of every good at time t. The law of motion of ζ_t is given by:

$$g_t = 1/(1 - \zeta_t) \tag{22}$$

$$\widetilde{g}_t = \log(g_t/g*) \tag{23}$$

$$\widetilde{g}_t = \rho_g \widetilde{g}_{t-1} + \epsilon_{g,t} \tag{24}$$

Fiscal policy is conducted subject to the government budget constraint:

$$\xi_t X_t + R_{t-1}\frac{B_{t-1}}{P_t} + \frac{M_{t-1}}{P_t} = \frac{T_t}{P_t} + \frac{M_t}{P_t} + \frac{B_t}{P_t} \tag{25}$$

6.1.4 The Log-Linearized System

First, define the log deviations of any variable Y from its trend value:

$$\widetilde{y}_t = \log Y_t - \log Y_t* \tag{26}$$

Then a log-linear approximation to the equilibrium conditions of the model is given by:

$$\widetilde{x}_t = E_t\widetilde{x}_{t+1} - \tau^{-1}(\widetilde{R}_t - E_t\widetilde{\pi}_t) + (1-\rho_g)\widetilde{g}_t + \rho_z\tau^{-1}\widetilde{z}_t \tag{27}$$

$$\widetilde{\pi}_t = \beta E_t\widetilde{\pi}_{t+1} + \kappa[\widetilde{x}_t - \widetilde{g}_t] \tag{28}$$

$$\widetilde{R}_t = \rho_R\widetilde{R}_{t-1} + (1-\rho_R)(\psi_1\widetilde{\pi}_t + \psi_2\widetilde{x}_t) + \epsilon_{R,t} \tag{29}$$

The standard deviations of $\epsilon_{z,t}$, $\epsilon_{g,t}$ and $\epsilon_{R,t}$ will be denoted σ_z, σ_g and σ_R.
The system of equations given by (27) to (29) can be solved numerically via a number of available algorithms to obtain a law of motion for the state variables of the system.
For this application, I have solved the system using the Gensys algorithm (Sims 2002). The law of motion for the states will form equation(1) of our state space system. The mapping from states to observables (i.e. equation(2)) is given by:

$$\Delta \log X_t = \log\gamma + \Delta\widetilde{x}_t + \widetilde{z}_t \tag{30}$$

$$\Delta \log P_t = \log \pi * + \widetilde{\pi}_t \tag{31}$$

$$\log R_t^a = 4[\log(\gamma\beta^{-1}) + \log \pi * + \widetilde{R}_t] \tag{32}$$

6.2 Simulated Data

I have simulated 200 observations of quarterly data from the model for the following parameter values [6]

parameter	value
$100 \log \pi*$	1
$100 \log \gamma$	0.4
$100 \log r*$	0.5
κ	0.3
τ	2
ϕ_1	1.5
ϕ_2	0.125
ρ_R	0.5
ρ_g	0.8
ρ_z	0.3
$100\sigma_z$	0.875
$100\sigma_g$	0.63
$100\sigma_R$	0.251

Figure 1 shows the simulated time series. The log interest rate is scaled by 100.

6.3 Prior Distributions

For the estimation, I assume that the priors for each of the 13 parameters are independently distributed. The parameter values are chosen such that the prior distributions are centered at the true values.

[6] Just as in Negro & Schorfheide (2004) the first three parameters and the standard deviations are scaled by 100 to convert them into percentages

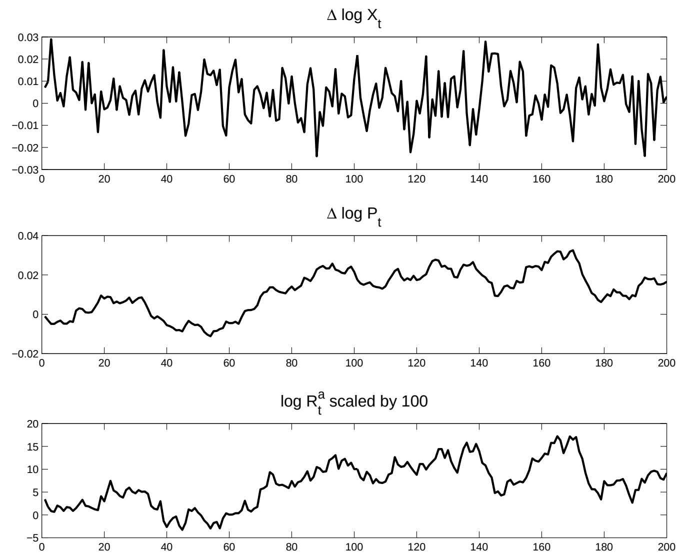

Figure 1: simulated data

parameter	distribution	parameter 1	parameter 2
$100 \log \pi*$	normal	0.5	0.25
$100 \log \gamma$	normal	1	0.5
$100 \log r*$	gamma	4	0.125
κ	gamma	4	0.075
τ	gamma	16	0.125
ϕ_1	gamma	36	0.04
ϕ_2	gamma	1.5625	0.08
ρ_R	beta	1	1
ρ_g	beta	1	1
ρ_z	beta	1	1
σ_z	inverse gamma	40	0.098
σ_g	inverse gamma	40	0.05
σ_R	inverse gamma	40	0.008

For the normal distributions, parameter 1 is the mean and parameter 2 the standard deviation. The values are mostly taken from Negro & Schorfheide (2004).

6.4 Estimation Results

The following part gives posterior estimates for the cases of

- no frequencies omitted

- frequencies lower than 0.31 omitted (corresponding to cycles of one every five years or longer)

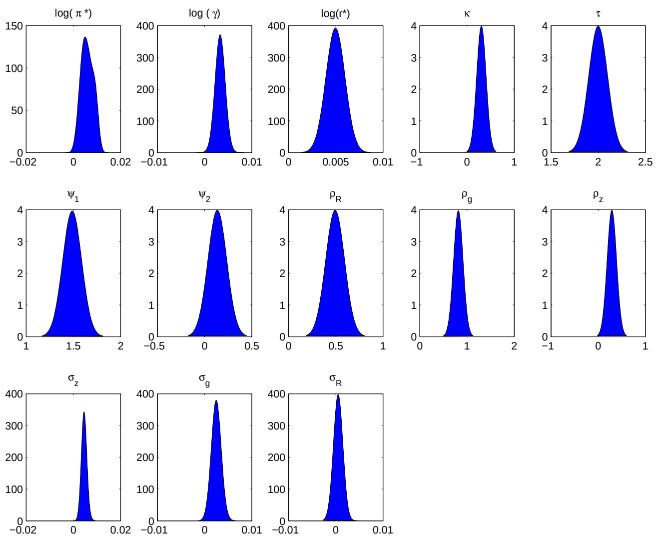

Figure 2: smoothed marginal posteriors, no frequencies omitted

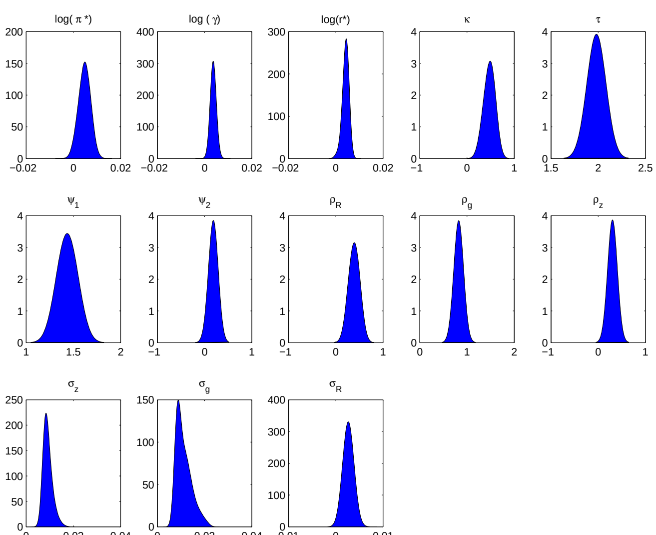

Figure 3: smoothed marginal posteriors, frequencies lower than 0.31 omitted

6.4.1 Discussion

As one can see from the pictures, dropping frequencies lower than business cycle frequencies, i.e. frequencies lower than 0.31, does not lead to substantially different posteriors, even though they are a bit wider. Further it should be noted that estimates for the case where no frequencies are dropped is naturally very close to the posterior estimates coming from a time domain estimation using the Kalman filter.
Dropping further frequencies, though, leads to substantially increased uncertainty (results not reported here). In general, one can see that while dropping low frequencies from the likelihood increases uncertainty, it does not seem to lead to biases in the means of the marginal posteriors.

7 Conclusion

This paper is proposing a variant of the Metropolis-Hastings based on the Whittle frequency domain approximation to the Gaussian likelihood function. The algorithm can be applied to issues such as checking the validity of inference based on detrended data, checking for model misspecification, seasonal adjustment and modelling of low frequency behavior of economic time series. These are issues that will be addressed in future research. Since this approach nests (an approximation to) the time domain estimation using the Kalman filter while not requiring any kind of filtering, it should be useful also for researchers who are interested in increasing the speed of their estimation routine (at the cost of using an approximation to the likelihood function).

References

Christiano, L. J. & Vigfusson, R. J. (2003), 'Maximum likelihood in the frequency domain: the importance of time-to-plan', *Journal of Monetary Economics* **50**(4), 789–815.

Cogley, T. (2001), 'Estimating and testing rational expectations models when the trend specification is uncertain', *Journal of Economic Dynamics and Control* **25**(10), 1485–1525.

Diebold, F. X., Ohanian, L. E. & Berkowitz, J. (1998), 'Dynamic equilibrium economies: A framework for comparing models and data', *Review of Economic Studies* **65**(3), 433–51.

Hansen, L. P. & Sargent, T. J. (1993), 'Seasonality and approximation errors in rational expectations models', *Journal of Econometrics* **55**(1-2), 21–55.

Hansen, L.-P. & Sargent, T. J. (2005), *Recursive Models of Dynamic Linear Economies*, Manuscript, NYU.

Harvey, A. (2003), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press.

Hurvich, C. M., Moulines, E. & Soulier, P. (2005), 'Estimating long memory in volatility', *Econometrica* **73**(4), 1283–1328.

Negro, M. D. & Schorfheide, F. (2004), 'Priors from general equilibrium models for vars', *International Economic Review* **45**(2), 643–673.

Sims, C. A. (2002), 'Solving linear rational expectations models', *Computational Economics* **20**(1-2), 1–20.

Smets, F. & Wouters, R. (2003), 'An estimated dynamic stochastic general equilibrium model of the euro area', *Journal of the European Economic Association* **1**(5), 1123–1175.