Bayesian estimation of DSGE models using time series data

Michael Nattinger

November, 2021

Outline

General Model

Consider a solved (e.g. by Dynare), linearized model $\mathcal{M}_\theta$ with parameterization θ. Given states $S_t = (x_t, z_t)'$ (endogenous and exogenous), $\mathcal{M}_\theta$ provides a potentially stochastic linear law of motion for the state variables:

$$S_t = A(\theta)S_{t-1} + Bv_t \tag{Transition}$$

For a variety of reasons, we may want to estimate the model. Let Y_t be a vector of observables. We observe:

$$Y_t = DS_t + w_t, \tag{Measurement}$$

where D is typically a matrix defined such that DS_t will be a vector of model variables (e.g. $(q_t, k_t)'$) (and therefore does not depend on θ). We typically assume multivariate normal errors,

$$\begin{pmatrix} v_t \\ w_t \end{pmatrix} \sim_{iid} N\left(0, \begin{bmatrix} Q(\theta) & 0 \\ 0 & R \end{bmatrix}\right).$$

Estimating θ

- In general, we do not observe all states or shocks.
- However, given (Transition), (Measurement) equations we can apply the 'Kalman filter' and estimate $\{\hat{v}_t, \hat{w}_t, \hat{S}_t\}$.
- KF also yields a likelihood, $L(Y|\theta)$, since we have estimates of the shocks and know their distributions.
- Can we use ML to estimate θ? Yes, but also no...
  - Time series are usually short (50-500 observations)
  - Time series data has peculiarities that make proper frequentist econometrics hard (autocorrelated errors etc.)
  - Bayesian macroeconometric models typically perform much better out-of-sample than frequentist

Bayesian Posterior

What is Bayesian econometrics? Frequentist but with a *prior* belief about the parameters $\pi(\theta)$.

- Bayes' rule: Posterior $\pi(\theta|Y) = \frac{L(Y|\theta)\pi(\theta)}{\int L(Y|\theta)\pi(\theta)d\theta}$
- Numerator: Great! Can compute $L(Y|\theta)$ for any θ, and $\pi(\theta)$ is our belief (so is known by definition).
- Denominator: ??? No analytical form in most cases. We want to ignore it.
- Solution: note that $\frac{\pi(\theta_1|Y)}{\pi(\theta_2|Y)} = \frac{L(Y|\theta_1)\pi(\theta_1)}{L(Y|\theta_2)\pi(\theta_2)}$.
    - Posterior ratio does not depend on $\int L(Y|\theta)\pi(\theta)d\theta$!!!

Metropolis-Hastings Algorithm

- We want to sample from the posterior distribution $\pi(\theta|Y)$ so we can find objects such as $\int \theta\pi(\theta|Y)d\theta$, and CIs.
- We do not know the posterior distribution because we can't calculate the pesky denominator
- Instead we use Markov Chain Monte Carlo (MCMC) method:
  - Compute posterior mode using numerical optimization routine $\theta_m = \arg\max_\theta \pi(\theta|Y) = \arg\max_\theta L(Y|\theta)\pi(\theta)$
  - Initialize $\theta^1 = \theta_m$. Define $\Sigma_m = (\mathcal{H}(L(Y|\theta)\pi(\theta))|_{\theta=\theta_m})^{-1}$.
  - Given θ^i, draw 'candidate' $\theta' \sim N(\theta^i, c\Sigma_m)$
  - Compute $w \sim U(0, 1)$.
  - Set $\theta^{i+1} = \begin{cases} \theta' \text{ if } \frac{\pi(\theta'|Y)}{\pi(\theta^i|Y)} > w \\ \theta^i \text{ otherwise} \end{cases}$.
- Resulting $\{\theta^i\}$ draws are weighted as if they were drawn from $\pi(\theta|Y)$.

Pros and Cons

Some pros:

- Works better than alternatives
- Dynare can do this automatically $\Rightarrow$ easy to implement
- Unobserved states at a particular point in time are estimated so can conduct counterfactuals, etc.

Some cons:

- Kalman filter requires linearized model
- Some people (especially at UW) dislike Bayesian priors
- Random walk MCMC needs lots of draws to properly sample distribution (often over a million)

Shocks & Measurement errors

- In general, need to write down model such that
  # observables $\leq$ # shocks
- More observables than shocks results in a likelihood of *zero*
  - Shocks (in model + measurement) account for unexpected movements in real-world variables
  - As many shocks as observables $\iff$ exact identification (shocks exactly account for unexpected changes in observables)
  - More observables than shocks: Model cannot account simultaneously for all movements in observables, likelihood is zero
- Can we have more shocks than unobservables? Yes. This method can handle it (without needing any extensions).

Measurement errors cont.

- In practice econometricians often apply measurement error onto every observable (analogous to residual)
    - Is this good practice? Almost feels like 'cheating', we want the model to predict reality well, not the measurement errors.
    - On the other hand, have you ever seen an OLS regression have no errors? No.
    - In some sense, maybe not the worst idea...
- Some nonlinear filters actually *require* measurement errors on every observable.
- In practice: usually write model and use as many observables in Y as shocks in model
    - Then sometimes add measurement error on each observable according to practitioner's preference/methodological necessity.