LEARNING AND NONLINEAR FILTERING IN ECONOMETRICS

Lars Peter Hansen   Nick Polson   Thomas J. Sargent

University of Chicago and New York University

April 23, 2010

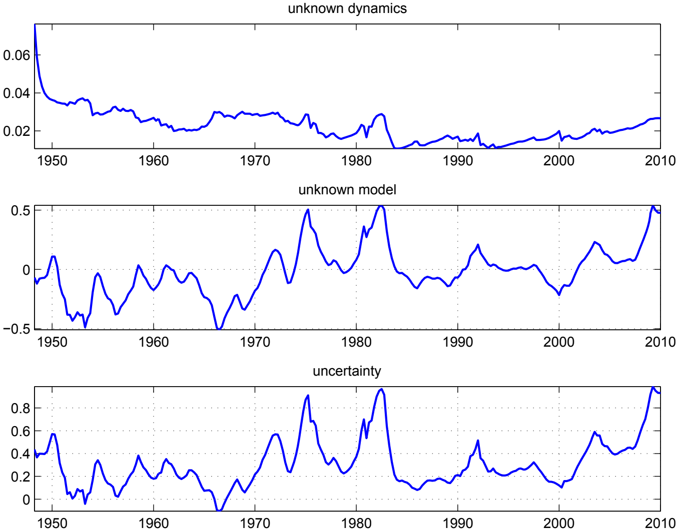
Econometric agents as econometricians

Explore the role of learning dynamics when

- learning is challenging
- economic agents are skeptical about model - expressed as ambiguity aversion or a concern about robustness

We aim to explore when learning on the part of economic agents is challenging even in the presence of substantial data histories and how these challenges contribute to dynamic evolution of economic variables. Hansen (Ely, AER, 2007), Sargent (Presidential address, AER, 2008).

Previous work



Previous applications: Limited to use of quasi-analytical filtering methods such as Kalman filtering or Wonham filtering.

This paper: We explore ways to improve particle filtering methods to use in models that entail agent-based learning.

Recall: Particle filtering methods are numerical, monte carlo-based techniques used to approximate Bayesian solutions to filtering and estimation problems given signal histories.

Why particle filtering?

Quasi-analytical recursive algorithms such as Kalman filtering (linear state and measurement equations) and Wonham filtering (discrete states) have limited applicability.

- Often underlying parameters are unknown.
- Alternative to discrete-state approximation.

Particle filtering methods become attractive alternatives to applications of quasi-analytical filtering methods.

What improvements?

- Use "sufficient statistics" when learning includes invariant parameters.
  - "Sufficient statistics" have recursive representations.
  - "Sufficient statistics" can depend on hidden states and even a subset of parameters.
- The decision problem directs the attention of the numerical approximation.
  - Investors care about some "tails" of distributions because of risk aversion.
  - Investors care where misspecification has the most pronounced consequences in terms of utility or continuation values when they are concerned about robustness.

Basic setup

We let X_t denote the underlying Markov state at date t and Y_t a vector of current period signals.

$$X_t = \begin{bmatrix} D_t \\ Z_t \end{bmatrix}$$

where D_t is observable to the decision maker and Z_t is not.

- Hidden state evolution:

  $$\Gamma(dz^*|y^*, x, \theta) = \gamma(z^*|y^*, x, \theta)\lambda(dz^*)$$

  which conditions on the current period signal.
- Invariant parameter: θ is an unknown parameter with "prior" $\pi(d\theta)$.
- Signal evolution:

  $$\nu(y^*|x, \theta)\eta(dy^*).$$

Recursive form of Bayes rule

- Hidden state evolution:

$$\Gamma(dz^*|y^*, x, \theta) = \gamma(z^*|y^*, x, \theta)\lambda(dz^*)$$

- Signal evolution:

$$\nu(y^*|x, \theta)\eta(dy^*).$$

Compute the *filtering distribution* of interest with next period density (relative to $\lambda(dx)\pi(d\theta)$):

$$q_{t+1}(z^*, \theta) \propto \int \gamma(z^*|Y_{t+1}, \theta)\nu(Y_{t+1}|D_t, z, \theta)q_t(z, \theta)\lambda(dz)$$

Then the next-period signal density given the signal history is:

$$\ell_t(y^*) = \int\int \nu(y^*|D_t, z, \theta)q_t(z, \theta)\lambda(dz)\pi(d\theta).$$

Replace hidden state/parameter (Z_t, θ) by density q_t.
Distribution associated with q_{t+1} is the target of the numerical method.

Sufficient statistics I

Consider a state vector S_t that is constructed recursively given an initial condition S_0 via:

$$S_{t+1} = \Phi(\theta, Y_{t+1}, Z_{t+1}, X_t, S_t).$$

This construction will be valuable in simulation provided that:

Assumption

The distribution of θ conditioned on S_t, Z_t, and the signal history $\mathcal{Y}_t$ satisfies

$$\rho_t(d\theta | S_t, Z_t, \mathcal{Y}_t) = \psi(d\theta | S_t).$$

for some prespecified ψ.

Sufficient statistics II

Special case i:

$$S_t = \theta$$

Special case ii: There exists a vector S_t such that

$$\rho_t(d\theta | S_t, Z_t, \mathcal{Y}_t) = \psi(d\theta | S_t)$$

where the statistic S has a recursive representation:

$$S_{t+1} = \Phi(Y_{t+1}, Z_{t+1}, X_t, S_t)$$

for $t = 0, 1, ...$ for some choice of S_0.

Sufficient statistics depend on hidden states. Storvik, Fearnhead, and Johannes-Polson. Kalman filtering with unknown parameters.

Intermediate case: "Sufficient statistics" for a subset of parameter values given the remaining subset. Dynamic factor models.

Alternative Markov law

Under a new Markov law, the state vector evolves as follows:

$$
\begin{aligned}
\tilde{\theta}_{t+1} &\sim \psi(\cdot|\tilde{S}_t) \\
\tilde{Y}_{t+1} &\sim \nu\left(\cdot|\tilde{X}_t, \tilde{\theta}_{t+1}\right) \\
\tilde{Z}_{t+1} &\sim \Gamma(\cdot|\tilde{Y}_{t+1}, \tilde{X}_t, \tilde{\theta}_{t+1}) \\
\tilde{S}_{t+1} &= \Phi\left(\tilde{\theta}_{t+1}, \tilde{Y}_{t+1}, \tilde{Z}_{t+1}, \tilde{X}_t, \tilde{S}_t\right).
\end{aligned}
$$

Parameter vector θ evolves over time as a device to replenish particles.

Algorithm

At date t there are N particles where a particle is specified as $(\tilde{Z}_t^{[i]}, \tilde{S}_t^{[i]})$.

1. Draw $\tilde{\theta}_{t+1}^{[i]}$ from $\psi\left(\theta|\tilde{S}_t^{[i]}\right)$
2. Construct weights

$$w_{t+1}^{[i]} = \frac{\nu\left(Y_{t+1}|D_t, \tilde{Z}_t^{[i]}, \tilde{\theta}_{t+1}^{[i]}\right)}{\sum_i^N \nu\left(Y_{t+1}|D_t, \tilde{Z}_t^{[i]}, \tilde{\theta}_{t+1}^{[i]}\right)},$$

   and draw $(\tilde{Z}_t^{[i]}, \tilde{S}_t^{[i]}, \tilde{\theta}_{t+1}^{[i]})$ from a multinomial distribution with probability $w_{t+1}^{[i]}$.
3. Draw $\tilde{Z}_{t+1}^{[i]}$ using density $\gamma\left(z^*|Y_{t+1}, D_t, Z_t^{[i]}, \tilde{\theta}_{t+1}^{[i]}\right)$.
4. Construct $\tilde{S}_{t+1} = \Phi\left(\tilde{S}_{t+1}^{[i]}, Y_{t+1}, Z_{t+1}^{[i]}, D_t, \tilde{X}_t^{[i]}, \tilde{\theta}_{t+1}^{[i]}\right)$.
5. Replace particle $(\tilde{Z}_t^{[i]}, \tilde{S}_t^{[i]})$ with $(\tilde{Z}_{t+1}^{[i]}, \tilde{S}_{t+1}^{[i]})$.

Partial smoothing

When feasible, it is advantageous to at least partially smooth our estimates of the states prior to simulation.
Construct a new state vector

$$\hat{X}_t \doteq \begin{bmatrix} Y_t \\ X_{t-1} \end{bmatrix}.$$

The unobservable component for the new state vector is Z_{t-1} whereas it previously was Z_t. The signal vector remains the same.

Build the distribution of the signal conditioned on the newly constructed state:

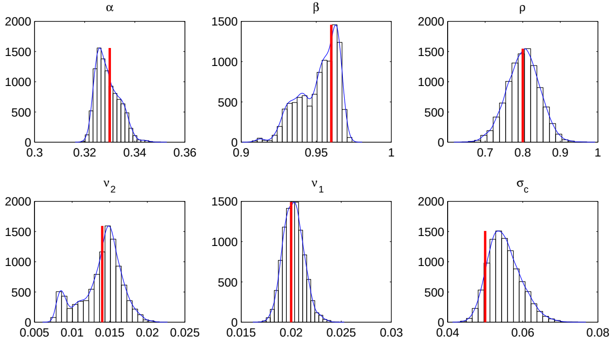
$$\hat{\eta}(y^* | \hat{X}_t, \theta) = \int \eta(y^* | D_t, z, \theta) \gamma(z | Y_t, X_{t-1}, \theta) \lambda(dz).$$

The new state evolution conditioned on the signal Y_{t+1} is

$$\hat{\gamma}(z | Y_{t+1}, \hat{X}_t, \theta) = \frac{\eta(Y_{t+1} | D_t, z, \theta) \gamma(z | Y_t, X_{t-1}, \theta)}{\int \eta(Y_{t+1} | D_t, \tilde{z}, \theta) \gamma(\tilde{z} | Y_t, X_{t-1}, \theta) \lambda(d\tilde{z})}.$$

Simple version of Brock-Mirman stochastic growth model

Restricted Kalman filtering model obtained as a solution to a stochastic growth model with logarithmic preferences. Key parameters: $0 < \alpha < 1$ governs Cobb-Douglas production function, $0 < \beta < 1$ is the subjective discount factor and $0 < \rho < 1$ is the AR parameter for the technology shock.



STOCHASTIC GROWTH MODEL CONTINUED

Learning dynamics for the parameters

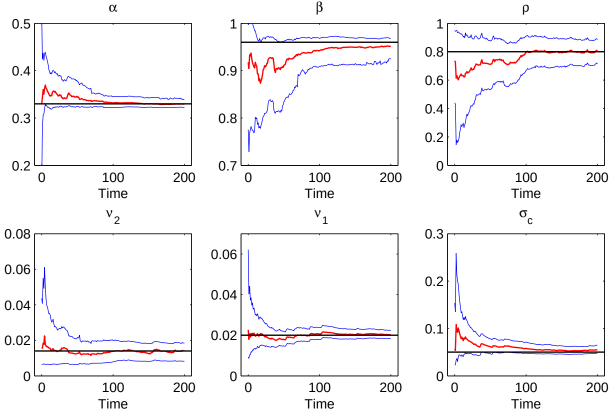


FIGURE: Posterior distribution dynamics, median, .05 and .95 quantiles.

Learning, robustness and volatility

The consumption growth rate: $\log C_{t+1} - \log C_t$ is one the signals, and the composite system of signals and states evolve as:

$$
\begin{aligned}
Y_{t+1} &= F + UX_{t,1} + Z_{t,2} G W_{t+1} \\
X_{t+1,1} &= A_1 X_{t,1} + Z_{t,2} B_1 W_{t+1} \\
Z_{t+1,2} &= (1 - A_2)\mu_2 + A_2 Z_{t,2} + B_2 W_{t+1}
\end{aligned}
$$

where the W_{t+1} is a composite shock vector for the entire system and is distributed as a multivariate standard normal and

$$
X_{t,1} \doteq \begin{bmatrix} D_t \\ Z_{t,1} \end{bmatrix}.
$$

$Z_{t,1}$ hidden growth state and $Z_{t,2}$ is a hidden volatility state where $|Z_{t,2}|$ is a measure of volatility.

ROBUSTNESS AND ESTIMATION

Observations:

- Models with long-run macroeconomic risk models like those of Bansal and Yaron (JF) introduce hard to detect growth components in the macroeconomic dynamics. Question: Where does the investor confidence come from?
- We study environments in which investors struggle in making inferences about the underlying growth the in economy. Apply robust decision theory in which investors treat models as approximations and engage in robust estimation.
- Made tractable through the application of exponential tilting of distributions that are important to the decision-maker.
- Previous research used the Wonham filter (Hansen, AER) or the Kalman filter augmented to accommodate discrete model selection (Hansen-Sargent, Fragile beliefs). No stochastic volatility as in the model that I just presented.

How do we implement robustness?

Use a recursive formulation given in Hansen-Sargent (JET).

- Compute continuation values conditioned on hidden states and parameters. The continuation values are quadratic in our example economy and computed by solving a Riccati equation.
- Use continuation values to determine directions of misspecification that cause the most concern for the investors.
- Implement this through the use of “relative entropy penalization” that results in exponential tilting of distributions based on the continuation values.
- The distorted distributions have implications for equilibrium asset prices.

Modify filtering methods based on continuation values

- Approximate the exponentially tilted distributions.
- Similar to:
    - ‘Risk sensitive particle filtering’ of Thurn-Langford-Verma. Use the risk function from statistical decision theory to redirect the particle approximation.
    - Donsker-Veradhan theory of Large Deviations applied to Markov processes. Characterize large deviation behavior characterized via a distorted distribution that emerges from an optimization problem.
    - Rare event simulation methods discussed by Bucklew.

Use investors decision problem to determine where to focus the numerical accuracy of the particle filter.

Distort state and signal evolution and hence the filtered distributions using methods in Hansen and Scheinkman- Econometrica.

HIDDEN STATE DENSITIES FOR MACRO GROWTH PROCESS

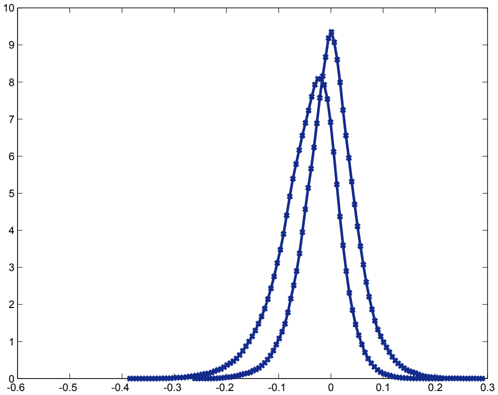


FIGURE: Original and (modestly) distorted densities

Hidden state densities for macro growth process

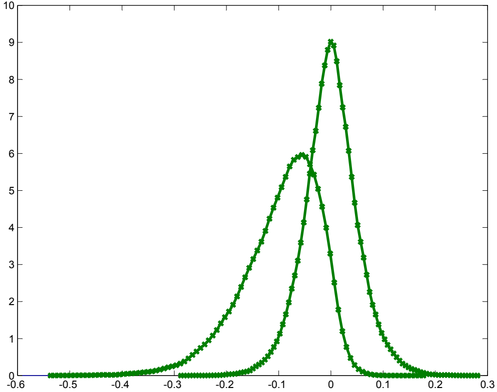


FIGURE: Original and (substantially) distorted densities

A CONVENIENT DISTORTION

Follow Hansen-Schienkman (Econometrica) by using a positive eigenfunction to build a change in probability measure that preserves the Markov structure.

Solve

$$\exp(\delta)\hat{h}(x,\theta) = E\left[\hat{h}(X_{t+1},\theta)\exp\left[-\frac{1}{\xi}(\log C_{t+1} - \log C_t)\right]|X_t = x, \theta\right]$$

where $\hat{\upsilon}(x,\theta) + \log c$ is the value function and

$$\hat{h}(x,\theta) = \exp\left[-\frac{1}{\xi}\hat{\upsilon}(x,\theta)\right].$$

The positive random variable

$$\exp(-\delta)\frac{\hat{h}(X_{t+1},\theta)}{\hat{h}(X_t,\theta)}\exp\left[-\frac{1}{\xi}(\log C_{t+1} - \log C_t)\right].$$

is the Randon-Nikodym derivative for the changing the transition law for the Markov process.

FILTER DISTORTION

PROPOSITION

If the joint distorted prior for Z_0, θ is proportional to $\hat{h}(D_0, z, \theta) q_0(z, \theta) \lambda(dz) \pi(d\theta)$, then

$$\exp(\delta t) \hat{q}_t(z, \theta) \propto \hat{h}(D_t, z, \theta) q_t(z, \theta)$$

for all $t \geq 0$ where δ depends on θ and the constant of proportionality depends only on the signal history.

$\hat{q}_t$ and q_t are central ingredients in characterizing equilibrium outcomes.

Conclusions

We explore modifications of particle filtering that are needed to study learning dynamics in economic models outside the realm of quasi-analytical filtering methods.

- Use "sufficient statistics" can depend on unknown states and parameters.
- Use the decision problem of investors to direct the numerical approximation.