Modelling Information, Learning and Expectations in Macroeconomics

Lecture 13

New York University

Spring 2008

Information, Learning and Expectations in Macroeconomics

Today:

Review:

- I Projections, their properties and the Kalman filter
- II Models of imperfect information
- III Recursive Least Squares learning

Unresolved issues/research wish list

The Projection Theorem

If $\mathcal{M}$ is a closed subspace of the Hilbert Space $\mathcal{H}$ and x$\in \mathcal{H}$, then
(i) there is a unique element $\widehat{x} \in \mathcal{M}$ such that

$$\|x - \widehat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\|$$

and
(ii) $\widehat{x} \in \mathcal{M}$ and $\|x - \widehat{x}\| = \inf_{y \in \mathcal{M}} \|x - y\|$ if and only if $\widehat{x} \in \mathcal{M}$ and $(x - \widehat{x}) \in \mathcal{M}^{\perp}$ where $\mathcal{M}^{\perp}$ is the orthogonal complement to $\mathcal{M}$ in $\mathcal{H}$.
The element $\widehat{x}$ is called the orthogonal projection of x onto $\mathcal{M}$.

Properties of projections

Let $\mathcal{H}$ be a Hilbert space and and let $P_{\mathcal{M}}$ be a projection mapping onto a closed subspace $\mathcal{M}$. Then
(i) each $x \in \mathcal{H}$ has a unique representation as a sum of an element in $\mathcal{M}$ and an element in $\mathcal{M}^{\perp}$, i.e.

$$x = P_{\mathcal{M}}x + (I - P_{\mathcal{M}})x$$

(ii) $x \in \mathcal{M}$ if and only if $P_{\mathcal{M}}x = x$
(iii) $x \in \mathcal{M}^{\perp}$ if and only if $P_{\mathcal{M}}x = 0$
(iv) $\mathcal{M}_1 \subseteq \mathcal{M}_2$ if and only if $P_{\mathcal{M}1}P_{\mathcal{M}2}x = P_{\mathcal{M}1}x$
(v) $\|x\|^2 = \|P_{\mathcal{M}}x\|^2 + \|(I - P_{\mathcal{M}})x\|^2$

Why is this useful?

(i) We can decompose actual outcomes of inflation and output as a sum of two orthogonal components: The projection and an orthogonal error, which implies that agents do not make systematic mistakes when forming expectations (given the specified information set).
(ii) State will be revealed if and only if the space spanned by observables, that is, $\mathcal{M}$, also span the state.
(iii) Actual state (or future outcomes) are orthogonal to the space spanned by observables if and only if projections onto space spanned by observables are zero.
(iv) If you let $\mathcal{M}_1$ be the history of observable variables up to time 1, $\mathcal{M}_2$ the history of observable variables up to time 2, then property (iv) is the law of iterated expectations.
(v) The variance of the actual state (or future outcomes) can be written as a sum of the variance of the projections and the variance of the orthogonal errors.

Using orthogonality to derive projection formula

Inner product definition

$$\langle Y, X \rangle = E\left(YX'\right)$$

Projection of Y on X given by

$$\mathcal{P}_X Y = AX$$

such that

$$\begin{aligned} \left\langle Y - AX, X' \right\rangle &= 0 \\ E\left(YX' - AXX'\right) &= 0 \end{aligned}$$

solving orthogonality condition gives

$$A = E\left[YX'\right]\left(E\left[XX'\right]\right)^{-1}$$

Projections onto uncorrelated variables

Let $\left\{\widetilde{Z}_j\right\}_{j=1}^{t}$ denote a sequence of uncorrelated mean zero variables so that

$$E\left[\widetilde{Z}_t\widetilde{Z}_{t-s}\right] = 0 : s \neq 0$$

then

$$\mathcal{P}\left(X_t|\widetilde{Z}_t, \widetilde{Z}_{t-1}, ...\widetilde{Z}_1\right) = \mathcal{P}\left(X_t|\widetilde{Z}_t\right)+\mathcal{P}\left(X_t|\widetilde{Z}_{t-1}\right)+...+\mathcal{P}\left(X_t|\widetilde{Z}_1\right)$$

Gram-Schmidt orthogonolization

Let the matrix $Y \ \ (m \times n)$ have columns $\mathbf{y}_1, \mathbf{y}_2, ...\mathbf{y}_n$.

$$Y = \left[\begin{array}{cccc} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_n \end{array} \right]$$

We want to construct a new matrix $\widetilde{Y}$ with the same column space as Y that has orthogonal columns so that $\widetilde{Y}\widetilde{Y}'$ is a diagonal matrix. Keep the first column as it can be chosen arbitrarily. Define the new second column $\widetilde{\mathbf{y}}_2$ by subtracting the projection of $\mathbf{y}_2$ on $\mathbf{y}_1$ from $\mathbf{y}_2$

$$\widetilde{\mathbf{y}}_2 = \mathbf{y}_2 - \mathbf{y}_1 \left(\mathbf{y}_1' \mathbf{y}_1 \right)^{-1} \mathbf{y}_1' \mathbf{y}_2$$

or

$$\widetilde{\mathbf{y}}_2 = \left(I - \mathcal{P}_{y_1} \right) \mathbf{y}_2$$

and then subtract the projection of $\mathbf{y}_3$ on $[\mathbf{y}_1 \ \ \mathbf{y}_2]$ from $\mathbf{y}_3$ to construct $\widetilde{\mathbf{y}}_3$ and so on.

The Kalman filter

$$\begin{aligned} X_t &= A_{t-1}X_{t-1} + C_t\mathbf{u}_t \\ Z_t &= D_tX_t + \mathbf{v}_t \end{aligned}$$

Prior $X_{0|0}$ with variance $P_{0|0}$ defined as

$$E\left(X_{0|0} - X_0\right)\left(X_{0|0} - X_0\right)' = P_{0|0}$$

and we further assumed that $X_{0|0}$ is uncorrelated with the shock processes $\{\mathbf{u}_t\}$ and $\{\mathbf{v}_t\}$. Find the K_t in

$$X_{t|t} = A_tX_{t-1|t-1} + K_t\left(Z_t - D_tX_{t|t-1}\right)$$

so that state estimation error is minimised.

The Kalman filter formulas

The prior error covariance matrix update equation

$$\begin{aligned} P_{t|t-1} &= A_{t-1}(P_{t-1|t-2} - \\ &\quad P_{t-1|t-2}D'_{t-1}\left(D_{t-1}P_{t-1|t-2}D'_{t-1} + \Sigma_{vv,t-1}\right)^{-1} \times \\ &\quad D_{t-1}P_{t-1|t-2})A'_{t-1} \\ &\quad + C_t C'_t \end{aligned}$$

The Kalman gain

$$K_t = P_{t|t-1}D'_t\left(D_t P_{t|t-1}D'_t + \Sigma_{vv,t}\right)^{-1}$$

K_t and $P_{t|t-1}$ does not depend on observations and $P_{t|t-1}$ is bounded from below and above (for stationary models)

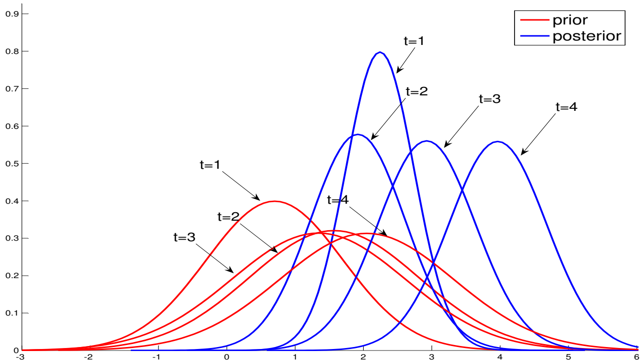

Figure: Propagation of prior and posterior distributions

Kalman filter representations

The Kalman filter can be thought of as adding several projections onto orthogonal observations.
True filtered estimates representation

$$X_{t|t} = A_t X_{t-1|t-1} + K_t \left(Z_t - D_t X_{t|t-1} \right)$$

and the innovation representation

$$\begin{aligned} \widetilde{Z}_t &= Z_t - D_t X_{t|t-1} \\ X_{t+1|t} &= A_{t-1} X_{t|t-1} + A_{t-1} K_t \widetilde{Z}_t \end{aligned}$$

are equivalent, but used for different purposes.

Part II: Models of imperfect information

Asymmetric information: One agents' information set subset of other agents information set, state is easy to define.

- LQ Optimal Control application: Central bank extracts information from bond market

Two tricks to remember:

1. Signal extraction from endogenous variables

$$X_{t|t} = A_t X_{t-1|t-1} + K_t \left(Z_t - D_t X_{t|t-1} - F_t X_{t|t} \right)$$

   Keep the "surprise" interpretation of term in brackets

2. Set up fixed point problem by using that the model must nest full information solution if $X_{t|t} = X_t$ by coincidence.

Models with private information

Private information and strategic interaction

- Every agent has his own "window to the world" but no agent is better informed than others
- Individual pay offs depend on (average) action taken by others.

Framework can explain:

- Over-reactions to public information (Morris and Shin AER 2002)
- Macroeconomic inertia: Woodford (2002) and Nimark JME (2008)

The principal modeling difficulty: The infinite regress of "forecasting the forecasts of others" (Townsend 1983) and no natural finite dimensional state

Strategies for solving private information models

- Lagged revelation of shocks and projection methods: Townsend (1983), Singleton (1987)
  - Not always realistic and can result in weird (kinked) IRF (Bacchetta and Van Wincoop (2006))
- Finite horizon: Allen, Morris and Shin (2006)
- Static choices: Woodford (2002), Morris and Shin (2002)
- Dynamic choices: Nimark (2007)
  - Use common knowledge of rationality and decreasing impact of expectations as order increases

Part III: Learning

Boundedly rational agents behave as econometricians when they form expectations

- Early work motivated by justifying rational expectations equilibria
- Can be an equilibrium selection device if there is a multiplicity of equilibria
- Can provide (yet another) explanation of:
  - macro-economic inertia
  - predictable stock market excess returns

Learning the language of learning

Two concepts:

- Perceived Law of Motion (PLM)
  - A parameterised equation describing how agents form expectations
- Actual Law of Motion (ALM)
  - An equation describing how the endogenous variables actually evolve
  - The ALM is a function of the PLM
- REE $\Longrightarrow$ PLM=ALM

The recursive learning mechanism

$$
\begin{aligned}
\phi_t &= \phi_{t-1} + \gamma_t R_t^{-1} z_{t-1} \left(p_t - \phi'_{t-1} z_{t-1} \right) \\
R_t &= R_{t-1} + \gamma_t \left(z_{t-1} z'_{t-1} - R_{t-1} \right)
\end{aligned}
$$

Decreasing gain: $\gamma_t = t^{-1}$ Equivalent to recursive least squares

Constant gain: $\gamma_t = \gamma \forall t \Rightarrow$ Less weight put on older observations which makes sense if there are structural change in the model

Unresolved (and perhaps interesting) questions

How can we discipline the choice of the precision of information in our models?

Can publicness of signals be determined endogenously?

What would it take to make full information rational expectations economics embarrassing?

That's it.