Macroeconomia II
Learning in DSGE models

George McCandless
UCEMA

Spring 2009

Learning

- How do people form expectations about the future
- Many economists are uneasy with rational expectations
  - Especially for individual decisions
  - Demands a lot of knowledge
  - More knowledge than most people might have
  - How much processing ability do individuals have
- Bounded Rationality are an alternative
- What does Bounded Rationality mean
  - people are only partly rational in their decisions?
  - people don't have full information when making their decisions?
  - decisions are too hard to make rationally (too much processing and too much information)
  - people only preceive the world with errors
  - people use learning processes to try to improve their forecasting

Learning processes

- Bayesian updating
  - use new information mixed with priors to get new forecasts
- Kalman filter
  - Have underlying state space model
  - use available data to estimate model

- update with each new observation

- Least squares learning (a special case of Kalman filtering)
  - Use a linear model and estimate coefficients by least squares
  - can be done recursively as new data is added
  - can have fixed or decreasing gain
    * can can be thought of as the weight given to new data
    * decreasing gain like OLS
    * constant gain is like OLS with forgetting (older data is less relevant)
    * Forgetting is good in models with regime changes

Least squares learning

- Assume that people behave as if they have an OLS model for forecasting
- Expectation variables are forecast with this model
- Model is build on old and current data
  - For parameter estimates
- Model is updated each period when new data arrives
- IMPORTANT RESULT: in many cases least squares learning converges to rational expectations
  - Marcet and Sargent (1988)
  - They use a continuous approximation of the descrete model

Least squares learning

- Assume that the world works as if

$$y_t = x_t \varphi_t + \varepsilon_t,$$

- where
  - y_t is a vector of endogenous variables,
  - x_t is the history up to moment t of the exogenous variables (that could include past values of y_t)
  - φ_t is the estimate of the coefficients of the model using data up to time $t-1$ and ε_t is a vector of error terms.
- Define $X_t = [x_t, x_{t-1}, ..., x_0]'$ and $Y_t = [y_t, y_{t-1}, ..., y_0]'$

- Ordinary least squares estimate of the coefficients, φ_t, is

$$\varphi_t = \left(X_t'X_t\right)^{-1} X_t'Y_t.$$

- For forecasting, the Y_t variables need to be one step ahead of the X_t variables

- For doing a model, we want a recursive way of doing OLS

Recursive Least Squares

- One can write the history vectors as

$$X_t = \begin{bmatrix} X_{t-1} \\ x_t \end{bmatrix}$$

- and

$$Y_t = \begin{bmatrix} Y_{t-1} \\ y_t \end{bmatrix}.$$

- Then

$$X_t'X_t = \begin{bmatrix} X_{t-1}' & x_t' \end{bmatrix} \begin{bmatrix} X_{t-1} \\ x_t \end{bmatrix} = X_{t-1}'X_{t-1} + x_t'x_t$$

- and

$$X_t'Y_t = \begin{bmatrix} X_{t-1}' & x_t' \end{bmatrix} \begin{bmatrix} Y_{t-1} \\ y_t \end{bmatrix} = X_{t-1}'Y_{t-1} + x_t'y_t$$

Recursive Least Squares

- OLS can be written as

$$\varphi_t = \left(X_{t-1}'X_{t-1} + x_t'x_t\right)^{-1} \left(X_{t-1}'Y_{t-1} + x_t'y_t\right)$$

- Problem is how to find a useful (for recursive) expression of

$$\left(X_{t-1}'X_{t-1} + x_t'x_t\right)^{-1}$$

Useful trick (1)

- When $a'b$ is a rank one matrix
  - this happens when a and b are vectors

- The inverse of the matrix $I + a'b$ can be written as

$$\left[I + a'b\right]^{-1} = I + ca'b$$

where c is the scalar

$$c = -\frac{1}{1 + ba'}$$

Useful trick (2)

- multiply $[I + ab']^{-1}$ by a non-singluar matrix B^{-1} to get

$$B^{-1}[I + a'b]^{-1} = [[I + a'b]\,B]^{-1} = [B + a'bB]^{-1}.$$

Using the formula above

$$B^{-1}[I + a'b]^{-1} = B^{-1}[I + ca'b] = B^{-1} + cB^{-1}a'b.$$

Combine

$$[B + a'bB]^{-1} = B^{-1} + cB^{-1}a'b.$$

Define the vector as $f = bB$, and substitute

$$[B + a'f]^{-1} = B^{-1} + cB^{-1}a'fB^{-1}$$

where the scalar c is now

$$c = -\frac{1}{1 + fB^{-1}a'}$$

For OLS

- The inverse of the X'X matrix is

$$\begin{aligned} & \left(X'_{t-1}X_{t-1} + x'_t x_t\right)^{-1} \\ = \; & \left(X'_{t-1}X_{t-1}\right)^{-1} + \mathbf{c}\left(X'_{t-1}X_{t-1}\right)^{-1} x'_t x_t \left(X'_{t-1}X_{t-1}\right)^{-1} \end{aligned}$$

where

$$\mathbf{c} = -\frac{1}{1 + x_t\left(X'_{t-1}X_{t-1}\right)^{-1} x'_t}$$

- Put this into the OLS equation

$$\begin{aligned} \varphi_t &= \left(X'_t X_t\right)^{-1} X'_t Y_t \\ &= \left(X'_{t-1}X_{t-1} + x'_t x_t\right)^{-1}\left(X'_{t-1}Y_{t-1} + x'_t y_t\right) \end{aligned}$$

- after some algebra get

$$\varphi_t = \varphi_{t-1} + \frac{\left(X'_{t-1}X_{t-1}\right)^{-1} x'_t}{1 + x_t\left(X'_{t-1}X_{t-1}\right)^{-1} x'_t}\left(y_t - x_t\varphi_{t-1}\right)$$

Recursive OLS (decreasing gain)

- Define $P_t = \left(X'_t X_t\right)^{-1}$,

- write above equation as

$$\varphi_t = \varphi_{t-1} + \frac{P_{t-1}x_t'}{1 + x_t P_{t-1} x_t'} \left(y_t - x_t \varphi_{t-1}\right)$$

- with the updating rule for P_t of

$$P_t = \left[I - \frac{P_{t-1}x_t'}{1 + x_t P_{t-1} x_t'} x_t\right] P_{t-1}$$

- This is the decreasing gain recursive OLS formula
- Begin with some P_0 and φ_0 and update using this formula and the data x_t and y_t in each period
- Here P_0 and φ_0 are like "priors"

Putting this into the Hansen model

- Hansen's basic model is

$$\begin{aligned}
0 &= \widetilde{C}_t - E_t \widetilde{C}_{t+1} + \beta \overline{r} E_t \widetilde{r}_{t+1} \\
0 &= \widetilde{Y}_t - \frac{\widetilde{H}_t}{1 - \overline{H}} - \widetilde{C}_t \\
0 &= \overline{Y}\widetilde{Y}_t - \overline{C}\widetilde{C}_t + \overline{K}\left[(1-\delta)\widetilde{K}_t - \widetilde{K}_{t+1}\right] \\
0 &= \widetilde{\lambda}_t + \theta \widetilde{K}_t + (1+\theta)\widetilde{H}_t - \widetilde{Y}_t \\
0 &= \widetilde{Y}_t - \widetilde{K}_t - \widetilde{r}_t
\end{aligned}$$

$$\widetilde{\lambda}_t = \gamma \widetilde{\lambda}_{t-1} + \widetilde{\varepsilon}_t.$$

- We assume that the expected variables $E_t \widetilde{C}_{t+1}$ and $E_t \widetilde{r}_{t+1}$ are found using current data, or

$$\begin{bmatrix} E_t \widetilde{C}_{t+1} \\ E_t \widetilde{r}_{t+1} \end{bmatrix} = \left[\varphi_{t-1}\right] \begin{bmatrix} \widetilde{K}_{t+1} \\ \widetilde{Y}_t \end{bmatrix} = \begin{bmatrix} \varphi_{11}^1 & \varphi_{12}^1 \\ \varphi_{21}^1 & \varphi_{22}^1 \end{bmatrix} \begin{bmatrix} \widetilde{K}_{t+1} \\ \widetilde{Y}_t \end{bmatrix},$$

Least square updating

- given some initial values for P_0 and φ_0
- they are updated using the OLS recursive formula

$$\varphi_t = \varphi_{t-1} + \frac{P_{t-1}x_t'}{1 + x_t P_{t-1} x_t'} \left(y_t - x_t \varphi_{t-1}\right)$$

and

$$P_t = \left[I + \frac{P_{t-1}x_t'}{1 + x_t P_{t-1} x_t'} x_t\right] P_{t-1}$$

- The model has two parts
  - the linear model to be solved each period
    - which is backward looking because the expectations are estimated on past values
  - and updating part that uses the data generated by the model to update the values of P_t and φ_t
  - the new value of φ_t is used in the next period's solution of the model

.The Hansen model

- Written is a state space version, where

$$x_t = \begin{bmatrix} K_{t+1} \\ H_t \\ Y_t \\ C_t \\ r_t \\ E_t \widetilde{C}_{t+1} \\ E_t \widetilde{r}_{t+1} \\ \widetilde{\lambda}_t \end{bmatrix}$$

- The state space version of the log-linear model can be written as

$$A_t\left(\varphi_{t-1}\right) x_t = B_t\left(\varphi_{t-1}\right) x_{t-1} + C\varepsilon_t$$

- Only backward looking

.The Hansen model

- If $A_t\left(\varphi_{t-1}\right)$ is invertible, the model is solved as

$$x_t = \left[A_t\left(\varphi_{t-1}\right)\right]^{-1} B_t\left(\varphi_{t-1}\right) x_{t-1} + \left[A_t\left(\varphi_{t-1}\right)\right]^{-1} C\varepsilon_t$$

- where

$$A_t = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & -1 & \beta\overline{r} & 0 \\ 0 & -\frac{1}{1-\overline{H}} & 1 & -1 & 0 & 0 & 0 & 0 \\ -\overline{K} & 0 & \overline{Y} & -\overline{C} & 0 & 0 & 0 & 0 \\ 0 & 1-\theta & -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ -\varphi_{11}^1(t-1) & 0 & -\varphi_{12}^1(t-1) & 0 & 0 & 1 & 0 & 0 \\ -\varphi_{21}^1(t-1) & 0 & -\varphi_{22}^1(t-1) & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

.The Hansen model

$$B_t = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -(1-\delta)\overline{K} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\theta & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma \end{bmatrix}$$

and

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}'.$$

.The Hansen model

- The rational expectations parameters for the OLS forecasting equation for this model are

$$\varphi = \begin{bmatrix} .5130 & .2614 \\ -1.007 & .9662 \end{bmatrix}$$

- These are found from the linear plans from a rational expectations solution
- A policy function is of the form

$$\begin{bmatrix} x_t^d \\ x_t^e \end{bmatrix} = \begin{bmatrix} d^k & d^\lambda \\ e^k & e^\lambda \end{bmatrix} \begin{bmatrix} k_t \\ \lambda_t \end{bmatrix}$$

- In a rational expectations model, a forecast for the expected variables $E_t\widetilde{C}_{t+1}$ and $E_t\widetilde{r}_{t+1}$ are

$$\begin{bmatrix} E_t\widetilde{C}_{t+1} \\ E_t\widetilde{r}_{t+1} \end{bmatrix} = \begin{bmatrix} c^k & c^\lambda \\ r^k & r^\lambda \end{bmatrix} \begin{bmatrix} E_t\widetilde{K}_{t+1} \\ E_t\widetilde{\lambda}_{t+1} \end{bmatrix}$$

.The Hansen model

- But $E_t\widetilde{K}_{t+1} = \widetilde{K}_{t+1}$ and $E_t\widetilde{\lambda}_{t+1} = \gamma\widetilde{\lambda}_t$ so these are found from

$$\begin{bmatrix} \widetilde{K}_{t+1} \\ \gamma\widetilde{\lambda}_t \end{bmatrix} = \begin{bmatrix} k^k & k^\lambda \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} \widetilde{K}_t \\ \widetilde{\lambda}_t \end{bmatrix}.$$

- Combining these two give

$$\begin{bmatrix} E_t\widetilde{C}_{t+1} \\ E_t\widetilde{r}_{t+1} \end{bmatrix} = \begin{bmatrix} c^k & c^\lambda \\ r^k & r^\lambda \end{bmatrix} \begin{bmatrix} k^k & k^\lambda \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} \widetilde{K}_t \\ \widetilde{\lambda}_t \end{bmatrix}$$

- the vector $\begin{bmatrix} \widetilde{K}_t \\ \widetilde{\lambda}_t \end{bmatrix}$ of states can be calculated from the policy functions for $\widetilde{K}_t$ and $\widetilde{Y}_t$ as

$$\begin{bmatrix} \widetilde{K}_{t+1} \\ \widetilde{Y}_t \end{bmatrix} = \begin{bmatrix} k^k & k^\lambda \\ y^k & y^\lambda \end{bmatrix} \begin{bmatrix} \widetilde{K}_t \\ \widetilde{\lambda}_t \end{bmatrix}$$

.

- or, by taking the inverse, as

$$\begin{bmatrix} \widetilde{K}_t \\ \widetilde{\lambda}_t \end{bmatrix} = \begin{bmatrix} k^k & k^\lambda \\ y^k & y^\lambda \end{bmatrix}^{-1} \begin{bmatrix} \widetilde{K}_{t+1} \\ \widetilde{Y}_t \end{bmatrix}$$

- The final OLS coefficients come from

$$\begin{bmatrix} E_t\widetilde{C}_{t+1} \\ E_t\widetilde{r}_{t+1} \end{bmatrix} = \begin{bmatrix} c^k & c^\lambda \\ r^k & r^\lambda \end{bmatrix} \begin{bmatrix} k^k & k^\lambda \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} k^k & k^\lambda \\ y^k & y^\lambda \end{bmatrix}^{-1} \begin{bmatrix} \widetilde{K}_{t+1} \\ \widetilde{Y}_t \end{bmatrix}$$

- or

$$\begin{bmatrix} \varphi^1_{11} & \varphi^1_{12} \\ \varphi^1_{21} & \varphi^1_{22} \end{bmatrix} = \begin{bmatrix} c^k & c^\lambda \\ r^k & r^\lambda \end{bmatrix} \begin{bmatrix} k^k & k^\lambda \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} k^k & k^\lambda \\ y^k & y^\lambda \end{bmatrix}^{-1}$$

.The Hansen Model

- Marcet and Sargent show that if the coefficients of the OLS forecasting rule are in a neighborhood of φ, they converge to φ.
- In practice, the neighborhood can be pretty big.
- Let the initial coefficients be

$$\begin{bmatrix} \varphi^1_{11}(0) & \varphi^1_{12}(0) \\ \varphi^1_{21}(0) & \varphi^1_{22}(0) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$P = \begin{bmatrix} P^1_{11}(0) & P^1_{12}(0) \\ P^1_{21}(0) & P^1_{22}(0) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

.The Hansen Model

- Unfortunately, learning with recursive OLS with declining gain can be very slow
- After 200,000 periods, the estimates at the end of the $200,000$ periods φ are equal to

$$\varphi = \begin{bmatrix} 0.6803 & 0.2750 \\ -1.0058 & 0.9597 \end{bmatrix}$$

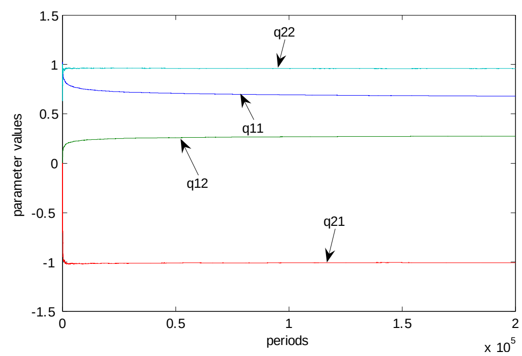

Figure 1: Parameter values over 200,000 periods: $\lambda = 1$

- and the updating matrix P is equal to

$$P = \begin{bmatrix} .00009069 & -.00005938 \\ -.00005938 & .00005983 \end{bmatrix}$$

- Further adjustments will be slow because the updating matrix is so small

The Hansen Model with memory

- see how slowly the coefficients converge

Learning with forgetting

- According to Lindoff adding "forgetting" to recursive least squares estimation is simple.
- Choose a λ where $0 < \lambda < 1$ and adjust the updating rule to be

$$P_{t+1}^{-1} = \lambda P_t^{-1} + x_{t+1}' x_{t+1}.$$

- Asymptotically, this is equivalent to a weighted least squares estimation of the form

$$\widehat{\varphi}_t = \left(\sum_{k=1}^{t} \lambda^{t-k} x_k' x_k \right)^{-1} \left(\sum_{k=1}^{t} \lambda^{t-k} x_k' y_k \right).$$

- The weights are smaller on older data and are relatively large on new data
- This kind of updating rule is good if one suspects that there has been a regime change

Learning with forgetting

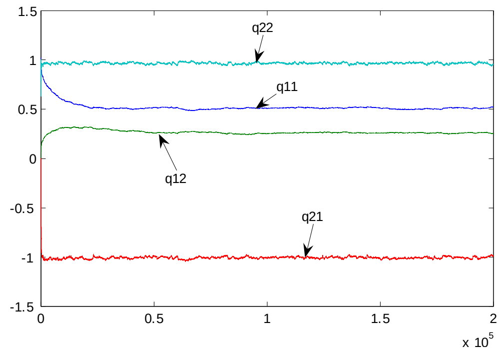

Figure 2: Parameter values over 200,000 periods: $\lambda = .999$

- Adding a forgetting factor of λ results in a new updating function of

$$\varphi_t = \varphi_{t-1} + \frac{P_{t-1}x_t'}{\lambda + x_t P_{t-1} x_t'} \left[y_t - x_t \varphi_{t-1} \right]$$

- and the updating equation for P is

$$P_t = \frac{1}{\lambda} \left[I + \frac{P_{t-1}x_t}{\lambda + x_t' P_{t-1} x_t} x_t' \right] P_{t-1}.$$

- Notice that $1/\lambda$ in the updating equation for P is greater than one
- This keeps P from shrinking too fast
- Values of λ between .999 and .95 are frequently used
- Even at the lower end of this range, the model can give weird results

Learning with forgetting

- Run same economy with $\lambda = .999$
- Coefficients converge faster (and in distribution) to the rational expectations values