Models with Occasionally Binding Constraints

Lecture Notes for a course at DIW

November 2 and 3, 2017

Matteo Iacoviello
Federal Reserve Board

Course overview

We will discuss broad topics

- Solving DSGE models, with particular attention to their nonlinearities
- Estimating DSGE models with OBC

DSGE Models: The Simplest Example

RBC Model with zero Capital Depreciation and Fixed Labor Supply
The planner's problem can be written as:

$$\max E_t \left(\sum_{s=t}^{\infty} \beta^{s-t} \log C_s \right)$$

subject to

$$K_t - K_{t-1} = A_t^{1-\alpha} K_{t-1}^{\alpha} - C_t \tag{1}$$

Optimal consumption implies

$$\frac{1}{C_t} = \beta E_t \left(\frac{1}{C_{t+1}} \left(\alpha \left(\frac{A_{t+1}}{K_t} \right)^{1-\alpha} + 1 \right) \right) \tag{2}$$

Assume technology follows an AR(1) process in logs, that is

$$\log A_t = \rho \log A_{t-1} + \log u_t \tag{3}$$

where ρ is the autocorrelation of the shock. $\log u$ has mean zero, finite variance.

Solution with Linearization

The system made by (1) to (3) is a non-linear system with rational expectations. We usually solve them in the following steps

- Find the steady state.

$$
\begin{aligned}
\log A &= 0- > A = 1 \\
C &= K^{\alpha} \\
1-\beta &= \alpha\beta\left(\frac{1}{K}\right)^{1-\alpha} - > \left(\frac{\alpha\beta}{1-\beta}\right)^{\frac{1}{1-\alpha}} = K
\end{aligned}
$$

- Linearize model equations around the steady state, letting $x_t \equiv \frac{X_t - X_{SS}}{X_{SS}}$

equation 1

$$C_t = A_t^{1-\alpha} K_{t-1}^{\alpha} - K_t + K_{t-1}$$

$$\begin{aligned}
\log C_t &= \log\left(A_t^{1-\alpha} K_{t-1}^{\alpha} - K_t + K_{t-1}\right) \\
&\quad \text{take total differential around steady state} \\
\frac{1}{C} dC_t &= \frac{1}{C}\left(\left(1-\alpha\right) A^{-\alpha} K^{\alpha} dA_t + \alpha A^{1-\alpha} K^{\alpha-1} dK_{t-1} - dK_t + dK_{t-1}\right) \\
c_t &= \left(1-\alpha\right) a_t + \alpha k_{t-1} - \frac{K}{C} k_t + \frac{K}{C} k_{t-1}
\end{aligned}$$

- equation 2

$$
\begin{aligned}
E_t\left(\frac{C_{t+1}}{C_t}\right) &= \beta E_t\left(\alpha\left(\frac{A_{t+1}}{K_t}\right)^{1-\alpha}+1\right)
\end{aligned}
$$

steady state of both sides is 1

$$
\begin{aligned}
E_t c_{t+1} - c_t &= \alpha\left(1-\alpha\right)\left(\frac{1}{K}\right)^{1-\alpha}\left(E_t a_{t+1} - k_t\right) \\
0 &= -E_t c_{t+1} + c_t + \frac{\left(1-\alpha\right)\left(1-\beta\right)}{\beta}\left(E_t a_{t+1} - k_t\right)
\end{aligned}
$$

- equation 3

$$a_t = \rho a_{t-1} + u_t$$

Taking Stock

This is a dynamic system of 3 equations in 3 unknowns. To use a more compact notation, we prefer to write it in the following form

$$\begin{aligned} 0 &= E_t\left[\mathbf{F}\mathbf{x}_{t+1} + \mathbf{G}\mathbf{x}_t + \mathbf{H}\mathbf{x}_{t-1} + \mathbf{L}\mathbf{z}_{t+1} + \mathbf{M}\mathbf{z}_t\right] && (4) \\ \mathbf{z}_t &= \mathbf{N}\mathbf{z}_{t-1} + \mathbf{e}_t && (5) \end{aligned}$$

where

- $\mathbf{x}_t$ is the vector collecting all the endogenous variables of the model.
- $\mathbf{z}_t$ collects all the exogenous stochastic processes.

In our above example

$$\mathbf{x} = \begin{bmatrix} c \\ k \end{bmatrix}, \mathbf{z} = [a]$$

and $\mathbf{F} = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} -1 & -\frac{\alpha\beta}{1-\beta} \\ 1 & -\frac{(1-\alpha)(1-\beta)}{\beta} \end{bmatrix}, \mathbf{H} = \begin{bmatrix} 0 & \frac{\alpha}{1-\beta} \\ 0 & 0 \end{bmatrix},$

$\mathbf{L} = \begin{bmatrix} 0 \\ \frac{(1-\alpha)(1-\beta)}{\beta} \end{bmatrix}, \mathbf{M} = \begin{bmatrix} 1-\alpha \\ 0 \end{bmatrix}, \mathbf{N} = [\rho]$

To summarize, a linearized DSGE model can be written in the following form

$$\begin{aligned} 0 &= E_t\left[\mathbf{F}\mathbf{x}_{t+1} + \mathbf{G}\mathbf{x}_t + \mathbf{H}\mathbf{x}_{t-1} + \mathbf{L}\mathbf{u}_{t+1} + \mathbf{M}\mathbf{u}_t\right] \\ \mathbf{z}_t &= \mathbf{N}\mathbf{z}_{t-1} + \mathbf{e}_t \end{aligned}$$

The recursive equilibrium law of motion describes endogenous variables as function of the state:

$$\mathbf{x}_t = \mathbf{P}\mathbf{x}_{t-1} + \mathbf{Q}\mathbf{z}_t \tag{6}$$

i.e., matrices **P**, **Q** such that the equilibrium is described by these rules.

- Finally, what we do is to plug the matrices in (4) and (5) in a computer, to obtain (6) .

In our toy example above, set $\alpha = 0.33, \beta = 0.99, \rho = 0.98$. Then (from `runrbc.m`)

$$\begin{bmatrix} c_t \\ k_t \end{bmatrix} = \begin{bmatrix} 0 & 0.6589 \\ 0 & 0.9899 \end{bmatrix} \begin{bmatrix} c_{t-1} \\ k_{t-1} \end{bmatrix} + \begin{bmatrix} 0.1755 \\ 0.0151 \end{bmatrix} [z_t]$$

Practice.

Let problem for planner be: $\max E_t \left(\sum_{s=t}^{\infty} \beta^{s-t} \left(\log C_s + \tau \log \left(1 - L_t \right) \right) \right)$
subject to $K_t - K_{t-1} = A_t^{1-\alpha} K_{t-1}^{\alpha} L_t^{1-\alpha} - C_t$
let $\tau = 1$.

1. Derive planner's first order conditions.
2. Find analytical steady state.
3. Solve for decision rules of the form $\mathbf{x}_t = \mathbf{P}\mathbf{x}_{t-1} + \mathbf{Q}\mathbf{z}_t$.
4. Compare coefficients of P and Q with the ones obtained in the model with fixed labor supply.

Solving Models with Occasionally Binding Constraints

Occasionally binding constraints arise in many economic applications. Examples include:
Models with limitations on the mobility of factors of production;

Models with heterogenous agents and constraints on the financial assets available to agents;

Models with a zero lower bound on the nominal interest rate;

Models with inventory management.

Why is a Toolkit Needed?

Encompassing realistic features to improve model fit in empirically driven applications may quickly raise the number of state variables. This may render standard global solution methods, such as dynamic programming, infeasible.

An alternative that has been used in practice, especially in applications that deal with the zero lower bound on policy rates, is to use a piece-wise perturbation approach.

This approach has the distinct advantage of delivering a solution for models with a large number of state variables. Furthermore, it can be easily extended to encompass multiple occasionally binding constraints.

1) Present a toolbox that extends Dynare to use this solution technique.
2) Can gauge performance of this approach relative to other solution methods (more accurate but slower, for instance).

The Solution Method

- The linearized system of necessary conditions for an equilibrium of a baseline DSGE model can be expressed as:

$$\mathcal{A}_1 E_t X_{t+1} + \mathcal{A}_0 X_t + \mathcal{A}_{-1} X_{t-1} + \mathcal{B} u_t = 0. \tag{M1}$$

Where X_t are variables in deviation from non-stochastic steady state. There are situations however (away from ss) when one (or more) of the equilibrium conditions may not hold, and is replaced by another one. When the "starred" system applies, express the system as:

$$\mathcal{A}_1^* E_t X_{t+1} + \mathcal{A}_0^* X_t + \mathcal{A}_{-1}^* X_{t-1} + \mathcal{B}^* u_t + \mathcal{C}^* = 0 \tag{M2}$$

where C^* is a vector of constants.

- Both systems are linearized around the same point – same X across systems.

The Solution Method

- When the baseline model applies (M1), we use standard methods to express solution as:

$$X_t = \mathcal{P}X_{t-1} + \mathcal{Q}u_t. \tag{M1\_DR}$$

- If the starred model applies (M2), *shoot back* towards the initial condition starting from the last period before we return to M1.
- Main idea: suppose that M2 applies in t, but M1 is expected to apply in all future periods $t+1$, the decision rule in t is:

$$\underset{\text{M2}}{\mathcal{A}_1^*}\ \underset{\text{M1\_DR}}{\mathcal{P}X_t} + \mathcal{A}_0^* X_t + \underset{\text{M2}}{\mathcal{A}_{-1}^*} X_{t-1} + \mathcal{C}^* = 0,$$

$$X_t = -\left(\mathcal{A}_1^*\mathcal{P} + \mathcal{A}_0^*\right)^{-1}\left(\mathcal{A}_{-1}^* X_{t-1} + \mathcal{B}^* u_t + \mathcal{C}^*\right)$$

- One can proceed in a similar fashion to construct the time-varying decision rules when M2 applies for multiple periods.
- In each period in which M2 applies, the expectation of how long one expects to stay in M2 affects the value of X_t today

The Solution Method

The search for the appropriate time-varying decision rules implies that for each set of shocks at a point in time one needs to calculate the expected future duration of each "regime."

Truncate the simulation at an arbitrary point and reject the truncation if the solution implies that the model has not returned to the reference regime by that point.

Start with a guess of the expected durations that is based on the linear solution. Update the guess based on where the conditions of system 1 are violated using the piece-wise linear method until no violation remains.

An Example

- To fix ideas, let's first consider a simple, forward-looking, linear model:

$$
\begin{aligned}
q_t &= \beta(1-\rho)E_t q_{t+1} + \rho q_{t-1} - \sigma r_t + u_t \\
r_t &= \max\left(\underline{r}, \phi q_t\right)
\end{aligned}
$$

  where u_t is an *iid* shock.

- The general solution for q_t takes the form

$$
\begin{aligned}
q_t &= \varepsilon_{qq,t} q_{t-1} + \varepsilon_{qu,t} u_t + c_{q,t} \\
r_t &= \varepsilon_{rr,t} q_{t-1} + \varepsilon_{ru,t} u_t + c_{r,t}
\end{aligned}
$$

- In turn, the ε are functions of q_{t-1} and u_t.
- How do we find the solution?

An Example: Solution ignoring the constraint

Ignore the constraint first

$$
\begin{aligned}
q_t &= \frac{\beta(1-\rho)}{1+\sigma\phi} E_t q_{t+1} + \frac{\rho}{1+\sigma\phi} q_{t-1} + \frac{1}{1+\sigma\phi} u_t \\
q_t &= a E_t q_{t+1} + b q_{t-1} + c u_t
\end{aligned}
$$

Find solution (method of undetermined coefficients)

$$
\begin{aligned}
q_t &= \varepsilon_q q_{t-1} + \varepsilon_u u_t \text{ (guess)} \\
E_t q_{t+1} &= \varepsilon_q q_t \text{ (expectation given guess)} \\
a E_t q_{t+1} &= a \varepsilon_q q_t = a \varepsilon_q^2 q_{t-1} + a \varepsilon_q \varepsilon_u u_t
\end{aligned}
$$

Match coefficients

$$
\underbrace{e q_{t-1} + \varepsilon_u u_t}_{q_t} = \underbrace{a \varepsilon_q^2 q_{t-1} + a \varepsilon_q \varepsilon_u u_t}_{a E_t q_{t+1}} + b q_{t-1} + c u_t
$$

so that (after picking the "stable" root)

$$
\varepsilon_q = a \varepsilon_q^2 + b, \; \varepsilon_u = a \varepsilon_q \varepsilon_u + c
$$

$$
\varepsilon_q = \left(1 - \sqrt{1 - 4ab}\right) / 2a, \varepsilon_u = c / \left(1 - a \varepsilon_q\right)
$$

Plug some numbers

$$
\begin{aligned}
\beta &= 0.99 \\
\phi &= 1 \\
\rho &= 0.5 \\
\sigma &= 1 \\
\underline{r} &= -0.02
\end{aligned}
$$

In this case (see `runsimmodelsimple.m`)

$$
\begin{aligned}
\varepsilon_q &= 0.2677 \\
\varepsilon_u &= 0.5355
\end{aligned}
$$

so that

$$q_t = r_t = 0.2677 q_{t-1} + 0.5355 u_t$$

Is this solution always correct? Consider the case of a large negative shock to u in period 1. If $q_0 = 0$, any u_t such that

$$\begin{aligned} r^* &= 0.5355u^* < -0.02 \\ u^* &< -0.0373 \end{aligned}$$

will violate constraint.
Suppose for instance $u_1 = -0.2$. Ignoring constraint, solution is

$$r_t = 0.2677r_{t-1} + 0.5355u_t$$

$$\begin{aligned} r_1 &= -0.5355 * 0.2 = -0.1071 \\ r_2 &= 0.2677r_1 = -0.0287 \\ r_3 &> -0.02 \end{aligned}$$

Hence ignoring the constraint r_t would be below -0.02 for 2 periods. Moreover, there is a feedback loop. Higher values of r imply lower q, which implies lower desired values of r, so r can end up being at $\underline{r}$ for longer

We use a guess and verify method to determine how long the constraint will bind. We start by guessing durations that are based on the linear solution that ignores the constraint. Iterate until convergence.

So the first guess is going to be 2 periods.

–> Suppose we guess that r remains at ϕ for $t\_low = 2$ periods.
–> Because r is not going to be low as guessed in linear solution ....
–> q will fall more than if r did not hit the constraint ...
–> and r might in turn stay at its lowest bound ϕ more than t_low periods.

In all interesting cases, first guess is not last guess, since dynamics of system depend on feedback loop between duration of constraint and endogenous reaction of variables to constraint. In the example above, one can think of a New Keynesian model at the ZLB.

Now cast system using our general notation (use $\beta' = \beta(1-\rho)$):

$$\begin{aligned} q_t &= \beta' E_t q_{t+1} + \rho q_{t-1} - \sigma r_t + u_t \\ r_t &= \max\left(\underline{r}, \phi q_t\right) \end{aligned}$$

$$\mathcal{A}_1 E_t X_{t+1} + \mathcal{A}_0 X_t + \mathcal{A}_{-1} X_{t-1} + \mathcal{B} u_t = 0. \tag{M1}$$

$$\begin{bmatrix} -\beta' & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ r_1 \end{bmatrix} + \begin{bmatrix} 1 & \sigma \\ -\phi & 1 \end{bmatrix} \begin{bmatrix} q \\ r \end{bmatrix} + \begin{bmatrix} -\rho & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} ql \\ rl \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} u = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

and

$$\mathcal{A}_1^* E_t X_{t+1} + \mathcal{A}_0^* X_t + \mathcal{A}_{-1}^* X_{t-1} + \mathcal{B}^* u_t = -\mathcal{C}^* \tag{M2}$$

$$\begin{bmatrix} -\beta' & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ r_1 \end{bmatrix} + \begin{bmatrix} 1 & \sigma \\ 0 & 1 \end{bmatrix} \begin{bmatrix} q \\ r \end{bmatrix} + \begin{bmatrix} -\rho & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} ql \\ rl \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix} u = \begin{bmatrix} 0 \\ \bar{r} \end{bmatrix}$$

and

$$X_t = \mathcal{P} X_{t-1} + \mathcal{Q} u_t \tag{M1\_DR}$$

$$\begin{bmatrix} q_t \\ r_t \end{bmatrix} = \begin{bmatrix} \varepsilon_q & 0 \\ \varepsilon_q & 0 \end{bmatrix} \begin{bmatrix} q_{t-1} \\ r_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_u \\ \varepsilon_u \end{bmatrix} u$$

We guess that in t=3 normal system applies. Hence need to find solution in t=1, 2, given the shock taking place in period 1, and knowing X_0. In that case, the solution in t=2 should satisfy

$$\begin{aligned} X_2 &= -\left(\mathcal{A}_1^*\mathcal{P}+\mathcal{A}_0^*\right)^{-1}\left(\mathcal{A}_{-1}^*X_1+\mathcal{B}^*u_2+\mathcal{C}^*\right) \\ &= P_2X_1+Q_2u_2+C_2 \end{aligned}$$

where

$$\begin{aligned} P_2 &= -\left(\mathcal{A}_1^*\mathcal{P}+\mathcal{A}_0^*\right)^{-1}\mathcal{A}_{-1}^*, \\ Q_2 &= -\left(\mathcal{A}_1^*\mathcal{P}+\mathcal{A}_0^*\right)^{-1}\mathcal{B}^*u_2, \quad C_2 = -\left(\mathcal{A}_1^*\mathcal{P}+\mathcal{A}_0^*\right)^{-1}C^* \end{aligned}$$

I plug the numbers now. Using

$$\begin{aligned} \mathcal{A}_1^* &= \begin{bmatrix} -0.495 & 0 \\ 0 & 0 \end{bmatrix}, \mathcal{P} = \begin{bmatrix} 0.2677 & 0 \\ 0.2677 & 0 \end{bmatrix}, \mathcal{A}_0^* = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\ \mathcal{B}^* &= \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mathcal{C}^* = \begin{bmatrix} 0 \\ 0.02 \end{bmatrix} \end{aligned}$$

I get

$$X_2 = \underbrace{\begin{bmatrix} 0.5706 & 0 \\ 0 & 0 \end{bmatrix}}_{P_2} X_1 + \underbrace{\begin{bmatrix} .023 \\ -0.02 \end{bmatrix}}_{C_2}$$

We do not know yet X_1 the solution in period 1. Assuming $(M2)$ applies in $t = 1$ and is expected to apply in $t = 2$, the solution in 1 is

$$\mathcal{A}_1^* \left(\mathcal{P}_2 X_1 + C_2\right) + \mathcal{A}_0^* X_1 + A_{-1}^* X_0 + B u_1 + \mathcal{C}^* = 0,$$

$$X_1 = -\left(\mathcal{A}_1^* \mathcal{P}_2 + \mathcal{A}_0^*\right)^{-1} \left(\mathcal{B}^* u_1 + \mathcal{C}^* + A_1^* C_2 + A_{-1}^* X_0\right)$$

$$X_1 = P_1 X_0 + Q_1 u_1 + C_1$$

where

$$\begin{aligned} \mathcal{P}_1 &= -\left(\mathcal{A}_1^* \mathcal{P}_2 + \mathcal{A}_0^*\right)^{-1} A_{-1}^* \\ Q_1 &= -\left(\mathcal{A}_1^* \mathcal{P}_2 + \mathcal{A}_0^*\right)^{-1} \mathcal{B}^* \\ \mathcal{C}_1 &= -\left(\mathcal{A}_1^* \mathcal{P}_2 + \mathcal{A}_0^*\right)^{-1} \left(\mathcal{C}^* + A_1^* \mathcal{C}_2\right) \end{aligned}$$

After plugging in all the numbers, assuming $X_0 = 0$, we get

$$X_1 = \begin{bmatrix} -0.235\,00 \\ -0.02 \end{bmatrix}$$

So far we went backwards from the last period in which regime 2 applies to the first.
Now we go forward. Plug X_1 back into solution for X_2 and get

$$X_2 = \begin{bmatrix} 0.570\,64 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -0.235\,00 \\ -0.02 \end{bmatrix} + \begin{bmatrix} .023 \\ -0.02 \end{bmatrix} = \begin{bmatrix} -0.111\,1 \\ -0.02 \end{bmatrix}$$

and now plug X_2 into X_3

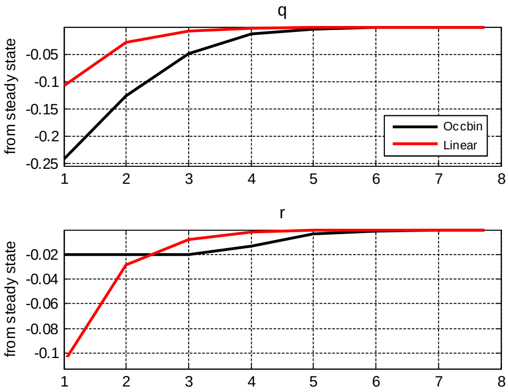
$$\begin{aligned} X_3 &= \mathcal{P} X_2 \\ X_3 &= \begin{bmatrix} -0.0297 \\ -0.0297 \end{bmatrix} \end{aligned}$$

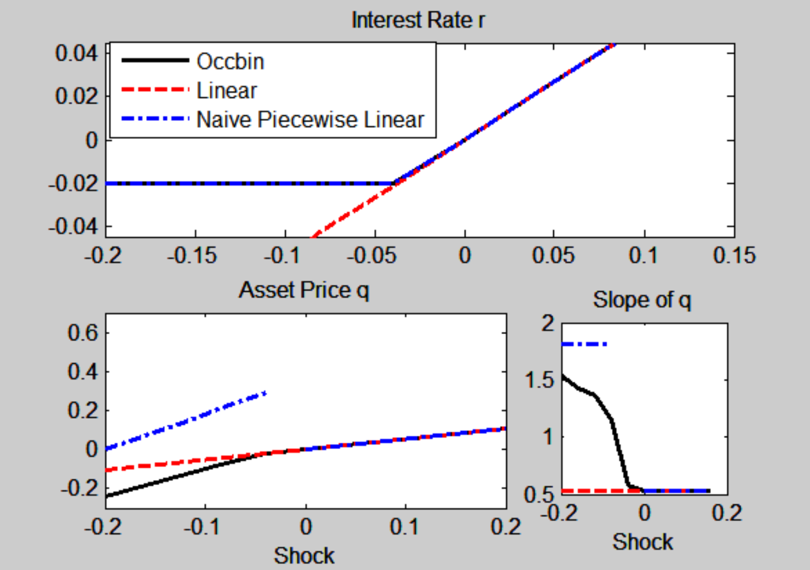
which violates constraint in 3.

- Note the need to update guess.
- We guessed that starred system $(M2)$ applies in 1 and 2 and that the normal applies in 3. Based on this guess, the starred system applies in 3.
- Hence we update the guess that starred system applies for 3 periods.
- Redo the whole thing again until the guessed duration in the starred regime coincides with the actual duration.

In the next step, we assume that the normal system applies in 4 but the starred applies in 1, 2 and 3, solve for P_3, Q_3 and C_3, use them to compute X_2 and X_1, and go back to see if X_3 satisfies the constraints (it does, I have checked it myself)

Impulse Responses to a $u_1 = -0.2$ shock




Interest Rate r
Occbin
Linear
Naive Piecewise Linear
0.04
0.02
0
-0.02
-0.04
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
Asset Price q
0.6
0.4
0.2
0
-0.2
-0.2
-0.1
0
0.1
0.2
Shock
Slope of q
2
1.5
1
0.5
-0.2
0
0.2
Shock

Example 1: Borrowing Constraint Model

- To check if method is accurate, we apply it to models for which we can compute a a full non-linear solution to arbitrary precision using dynamic programming methods.
- Random endowment y_t can be used as collateral

$$u = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \log (c_t) \right\}$$

$$c_t = y_t + b_t - 1.05 b_{t-1}$$

$$b_t \leq 2 y_t \tag{c1}$$

$$\log (y_t) = \rho \log (y_{t-1}) + \sigma \sqrt{1 - \rho^2} \varepsilon_t$$

$$\varepsilon_t \tilde{} N(0,1),\ \sigma = 0.03,\ \rho = 0.9$$

- We look at how solution method handles cases when increases in y_t are large enough so that constraint is not binding. We try $\beta = 0.94$ and $\beta = 0.949$
- Here: constraint $(c1)$ BINDS in *normal* times.
- Examples:runsim_borrcon_test1_compare_linear.m and runsim_borrcon_test1_compare_global.m in the folder borrcon_global

Example 2: RBC with Irreversible Capital

- Investment cannot fall below a given threshold

$$u = E_0 \left\{ \sum_{t=0}^{\infty} 0.96^t \log(c_t) \right\}$$

$$c_t + k_t - 0.9k_{t-1} = A_t k_{t-1}^{0.33}$$

$$k_t - 0.9k_{t-1} \geq \phi k_{t-1} \tag{c2}$$

$$\log(A_t) = 0.9 \log(A_{t-1}) + \sigma\sqrt{1-\rho^2}\varepsilon_t$$

$$\varepsilon_t \tilde{} N(0,1), \ \sigma = 0.03, \ \rho = 0.9$$

where $\phi > 0$.

- Here: constraint $(c2)$ DOES NOT bind in *normal* times.

Example 3: Borrowing and Housing

$$U = E_0 \sum_{t=0}^{\infty} \beta^t \left(\log c_t + j \log h_t \right)$$

$$\begin{aligned} c_t + q_t h_t &= y + b_t - R b_{t-1} + q_t h_{t-1} \left(1 - \delta\right) \\ b_t &\leq m q_t h_t \\ \log q_t &= \rho \log q_{t-1} + v_t \end{aligned}$$

Here the FOCs would be

$$\mu_t \left(b_t - m q_t h_t \right) = 0$$

$$u' \left(c_t \right) = \beta R E_t u' \left(c_{t+1} \right) + \mu_t$$

$$q_t u' \left(c_t \right) = u' \left(h_t \right) + \beta \left(1 - \delta \right) E_t q_{t+1} u' \left(c_{t+1} \right) + \mu_t m q_t$$

Assuming $\beta R < 1$, here the borrowing constraint binds in normal times.

Structure of Solution Programs (Dynare)

The programs we devised take as input two Dynare model files.
One .mod file specifies the normal M1 model from which we calculate

$$A_1 E_t X_{t+1} + A_0 X_t + A_{-1} X_{t-1} = 0.$$

The other .mod file specifies the starred model M2 with the occasionally binding constraint inverted (binding if it was not binding in the reference model, or not binding if it was binding in the reference model). This .mod file yields

$$A_1^* E_t X_{t+1} + A_0^* X_t + A_{-1}^* X_{t-1} + C^* = 0.$$

We use the analytical derivatives computed by Dynare to construct $A_1, A_0, A_{-1}, A_1^*, A_0^*, A_{-1}^*$, and C^*.

M1: hp.mod

y=1;
c+q*h=y+b-R*b(-1)+q*h(-1)*(1-δ);
b=M*q*h;
lb=1/c-β*R/c(+1);
q/c=j/h+β*(1-δ)*q(+1)/c(+1)+lb*M*q;
lev=b/(M*q*h)-1;
log(q)=ρ*log(q(-1))+u;

M2: hpnotbinding.mod

y=1;
c+q*h=y+b-R*b(-1)+q*h(-1)*(1-δ);
lb=0;
lb=1/c-β*R/c(+1));
q/c=j/h+β*(1-δ)*q(+1)/c(+1)+lb*M*q;
lev=b/(M*q*h)-1;
log(q)=ρ*log(q(-1))+u;

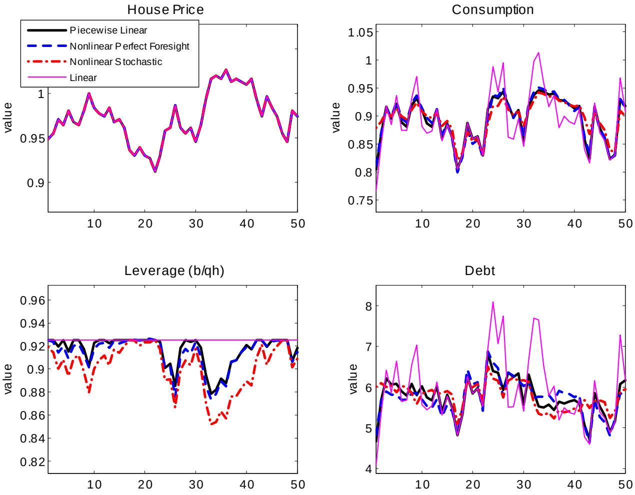
The main file `runsim_hp.m` contains

1. mod files: `model1 = 'hp'; model2 = 'hpnotbinding';`
2. constraint violation triggers switch to m2: `constraint='lb<-lb_ss';`
3. constraint violation triggers switch to m1: `constraint_relax1='lev>0'`
4. solve model1 to obtain M1_DR; write model2 in M2 form
5. given shocks, check if model1 assumptions are violated, if so, look for solution

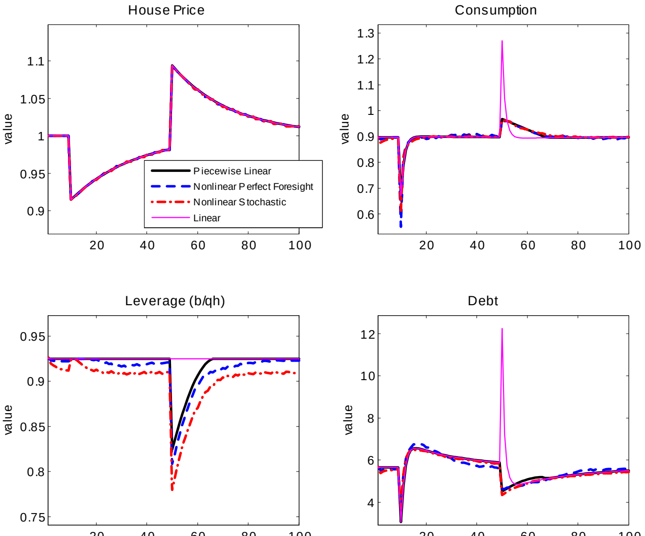
function [zdata zdataconcatenated oobase_ Mbase_] = ...
solve_one_constraint(model1,model2,...
constraint, constraint_relax,...
shockssequence,irfshock,nperiods,maxiter)

- `model1, model2`: dynare mod files containing (linear or nonlinear) model equations
- `constraint, constraint_relax`: strings with constraints that have to be verified
  constraint defines the first constraint
  if constraint is true, solution switches to model2
  but if constraint_relax is true, solution reverts to model1
- `shockssequence`: `sequence of innovations under which one wants to solve model`
  e.g. `randn(100,1)*`σ_j for simulation
  or `[ 1; zeros(50,1) ]` for impulse responses

Simulations - Borrowing and Housing Model



IRF - Borrowing and Housing Model



Accuracy – Borrowing and Housing Model

	Log Consumption		Correlations		$\frac{b}{qh}$	ΔWelf.
	st.dv	skewness	$\ln q, \ln c$	$\ln q, \frac{b}{qh}$	mean	
Linear	6.1%	0.03	0.40	0.00	0.925	0.18%
Occbin	4.5%	-1.17	0.54	-0.60	0.911	0.02%
Nonlin.perf.fores.	4.6%	-1.20	0.53	-0.58	0.910	0.01%
Nonlinear stoch.	3.7%	-1.30	0.65	-0.71	0.896	—

Nonlinear models

Nonlinear structural models can be described by:

1. The model's equilibrium conditions and FOCs

$$\Gamma\left(E_t z_{t+1}, z_t, z_{t-1}, \varepsilon_t\right) = 0$$

   where z_t includes s_t and c_t, "states" and "controls".
   $E_t z_{t+1}$ is an unknown object!

2. The solution is a set of policy functions ζ

$$z_t = \zeta\left(z_{t-1}, \varepsilon_t\right)$$

   such that, for any value of $\left(z_{t-1}, \varepsilon_t\right)$

$$F\left(z_{t-1}, \varepsilon_t\right) \equiv \Gamma\left(E_t \zeta\left(z_t, \varepsilon_{t+1}\right), \zeta\left(z_{t-1}, \varepsilon_t\right), z_{t-1}, \varepsilon_t\right) = 0$$

One Example

- Example of DSGE model we want to solve:

$$C_t^{-\gamma} = \beta \delta_t R_t E_t \left(C_{t+1}^{-\gamma} / \Pi_{t+1} \right) \tag{1}$$

$$w_t = N_t^{\eta} C_t^{\gamma} \tag{2}$$

$$\frac{N_t}{C_t^{\gamma}} \left(\psi \left(\Pi_t - 1 \right) \Pi_t - (1-\theta) - \theta w_t \right) = \beta \delta_t E_t \left(\frac{N_{t+1}}{C_{t+1}^{\gamma}} \psi \left(\Pi_{t+1} - 1 \right) \Pi_{t+1} \right) \tag{3}$$

$$N_t = C_t + \frac{\psi}{2} \left(\Pi_t - 1 \right)^2 N_t \tag{4}$$

$$R_t = \max \left(1, \Pi_t^{\phi} / \beta \right) \tag{5}$$

- Solution is a set of policy functions $C_t = C(\delta_t), R_t = R(\delta_t), w_t = w(\delta_t), \Pi_t = \Pi(\delta_t), N_t = N(\delta_t)$ such that (1) to (5) hold for every value of δ_t.

Computing the Likelihood

- Assume we solve the model using Occbin. Model solution is:

$$X_t = \mathbf{P}(X_{t-1}, \epsilon_t)X_{t-1} + \mathbf{D}(X_{t-1}, \epsilon_t) + \mathbf{Q}(X_{t-1}, \epsilon_t)\epsilon_t$$

- In terms of observables, through observation equation $Y_t = \mathbf{H}X_t$,:

$$Y_t = \mathbf{HP}(X_{t-1}, \epsilon_t)X_{t-1} + \mathbf{HD}(X_{t-1}, \epsilon_t) + \mathbf{HQ}(X_{t-1}, \epsilon_t)\epsilon_t$$

- We initialize X_0 and recursively solve for ϵ_t, given X_{t-1} and current Y_t.
- Given that ϵ_t is $NID(0, \Sigma)$, a change in variables argument implies that the log likelihood l for $Y^T \equiv \{Y_t\}_{t=1}^T$ given parameters can be derived analytically as:

$$l = -\frac{T}{2}\log(\det\Sigma) - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t'\Sigma^{-1}\epsilon_t + \sum_{t=1}^{T}\log(|\det\frac{\partial\epsilon_t}{\partial Y_t}|)$$

- where $\frac{\partial\epsilon_t}{\partial Y_t} = inv\left(\mathbf{HQ}_t\right)$ is the Jacobian matrix of the transformation from the shocks to the observations

Intuition

- The key to compute the likelihood is to make sure that we can invert the policy functions for Y_t
- There is a long tradition in science dealing with the so-called *inverse problem* : the process of calculating from a set of observations the causal factors that produced them.
- Kollmann (2017) calls this procedure inversion filter (IF)
- This is different from Particle filters (PFs), which use Monte Carlo methods to infer latent states (An and Schorfheide (2007)), and are thus computationally slow.
- The Inversion filter is fast as it does not involve computing moments of states.
- One drawback: PFs can be used when there are less observables than shocks.

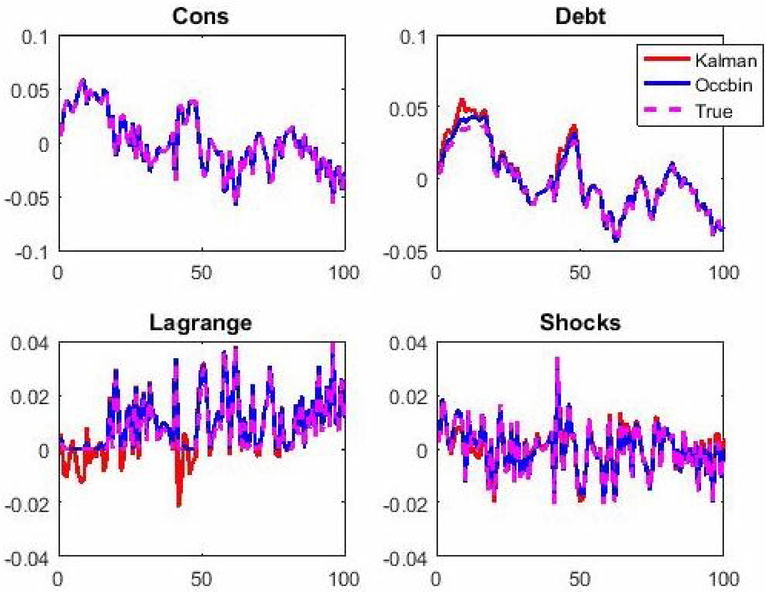
Performance of the Filter 1

Based on ongoing work with Guerrieri, Pablo Cuba Borda and Molin Zhong

- Consider Example 1: borrowing constraints model
- Generate data of $T = 100$ from fully nonlinear model with parameter values $\gamma = 1, \sigma = 0.01, R = 1.05, \beta = 0.945, \rho = 0.9, m = 1.$
- Fit the model using consumption as the only observable
  - with the occbin/inversion filter in JME 2017 paper....
  - ... and the kalman filter (assuming that the borrowing constraint is always binding).
- Next figure shows model variables. C and b are in deviations from ss. λ in levels.

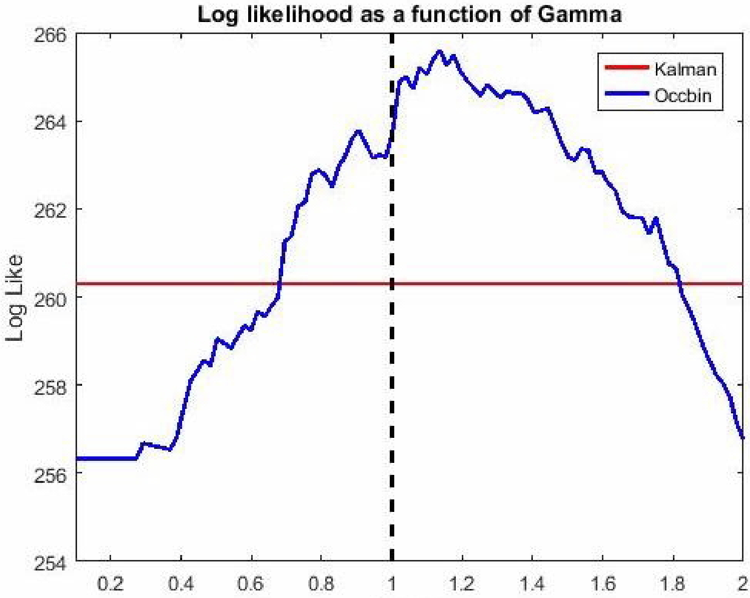
Performance of the Filter

- Occbin filter does well in terms of extracting out the latent states (debt, lagrange multiplier, and shocks).
- The Kalman filter does a bit worse, especially in time periods in which the constraint is binding.
- The occbin solution method + filter does pretty well on data generated from the fully nonlinear model. By neglecting completely the occasionally binding constraint, however, the filtering of the states is affected.


Cons
0.1
0.05
0
-0.05
-0.1
0
50
100
Debt
0.1
0.05
0
-0.05
Kalman
Occbin
True
Lagrange
0.04
0.02
0
-0.02
-0.04
Shocks

Performance of the Filter 2

- Next Figure shows log likelihood cut of γ.
- Here, vary the values of γ on grid between 0 to 2 (true γ is 1) while keeping all other parameter values at their true values.
- See how the likelihood changes.
- Neglecting occasionally binding constraint (kalman filter case) makes inference about γ impossible with just C as observable, because constraint is always binding.


Log likelihood as a function of Gamma
Log Like
Kalman
Occbin
266
264
262
260
258
256
254
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

Application to ZLB: the linearized CGG Model

$$
\begin{aligned}
y_t &= y_{t+1} - \phi(r_t - \pi_{t+1}) + aay_t \\
\pi_t &= \beta\pi_{t+1} + \lambda y_t + \varepsilon_p \\
rnot_t &= \phi_R rnot_{t-1} + (1-\phi_R)(\phi_p \pi_t + \phi_y y_t) + \varepsilon_r \\
rlong_t &= 0.10 r_t + 0.90 rlong_{t+1} \\
r_t &= rnot_t \text{ when ZLB does not bind} \\
aay_t &= \rho_y aay_{t-1} + \varepsilon_y,\ aap_t = \varepsilon_p,\ aar_t = \varepsilon_r
\end{aligned}
$$

at the ZLB only change one equation

$$
r_t = -(1/\beta - 1) \text{ when ZLB binds}
$$

also

$$
\begin{aligned}
constraint1 &= \prime r < (1-1/BETA)\prime \\
constraint\_relax1 &= \prime rnot > (1-1/BETA)\prime
\end{aligned}
$$

Applications

- To see how the method works in practice, consider running the following examples in occbin_estimation_web\cgg_rlong
    1. runsim_cgg_generate_fakedata.m (solves model, generates artificial time series)
    2. run_filter_cgg (solves the inversion problem and filters shocks)

Intro: A Baby Example

We have $n = 5$ independent observations from the normal distribution for the variable z_t, mean zero (known) and unit standard deviation (known).

$$z = [-0.5925, 0.3298, -0.9984, 1.8028, -0.5416]$$

What is the likelihood of observing this sample? From the formula for the density of an $n-$variate normal distribution with mean 0 and variance covariance matrix Σ, we have

$$L = (2\pi)^{-T/2} |\Sigma|^{-1/2} \exp\left[-\frac{1}{2} z' \Sigma^{-1} z\right]$$

$$L = (2\pi)^{-5/2} \exp\left(-\frac{0.5925^2 + 0.3298^2 + 0.9984^2 + 1.8028^2 + 0.5416^2}{2}\right) = .00082948$$

$$\ln L = \ln 8.2948 \times 10^{-4} = -7.0947$$

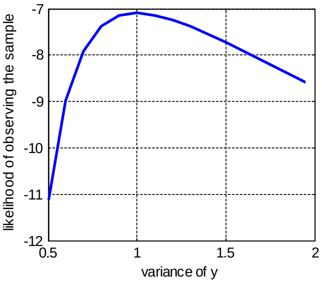
Estimation of DSGE models requires knowing (or remembering) all of the above, plus a lot of practice, and a few other things.

Why is it called Estimation? We did not Estimate Anything...

Treat the likelihood as a function of Σ. Then can write it as

$$L = L\left(z_T, \Sigma\right)$$

L : likelihood of observing particular sequence z_T as function of the parameter Σ.
Estimate of Σ is the value that maximizes the likelihood function above.
In this case, the likelihood function is the plot to left. Most often, with DSGE models, you end up with a likelihood that looks like the one on the right.



General Problem

1. Consider solution of a DSGE model
$$\begin{aligned} x_t &= F(\mu)\, x_{t-1} + G(\mu)\, v_t \\ z_t &= H' x_t \end{aligned}$$
$E(Gvv'G') = Q$. F and G are functions of vector of model parameters μ. $\dim(x) = m \times 1$, $\dim(z) = n \times 1$, $\dim(H) = m \times n$. Let
$P_{t|t-1} \equiv E\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)'$
2. We are interested in estimating unknown parameters in vector μ based on a sample observations about $z^T \equiv \{z_t\}_{t=1}^T$
3. Define as ML estimates of the model the values of μ that maximize the likelihood associated with a particular sample of realizations of z over time.
4. Likelihood associated with particular realization of z at time t as $L\left(z_t | z^{t-1}\right)$.
Sequence of conditional likelihoods $\left\{L\left(z_t | z^{t-1}\right)\right\}_{t=1}^T$ is independent over time, thus
$$L\left(z^T\right) = \Pi_{t=1}^T L\left(z_t | z^{t-1}\right)$$
5. $L\left(z^T\right)$ is the likelihood of our model. But, how do we compute $L\left(z_t | z^{t-1}\right)$?

Writing Down the Likelihood: All Variables are Observed (1)

- All the variables in x are "observed" $(z = x)$. In that case (provided #shocks=#observables), calculation of the likelihood proceeds as in standard econometric textbooks.
- Conditional on $\left\{x_j\right\}_{j=1}^{t-1}$, note that the optimal forecast of x_t is given by

$$\widehat{x}_t = Fx_{t-1}$$

so that the error in predicting x_t is

$$\widehat{e}_t = x_t - Fx_{t-1}$$

conditional likelihood associated with a realization of x_t can be assessed as the likelihood assigned to $\widehat{e}_t$ by its pdf

$$L\left(x_t | x^{t-1}\right) = p_e\left(\widehat{e}_t\right)$$

Writing Down the Likelihood: All Variables are Observed (2)

- The likelihood evaluation begins by inserting x_1 into its unconditional distribution, which is $N\left(0, Q\right)$, hence

$$L\left(x_1|\mu\right) = \left(2\pi\right)^{-m/2}\left|Q^{-1}\right|^{1/2}\exp\left[-\frac{1}{2}\left(x_1' Q^{-1} x_1\right)\right]$$

then, for $t = 2, ..., T$, we have

$$L\left(x_t|\mu\right) = \left(2\pi\right)^{-m/2}\left|Q^{-1}\right|^{1/2}\exp\left[-\frac{1}{2}\left(x_t - Fx_{t-1}\right)' Q^{-1}\left(x_t - Fx_{t-1}\right)\right]$$

- Finally the sample likelihood is the product of the individual likelihoods.

$$L\left(x|\mu\right) = \Pi_{t=1}^{T} L\left(x_t|\mu\right)$$

Writing Down the Likelihood: Some Variables are not Observed(1)

- The more complicated case is when $n < m$. In this case, one need to find a way to infer the value of x from observations on z. Inferring the values of x is essential to calculate the likelihood of z.
- To do so, we use in practice a particular algorithm (the Kalman filter) which is used to produce assessment of the conditional probability $L\left(z_t|z^{t-1}\right)$ associated with the time-t observation z_t, given the history of past realizations $z^{t-1} \equiv \left\{z_j\right\}_{j=1}^{t-1}$.
- Hidden states and observables are described by a state space system that is perturbed at each point by Gaussian shocks with zero mean and known covariances.
- The next slides sketch a recursive version of the Kalman filtering problem that allows computing recursively the likelihood of a model.

Writing Down the Likelihood: Some Variables are not Observed (2)

Computation is recursive: start at time 0, where we can calculate the following:

- $z_{1|0}$: Conditional expectation of z_1 (the vector of unobservables that enter the computation of the likelihood) given observations on z_0
- z_1: Actual observations on z_1
- $x_{1|0}$: Conditional expectation of x_1 given observations on z_0
  At time zero, we want to derive the best estimate of x_1.
- The key question is how to estimate x_1 given $x_{1|0}$ and z_1.

  Using $x_t = Fx_{t-1} + Gv_t$ and $P_{t|t-1} \equiv E\left(\left(x_t - x_{t|t-1}\right)\left(x_t - x_{t|t-1}\right)'\right)$

$$
\begin{aligned}
x_{1|0} &= 0 \\
P_{1|0} &= FP_{1|0}F' + Q
\end{aligned}
$$

- This way, one can construct associated values for observables z, given by:

$$
\begin{aligned}
z_{1|0} &= H'x_{1|0} = 0 \\
\Omega_{1|0} &= E\left[\left(z_1 - z_{1|0}\right)\left(z_1 - z_{1|0}\right)'\right] = H'P_{1|0}H \rightarrow \text{innovation covariance}
\end{aligned}
$$

Writing Down the Likelihood: Some Variables are not Observed (3)

- The two objects above are used to compute the likelihood function of z_1, which is a normal variable with mean $z_{1|0}$ and variance $\Omega_{1|0}$, that is $z_1 \tilde{} N\left(z_{1|0}, \Omega_{1|0}\right)$

$$L\left(z_1|\mu\right) = (2\pi)^{-m/2} \left|\Omega_{1|0}^{-1}\right|^{1/2} \exp\left[-\frac{1}{2}\left(z_1'\Omega_{1|0}^{-1}z_1\right)\right].$$

- Next, the values of $x_{1|0}$ and $P_{1|0}$ are updated to construct new updates of $x_{1|1} \equiv x_1$ and $P_{1|1} \equiv P_1$.

$$\begin{aligned} x_{1|1} &= \underset{\text{old value}}{x_{1|0}} + \underset{\text{Kalman gain}}{P_{1|0}H\Omega_{1|0}^{-1}} \underset{\text{prediction error}}{\left(z_1 - z_{1|0}\right)} \rightarrow \text{updated state estimate} \\ P_{1|1} &= P_{1|0} - P_{1|0}H\Omega_{1|0}^{-1}H'P_{1|0} \rightarrow \text{updated covariance estimate} \end{aligned}$$

Writing Down the Likelihood: Some Variables are not Observed (4)

The term $K_1 \equiv P_{1|0} H \Omega_{1|0}^{-1}$ denotes the Kalman gain matrix. It is a minimum-mean square estimator that yields the best prediction of x_1 given estimates of $x_{1|0}$, z_1 and $z_{1|0}$. It can be derived as follows; consider the covariance matrix of x_1, $P_{1|1}$

$$P_{1|1} = E\left[(\underset{\text{actual}}{x_1} - \underset{\text{mean}}{x_{1|1}}) \left(x_1 - x_{1|1} \right)' \right]$$

$$\text{use } x_{1|1} = x_{1|0} + K_1 \left(z_1 - z_{1|0} \right) \text{ for some } K_1 \text{ to be determined}$$

$$P_{1|1} = cov\left(x_1 - x_{1|0} - K_1 \left(z_1 - z_{1|0} \right) \right) = P_{1|1} = cov\left(x_1 - x_{1|0} - K_1 H' \left(x_1 - x_{1|0} \right) \right)$$

$$P_{1|1} = E\left[\left(I - K_1 H' \right) P_{1|0} \left(I - K_1 H' \right)' \right]$$

Minimize the expected value of the square of the magnitude of this vector.

$$\min_{K_1} trace\left(P_{1|1} \right) \rightarrow K_1 = P_{1|0} H \Omega_{1|0}^{-1}$$

Writing Down the Likelihood: Some Variables are not Observed (4)

Next, for every other period $t = 2, ..., T$, we have:

$$
\begin{aligned}
x_{t|t-1} &= Fx_{t-1} \\
P_{t|t-1} &= FP_{t|t-1}F' + Q \\
z_{t|t-1} &= H'x_{t|t-1} \\
\Omega_{t|t-1} &= H'P_{t|t-1}H \\
L\left(z_t|\mu\right) &= (2\pi)^{-m/2}\left|\Omega_{t|t-1}^{-1}\right|^{1/2}\exp\left[-\frac{1}{2}\left(\left(z_t - z_{t|t-1}\right)'\Omega_{t|t-1}^{-1}\left(z_t - z_{t|t-1}\right)\right)\right] \\
x_{t|t} &= x_{t|t-1} + P_{t|t-1}H\Omega_{t|t-1}^{-1}\left(z_t - z_{t|t-1}\right) \\
P_{t|t} &= P_{t|t-1} - P_{t|t-1}H\Omega_{t|t-1}^{-1}H'P_{t|t-1} \\
L\left(z|\mu\right) &= \Pi_{t=1}^{T}L\left(z_t|\mu\right)
\end{aligned}
$$

Bayesian Estimation

- **Classical estimation**: parameters treated as fixed but unknown, and likelihood function is interpreted as sampling distribution from data. Realizations of z interpreted as one possible realizations from $L(z|\mu)$. Inferences on μ are statements regarding probabilities associated with particular realizations of z given μ.
- **Bayesian estimation**: observations on z treated as given. Make inferences about distribution of μ conditional on z. Probabilistic interpretation of μ allows incorporating judgements on μ through prior distribution $\pi(\mu)$.
- From the definition of joint probability, we have that:

$$p(\mu, z) = L(z|\mu)\pi(\mu)$$

reversing the role of μ and z gives

$$p(z, \mu) = P(\mu|z)\, p(z).$$

Solving for $P(\mu|z)$ gives

$$\underbrace{P(\mu|z)}_{\text{posterior}} = \frac{L(z|\mu)\pi(\mu)}{p(z)} \propto \underbrace{L(z|\mu)}_{\text{likelihood}}\ \underbrace{\pi(\mu)}_{\text{prior}}$$

where $p(z)$ is a constant from the point of view of the distribution for μ

Applications to DSGE Models

We consider three examples of estimation. Dynare files are on the course webpage.

1. The toy model
2. A model with a Phillips curve (identification)
3. A richer dynamic-new-keynesian model

Do not attempt anything more complicated than this until you have fully mastered these examples.

1. Toy Model

- The model is described by

$$y_t = e_t$$

  where e_t is an exogenous iid shock with zero mean and unknown variance σ^2. We want to estimate σ. In principle, we can impose a prior on the distribution of σ and combine it with information from the data on y.
- See `basic_estimation`

2. Phillips curve

- Model is

$$\begin{aligned} \pi_t &= \beta E_t \pi_{t+1} + k y_t \\ y_t &= \rho y_{t-1} + e_t, \;\; e_t \sim N\left(0, \sigma^2\right) \end{aligned}$$

- Assume our only obervable is π_t. The solution to our model takes the form

$$\begin{aligned} x_t &= F x_{t-1} + G v_t \\ z_t &= H' x_t \end{aligned}$$

$$\begin{aligned} x_t &= \begin{bmatrix} \pi_t & y_t \end{bmatrix}', z_t = [\pi_t], v_t = [e_t] \\ F &= \begin{bmatrix} 0 & \frac{\kappa\rho}{1-\beta\rho} \\ 0 & \rho \end{bmatrix}, G = \begin{bmatrix} \frac{\kappa}{1-\beta\rho} \\ 1 \end{bmatrix}, H = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{aligned}$$

- Note the rational expectations solution for π_t

$$\pi_t = \rho \pi_{t-1} + \frac{\kappa}{1-\beta\rho} \varepsilon_t$$

- Estimation will recover ρ, and only one parameter among β, σ_ε and κ. ML will fail to recover separately estimates of $\kappa, \beta, \sigma_\varepsilon$. (see `toy_pc`)

3. New Keynesian Model

- Consider the following log-linear model for output y, inflation π and the nominal interest rate R. iid shocks are g and u

$$\begin{aligned} y_t &= E_t y_{t+1} - R_t + E_t \pi_{t+1} + g_t \\ \pi_t &= k y_t + \beta E_t \pi_{t+1} + u_t \\ R_t &= \phi \pi_t \end{aligned}$$

- Solution is (`toy_dnk`)

$$y_t = \frac{1}{1+\phi k} g_t - \frac{\phi}{1+\phi k} u_t, \quad \pi_t = \frac{k}{1+\phi k} g_t + \frac{1}{1+\phi k} u_t$$

- Estimation cannot recover β, and at most three parameters among σ_g, σ_u, ϕ and k. To think about why, the series are iid, and all that enters the likelihood function is their variance and their covariance.

$$var\,(y) = \frac{1}{(1+\phi k)^2} \sigma_g^2 + \frac{\phi^2}{(1+\phi k)^2} \sigma_u^2; \quad var\,(\pi) = \frac{k^2}{(1+\phi k)^2} \sigma_g^2 + \frac{1}{(1+\phi k)^2} \sigma_u^2$$

$$cov\,(y, \pi) = \frac{k}{(1+\phi k)^2} \sigma_g^2 - \frac{\phi}{(1+\phi k)^2} \sigma_u^2$$

MCMC

- Key computational problem: how to compute the distribution of μ, $P(\mu|z)$: standard Monte Carlo integration techniques cannot be used, because one cannot draw random numbers directly from $P(\mu|z)$.
- Typically, we use Markov Chain Monte Carlo (MCMC) techniques, in particular the Metropolis-Hastings algorithm which is a particular version of the MCMC algorithm. The idea of the algorithm is to explore the distribution and to weigh to outcomes appropriately.

See Chapter 9 in Dejong and Dave for more details.

Basic Idea of the Kalman Filter

- Consider the following system

$$
\begin{aligned}
s_t &= F s_{t-1} + G\omega_t, \omega_t \sim N(0, Q) && \text{(state\_equation)} \\
y_t &= H s_t + v_t, v_t \sim N(0, R) && \text{(measurement\_equation)}
\end{aligned}
$$

  y_t is observed, everything else is not

- Problem: want to compute the following objects

  likelihood function of y^T

  recover from the y^T the sequences of s^T,ω^T, v^T

- Note that the above implies

$$
y_t = HFs_{t-1} + HG\omega_t + v_t
$$

  in order to compute y_t, we need to know s_{t-1}, ω_t, v_t
  The Kalman filter offers a way to go compute s, ω, v, and therefore to compute the likelihood of y^T.

- Inputs $(y_t, H, F, G) - - >$ Output $(\omega^T, s^T, v^T, Q, R)$

The Workings of the filter

- Suppose we are in t, and know s^{t-1} as well as its covariance matrix Σ_{t-1}. At that point, we are interested in computing an estimate of s_t given surprises in y_t
$$s_t = E_{t-1}s_t + K_t\left(y_t - E_{t-1}y_t\right)$$
- Note that
$$s_t = E_{t-1}s_t + K_t\left(y_t - HE_{t-1}s_t\right)$$
- Think of the best K_t as solving a minimization problem. It tells you how much you want to change your estimate given a measurement. Intuitively, the more y_t changes, the more likely it is that you want to change s_t (K_t large)
- So, how do we derive K_t

Deriving K

- Start with estimates of $E_{t-1}\Sigma_t$, $E_{t-1}s_t$, together with y_t. Then

$$
\begin{aligned}
K_t &= E_{t-1}\Sigma_t H' \left(H E_{t-1}\Sigma_t H' + R\right)^{-1} \\
\Sigma_t &= E_{t-1}\Sigma_t - K_t H E_{t-1}\Sigma_t \\
s_t &= E_{t-1}s_t + K_t \left(y_t - H E_{t-1}s_t\right) \\
E_t\Sigma_{t+1} &= F\Sigma_t F' + GQG' \\
E_t s_{t+1} &= F s_t
\end{aligned}
$$

- At this point, we are in $t+1$ and we can start again.
- Typically $E_{t-1}\Sigma_t = \Sigma^*$, $E_{t-1}s_t = \bar{s}$
- The formula for K_t is the outcome of a minimization problem where we try to minimize the squared forecast error of s_t