V. Nonlinear state-space models

A. Extended Kalman filter
B. Particle filter
C. Solution and estimation of nonlinear dynamic stochastic general equilibrium models
   1. Motivation
\( \mathbf{x}_t \) = vector of exogenous variables
\( \mathbf{\varepsilon}_t \) = vector of exogenous disturbances

\[ f(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{\varepsilon}_{t+1}) = \mathbf{0} \]

(equation of motion for \( \mathbf{x}_t \))

\( \mathbf{z}_t \) = vector of endogenous variables

\[ E_t \mathbf{a}(\mathbf{z}_{t+1}, \mathbf{z}_t, \mathbf{x}_t) = \mathbf{0} \]

(equations derived from econ theory)
Approach we discussed earlier:
(1) Log-linearize system.

\[ AE_t z_{t+1} = B z_t + C x_t \]

\[ x_{t+1} = \Phi x_t + \varepsilon_{t+1} \]
(2) Find rational-expectations solution.  

predetermined component: $\mathbf{z}_{1t}$

$$\mathbf{z}_{1,t+1} = \mathbf{H}_{11} \mathbf{z}_{1t} + \mathbf{H}_{12} \mathbf{x}_t$$

forward-looking component: $\mathbf{z}_{2t}$

$$\mathbf{z}_{2t} = \mathbf{H}_{21} \mathbf{z}_{1t} + \mathbf{H}_{22} \mathbf{x}_t$$
(3) Recognize as state-space system.

\[ y_t = \text{observed elements of } \{z_t, x_t\} \]
\[ \xi_t = \text{unobserved elements of } \{z_t, x_t\} \]
\[ \xi_{t+1} = \Phi \xi_t + v_{t+1} \]
\[ y_t = a + H'\xi_t + w_t \]

(4) Estimate parameters by MLE or Bayesian methods.
Things we lose from linearization:
(1) Statistical representation of recessions. Recall that a discrete Markov chain can be viewed as VAR(1).
Things we lose from linearization:

(2) Economic characterization of risk aversion.

\[ 1 = E_t \left[ \frac{\beta U'(c_{t+1})(1+r_{j,t+1})}{U'(c_t)} \right] \]

for \( r_{j,t+1} \) the real return on any asset.

Finance: different assets have different expected returns due to covariance between \( r_{j,t+1} \) and \( c_{t+1} \).
$$1 = E_t \left[ \frac{\beta U'(c_{t+1})(1+r_{j,t+1})}{U'(c_t)} \right]$$

steady state:

$$1 = \frac{\beta U'(c)(1+r_j)}{U'(c)}$$

$$\beta(1 + r_j) = 1 \text{ for all } j$$
linearization around steady state

\[ U'(c_t) = E_t[\beta U'(c_{t+1})(1 + r_{j,t+1})] \]

\[ \approx (1 + r)\beta U''(c)E_t(c_{t+1} - c) \]

\[ + \beta U'(c)E_t(r_{j,t+1} - r) \]

same for all \( j \)
Things we lose from linearization:
(3) Role of changes in uncertainty, time-varying volatility.
(4) Behavior of economy when interest rate is at zero lower bound
\[ R_t = \min(R_t^*, \bar{R}) \]
Approaches to estimating nonlinear dynamic economic models.

Step 1: Find approximating nonlinear state-space representation using either
   (1) perturbation methods (e.g., Fernandez-Villaverde and Rubio-Ramirez), or
   (2) projection methods (e.g., Gust, Lopez-Salido, and Smith)

Step 2: Estimate parameters using particle filter or other nonlinear estimation (MLE or Bayesian)
V. Nonlinear state-space models

A. Extended Kalman filter
B. Unscented Kalman filter
C. Particle filter
D. Nonlinear DSGE’s
   1. Motivation
   2. Perturbation methods
Example:

\[
\max_{\{c_t, k_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \log c_t
\]

s.t. \(c_t + k_{t+1} = e^{z_t} k_t^\alpha + (1 - \delta)k_t\quad t = 1, 2, \ldots\)

\(z_t = \rho z_{t-1} + \sigma \varepsilon_t\quad t = 1, 2, \ldots\)

\(k_0, z_0\) given

\(\varepsilon_t \sim N(0, 1)\)
Approach: we will consider a continuum of economies indexed by $\sigma$ and study solutions as $\sigma \to 0$ (that is, as economy becomes deterministic). We seek decision rules of the form

\[ c_t = c(k_t, z_t; \sigma) \]

\[ k_{t+1} = k(k_t, z_t; \sigma) \]
Write F.O.C. as \( E_t a(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0 \)
\[
a_1(k_t, z_t; \sigma, \varepsilon_{t+1}) = \frac{1}{c(k_t, z_t; \sigma)} - \\
\beta \frac{\alpha k(k_t, z_t; \sigma)^{\alpha-1} \exp(\rho z_t + \sigma \varepsilon_{t+1})}{c(k_t, z_t; \sigma, \rho z_t + \sigma \varepsilon_{t+1}; \sigma)} \]
\[
a_2(k_t, z_t; \sigma, \varepsilon_{t+1}) = c(k_t, z_t; \sigma) + k(k_t, z_t; \sigma) \\
- e^{z_t} k_t^\alpha - (1 - \delta) k_t
\]
Zero-order approximation
(deterministic steady state)

\[ \sigma = 0 \]

\[ z_t = z = 0 \]

\[ k_t = k \]

\[ a(k, 0; 0) = 0 \]
\begin{align*}
a_1(k, 0; 0) &= 0 \\
\Rightarrow \quad \frac{1}{c} - \beta \frac{\alpha k^{\alpha - 1}}{c} &= 0 \\
\Rightarrow \quad 1 &= \beta \alpha k^{\alpha - 1} \\

a_2(k, 0; 0) &= 0 \\
\Rightarrow \quad c + k - k^\alpha - (1 - \delta)k \\
\Rightarrow \quad c &= k^\alpha - \delta k
\end{align*}
First-order approximation:

Since $E_t a(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0$ for all $k_t, z_t; \sigma$, it follows that

$E_t a_k(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0$

for $a_k(k_t, z_t; \sigma, \varepsilon_{t+1}) = \frac{\partial a(k_t, z_t; \sigma, \varepsilon_{t+1})}{\partial k_t}$

likewise

$E_t a_z(k_t, z_t; \sigma, \varepsilon_{t+1}) = E_t a_\sigma(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0$
\[ E_t \left\{ \frac{\partial a_1(k_t, z_t; \sigma, \varepsilon_{t+1})}{\partial k_t} \bigg|_{k_t=k, z_t=0, \sigma=0} \right\} = \]
\[ -\frac{1}{c^2} c_k - \frac{\beta \alpha(\alpha-1)k^{\alpha-2}}{c} k_k + \frac{\beta \alpha k^{\alpha-1}}{c^2} c_k k_k \]

Since \( c \) and \( k \) are known from previous step, setting this to zero gives us an equation in the unknowns \( c_k \) and \( k_k \) where for example
\[ c_k = \frac{\partial c(k_t, z_t; \sigma)}{\partial k_t} \bigg|_{k_t=k, z_t=0, \sigma=0} \]
\[ \frac{\partial a_2(k_t, z_t; \sigma)}{\partial k_t} \bigg|_{k_t=k, z_t=0, \sigma=0} = \]

\[ c_k + k_k - \alpha k^{\alpha-1} - (1 - \delta) \]

This is a second equation in \( c_k, k_k \), which together with the first can now be solved for \( c_k, k_k \) as a function of \( c \) and \( k \).
\[ E_t \left\{ \frac{\partial a_1(k_t,z_t;\sigma,\varepsilon_{t+1})}{\partial z_t} \bigg|_{k_t=k,z_t=0,\sigma=0} \right\} = \]
\[ -\frac{1}{c^2} c_z - \frac{\beta \alpha (\alpha - 1) k^{\alpha - 2}}{c} k_z - \frac{\beta \alpha k^{\alpha - 1}}{c} \rho \]
\[ + \frac{\beta \alpha k^{\alpha - 1}}{c^2} (c_k k_z + \rho c_z) \]
\[ \frac{\partial a_2(k_t,z_t;\sigma)}{\partial z_t} \bigg|_{k_t=k,z_t=0,\sigma=0} = \]
\[ c_z + k_z - k^\alpha \]
setting these to zero allows us to solve for \( c_z, k_z \)
\[
\frac{\partial a_1(k_t,z_t;\sigma,\varepsilon_{t+1})}{\partial \sigma} \bigg|_{k_t=k,z_t=0,\sigma=0} = \\
-\frac{1}{c^2} c_\sigma - \frac{\beta \alpha (\alpha-1) k^{\alpha-2}}{c} k_\sigma - \frac{\beta \alpha k^{\alpha-1} \varepsilon_{t+1}}{c} \\
+ \frac{\beta \alpha k^{\alpha-1}}{c^2} \left( c_k k_\sigma + \varepsilon_{t+1} c_z + c_\sigma \right)
\]

\[
\frac{\partial a_2(k_t,z_t;\sigma)}{\partial \sigma} \bigg|_{k_t=k,z_t=0,\sigma=0} = \\
c_\sigma + k_\sigma
\]
Taking expectations and setting to zero yields

\[
\frac{-1}{c^2} c_\sigma - \frac{\beta \alpha (\alpha - 1) k^{\alpha - 2}}{c} k_\sigma \\
+ \frac{\beta \alpha k^{\alpha - 1}}{c^2} (c_k k_\sigma + c_\sigma) = 0 \\
c_\sigma + k_\sigma = 0
\]

which has solution \( c_\sigma = k_\sigma = 0 \)

\( \Rightarrow \) volatility, risk aversion play no role in first-order approximation
Now that we’ve calculated derivatives, we have the approximate solutions
\[ c(k_t, z_t; \sigma) \approx c + c_k(k_t - k) + c_{zz} z_t + c_{\sigma} \sigma \]
\[ k(k_t, z_t; \sigma) \approx k + k_k(k_t - k) + k_{zz} z_t + k_{\sigma} \sigma \]
where we showed that \( c_{\sigma} = k_{\sigma} = 0 \)
Thus, first-order perturbation is a way to find linearization or log-linearization
But we don’t have to stop here. Since 

\[ E_t a(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0 \]

for all \( k_t, z_t, \sigma \), second derivatives with respect to \((k_t, z_t; \sigma)\) also have to be zero.
Differentiate each of the 6 equations

\[ E_t a_k(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0 \]
\[ E_t a_z(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0 \]
\[ E_t a_\sigma(k_t, z_t; \sigma, \varepsilon_{t+1}) = 0 \]

with respect to \( k_t, z_t, \) and \( \sigma. \)
Gives 18 linear equations in the 12 unknowns
\[ \{c_{ij}, k_{ij}\} \quad i,j \in \{k,z,\sigma\} \]
with 6 equations redundant by symmetry of second derivatives (e.g., \( c_{kz} = c_{zk} \))
and where coefficients on \( c_{ij}, k_{ij} \) are known from previous step
We then have second-order approximation to decision functions,

\[
c(k_t, z_t; \sigma) \approx c + c'_1 s_t + (1/2)s_t' C_2 s_t
\]

\[
k(k_t, z_t; \sigma) \approx k + k'_1 s_t + (1/2)s_t' K_2 s_t
\]

\[
c'_1 = \begin{bmatrix}
  c_k & c_z & 0
\end{bmatrix}
\]

\[
s_t = \begin{bmatrix}
  (k_t - k) & z_t & \sigma
\end{bmatrix}
\]

\[
k'_1 = \begin{bmatrix}
  k_k & k_z & 0
\end{bmatrix}
\]
\[ \mathbf{C}_2 = \begin{bmatrix} c_{kk} & c_{k_z} & c_{k\sigma} \\ c_{z_k} & c_{zz} & c_{z\sigma} \\ c_{\sigma k} & c_{\sigma z} & c_{\sigma\sigma} \end{bmatrix} \]

\[ \mathbf{K}_2 = \begin{bmatrix} k_{kk} & k_{k_z} & k_{k\sigma} \\ k_{z_k} & k_{zz} & k_{z\sigma} \\ k_{\sigma k} & k_{\sigma z} & k_{\sigma\sigma} \end{bmatrix} \]
\[ c(k_t, z_t; \sigma) \approx c + c'_1 s_t + (1/2)s'_t C_2 s_t \]

\[ s_t = \begin{bmatrix} (k_t - k) & z_t & \sigma \end{bmatrix} \]

Note: term on \( \sigma^2 \) in \( s'_t C_2 s_t \) acts like another constant reflecting precautionary behavior left out of certainty-equivalence steady-state \( c \)
We could in principle continue to as high an order approximation as we wanted.
V. Nonlinear state-space models

D. Nonlinear DSGE’s
   1. Motivation
   2. Perturbation methods
   3. Illustration

Using a particle filter to estimate a DSGE with second-order perturbation approximation (Fernandez-Villaverde and Rubio-Ramirez, REStud, 2007)
\[ C_t + I_t = A_t K_t^K L_t^{1-a} \]
\[ K_{t+1} = (1 - \delta) K_t + U_t I_t \]
\[ \log A_t = \zeta + \log A_{t-1} + \sigma_{at} \epsilon_{at} \]
\[ \log U_t = \theta + \log U_{t-1} + \sigma_{vt} \epsilon_{vt} \]
\[ \log \sigma_{at} = (1 - \lambda_a) \log \bar{\sigma}_a \]
\[ + \lambda_a \log \sigma_{a,t-1} + \tau_a \eta_{at} \]
\[ \log \sigma_{vt} = (1 - \lambda_v) \log \bar{\sigma}_v \]
\[ + \lambda_v \log \sigma_{v,t-1} + \tau_v \eta_{vt} \]
$$E_0 \sum_{t=0}^{\infty} \beta^t \{ e^{dt} \log C_t + \psi \log (1 - L_t) \}$$

d_t = \rho d_{t-1} + \sigma_{dt} \epsilon_{dt}$$

$$\log \sigma_{dt} = (1 - \lambda_d) \log \overline{\sigma}_d + \lambda_d \log \sigma_{d,t-1} + \tau_d \eta_{dt}$$
\[ \mathbf{v}_t = (\varepsilon_{at}, \varepsilon_{vt}, \varepsilon_{dt}, \eta_{at}, \eta_{vt}, \eta_{dt})' \]

\[ \mathbf{v}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_6) \]

\[ \Omega = \text{diag}\{\bar{\sigma}_a^2, \bar{\sigma}_v^2, \bar{\sigma}_d^2, \tau_a^2, \tau_v^2, \tau_d^2\} \]

perturbation method: Continuum of economies with variance \( \chi \Omega \), take expansion around \( \chi = 0 \)
Transformations to find steady-state representation:

\[ Z_t = A_{t-1}^{1/(1-\alpha)} U_{t-1}^{\alpha/(1-\alpha)} \]
\[ \tilde{Y}_t = Y_t/Z_t, \quad \tilde{C}_t = C_t/Z_t, \quad \tilde{I}_t = I_t/Z_t \]
\[ \tilde{U}_t = U_t/U_{t-1}, \quad \tilde{A}_t = A_t/A_{t-1}, \quad \tilde{K}_t = K_t/Z_t U_{t-1} \]
\[ \tilde{k} = \log \text{ of steady-state value for } \tilde{K} \]
\[ \hat{k}_t = \log \tilde{K}_t - \tilde{k} \]
state vector for economic model:
\[ \tilde{s}_t = (\tilde{k}_t, \varepsilon_{at}, \varepsilon_{vt}, \varepsilon_{dt}, d_{t-1}, \sigma_{at} - \bar{\sigma}_a, \sigma_{vt} - \bar{\sigma}_v, \sigma_{dt} - \bar{\sigma}_d)' \]

second-order perturbation:
\[ \tilde{k}_{t+1} = \psi'_{k1} \tilde{s}_t + (1/2)\tilde{s}'_t \Psi_{k2} \tilde{s}_t + \psi_{k0} \]
\[ \tilde{i}_t = \psi'_{i1} \tilde{s}_t + (1/2)\tilde{s}'_t \Psi_{i2} \tilde{s}_t + \psi_{i0} \]
\[ \tilde{\ell}_t = \psi'_{\ell1} \tilde{s}_t + (1/2)\tilde{s}'_t \Psi_{\ell2} \tilde{s}_t + \psi_{\ell0} \]
\[ \psi_{j0} \text{ reflects precautionary effects} \]
However, we will observe actual GDP growth per capita
\[ \Delta \log Y_t = \Delta \log \tilde{Y}_t \]
\[ + \frac{1}{1-\alpha} (\Delta \log A_{t-1} + \alpha \Delta \log U_{t-1}) + \sigma_y \varepsilon_{yt} \]
\[ = h_y(\tilde{s}_t, \tilde{s}_{t-1}) + \sigma_y \varepsilon_{yt} \]
\[ \varepsilon_{yt} = \text{measurement error} \]
Also observe real gross investment per capita ($I_t$), hours worked per capita ($\ell_t$), and relative price of investment goods $P_t$

$$\Delta \log I_t = h_i(\tilde{s}_t, \tilde{s}_{t-1}) + \sigma_{i\varepsilon} \varepsilon_{it}$$

$$\log \ell_t = h_\ell(\tilde{s}_t, \tilde{s}_{t-1}) + \sigma_{\ell\varepsilon} \varepsilon_{\ell t}$$

$$\Delta \log P_t = -\Delta \log U_t$$
\[ \hat{s}_t = \left( k_t, \varepsilon_{at}, \varepsilon_{vt}, \varepsilon_{dt}, d_{t-1}, \sigma_{at} - \bar{\sigma}_a, \sigma_{vt} - \bar{\sigma}_v, \sigma_{dt} - \bar{\sigma}_d \right)'
\]
\[ v_t = (\varepsilon_{at}, \varepsilon_{vt}, \varepsilon_{dt}, \eta_{at}, \eta_{vt}, \eta_{dt})'
\]
\[ S_t = (\hat{s}_t', \hat{s}_{t-1}')
\]

state equation

\[ S_t = f(S_{t-1}, v_t)
\]
\[ f_1(S_{t-1}, v_t) = \psi_{k1}' \hat{s}_t + (1/2)\hat{s}_t' \Psi_{k2} \hat{s}_t + \psi_{k0}
\]
\begin{align*}
f_2(S_{t-1}, v_t) &= \varepsilon_{at} \\
\vdots \\
f_5(S_{t-1}, v_t) &= \rho d_{t-2} + \sigma_{d,t-1} \varepsilon_{d,t-1} \\
f_6(S_{t-1}, v_t) &= \exp[(1 - \lambda_a) \log \bar{\sigma}_a \\
&\quad + \lambda_a \log \sigma_{a,t-1} + \tau_a \eta_{at}] - \bar{\sigma}_a \\
\vdots \\
f_{9-16}(S_{t-1}, v_t) &= \tilde{s}_{t-1}
\end{align*}
\[ y_t = (\Delta \log Y_t, \Delta \log I_t, \log \ell_t, \Delta \log P_t)' \]

observation equation:
\[ y_t = h(S_t) + w_t \]
According to the set-up, $\varepsilon_{vt}$ is observed directly from the change in investment price each period

\[ \log U_t = \theta + \log U_{t-1} + \sigma_{vt} \varepsilon_{vt} \]

\[ \Delta \log P_t = -\Delta \log U_t \]
We only need to generate a draw for

\[ \mathbf{v}_{1t} = (\varepsilon_{at}, \varepsilon_{dt}, \eta_{at}, \eta_{vt}, \eta_{dt})' \]

in order to have a value for \( \sigma_{vt} \) and value for \( \varepsilon_{vt} \)

\[ \varepsilon_{vt} = -\frac{\Delta \log P_t + \theta}{\sigma_{vt}} \]
Initialization:

\[ S_t = f(S_{t-1}, v_t) \]

One approach is to set

\[ S_{-N} = 0, \text{ draw } v_{-N+1}, v_{-N+2}, \ldots, v_0 \]

from \( N(0, I_6) \) to obtain \( D \) draws (particles) for \( \{ S_0^{(i)} \}_{i=1}^D \)
Estimation using bootstrap particle filter

As of date $t$ we have calculated a set

$$\Lambda_{t}^{(i)} = \{S_{t}^{(i)}, S_{t-1}^{(i)}, \ldots, S_{0}^{(i)}\}$$

for $i = 1, \ldots, D$

To update for $t + 1$ we do the following:
Step 1: generate $v_{1,t+1}^{(i)} \sim N(0, I_5)$ for $i = 1, \ldots, D$

Step 2: generate $S_{t+1}^{(i)} = f(S_t^{(i)}, v_{t+1}^{(i)})$

except for the third element $\varepsilon_{v,t+1}^{(i)}$

Step 3: calculate

$w_{t+1}^{(i)} = y_{t+1} - h(S_{t+1}^{(i)})$

and set third element of $S_{t+1}^{(i)}$ equal to fourth element of $w_{t+1}^{(i)}$, $\varepsilon_{v,t+1}^{(i)} = -\frac{\Delta \log P_{t+1} + \theta}{\sigma_{v,t+1}^{(i)}}$
Step 4: calculate

\[ \tilde{\omega}_{t+1}^{(i)} = (2\pi)^{-4/2} |\mathbf{D}_{t+1}^{(i)}| \]

\[ \times \exp\left( -(1/2) \begin{bmatrix} \mathbf{w}_{t+1}^{(i)} \end{bmatrix} [\mathbf{D}_{t+1}^{(i)}]^{-1} \begin{bmatrix} \mathbf{w}_{t+1}^{(i)} \end{bmatrix} \right) \]

\[ \mathbf{D}_{t+1}^{(i)} = \begin{bmatrix}
\sigma_{y\varepsilon}^2 & 0 & 0 & 0 \\
0 & \sigma_{i\varepsilon}^2 & 0 & 0 \\
0 & 0 & \sigma_{\ell\varepsilon}^2 & 0 \\
0 & 0 & 0 & \left[ \sigma_{v,t+1}^{(i)} \right]^2
\end{bmatrix} \]
Step 5: Contribution to likelihood is

\[ \hat{p}(y_{t+1} | \Omega_t) = D^{-1} \sum_{i=1}^{D} \tilde{\omega}_{t+1}^{(i)} = \overline{\omega}_{t+1} \]

Step 6: Calculate \( \hat{\omega}_{t+1} = \tilde{\omega}_{t+1}^{(i)}/\overline{\omega}_{t+1} \)

and resample

\[ \Lambda_{t+1}^{(j)} = \begin{cases} 
\Lambda_{t+1}^{(1)} & \text{with probability } \hat{\omega}_{t+1}^{(1)} \\
\vdots & \\
\Lambda_{t+1}^{(D)} & \text{with probability } \hat{\omega}_{t+1}^{(D)} 
\end{cases} \]
Structural parameters:
\[ \theta = (\alpha, \delta, \rho, \beta, \psi, \theta, \zeta, \tau_a, \tau_v, \tau_d, \overline{\sigma}_a, \overline{\sigma}_v, \overline{\sigma}_d, \lambda_a, \lambda_v, \lambda_d, \sigma_{y\varepsilon}, \sigma_{i\varepsilon}, \sigma_{\ell\varepsilon})' \]

Fernandez-Villaverde and Rubio-Ramirez estimate \( \theta \) by maximizing
\[ \hat{\mathcal{L}}(\theta) = \sum_{t=1}^{T} \hat{p}(y_t|\Omega_{t-1};\theta) \]