Solution Methods

Jesús Fernández-Villaverde

University of Pennsylvania

March 16, 2016

Functional equations

- A large class of problems in macroeconomics search for a function d that solves a *functional equation*:

$$\mathcal{H}(d) = \mathbf{0}$$

- More formally:

  1. Let J^1 and J^2 be two functional spaces and let $\mathcal{H} : J^1 \rightarrow J^2$ be an operator between these two spaces.
  2. Let $\Omega \subseteq \mathbb{R}^l$.
  3. Then, we need to find a function $d : \Omega \rightarrow \mathbb{R}^m$ such that $\mathcal{H}(d) = \mathbf{0}$.

- Notes:

  1. Regular equations are particular examples of functional equations.
  2. $\mathbf{0}$ is the space zero, different in general that the zero in the reals.

Example I

- Let's go back to our basic stochastic neoclassical growth model:

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t, l_t\right)$$

$$c_t + k_{t+1} = e^{z_t} A k_t^{\alpha} l_t^{1-\alpha} + \left(1-\delta\right) k_t, \ \forall \ t > 0$$

$$z_t = \rho z_{t-1} + \sigma \varepsilon_t, \ \varepsilon_t \sim \mathcal{N}(0,1)$$

$$\log \sigma_t = \left(1-\rho_\sigma\right) \log \sigma + \rho_\sigma \log \sigma_{t-1} + \left(1-\rho_\sigma^2\right)^{\frac{1}{2}} \eta u_t$$

- The first order condition:

$$u'\left(c_t, l_t\right) = \beta \mathbb{E}_t \left\{ u'\left(c_{t+1}, l_{t+1}\right) \left(1 + \alpha e^{z_{t+1}} A k_{t+1}^{\alpha-1} l_{t+1}^{1-\alpha} - \delta\right) \right\}$$

- Where is the stochastic volatility?

Example II

- Define:

$$x_t = (k_t, z_{t-1}, \log \sigma_{t-1}, \varepsilon_t, u_t)$$

There is a decision rule (a.k.a. policy function) that gives the optimal choice of consumption and capital tomorrow given the states today:

$$d = \begin{cases} d^1(x_t) = l_t \\ d^2(x_t) = k_{t+1} \end{cases}$$

- From these two choices, we can find:

$$c_t = e^{z_t} A k_t^{\alpha} \left(d^1(x_t) \right)^{1-\alpha} + (1-\delta) k_t - d^2(x_t)$$

Example III

- Then:

$$\mathcal{H} = u'\left(c_t, d^1\left(x_t\right)\right)$$
$$-\beta\mathbb{E}_t \left\{ \begin{array}{c} u'\left(c_{t+1}, d^1\left(x_{t+1}\right)\right) * \\ \left(1 + \alpha e^{z_{t+1}} A d^2\left(x_t\right)^{\alpha-1} d^1\left(x_{t+1}\right)^{1-\alpha} - \delta\right) \end{array} \right\} = 0$$

- If we find g, and a transversality condition is satisfied, we are done!

Example IV

- There is a recursive problem associated with the previous sequential problem:

$$V(x_t) = \max_{k_{t+1}, l_t} \left\{ u(c_t, l_t) + \beta \mathbb{E}_t V(x_{t+1}) \right\}$$

$$c_t + k_{t+1} = e^{z_t} A k_t^{\alpha} l_t^{1-\alpha} + (1-\delta) k_t, \ \forall \ t > 0$$

$$z_t = \rho z_{t-1} + \sigma \varepsilon_t, \ \varepsilon_t \sim \mathcal{N}(0,1)$$

$$\log \sigma_t = (1-\rho_\sigma) \log \sigma + \rho_\sigma \log \sigma_{t-1} + \left(1-\rho_\sigma^2\right)^{\frac{1}{2}} \eta u_t$$

- Then:

$$d(x_t) = V(x_t)$$

and

$$\widetilde{\mathcal{H}}(d) = d(x_t) - \max_{k_{t+1}, l_t} \left\{ u(c_t, l_t) + \beta \mathbb{E}_t d(x_{t+1}) \right\} = \mathbf{0}$$

How do we solve functional equations?

- General idea: substitute $d(x)$ by $d^n(x,\theta)$ where θ is an $n-$dim vector of coefficients to be determined.

- Two Main Approaches:

1. Perturbation methods:

$$d^n(x,\theta) = \sum_{i=0}^{n} \theta_i (x - x_0)^i$$

We use implicit-function theorems to find θ_i.

2. Projection methods:

$$d^n(x,\theta) = \sum_{i=0}^{n} \theta_i \Psi_i(x)$$

We pick a basis $\{\Psi_i(x)\}_{i=0}^{\infty}$ and "project" $\mathcal{H}(\cdot)$ against that basis.

Comparison with traditional solution methods

- Linearization (or loglinearization): equivalent to a first-order perturbation.

- Linear-quadratic approximation: equivalent (under certain conditions) to a first-order perturbation.

- Parameterized expectations: a particular example of projection.

- Value function iteration:it can be interpreted as an iterative procedure to solve a particular projection method. Nevertheless, I prefer to think about it as a different family of problems.

- Policy function iteration: similar to VFI.

Advantages of the functional equation approach

- Generality: abstract framework highlights commonalities across problems.
- Large set of existing theoretical and numerical results in applied math.
- It allows us to identify more clearly issue and challenges specific to economic problems (for example, importance of expectations).
- It allows us to deal efficiently with nonlinearities.

Perturbation: motivation

- Perturbation builds a Taylor-series approximation of the exact solution.
- Very accurate around the point where the approximation is undertakes.
- Often, surprisingly good global properties.
- Only approach that handle models with dozens of state variables.
- Relation between uncertainty shocks and the curse of dimensionality.

References

- General:
  1. *A First Look at Perturbation Theory* by James G. Simmonds and James E. Mann Jr.
  2. *Advanced Mathematical Methods for Scientists and Engineers: Asymptotic Methods and Perturbation Theory* by Carl M. Bender, Steven A. Orszag.
- Economics:
  1. Perturbation Methods for General Dynamic Stochastic Models" by Hehui Jin and Kenneth Judd.
  2. Perturbation Methods with Nonlinear Changes of Variables" by Kenneth Judd.
  3. A gentle introduction: "Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function" by Martín Uribe and Stephanie Schmitt-Grohe.

Our RBC-SV model

- Let me come back to our RBC-SV model.
- Three changes:
  1. Eliminate labor supply and have a log utility function for consumption.
  2. Full depreciation.
  3. $A = 1$.
- Why?

Environment

$$\max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \log c_t$$

$$\text{s.t. } c_t + k_{t+1} = e^{z_t} k_t^{\alpha}, \ \forall \ t > 0$$

$$z_t = \rho z_{t-1} + \sigma \varepsilon_t$$

$$\log \sigma_t = (1 - \rho_\sigma) \log \sigma + \rho_\sigma \log \sigma_{t-1} + \left(1 - \rho_\sigma^2\right)^{\frac{1}{2}} \eta u_t$$

- Equilibrium conditions:

$$\frac{1}{c_t} = \beta \mathbb{E}_t \frac{1}{c_{t+1}} \alpha e^{z_{t+1}} k_{t+1}^{\alpha - 1}$$

$$c_t + k_{t+1} = e^{z_t} k_t^{\alpha}$$

$$z_t = \rho z_{t-1} + \sigma \varepsilon_t$$

$$\log \sigma_t = (1 - \rho_\sigma) \log \sigma + \rho_\sigma \log \sigma_{t-1} + \left(1 - \rho_\sigma^2\right)^{\frac{1}{2}} \eta u_t$$

Solution

- Exact solution (found by "guess and verify"):

$$c_t = (1 - \alpha\beta)\, e^{z_t} k_t^{\alpha}$$
$$k_{t+1} = \alpha\beta e^{z_t} k_t^{\alpha}$$

- Note that this solution is the same than in the model without stochastic volatility.

- Intuition.

- However, the dynamics of the model is affected by the law of motion for z_t.

Steady state

- Steady state is also easy to find:

$$
\begin{aligned}
k &= (\alpha\beta)^{\frac{1}{1-\alpha}} \\
c &= (\alpha\beta)^{\frac{\alpha}{1-\alpha}} - (\alpha\beta)^{\frac{1}{1-\alpha}} \\
z &= 0 \\
\log\sigma &= \log\sigma
\end{aligned}
$$

- Steady state in more general models.

The goal

- Define $x_t = (k_t, z_{t-1}, \log \sigma_{t-1}, \varepsilon_t, u_t; \lambda)$.
- Role of λ: perturbation parameter.
- We are searching for decision rules:

$$d = \begin{cases} c_t = c(x_t) \\ k_{t+1} = k(x_t) \end{cases}$$

- Then, we have a system of functional equations:

$$\frac{1}{c(x_t)} = \beta \mathbb{E}_t \frac{1}{c(x_{t+1})} \alpha e^{z_{t+1}} k(x_t)^{\alpha-1}$$

$$c(x_t) + k(x_t) = e^{z_t} k_t^{\alpha}$$

$$z_t = \rho z_{t-1} + \sigma_t \lambda \varepsilon_t$$

$$\log \sigma_t = (1-\rho_\sigma) \log \sigma + \rho_\sigma \log \sigma_{t-1} + \left(1-\rho_\sigma^2\right)^{\frac{1}{2}} \eta \lambda u_t$$

Taylor's theorem

- We will build a local approximation around $x = (k, 0, \log \sigma, 0, 0; \lambda)$.
- Given equilibrium conditions:

$$\mathbb{E}_t \left(\frac{1}{c(x_t)} - \beta \frac{1}{c(x_{t+1})} \alpha e^{z_{t+1}} k(x_t)^{\alpha-1} \right) = 0$$
$$c(x_t) + k(x_t) - e^{z_t} k_t^{\alpha} = 0$$

We will take derivatives with respect to $(k, z, \log \sigma, \varepsilon, u; \lambda)$ and evaluate them around $x = (k, 0, \log \sigma, 0, 0; \lambda)$.

- Why?

- Apply Taylor's theorem and a version of the implicit-function theorem.

Taylor series expansion I

$$
\begin{aligned}
c_t = & \; c\left(k_t, z_{t-1}, \log \sigma_{t-1}, \varepsilon_t, u_t; 1\right)\big|_{k,0,0} = c\left(x\right) \\
& + c_k\left(x\right)\left(k_t - k\right) + c_z\left(x\right) z_{t-1} + c_{\log \sigma}\left(x\right) \log \sigma_{t-1} + \\
& + c_\varepsilon\left(x\right) \varepsilon_t + c_u\left(x\right) u_t + c_\lambda\left(x\right) \\
& + \frac{1}{2} c_{kk}\left(x\right)\left(k_t - k\right)^2 + \frac{1}{2} c_{kz}\left(x\right)\left(k_t - k\right) z_{t-1} + \\
& + \frac{1}{2} c_{k \log \sigma}\left(x\right)\left(k_t - k\right) \log \sigma_{t-1} + \ldots
\end{aligned}
$$

Taylor series expansion II

$$
\begin{aligned}
k_t = \; & k\left(k_t, z_{t-1}, \log \sigma_{t-1}, \varepsilon_t, u_t; 1\right)\big|_{k,0,0} = k\left(x\right) \\
& + k_k\left(x\right)\left(k_t - k\right) + k_z\left(x\right) z_{t-1} + k_{\log \sigma}\left(x\right) \log \sigma_{t-1} + \\
& + k_\varepsilon\left(x\right) \varepsilon_t + k_u\left(x\right) u_t + k_\lambda\left(x\right) \\
& + \frac{1}{2} k_{kk}\left(x\right)\left(k_t - k\right)^2 + \frac{1}{2} k_{kz}\left(x\right)\left(k_t - k\right) z_{t-1} + \\
& + \frac{1}{2} k_{k \log \sigma}\left(x\right)\left(k_t - k\right) \log \sigma_{t-1} + \ldots
\end{aligned}
$$

Comment on notation

- From now on, to save on notation, I will write

$$F(x_t) = \mathbb{E}_t \begin{bmatrix} \frac{1}{c(x_t)} - \beta \frac{1}{c(x_{t+1})} \alpha e^{z_{t+1}} k(x_t)^{\alpha-1} \\ c(x_t) + k(x_t) - e^{z_t} k_t^{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- I will take partial derivatives of $F(x_t)$ and evaluate them at the steady state

- Do these derivatives exist?

First-order approximation

- We take first-order derivatives of $F(x_t)$ around x.
- Thus, we get:

$$DF(x) = 0$$

- A matrix quadratic system.
- Why quadratic? Stable and unstable manifold.
- However, it has a nice recursive structure that we can and should exploit.

Solving the system II

- Procedures to solve quadratic systems:

  1. Blanchard and Kahn (1980) .
  2. Uhlig (1999).
  3. Sims (2000).
  4. Klein (2000).

- All of them equivalent.

Properties of the first-order solution

- Coefficients associated with λ are zero.
- Coefficients associated with $\log \sigma_{t-1}$ are zero.
- Coefficients associated with u_t are zero.
- In fact, up to first-order, stochastic volatility is irrelevant.
- This result recovers traditional macroeconomic approach to fluctuations.

Interpretation

- No precautionary behavior.
- Difference between risk-aversion and precautionary behavior. Leland (1968), Kimball (1990).
- Risk-aversion depends on the second derivative (concave utility).
- Precautionary behavior depends on the third derivative (convex marginal utility).

Second-order approximation

- We take first-order derivatives of $F(x_t)$ around x.
- Thus, we get:
$$D^2 F(x) = 0$$
- We substitute the coefficients that we already know.
- A matrix linear system.
- It also has a recursive structure, but now it is less crucial to exploit it.

Properties of the second-order solution

- Coefficients associated with λ are zero, but not coefficients associated with λ^2.
- Coefficients associated with $(\log \sigma_{t-1})^2$ are zero, but not coefficients associated with $\varepsilon_t \log \sigma_{t-1}$.
- Coefficients associated with u_t^2 are zero, but not coefficients associated with $\varepsilon_t u_t$.
- Thus, up to second-order, stochastic volatility matters.
- However, we cannot compute impulse-response functions to volatility shocks.

Third-order approximation

- We take third-order derivatives of $F(x_t)$ around x.
- Thus, we get:

$$D^3 F(x) = 0$$

- We substitute the coefficients that we already know.
- Still a matrix linear system with a recursive structure.
- Memory management considerations.

Computer

- In practice you do all this approximations with a computer:
  1. First-,second-, and third- order: `Dynare`.
  2. Higher order: `Mathematica`, `Dynare++`.
- Burden: analytical derivatives.
- Why are numerical derivatives a bad idea?
- Alternatives: automatic differentiation?

An alternative: projection

- Remember that we want to solve a functional equations of the form:

$$\mathcal{H}(d) = \mathbf{0}$$

for an unknown decision rule d.

- Projection methods solve the problem by specifying:

$$d^n(x, \theta) = \sum_{i=0}^{n} \theta_i \Psi_i(x)$$

We pick a basis $\{\Psi_i(x)\}_{i=0}^{\infty}$ and "project" $\mathcal{H}(\cdot)$ against that basis to find the θ_i's.

- We work with linear combinations of basis functions because theory of nonlinear approximations is not as developed as the linear case.

Algorithm

1. Define n known linearly independent functions $\psi_i : \Omega \to \Re^m$, where $n < \infty$. We call the $\psi_0(\cdot), \psi_2(\cdot), ..., \psi_n(\cdot)$ the *basis functions*.
2. Define a vector of coefficients $\theta = [\theta_0, \theta_\cdot, ..., \theta_n]$.
3. Define a combination of the basis functions and the θ's:

$$d^n(\cdot|\theta) = \sum_{i=0}^{n} \theta_i \psi_n(\cdot)$$

4. Plug $d^n(\cdot|\theta)$ into $H(\cdot)$ to find the *residual equation*:

$$R(\cdot|\theta) = \mathcal{H}(d^n(\cdot|\theta))$$

5. Find $\widehat{\theta}$ that make the residual equation as close to $\mathbf{0}$ as possible given some objective function $\rho : J^1 \times J^1 \to J^2$:

$$\widehat{\theta} = \arg\min_{\theta \in \Re^n} \rho(R(\cdot|\theta), \mathbf{0})$$

Models with heterogeneous agents

- Obviously, we cannot review here all the literature on solution methods for models with heterogeneous agents.
- Particular example of Bloom et al. (20012)
- Based on Krusell and Smith (1998) and Kahn and Thomas (2008)
- FOCs:

$$w_t = \phi c_t n_t^{\xi}$$
$$m_t = \beta \frac{c_t}{c_{t+1}}$$

Original formulation of the recursive problem

- Remember:

$$
\begin{aligned}
& V\left(k, n_{-1}, z; A, \sigma^{Z}, \sigma^{A}, \Phi\right) \\
= \max_{i,n} & \left\{ \begin{array}{c} y - w\left(A, \sigma^{Z}, \sigma^{A}, \Phi\right) n - i \\ -AC^{k}\left(k, k'\right) - AC^{n}\left(n_{-1}, n\right) \\ +\mathbb{E} m V\left(k', n, z'; A', \sigma^{Z\prime}, \sigma^{A\prime}, \Phi'\right) \end{array} \right\} \\
& \text{s.t. } \Phi' = \Gamma\left(A, \sigma^{Z}, \sigma^{A}, \Phi\right)
\end{aligned}
$$

Alternative formulation of the recursive problem

- Then, we can rewrite:

$$\widetilde{V}\left(k, n_{-1}, z; A, \sigma^{Z}, \sigma^{A}, \Phi\right)$$

$$= \max_{i,n} \left\{ \begin{array}{c} \frac{1}{c}\left(y - \phi c n^{1+\xi} - i - AC^{k}\left(k, k'\right) - AC^{n}\left(n_{-1}, n\right)\right) \\ +\beta \mathbb{E} \widetilde{V}\left(k', n, z'; A', \sigma^{Z\prime}, \sigma^{A\prime}, \Phi'\right) \end{array} \right\}$$

$$\text{s.t. } \Phi^{m\prime} = \Gamma\left(A, \sigma^{Z}, \sigma^{A}, \Phi\right)$$

where

$$\widetilde{V} = \frac{1}{c} V$$

- We can apply a version of K-S algorithm.

K-S algorithm I

- We approximate:

$$\widetilde{V}\left(k, n_{-1}, z; A, \sigma^Z, \sigma^A, \Phi^m\right)$$

$$= \max_{i,n} \left\{ \begin{array}{c} \frac{1}{c}\left(y - \phi c n^{1+\xi} - i - AC^k\left(k, k'\right) - AC^n\left(n_{-1}, n\right)\right) \\ +\beta \mathbb{E} \widetilde{V}\left(k', n, z'; A', \sigma^{Z\prime}, \sigma^{A\prime}, \Phi^{m\prime}\right) \end{array} \right\}$$

$$\text{s.t. } \Phi^{m\prime} = \Gamma\left(A, \sigma^Z, \sigma^A, \Phi^m\right)$$

- Forecasting rules for $\frac{1}{c}$ and Γ.
- Since they are aggregate rules, a common guess is of the form:

$$\begin{aligned} \log \frac{1}{c} &= \alpha_1\left(A, \sigma^Z, \sigma^A\right) + \alpha_2\left(A, \sigma^Z, \sigma^A\right) K \\ \log K' &= \alpha_3\left(A, \sigma^Z, \sigma^A\right) + \alpha_4\left(A, \sigma^Z, \sigma^A\right) K \end{aligned}$$

K-S algorithm II

1. Guess initial values for $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$.
2. Given guessed forecasting rules, solve value function of an individual firm.
3. Simulate the economy for a large number of periods, computing $\frac{1}{c}$ and K'.
4. Use regression to update forecasting rules.
5. Iterate until convergence

Extensions

- I presented the plain vanilla K-S algorithm.
- Many recent developments.
- Check Algan, Allais, Den Haan, and Rendahl (2014).