Nonlinear Solution Methods in Economics

(Lectures on Solution Methods for Economists III)

Jesús Fernández-Villaverde[1] and Pablo Guerrón[2]

March 8, 2022

[1]University of Pennsylvania

[2]Boston College

Introduction

Functional equations

- A large class of problems in economics 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 $\mathcal{H} : J^1 \to J^2$ be an operator between these two spaces.
  2. Let $\Omega \subseteq \mathbb{R}^l$.
  3. Then, we need to find a function $d : \Omega \to \mathbb{R}^m$ 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: decision rules

- Take the basic stochastic neoclassical growth model:

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

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

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

- The first order condition:

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

Example I: decision rules

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

$$d = \begin{cases} d^1(k_t, z_t) = c_t \\ d^2(k_t, z_t) = k_{t+1} \end{cases}$$

- Then:

$$\mathcal{H} = u'\left(d^1(k_t, z_t)\right)$$
$$-\beta \mathbb{E}_t \left\{ u'\left(d^1\left(d^2(k_t, z_t), z_{t+1}\right)\right)\left(1 + \alpha e^{z_{t+1}}\left(d^2(k_t, z_t)\right)^{\alpha-1} - \delta\right)\right\} = 0$$

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

Example II: conditional expectations

- Let us go back to our Euler equation:

$$u'\left(c_t\right)-\beta\mathbb{E}_t\left\{u'\left(c_{t+1}\right)\left(1+\alpha e^{z_{t+1}}k_{t+1}^{\alpha-1}-\delta\right)\right\}=0$$

- Define now:

$$d=\begin{cases} d^1\left(k_t,z_t\right)=c_t \\ d^2\left(k_t,z_t\right)=\mathbb{E}_t\left\{u'\left(c_{t+1}\right)\left(1+\alpha e^{z_{t+1}}k_{t+1}^{\alpha-1}-\delta\right)\right\}\end{cases}$$

- Why? Example: ZLB.
- Then:

$$\mathcal{H}\left(d\right)=u'\left(d^1\left(k_t,z_t\right)\right)-\beta d^2\left(k_t,z_t\right)=\mathbf{0}$$

Example III: value functions

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

$$V(k_t, z_t) = \max_{k_{t+1}} \{u(c_t) + \beta \mathbb{E}_t V(k_{t+1}, z_{t+1})\}$$

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

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

- Then:

$$d(k_t, z_t) = V(k_t, z_t)$$

and

$$\mathcal{H}(d) = d(k_t, z_t) - \max_{k_{t+1}} \{u(c_t) + \beta \mathbb{E}_t d(k_{t+1}, z_{t+1})\} = \mathbf{0}$$

How do we solve functional equations?

- General idea: substitute $d(x)$ by $d^n(x, \theta)$ where θ is an $n - \text{dim}$ 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 to the utility function: 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.

Nonlinearities

- Most dynamic models are nonlinear.
- Common practice: solve and estimate a linearized version with Gaussian shocks.
- Aruoba, Fernández-Villaverde, Rubio-Ramírez, 2005: stochastic neoclassical growth model is nearly linear for the benchmark calibration.
- However, we want to depart from this basic framework.
- We will present three examples.

Three Examples

Example I: recursive preferences

- Recursive preferences (Kreps-Porteus-Epstein-Zin-Weil) have become a popular way to account for asset pricing observations.
- Natural separation between IES and risk aversion.
- Example of a more general set of preferences in macroeconomics.
- Consequences for business cycles, welfare, and optimal policy design.
- Link with robust control and with preference for the timing of revelation of uncertainty.

Model

- Basic stochastic neoclassical growth model with recursive preferences

$$U_t = \left[c_t^{\frac{1-\gamma}{\theta}} + \beta \underbrace{\left(\mathbb{E}_t U_{t+1}^{1-\gamma} \right)^{\frac{1}{\theta}}}_{\text{Risk-adjustment operator}} \right]^{\frac{\theta}{1-\gamma}}$$

$$\begin{aligned} c_t + k_{t+1} &= e^{z_t} k_t^{\alpha} + (1-\delta) k_t, \ \forall \ t > 0 \\ z_t &= \rho z_{t-1} + \sigma \varepsilon_t, \ \varepsilon_t \sim \mathcal{N}(0,1) \end{aligned}$$

where:

$$\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}.$$

Approximating the solution

The Term Structure of Interest Rates in a DSGE Model with Recursive Preferences.

1. **None** of the terms in the first-order approximation depend on γ.

2. **None** of the terms in the second-order approximation depend on γ, except for constants that captures precautionary behavior.

3. In the third-order approximation, we have time-varying terms that depend on γ.

Important points

- Moreover:

  1. Cubic terms are quantitatively important.

  2. The mean of the ergodic distributions of the endogenous variables and the deterministic steady state values are quite different. Key for calibration.

The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications.

Example II: volatility shocks

- Widespread evidence of time-varying volatility in time series.
- Basic stochastic neoclassical growth model with recursive preferences

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

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

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

$$\log \sigma_t = (1 - \rho_{\sigma_r}) \log \sigma_r + \rho_\sigma \log \sigma_{t-1} + \eta_r u_t, \ u_t \sim \mathcal{N}(0,1)$$

- Risk Matters: The Real Effects of Volatility Shocks.
- Fiscal Volatility Shocks.

Solving the model

- We are interested on the effects of a volatility increase: a positive shock to u_t while $\varepsilon_t = 0$.
- We need to obtain a *third* approximation of the policy functions:

1. A first-order approximation satisfies a certainty equivalence principle. Only level shocks ε_t appear.
2. A second-order approximation only captures volatility indirectly via cross products $\varepsilon_t u_t$.
3. In the third order, volatility shocks, u_t enter as independent arguments.

Example III: zero lower bound

- Nonlinear Adventures at the ZLB.
- Representative household

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \left(\prod_{i=0}^{t} \beta_i \right) \left\{ \log c_t - \psi \frac{l_t^{1+\vartheta}}{1+\vartheta} \right\}$$

where

$$\beta_{t+1} = \beta^{1-\rho_b} \beta_t^{\rho_b} \exp\left(\sigma_b \varepsilon_{b,t+1}\right), \ \varepsilon_{b,t+1} \sim \mathcal{N}(0,1)$$

$$c_t + \frac{b_{t+1}}{p_t} = w_t l_t + R_{t-1} \frac{b_t}{p_t} + T_t + F_t$$

- Final good producer:

$$y_t = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

Example III: intermediate firm

- Technology

$$y_{it} = A_t l_{it}$$

where:

$$A_t = A^{1-\rho_a} A_{t-1}^{\rho_a} \exp\left(\sigma_a \varepsilon_{a,t}\right), \; \varepsilon_{a,t} \sim \mathcal{N}(0,1)$$

- Calvo pricing without indexation:

$$\max_{p_{it}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \theta^{\tau} \left(\prod_{i=0}^{\tau} \beta_{t+i} \right) \frac{\lambda_{t+\tau}}{\lambda_t} \left(\frac{p_{it}}{p_{t+\tau}} - mc_{t+\tau} \right) y_{it+\tau}$$

$$\text{s.t. } \; y_{it} = \left(\frac{p_{it}}{p_t} \right)^{-\varepsilon} y_t$$

Example III: government policy

- Taylor rule:

$$R_t = \max\left[Z_t, 1\right]$$

$$Z_t = R^{1-\rho_r} R_{t-1}^{\rho_r} \left[\left(\frac{\Pi_t}{\Pi} \right)^{\phi_\pi} \left(\frac{y_t}{y} \right)^{\phi_y} \right]^{1-\rho_r} \exp\left(\sigma_m \varepsilon_{m,t}\right), \; \varepsilon_{m,t} \sim \mathcal{N}(0,1)$$

- Lump-sum transfers finance

$$g_t = s_{g,t} y_t$$

$$s_{g,t} = s_g^{1-\rho_g} s_{g,t-1}^{\rho_g} \exp\left(\sigma_g \varepsilon_{g,t}\right), \; \varepsilon_{g,t} \sim \mathcal{N}(0,1)$$

More About Nonlinearities

More About nonlinearities I

- The previous examples are not exhaustive.
- Unfortunately, linearization eliminates phenomena of interest:

    1. Asymmetries.
    2. Threshold effects.
    3. Precautionary behavior.
    4. Big shocks.
    5. Convergence away from the steady state.
    6. And many others....

More about nonlinearities II

Linearization limits our study of dynamics:

1. Zero bound on the nominal interest rate.
2. Finite escape time.
3. Multiple steady states.
4. Limit cycles.
5. Subharmonic, harmonic, or almost-periodic oscillations.
6. Chaos.

More about nonlinearities III

- Moreover, linearization induces an approximation error.
- This is worse than you may think.
  1. Theoretical arguments:
     1.1 Second-order errors in the approximated policy function imply first-order errors in the loglikelihood function.
     1.2 As the sample size grows, the error in the likelihood function also grows and we may have inconsistent point estimates.
     1.3 Linearization complicates the identification of parameters.
  2. Computational evidence.

Arguments against nonlinearities

1. Theoretical reasons: we know way less about nonlinear and non-Gaussian systems.
2. Computational limitations.
3. Bias.

Mark Twain

To a man with a hammer, everything looks like a nail.

Teller's Law

A state-of-the-art computation requires 100 hours of CPU time on the state-of-the art computer, independent of the decade.