Perturbation Methods I: Basic Results

(Lectures on Solution Methods for Economists V)

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

November 21, 2021

[1]University of Pennsylvania

[2]Boston College

Introduction

Introduction

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

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

  for an unknown decision rule d.

- Perturbation solves the problem by specifying:

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

- We use implicit-function theorems to find coefficients θ_i's.

- Inherently local approximation. Often good global properties.

Motivation

- Many complicated mathematical problems have:
  1. either a particular case
  2. or a related problem.

  that is easy to solve.
- Often, we can use the solution of the simpler problem as a building block of the general solution.
- Very successful in physics.
- Sometimes perturbation is known as asymptotic methods.

A simple example

- Imagine we want to compute $\sqrt{26}$ by hand.
- We do not remember how to do it.
- But, we note that

$$\sqrt{26} = \sqrt{25 * 1.04} = \sqrt{25} * \sqrt{1.04} = 5 * \sqrt{1.04} \approx 5 * 1.02 = 5.1$$

- Exact solution: $\sqrt{26} = 5.09902$.
- More in general:

$$\sqrt{x} = \sqrt{y^2 * (1+\varepsilon)} = y * \sqrt{(1+\varepsilon)} \approx y * (1+\theta)$$

- Accuracy depends on how big ε is.

Applications in economics

- Judd and Guu (1993) showed how to apply it to economic problems.
- Recently, perturbation methods have been gaining much popularity.
- In particular, second- and third-order approximations are easy to compute and notably improve accuracy.
- Perturbation theory is the generalization of the well-known linearization strategy.
- Hence, we can use much of what we already know about linearization.

Regular versus singular perturbations

- Regular perturbation: a *small* change in the problem induces a *small* change in the solution.
- Singular perturbation: a *small* change in the problem induces a *large* change in the solution.
- Example: excess demand function.
- Most problems in economics involve regular perturbations.
- Sometimes, however, we can have singularities. Example: introducing a new asset in an incomplete market model.

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.

An Economics Application

Stochastic neoclassical growth model

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

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

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

- Note: full depreciation.
- Equilibrium conditions:

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

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

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

Solution and steady state

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

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

- Steady state is also easy to find:

$$k = (\alpha\beta)^{\frac{1}{1-\alpha}}$$
$$c = (\alpha\beta)^{\frac{\alpha}{1-\alpha}} - (\alpha\beta)^{\frac{1}{1-\alpha}}$$
$$z = 0$$

- Steady state in more general models.

The goal

- We are searching for decision rules:

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

- Then, we have:

$$\frac{1}{c(k_t, z_t)} = \beta \mathbb{E}_t \frac{\alpha e^{\rho z_t + \sigma \varepsilon_{t+1}} k(k_t, z_t)^{\alpha - 1}}{c(k(k_t, z_t), \rho z_t + \sigma \varepsilon_{t+1})}$$

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

- This is a system of functional equations.

A perturbation solution

- Rewrite the problem in terms of perturbation parameter λ.
- Different possibilities for λ. For this case, I pick:

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

1. When $\lambda = 1$, stochastic case.
2. When $\lambda = 0$, deterministic case (with $z_0 = 0$ and then $e^{z_t} = 1$).

- Now we are searching for the decision rules:

$$c_t = c\left(k_t, z_t; \lambda\right)$$
$$k_{t+1} = k\left(k_t, z_t; \lambda\right)$$

Taylor's theorem

- We will build a local approximation around $(k, 0; 0)$.
- Given equilibrium conditions:

$$\mathbb{E}_t\left(\frac{1}{c\left(k_t, z_t; \lambda\right)} - \beta \frac{\alpha e^{\rho z_t + \lambda \sigma \varepsilon_{t+1}} k\left(k_t, z_t; \lambda\right)^{\alpha-1}}{c\left(k\left(k_t, z_t; \lambda\right), \rho z_t + \lambda \sigma \varepsilon_{t+1}; \lambda\right)}\right) = 0$$

$$c\left(k_t, z_t; \lambda\right) + k\left(k_t, z_t; \lambda\right) - e^{z_t} k_t^{\alpha} = 0$$

  We will take derivatives with respect to k_t, z_t, and λ and evaluate them around $(k, 0; 0)$.

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

Asymptotic expansion I

$$
\begin{aligned}
c_t &= c\left(k_t, z_t; 1\right)\Big|_{k,0,0} = c\left(k,0;0\right) \\
&+ c_k\left(k,0;0\right)\left(k_t - k\right) + c_z\left(k,0;0\right) z_t + c_\lambda\left(k,0;0\right) \\
&+ \frac{1}{2} c_{kk}\left(k,0;0\right)\left(k_t - k\right)^2 + \frac{1}{2} c_{kz}\left(k,0;0\right)\left(k_t - k\right) z_t \\
&+ \frac{1}{2} c_{k\lambda}\left(k,0;0\right)\left(k_t - k\right) + \frac{1}{2} c_{zk}\left(k,0;0\right) z_t \left(k_t - k\right) \\
&+ \frac{1}{2} c_{zz}\left(k,0;0\right) z_t^2 + \frac{1}{2} c_{z\lambda}\left(k,0;0\right) z_t \\
&+ \frac{1}{2} c_{\lambda k}\left(k,0;0\right)\left(k_t - k\right) + \frac{1}{2} c_{\lambda z}\left(k,0;0\right) \lambda z_t \\
&+ \frac{1}{2} c_{\lambda^2}\left(k,0;0\right) + \ldots
\end{aligned}
$$

Asymptotic expansion II

$$
\begin{aligned}
k_{t+1} \quad = \quad & k\left(k_t, z_t; 1\right)\big|_{k,0,0} = k\left(k, 0; 0\right) \\
& + k_k\left(k, 0; 0\right)\left(k_t - k\right) + k_z\left(k, 0; 0\right) z_t + k_\lambda\left(k, 0; 0\right) \\
& + \frac{1}{2} k_{kk}\left(k, 0; 0\right)\left(k_t - k\right)^2 + \frac{1}{2} k_{kz}\left(k, 0; 0\right)\left(k_t - k\right) z_t \\
& + \frac{1}{2} k_{k\lambda}\left(k, 0; 0\right)\left(k_t - k\right) + \frac{1}{2} k_{zk}\left(k, 0; 0\right) z_t\left(k_t - k\right) \\
& + \frac{1}{2} k_{zz}\left(k, 0; 0\right) z_t^2 + \frac{1}{2} k_{z\lambda}\left(k, 0; 0\right) z_t \\
& + \frac{1}{2} k_{\lambda k}\left(k, 0; 0\right)\left(k_t - k\right) + \frac{1}{2} k_{\lambda z}\left(k, 0; 0\right) z_t \\
& + \frac{1}{2} k_{\lambda^2}\left(k, 0; 0\right) + \ldots
\end{aligned}
$$

Comment on notation

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

$$F\left(k_t, z_t; \lambda\right) = \mathbb{E}_t \left[\begin{array}{c} \frac{1}{c(k_t, z_t; \lambda)} - \beta \frac{\alpha e^{\rho z_t + \lambda \sigma \varepsilon_{t+1}} k(k_t, z_t; \lambda)^{\alpha - 1}}{c(k(k_t, z_t; \lambda), \rho z_t + \lambda \sigma \varepsilon_{t+1}; \sigma)} \\ c\left(k_t, z_t; \lambda\right) + k\left(k_t, z_t; \lambda\right) - e^{z_t} k_t^{\alpha} \end{array} \right] = \left[\begin{array}{c} 0 \\ 0 \end{array} \right]$$

- Note that:

$$\begin{aligned} F\left(k_t, z_t; \lambda\right) &= \mathcal{H}\left(c_t, c_{t+1}, k_t, k_{t+1}, z_t; \lambda\right) \\ &= \mathcal{H}\left(c\left(k_t, z_t; \lambda\right), c\left(k\left(k_t, z_t; \lambda\right), z_{t+1}; \lambda\right), k_t, k\left(k_t, z_t; \lambda\right), z_t; \lambda\right) \end{aligned}$$

- I will use $\mathcal{H}_i$ to represent the partial derivative of $\mathcal{H}$ with respect to the i component and drop the evaluation at the steady state of the functions when we do not need it.

First-order approximation

- We take derivatives of $F(k_t, z_t; \lambda)$ around k, 0, and 0.
- With respect to k_t:

$$F_k(k, 0; 0) = 0$$

- With respect to z_t:

$$F_z(k, 0; 0) = 0$$

- With respect to λ:

$$F_\lambda(k, 0; 0) = 0$$

Solving the system I

- Remember that:

$$F\left(k_t, z_t; \lambda\right) =$$
$$\mathcal{H}\left(c\left(k_t, z_t; \lambda\right), c\left(k\left(k_t, z_t; \lambda\right), z_{t+1}; \lambda\right), k_t, k\left(k_t, z_t; \lambda\right), z_t; \lambda\right) = 0$$

- Because $F\left(k_t, z_t; \lambda\right)$ must be equal to zero for any possible values of k_t, z_t, and λ, the derivatives of any order of F must also be zero.

- Then:

$$F_k\left(k, 0; 0\right) = \mathcal{H}_1 c_k + \mathcal{H}_2 c_k k_k + \mathcal{H}_3 + \mathcal{H}_4 k_k = 0$$
$$F_z\left(k, 0; 0\right) = \mathcal{H}_1 c_z + \mathcal{H}_2\left(c_k k_z + c_z \rho\right) + \mathcal{H}_4 k_z + \mathcal{H}_5 = 0$$
$$F_\lambda\left(k, 0; 0\right) = \mathcal{H}_1 c_\lambda + \mathcal{H}_2\left(c_k k_\lambda + c_\lambda\right) + \mathcal{H}_4 k_\lambda + \mathcal{H}_6 = 0$$

Solving the system II

- Note that:

$$F_k(k,0;0) = \mathcal{H}_1 c_k + \mathcal{H}_2 c_k k_k + \mathcal{H}_3 + \mathcal{H}_4 k_k = 0$$
$$F_z(k,0;0) = \mathcal{H}_1 c_z + \mathcal{H}_2 (c_k k_z + c_z \rho) + \mathcal{H}_4 k_z + \mathcal{H}_5 = 0$$

  is a quadratic system of four equations on four unknowns: c_k, c_z, k_k, and k_z.

- Procedures to solve quadratic systems:

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

- All of them equivalent.

- Why quadratic? Stable and unstable manifold.

Solving the system III

- Also, note that:

$$F_{\lambda}(k, 0; 0) = \mathcal{H}_1 c_{\lambda} + \mathcal{H}_2 (c_k k_{\lambda} + c_{\lambda}) + \mathcal{H}_4 k_{\lambda} + \mathcal{H}_6 = 0$$

  is a linear and homogeneous system in c_{λ} and k_{λ}.

- Hence:

$$c_{\lambda} = k_{\lambda} = 0$$

- This means the system is certainty equivalent.
- Interpretation$\Rightarrow$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).

Comparison with linearization

- After Kydland and Prescott (1982) a popular method to solve economic models has been the use of a LQ approximation of the objective function of the agents.
- Close relative: linearization of equilibrium conditions.
- When properly implemented linearization, LQ, and first-order perturbation are equivalent.
- Advantages of linearization:
  1. Theorems.
  2. Higher order terms.

Second-order approximation

- We take second-order derivatives of $F\left(k_t, z_t; \lambda\right)$ around $k, 0,$ and 0:

$$
\begin{aligned}
F_{kk}\left(k, 0; 0\right) &= 0 \\
F_{kz}\left(k, 0; 0\right) &= 0 \\
F_{k\lambda}\left(k, 0; 0\right) &= 0 \\
F_{zz}\left(k, 0; 0\right) &= 0 \\
F_{z\lambda}\left(k, 0; 0\right) &= 0 \\
F_{\lambda\lambda}\left(k, 0; 0\right) &= 0
\end{aligned}
$$

- We substitute the coefficients that we already know.
- A linear system of 12 equations on 12 unknowns (remember Young's theorem!). Why linear?
- Cross-terms on $k\lambda$ and $z\lambda$ are zero.
- More general result: all the terms in odd derivatives of λ are zero.

Correction for risk

- We have the term $\frac{1}{2}c_{\lambda^2}(k, 0; 0)$.
- Captures precautionary behavior.
- We do not have certainty equivalence any more!
- Important advantage of second order approximation.
- Changes ergodic distribution of states.

Higher-order terms

- We can continue the iteration for as long as we want.
- Great advantage of procedure: it is recursive!
- Often, a few iterations will be enough.
- The level of accuracy depends on the goal of the exercise:
    1. Welfare analysis: Kim and Kim (2001).
    2. Empirical strategies: Fernández-Villaverde, Rubio-Ramírez, and Santos (2006).

A Numerical Example

A numerical example

Parameter	β	α	ρ	σ
Value	0.99	0.33	0.95	0.01

- Steady State:

$$c = 0.388069 \quad k = 0.1883$$

- First-order components:

$$\begin{aligned} c_k\left(k,0;0\right) &= 0.680101 & k_k\left(k,0;0\right) &= 0.33 \\ c_z\left(k,0;0\right) &= 0.388069 & k_z\left(k,0;0\right) &= 0.1883 \end{aligned}$$

- Second-order components:

$$\begin{aligned} c_{kk}\left(k,0;0\right) &= -2.41990 & k_{kk}\left(k,0;0\right) &= -1.1742 \\ c_{kz}\left(k,0;0\right) &= 0.680099 & k_{kz}\left(k,0;0\right) &= 0.33 \\ c_{zz}\left(k,0;0\right) &= 0.388064 & k_{zz}\left(k,0;0\right) &= 0.1883 \\ c_{\lambda^2}\left(k,0;0\right) &= 0 & k_{\lambda^2}\left(k,0;0\right) &= 0 \end{aligned}$$

- $c_\lambda\left(k,0;0\right) = k_\lambda\left(k,0;0\right) = c_{k\lambda}\left(k,0;0\right) = k_{k\lambda}\left(k,0;0\right) = c_{z\lambda}\left(k,0;0\right) = k_{z\lambda}\left(k,0;0\right) = 0.$

Comparison

$$c_t = 0.6733 e^{z_t} k_t^{0.33}$$

$$c_t \simeq 0.388069 + 0.680101 \left(k_t - k\right) + 0.388069 z_t$$

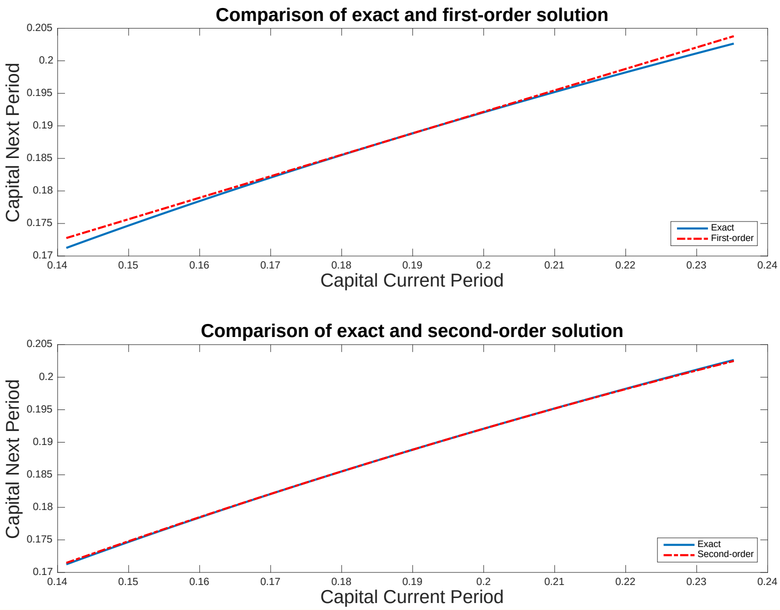
$$-\frac{2.41990}{2} \left(k_t - k\right)^2 + 0.680099 \left(k_t - k\right) z_t + \frac{0.388064}{2} z_t^2$$

and:

$$k_{t+1} = 0.3267 e^{z_t} k_t^{0.33}$$

$$k_{t+1} \simeq 0.1883 + 0.33 \left(k_t - k\right) + 0.1883 z_t$$

$$-\frac{1.1742}{2} \left(k_t - k\right)^2 + 0.33 \left(k_t - k\right) z_t + \frac{0.1883}{2} z_t^2$$


Comparison of exact and first-order solution
Capital Next Period
Capital Current Period
Exact
First-order
0.205
0.2
0.195
0.19
0.185
0.18
0.175
0.17
0.14
0.15
0.16
0.17
0.18
0.19
0.2
0.21
0.22
0.23
0.24
Comparison of exact and second-order solution
Capital Next Period
Capital Current Period
Exact
Second-order

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?

Local properties of the solution I

- Perturbation is a local method.
- It approximates the solution around the deterministic steady state of the problem.
- It is valid within a radius of convergence.

Local properties of the solution II

- What is the radius of convergence of a power series around x? An $r \in \mathbb{R}_{+}^{\infty}$ such that $\forall x'$, $|x' - z| < r$, the power series of x' will converge.

A Remarkable Result from Complex Analysis

The radius of convergence is always equal to the distance from the center to the nearest point where the decision rule has a (non-removable) singularity. If no such point exists then the radius of convergence is infinite.

- Singularity here refers to poles, fractional powers, and other branch powers or discontinuities of the functional or its derivatives.

Local properties of the solution III

- Holomorphic functions are analytic:

  1. A function is holomorphic at a point x if it is differentiable at every point within some open disk centered at x.

  2. A function is analytic at x if in some open disk centered at x it can be expanded as a convergent power series:

  $$f(z) = \sum_{n=0}^{\infty} \theta_n (z - x)^n$$

- Distance is in the complex plane.

- Often, we can check numerically that perturbations have good non-local behavior.

- However: problem with boundaries.

Non-local accuracy test

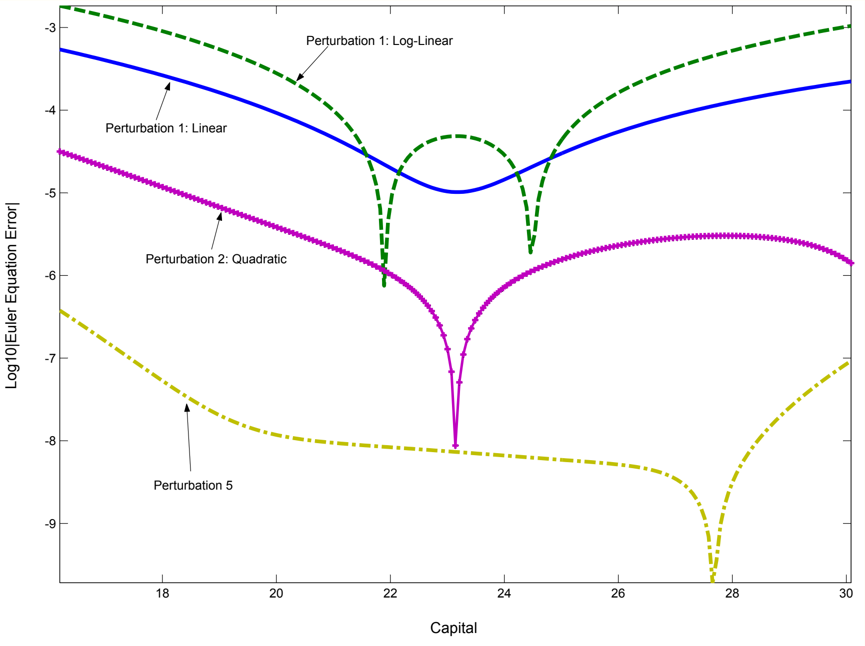
- Proposed by Judd (1992) and Judd and Guu (1997).
- Given the Euler equation:

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

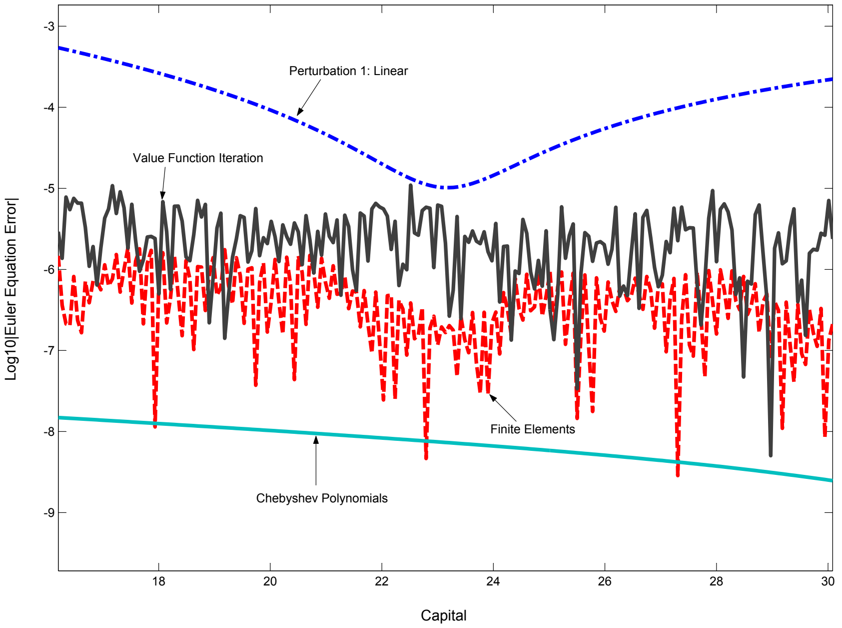
  we can define:

$$EE^i\left(k_t\, , z_t\right) \equiv 1 - c^i\left(k_t\, , z_t\right)\mathbb{E}_t\left(\frac{\alpha e^{z_{t+1}} k^i(k_t, z_t)^{\alpha-1}}{c^i\left(k^i(k_t, z_t), z_{t+1}\right)}\right)$$

- Units of reporting.
- Interpretation.



Perturbation 1: Log-Linear
Perturbation 1: Linear
Perturbation 2: Quadratic
Perturbation 5
Log10|Euler Equation Error|
Capital
-3
-4
-5
-6
-7
-8
-9
18
20
22
24
26
28
30


Perturbation 1: Linear
Value Function Iteration
Finite Elements
Chebyshev Polynomials
Log10|Euler Equation Error|
Capital
-3
-4
-5
-6
-7
-8
-9
18
20
22
24
26
28
30