Perturbation Methods I: Basic Results

(Lectures on Solution Methods for Economists V)

Jesús Fernández-Villaverde,\textsuperscript{1} Pablo Guerrón,\textsuperscript{2} and David Zarruk Valencia\textsuperscript{3}
November 18, 2019

\textsuperscript{1}University of Pennsylvania

\textsuperscript{2}Boston College

\textsuperscript{3}ITAM
Introduction
Remember that we want to solve a functional equations of the form:

\[ \mathcal{H}(d) = 0 \]

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 \( \theta_i \)'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

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

• Exact solution: $\sqrt{26} = 5.09902$.

• More in general:

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

• Accuracy depends on how big $\varepsilon$ 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.


• Economics:


2. “Perturbation Methods with Nonlinear Changes of Variables” by Kenneth Judd.

An Economics Application
Stochastic neoclassical growth model

\[
\max_0^\infty \mathbb{E}_0 \sum_{t=0}^\infty \beta^t \log c_t \\
\text{s.t. } c_t + k_{t+1} = e^{z_t} k_t^\alpha, \forall t > 0 \\
z_t = \rho z_{t-1} + \sigma \varepsilon_t, \varepsilon_t \sim \mathcal{N}(0, 1)
\]

- 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:

\[ k = (\alpha\beta)^{1/\alpha} \]
\[ c = (\alpha\beta)^{\alpha/\alpha} - (\alpha\beta)^{1/\alpha} \]
\[ z = 0 \]

- 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 $\lambda$.

- Different possibilities for $\lambda$. 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:

  $$c_t = c(k_t, z_t; \lambda)$$
  $$k_{t+1} = k(k_t, z_t; \lambda)$$
Taylor’s theorem

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

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

\[
c(k_t, z_t; \lambda) + k(k_t, z_t; \lambda) - e^{\rho z_t} k_t^\alpha = 0
\]

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

- Why?

- Apply Taylor’s theorem and a version of the implicit-function theorem.
Asymptotic expansion I

\[ c_t = c(k, 0; 0) + c_k(k, 0; 0)(k_t - k) + c_z(k, 0; 0)z_t + c_\lambda(k, 0; 0) \]
\[ + \frac{1}{2}c_{kk}(k, 0; 0)(k_t - k)^2 + \frac{1}{2}c_{kz}(k, 0; 0)(k_t - k)z_t \]
\[ + \frac{1}{2}c_{k\lambda}(k, 0; 0)(k_t - k) + \frac{1}{2}c_{zk}(k, 0; 0)z_t(k_t - k) \]
\[ + \frac{1}{2}c_{zz}(k, 0; 0)z_t^2 + \frac{1}{2}c_{z\lambda}(k, 0; 0)z_t \]
\[ + \frac{1}{2}c_{\lambda k}(k, 0; 0)(k_t - k) + \frac{1}{2}c_{\lambda z}(k, 0; 0)\lambda z_t \]
\[ + \frac{1}{2}c_{\lambda^2}(k, 0; 0) + ... \]
\[ k_{t+1} = k(k_t, z_t; 1)|_{k,0,0} = k(k, 0; 0) \]
\[ + k_k(k, 0; 0)(k_t - k) + k_z(k, 0; 0)z_t + k_\lambda(k, 0; 0) \]
\[ + \frac{1}{2}k_{kk}(k, 0; 0)(k_t - k)^2 + \frac{1}{2}k_{kz}(k, 0; 0)(k_t - k)z_t \]
\[ + \frac{1}{2}k_{k\lambda}(k, 0; 0)(k_t - k) + \frac{1}{2}k_{z\lambda}(k, 0; 0)z_t(k_t - k) \]
\[ + \frac{1}{2}k_{zz}(k, 0; 0)z_t^2 + \frac{1}{2}k_{z\lambda}(k, 0; 0)z_t \]
\[ + \frac{1}{2}k_{\lambda k}(k, 0; 0)(k_t - k) + \frac{1}{2}k_{\lambda z}(k, 0; 0)z_t \]
\[ + \frac{1}{2}k_{\lambda\lambda}(k, 0; 0) + ... \]
Comment on notation

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

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

- Note that:

$$F(k_t, z_t; \lambda) = \mathcal{H}(c_t, c_{t+1}, k_t, k_{t+1}, z_t; \lambda)$$

$$= \mathcal{H}(c(k_t, z_t; \lambda), c(k_t, z_t; \lambda), k_t, k(k_t, z_t; \lambda), k_t, k(k_t, z_t; \lambda), z_t; \lambda)$$

- 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 $\lambda$:
  $$F_\lambda(k, 0; 0) = 0$$
Solving the system

- Remember that:

\[ F (k_t, z_t; \lambda) = \]
\[ \mathcal{H} (c (k_t, z_t; \lambda), c (k_t, z_t; \lambda), k_t, k (k_t, z_t; \lambda), z_t; \lambda) = 0 \]

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

- Then:

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

- 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 \( c_\lambda \) and \( k_\lambda \).

• Hence:

\[ c_\lambda = k_\lambda = 0 \]

• This means the system is certainty equivalent.

• Interpretation \( \Rightarrow \) no precautionary behavior.


• 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 \( F (k_t, z_t; \lambda) \) around \( k, 0, \) and \( 0: \)

\[
egin{align*}
F_{kk}(k, 0; 0) & = 0 \\
F_{kz}(k, 0; 0) & = 0 \\
F_{k\lambda}(k, 0; 0) & = 0 \\
F_{zz}(k, 0; 0) & = 0 \\
F_{z\lambda}(k, 0; 0) & = 0 \\
F_{\lambda\lambda}(k, 0; 0) & = 0
\end{align*}
\]

- 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 \( \lambda \) are zero.
Correction for risk

- We have the term $\frac{1}{2} \lambda^2 (k, 0; 0)$.

- 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:
A Numerical Example
A numerical example

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\beta$</th>
<th>$\alpha$</th>
<th>$\rho$</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.99</td>
<td>0.33</td>
<td>0.95</td>
<td>0.01</td>
</tr>
</tbody>
</table>

- Steady State:
  \[ c = 0.388069 \quad k = 0.1883 \]

- First-order components:
  \[
  c_k (k, 0; 0) = 0.680101 \quad k_k (k, 0; 0) = 0.33 \\
  c_z (k, 0; 0) = 0.388069 \quad k_z (k, 0; 0) = 0.1883 
  \]

- Second-order components:
  \[
  c_{kk} (k, 0; 0) = -2.41990 \quad k_{kk} (k, 0; 0) = -1.1742 \\
  c_{kz} (k, 0; 0) = 0.680099 \quad k_{kz} (k, 0; 0) = 0.33 \\
  c_{zz} (k, 0; 0) = 0.388064 \quad k_{zz} (k, 0; 0) = 0.1883 \\
  c_{\lambda^2} (k, 0; 0) = 0 \quad k_{\lambda^2} (k, 0; 0) = 0 
  \]

- \[
  c_{\lambda} (k, 0; 0) = k_{\lambda} (k, 0; 0) = c_{k\lambda} (k, 0; 0) = k_{k\lambda} (k, 0; 0) = c_{z\lambda} (k, 0; 0) = k_{z\lambda} (k, 0; 0) = 0. 
  \]
\[ c_t = 0.6733e^{z_t}k_t^{0.33} \]

\[ c_t \simeq 0.388069 + 0.680101(k_t - k) + 0.388069z_t \]

\[ -\frac{2.41990}{2}(k_t - k)^2 + 0.680099(k_t - k)z_t + \frac{0.388064z_t^2}{2} \]

and:

\[ k_{t+1} = 0.3267e^{z_t}k_t^{0.33} \]

\[ k_{t+1} \simeq 0.1883 + 0.33(k_t - k) + 0.1883z_t \]

\[ -\frac{1.1742}{2}(k_t - k)^2 + 0.33(k_t - k)z_t + \frac{0.1883z_t^2}{2} \]
• In practice you do all this approximations with a computer:

  1. First-, second-, and third- order: 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.
• 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:

\[
\frac{1}{c^i(k_t, z_t)} = E_t \left( \frac{\alpha e^{z_{t+1}} k^i(k_t, z_t)^{\alpha - 1}}{c^i(k^i(k_t, z_t), z_{t+1})} \right)
\]

we can define:

\[
EE^i(k_t, z_t) \equiv 1 - c^i(k_t, z_t) E_t \left( \frac{\alpha e^{z_{t+1}} k^i(k_t, z_t)^{\alpha - 1}}{c^i(k^i(k_t, z_t), z_{t+1})} \right)
\]

- Units of reporting.

- Interpretation.
Figure 8: Euler Equation Errors at $z = 0$, $\tau = 2 / \sigma = 0.007$

Capital

Log$_{10}$|Euler Equation Error|

Perturbation 1: Log-Linear
Perturbation 1: Linear
Perturbation 2: Quadratic
Perturbation 5
Figure 9: Euler Equation Errors at $z = 0, \tau = 2 / \sigma = 0.007$

Capital

Log$_{10}$|Euler Equation Error|

Perturbation 1: Linear
Value Function Iteration
Chebyshev Polynomials
Finite Elements