

Perturbation and Projection Methods for Solving DSGE Models

Lawrence J. Christiano

Discussion of projections taken from Christiano-Fisher, 'Algorithms for Solving Dynamic Models with Occasionally Binding Constraints', 2000, *Journal of Economic Dynamics and Control*.

Discussion of perturbations taken from Judd's textbook.

Outline

- A Toy Example to Illustrate the basic ideas.
 - Functional form characterization of model solution.
 - Use of Projections and Perturbations.
- Neoclassical model.
 - Projection methods
 - Perturbation methods
 - Make sense of the proposition, ‘to a first order approximation, can replace equilibrium conditions with linear expansion about nonstochastic steady state and solve the resulting system using certainty equivalence’

Simple Example

- Suppose that x is some exogenous variable and that the following equation implicitly defines y :

$$h(x, y) = 0, \text{ for all } x \in X$$

- Let the solution be defined by the ‘policy rule’, g :

$$y = g(x)$$

‘Error function’

- satisfying

$$R(x; g) \equiv h(x, g(x)) = 0$$

- for all $x \in X$

The Need to Approximate

- Finding the policy rule, g , is a big problem outside special cases
 - ‘Infinite number of unknowns (i.e., one value of g for each possible x) in an infinite number of equations (i.e., one equation for each possible x).’
- Two approaches:
 - projection and perturbation

Projection

- Find a parametric function, $\hat{g}(x; \gamma)$, where γ is a vector of parameters chosen so that it imitates the property of the exact solution, i.e., $R(x; g) = 0$ for all $x \in X$, as well as possible.
- Choose values for γ so that

$$\hat{R}(x; \gamma) = h(x, \hat{g}(x; \gamma))$$

- is close to zero for $x \in X$.
- The method is defined by how 'close to zero' is defined and by the parametric function, $\hat{g}(x; \gamma)$, that is used.

Projection, continued

- Spectral and finite element approximations
 - **Spectral functions:** functions, $\hat{g}(x; \gamma)$, in which each parameter in γ influences $\hat{g}(x; \gamma)$ for all $x \in X$
example:

$$\hat{g}(x; \gamma) = \sum_{i=0}^n \gamma_i H_i(x), \quad \gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix}$$

$H_i(x) = x^i$ ~ordinary polynomial (not computationally efficient)

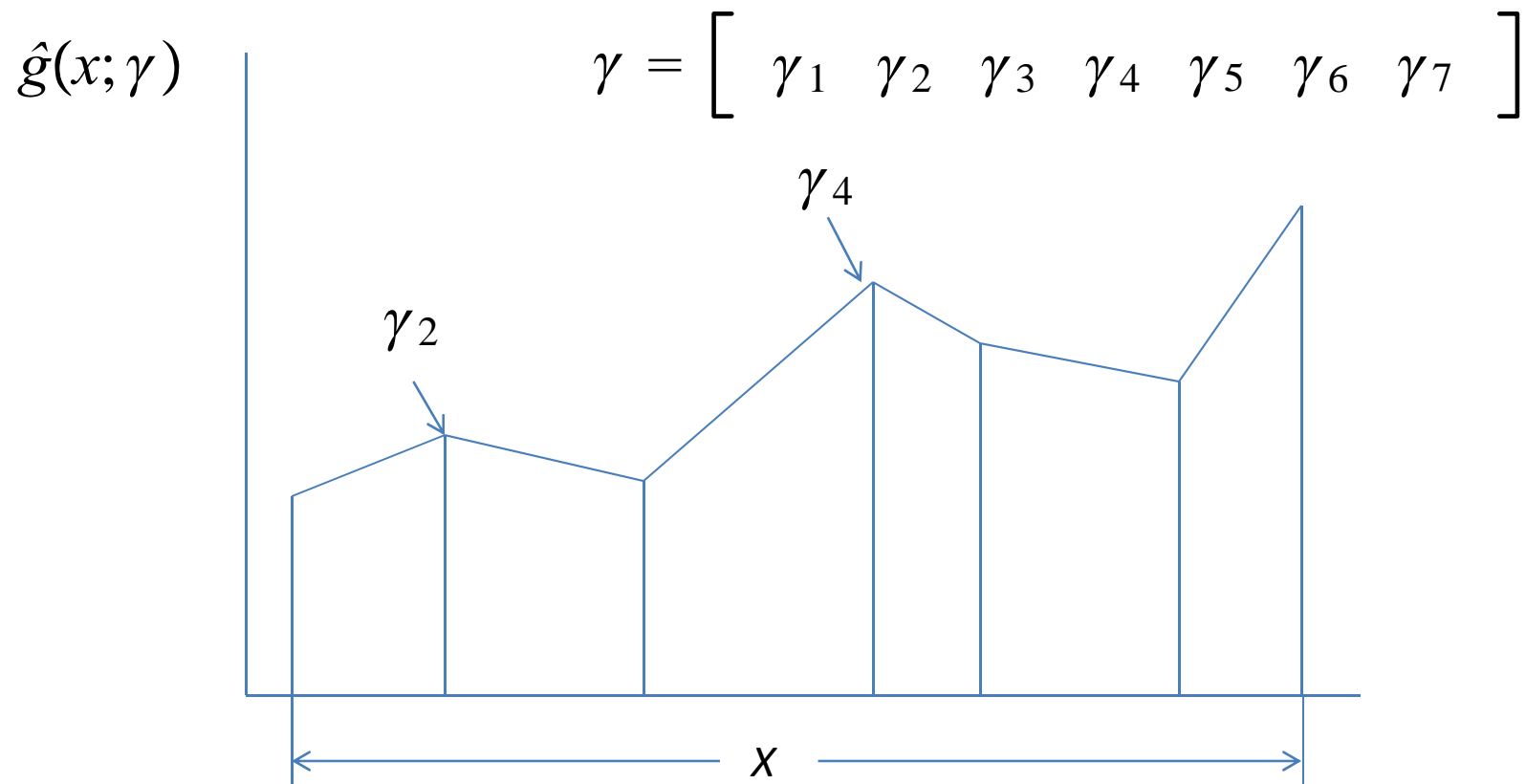
$$H_i(x) = T_i(\varphi(x)),$$

$T_i(z) : [-1, 1] \rightarrow [-1, 1]$, i^{th} order Chebyshev polynomial

$$\varphi : X \rightarrow [-1, 1]$$

Projection, continued

- Finite element approximations: functions, $\hat{g}(x; \gamma)$, in which each parameter in γ influences $\hat{g}(x; \gamma)$ over only a subinterval of $x \in X$



Projection, continued

- ‘Close to zero’: two methods
- **Collocation**, for n values of $x : x_1, x_2, \dots, x_n \in X$ choose n elements of $\gamma = \begin{bmatrix} \gamma_1 & \cdots & \gamma_n \end{bmatrix}$ so that

$$\hat{R}(x_i; \gamma) = h(x_i, \hat{g}(x_i; \gamma)) = 0, \quad i = 1, \dots, n$$

– how you choose the grid of x 's matters...

Example of Importance of Grid Points

- Here is an example, taken from a related problem, the problem of *interpolation*.
 - You get to evaluate a function on a set of grid points that you select, and you must guess the shape of the function between the grid points.

- Consider the function,

$$f(k) = \frac{1}{1 + k^2}, k \in [-5, 5]$$

- Next slide shows what happens when you select 11 equally-spaced grid points and interpolate by fitting a 10th order polynomial.
 - As you increase the number of grid points on a fixed interval grid, oscillations in tails grow more and more violent.
- Chebyshev approximation theorem: distribute more points in the tails (by selecting zeros of Chebyshev polynomial) and get convergence in sup norm.

How You Select the Grid Points Matters
Function Approximation with Fixed Interval Grid

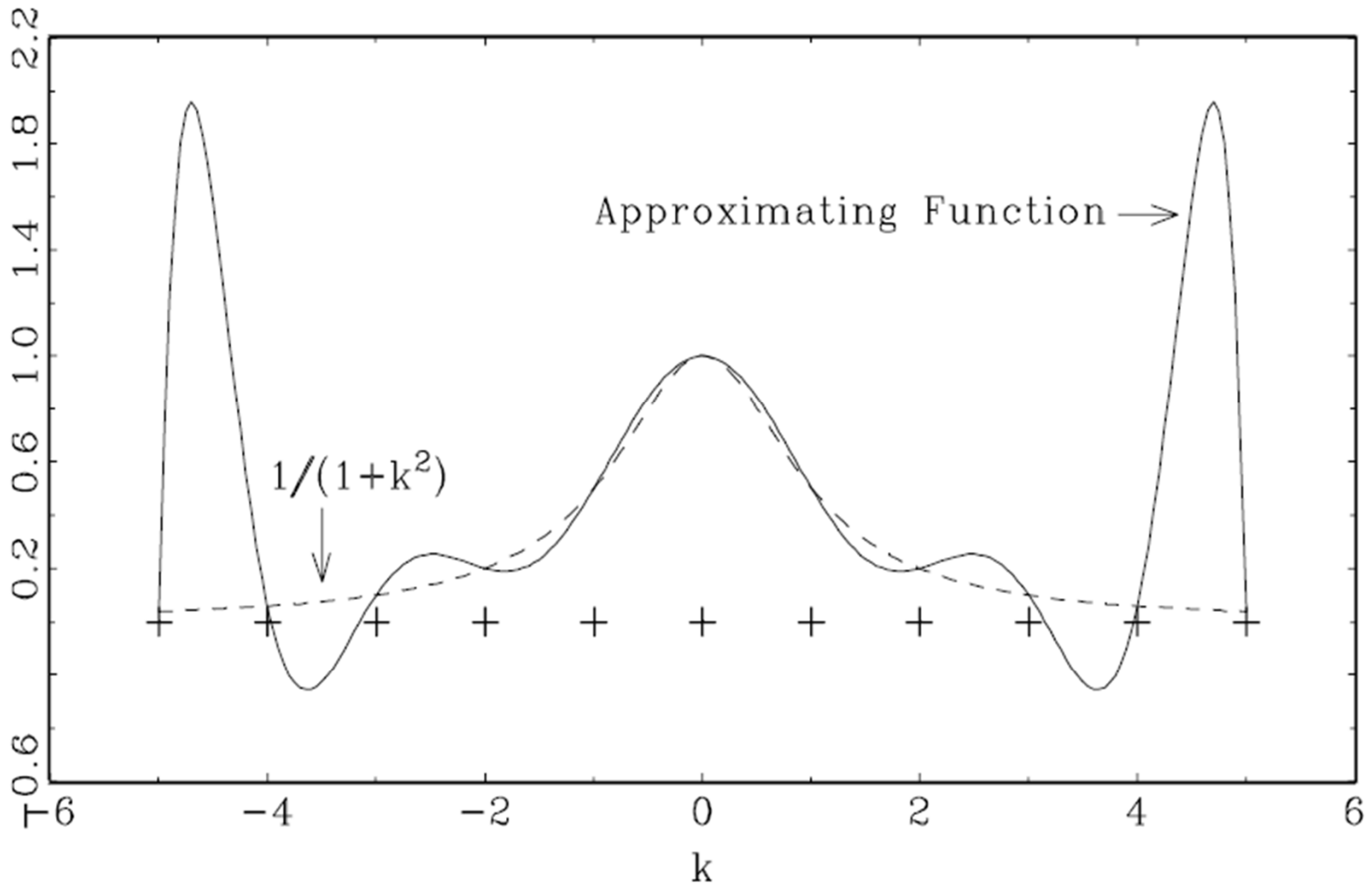
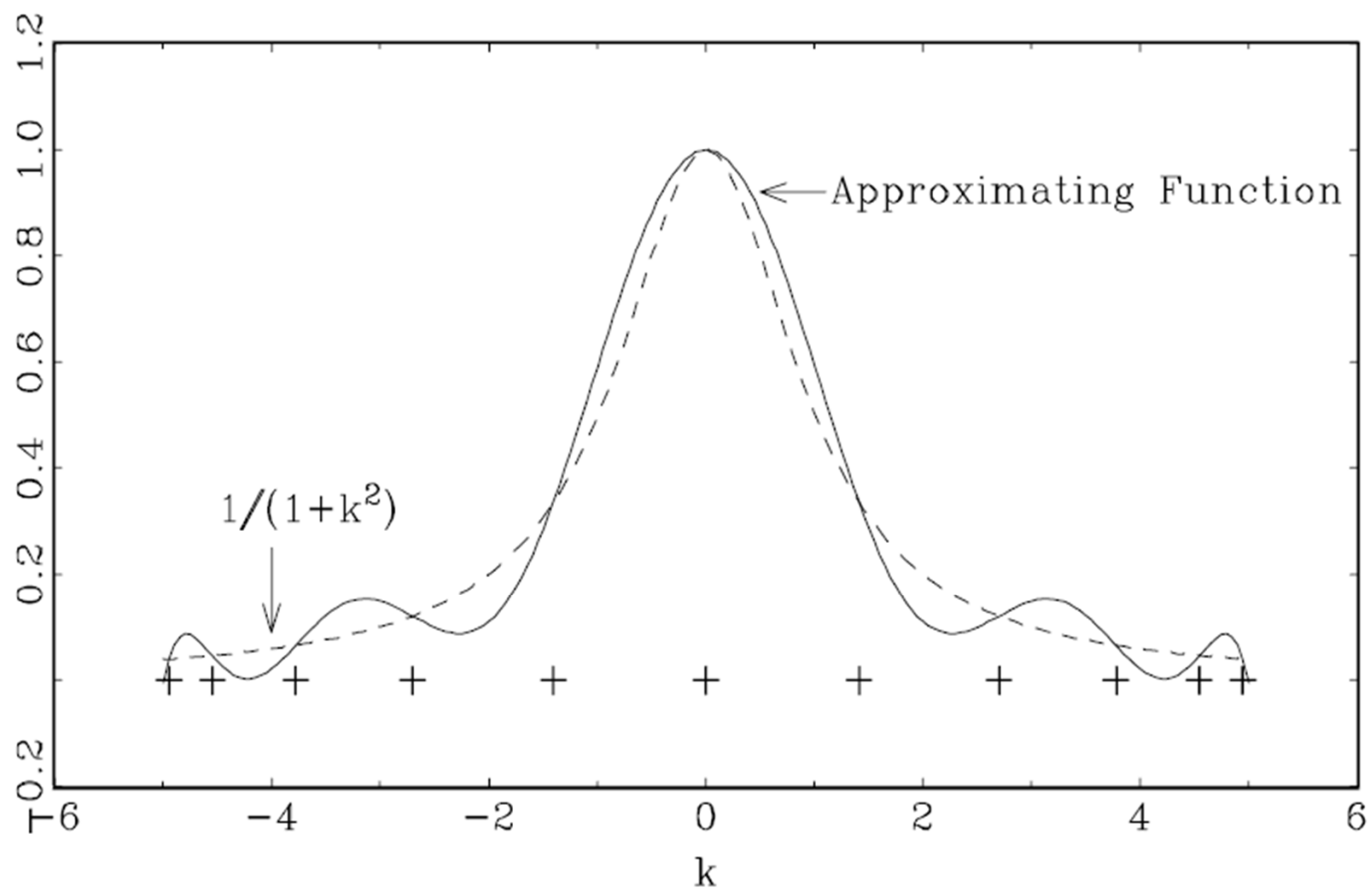


Figure from Christiano-Fisher, JEDC, 1990

Function Approximation with Chebyshev Zeros



Projection, continued

- ‘Close to zero’: two methods
- **Collocation**, for n values of $x : x_1, x_2, \dots, x_n \in X$ choose n elements of $\gamma = \begin{bmatrix} \gamma_1 & \cdots & \gamma_n \end{bmatrix}$ so that

$$\hat{R}(x_i; \gamma) = h(x_i, \hat{g}(x_i; \gamma)) = 0, \quad i = 1, \dots, n$$

– how you choose the grid of x 's matters...

- **Weighted Residual**, for $m > n$ values of $x : x_1, x_2, \dots, x_m \in X$ choose the n γ_i 's

$$\sum_{j=1}^m w_j^i h(x_j, \hat{g}(x_j; \gamma)) = 0, \quad i = 1, \dots, n$$

Perturbation

- Projection uses the ‘global’ behavior of the functional equation to approximate solution.
 - Problem: requires finding zeros of non-linear equations. Iterative methods for doing this are a pain.
 - Advantage: can easily adapt to situations the policy rule is not continuous or simply non-differentiable (e.g., occasionally binding zero lower bound on interest rate).
- Perturbation method uses local properties of functional equation and Implicit Function/Taylor’s theorem to approximate solution.
 - Advantage: can implement it using non-iterative methods.
 - Possible disadvantages:
 - may require enormously high derivatives to achieve a decent global approximation.
 - Does not work when there are important non-differentiabilities (e.g., occasionally binding zero lower bound on interest rate).

Perturbation, cnt'd

- Suppose there is a point, $x^* \in X$, where we know the value taken on by the function, g , that we wish to approximate:

$$g(x^*) = g^*, \text{ some } x^*$$

- Use the implicit function theorem to approximate g in a neighborhood of x^*
- Note:

$$R(x; g) = 0 \text{ for all } x \in X$$

→

$$R^{(j)}(x; g) \equiv \frac{d^j}{dx^j} R(x; g) = 0 \text{ for all } j, \text{ all } x \in X.$$

Perturbation, cnt'd

- Differentiate R with respect to x and evaluate the result at x^* :

$$R^{(1)}(x^*) = \frac{d}{dx} h(x, g(x))|_{x=x^*} = h_1(x^*, g^*) + h_2(x^*, g^*)g'(x^*) = 0$$

$$\rightarrow g'(x^*) = -\frac{h_1(x^*, g^*)}{h_2(x^*, g^*)}$$

- Do it again!

$$R^{(2)}(x^*) = \frac{d^2}{dx^2} h(x, g(x))|_{x=x^*} = h_{11}(x^*, g^*) + 2h_{12}(x^*, g^*)g'(x^*) + h_{22}(x^*, g^*)[g'(x^*)]^2 + h_2(x^*, g^*)g''(x^*).$$

→ Solve this linearly for $g''(x^*)$.

Perturbation, cnt'd

- Preceding calculations deliver (assuming enough differentiability, appropriate invertibility, a high tolerance for painful notation!), recursively:

$$g'(x^*), g''(x^*), \dots, g^{(n)}(x^*)$$

- Then, have the following Taylor's series approximation:

$$g(x) \approx \hat{g}(x)$$

$$\hat{g}(x) = g^* + g'(x^*) \times (x - x^*)$$

$$+ \frac{1}{2} g''(x^*) \times (x - x^*)^2 + \dots + \frac{1}{n!} g^{(n)}(x^*) \times (x - x^*)^n$$

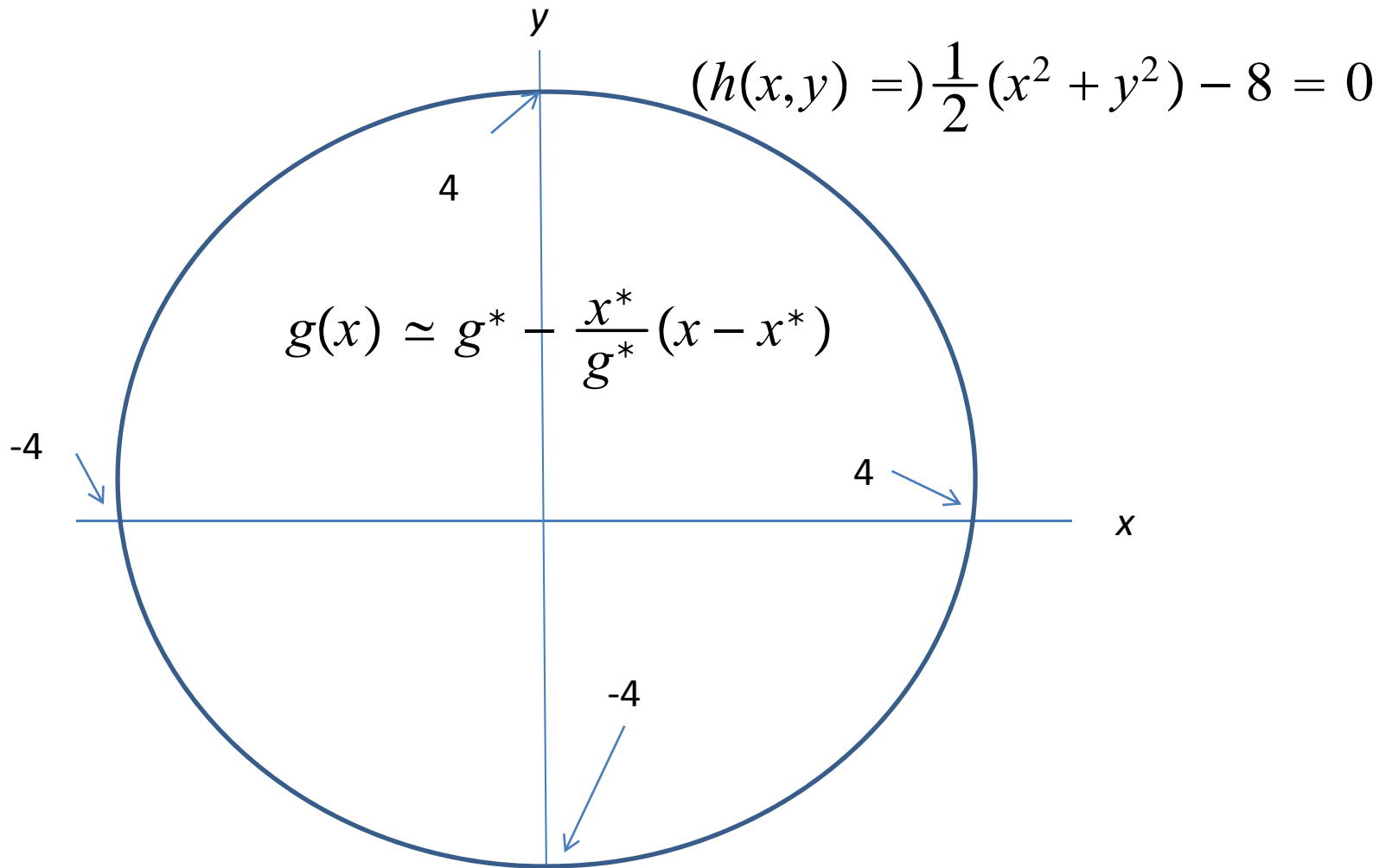
Perturbation, cnt'd

- Check....
- Study the graph of

$$R(x; \hat{g})$$

- over $x \in X$ to verify that it is everywhere close to zero (or, at least in the region of interest).

Example of Implicit Function Theorem



$$g'(x^*) = -\frac{h_1(x^*, g^*)}{h_2(x^*, g^*)} = -\frac{x^*}{g^*} \quad (h_2 \text{ had better not be zero!})$$

Neoclassical Growth Model

- Objective:

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t), \quad u(c_t) = \frac{c_t^{1-\gamma} - 1}{1-\gamma}$$

- Constraints:

$$c_t + \exp(k_{t+1}) \leq f(k_t, a_t), \quad t = 0, 1, 2, \dots$$

$$a_t = \rho a_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim E\varepsilon_t = 0, \quad E\varepsilon_t^2 = V_\varepsilon$$

$$f(k_t, a_t) = \exp(\alpha k_t) \exp(a_t) + (1 - \delta) \exp(k_t)$$

Efficiency Condition

$$E_t \left[u' \left(\overbrace{f(k_t, a_t) - \exp(k_{t+1})}^{c_{t+1}} \right) - \beta u' \left(\overbrace{f(k_{t+1}, \rho a_t + \sigma \varepsilon_{t+1}) - \exp(k_{t+2})}^{c_{t+1}} \right) \overbrace{f_K(k_{t+1}, \rho a_t + \sigma \varepsilon_{t+1})}^{\text{period } t+1 \text{ marginal product of capital}} \right] = 0.$$

$k_t, a_t \sim$ given numbers

- Here,

$\varepsilon_{t+1} \sim$ random variable

time t choice variable, k_{t+1}

- Parameter, σ , indexes a set of models, with the model of interest corresponding to

$$\sigma = 1$$

Solution

- A policy rule,

$$k_{t+1} = g(k_t, a_t, \sigma).$$

- With the property:

$$R(k_t, a_t, \sigma; g) \equiv E_t \left\{ u' \left(\overbrace{f(k_t, a_t) - \exp[g(k_t, a_t, \sigma)]}^{c_t} \right) \right. \\ \left. - \beta u' \left(\overbrace{f \left(\overbrace{g(k_t, a_t, \sigma)}^{k_{t+1}}, \overbrace{\rho a_t + \sigma \varepsilon_{t+1}}^{a_{t+1}} \right) - \exp \left[g \left(\overbrace{g(k_t, a_t, \sigma)}^{k_{t+1}}, \overbrace{\rho a_t + \sigma \varepsilon_{t+1}, \sigma}^{a_{t+1}} \right) \right]}^{c_{t+1}} \right) \right. \\ \left. \times f_K \left(\overbrace{g(k_t, a_t, \sigma)}^{k_{t+1}}, \overbrace{\rho a_t + \sigma \varepsilon_{t+1}}^{a_{t+1}} \right) \right\} = 0,$$

- for all a_t, k_t, σ .

Projection Methods

- Let

$$\hat{g}(k_t, a_t, \sigma; \gamma)$$

- be a function with finite parameters (could be either spectral or finite element, as before).

- Choose parameters, γ , to make

$$R(k_t, a_t, \sigma; \hat{g})$$

- as close to zero as possible, over a range of values of the state.
- use weighted residuals or Collocation.

Occasionally Binding Constraints

- Suppose we add the non-negativity constraint on investment:

$$\exp(g(k_t, a_t, \sigma)) - (1 - \delta) \exp(k_t) \geq 0$$

- Express problem in Lagrangian form and optimum is characterized in terms of equality conditions with a multiplier and with a complementary slackness condition associated with the constraint.
- Conceptually straightforward to apply preceding method. For details, see Christiano-Fisher, 'Algorithms for Solving Dynamic Models with Occasionally Binding Constraints', 2000, *Journal of Economic Dynamics and Control*.
 - This paper describes alternative strategies, based on parameterizing the expectation function, that may be easier, when constraints are occasionally binding constraints.

Perturbation Approach

- Straightforward application of the perturbation approach, as in the simple example, requires knowing the value taken on by the policy rule at a point.
- The overwhelming majority of models used in macro do have this property.
 - In these models, can compute non-stochastic steady state without any knowledge of the policy rule, g .
 - Non-stochastic steady state is k^* such that

$$k^* = g \left(k^*, \underbrace{0}_{a=0 \text{ (nonstochastic steady state in no uncertainty case)}}, \underbrace{0}_{\sigma=0 \text{ (no uncertainty)}} \right)$$

– and

$$k^* = \log \left\{ \left[\frac{\alpha\beta}{1 - (1 - \delta)\beta} \right]^{\frac{1}{1-\alpha}} \right\}.$$

Perturbation

- Error function:

$$R(k_t, a_t, \sigma; g) \equiv E_t \left\{ u' \left(\overbrace{f(k_t, a_t) - \exp[g(k_t, a_t, \sigma)]}^{c_t} \right) \right. \\ \left. - \beta u' \left[\overbrace{f(g(k_t, a_t, \sigma), \rho a_t + \sigma \varepsilon_{t+1}) - \exp[g(g(k_t, a_t, \sigma), \rho a_t + \sigma \varepsilon_{t+1}, \sigma)]}^{c_{t+1}} \right] \right. \\ \left. \times f_K(g(k_t, a_t, \sigma), \rho a_t + \sigma \varepsilon_{t+1}) \right\} = 0,$$

– for all values of k_t, a_t, σ .

- So, all order derivatives of R with respect to its arguments are zero (assuming they exist!).

Four (Easy to Show) Results About Perturbations

- Taylor series expansion of policy rule:

$$g(k_t, a_t, \sigma) \simeq \overbrace{k + g_k(k_t - k) + g_a a_t + g_\sigma \sigma}^{\text{linear component of policy rule}}$$

$$+ \overbrace{\frac{1}{2} [g_{kk}(k_t - k)^2 + g_{aa}a_t^2 + g_{\sigma\sigma}\sigma^2] + g_{ka}(k_t - k)a_t + g_{k\sigma}(k_t - k)\sigma + g_{a\sigma}a_t\sigma + \dots}^{\text{second and higher order terms}}$$

- $g_\sigma = 0$: to a first order approximation, ‘certainty equivalence’
- All terms found by solving linear equations, except coefficient on past endogenous variable, g_k , which requires solving for eigenvalues
- To second order approximation: slope terms certainty equivalent –

$$g_{k\sigma} = g_{a\sigma} = 0$$
- Quadratic, higher order terms computed recursively.

First Order Perturbation

- Working out the following derivatives and evaluating at $k_t = k^*, a_t = \sigma = 0$

$$R_k(k_t, a_t, \sigma; g) = R_a(k_t, a_t, \sigma; g) = R_\sigma(k_t, a_t, \sigma; g) = 0$$

- Implies:

'problematic term'

Source of certainty equivalence
In linear approximation

$$R_k = u''(f_k - e^g g_k) - \beta u' f_{Kk} g_k - \beta u''(f_k g_k - e^g g_k^2) f_K = 0$$

$$R_a = u''(f_a - e^g g_a) - \beta u' [f_{Kk} g_a + f_{Ka} \rho] - \beta u''(f_k g_a + f_a \rho - e^g [g_k g_a + g_a \rho]) f_K = 0$$

$$R_\sigma = -[u' e^g + \beta u''(f_k - e^g g_k) f_K] g_\sigma = 0$$

Absence of arguments in these functions reflects they are evaluated in $k_t = k^*, a_t = \sigma = 0$

Technical notes for following slide

$$u''(f_k - e^g g_k) - \beta u' f_{Kk} g_k - \beta u''(f_k g_k - e^g g_k^2) f_K = 0$$

$$\frac{1}{\beta}(f_k - e^g g_k) - u' \frac{f_{Kk}}{u''} g_k - (f_k g_k - e^g g_k^2) f_K = 0$$

$$\frac{1}{\beta} f_k - \left[\frac{1}{\beta} e^g + u' \frac{f_{Kk}}{u''} + f_k f_K \right] g_k + e^g g_k^2 f_K = 0$$

$$\frac{1}{\beta} \frac{f_k}{e^g f_K} - \left[\frac{1}{\beta f_K} + \frac{u'}{u''} \frac{f_{Kk}}{e^g f_K} + \frac{f_k}{e^g} \right] g_k + g_k^2 = 0$$

$$\frac{1}{\beta} - \left[1 + \frac{1}{\beta} + \frac{u'}{u''} \frac{f_{Kk}}{e^g f_K} \right] g_k + g_k^2 = 0$$

- Simplify this further using:

βf_K ~steady state equation

$$f_K = \alpha K^{\alpha-1} \exp(a) + (1 - \delta), \quad K \equiv \exp(k)$$

$$= \alpha \exp[(\alpha - 1)k + a] + (1 - \delta)$$

$$f_k = \alpha \exp[\alpha k + a] + (1 - \delta) \exp(k) = f_K e^g$$

$$f_{Kk} = \alpha(\alpha - 1) \exp[(\alpha - 1)k + a]$$

$$f_{KK} = \alpha(\alpha - 1) K^{\alpha-2} \exp(a) = \alpha(\alpha - 1) \exp[(\alpha - 2)k + a] = f_{Kk} e^{-g}$$

- to obtain polynomial on next slide.

First Order, cont'd

- Rewriting $R_k = 0$ term:

$$\frac{1}{\beta} - \left[1 + \frac{1}{\beta} + \frac{u'}{u''} \frac{f_{KK}}{f_K} \right] g_k + g_k^2 = 0$$

- There are two solutions, $0 < g_k < 1$, $g_k > \frac{1}{\beta}$
 - Theory (see Stokey-Lucas) tells us to pick the smaller one.
 - In general, could be more than one eigenvalue less than unity: multiple solutions.
- Conditional on solution to g_k , g_a solved for linearly using $R_a = 0$ equation.
- These results all generalize to multidimensional case

Numerical Example

- Parameters taken from Prescott (1986):

$$\beta = 0.99, \gamma = 2(20), \alpha = 0.36, \delta = 0.02, \rho = 0.95, V_\varepsilon = 0.01^2$$

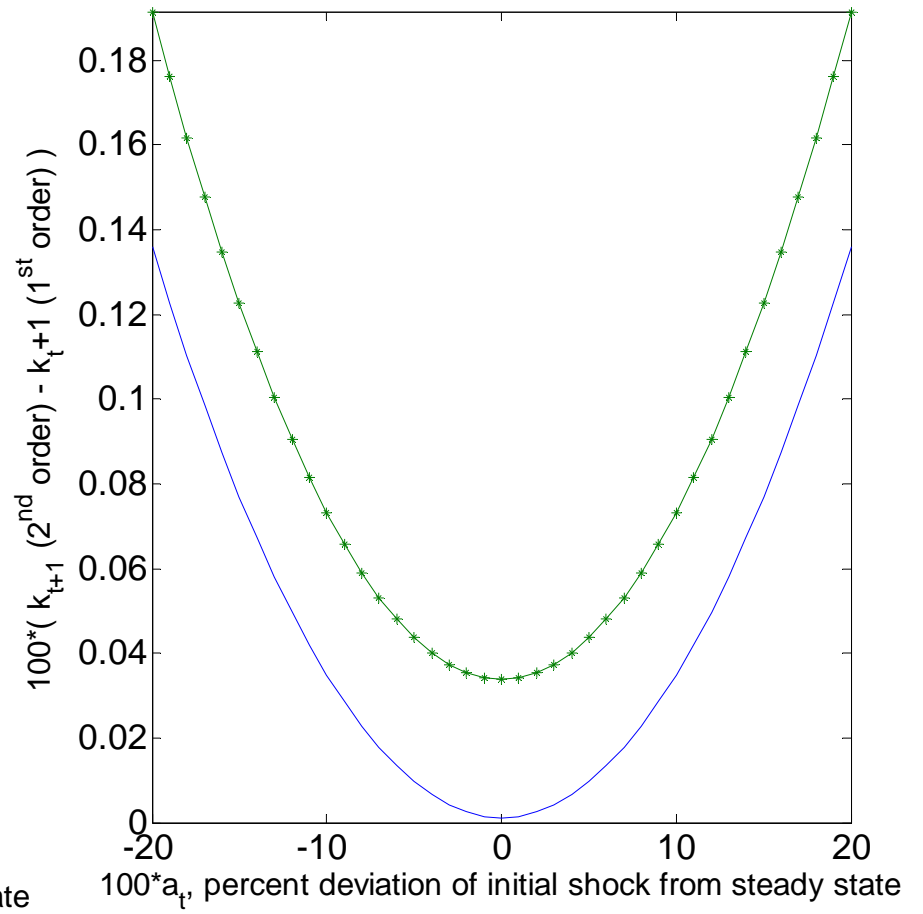
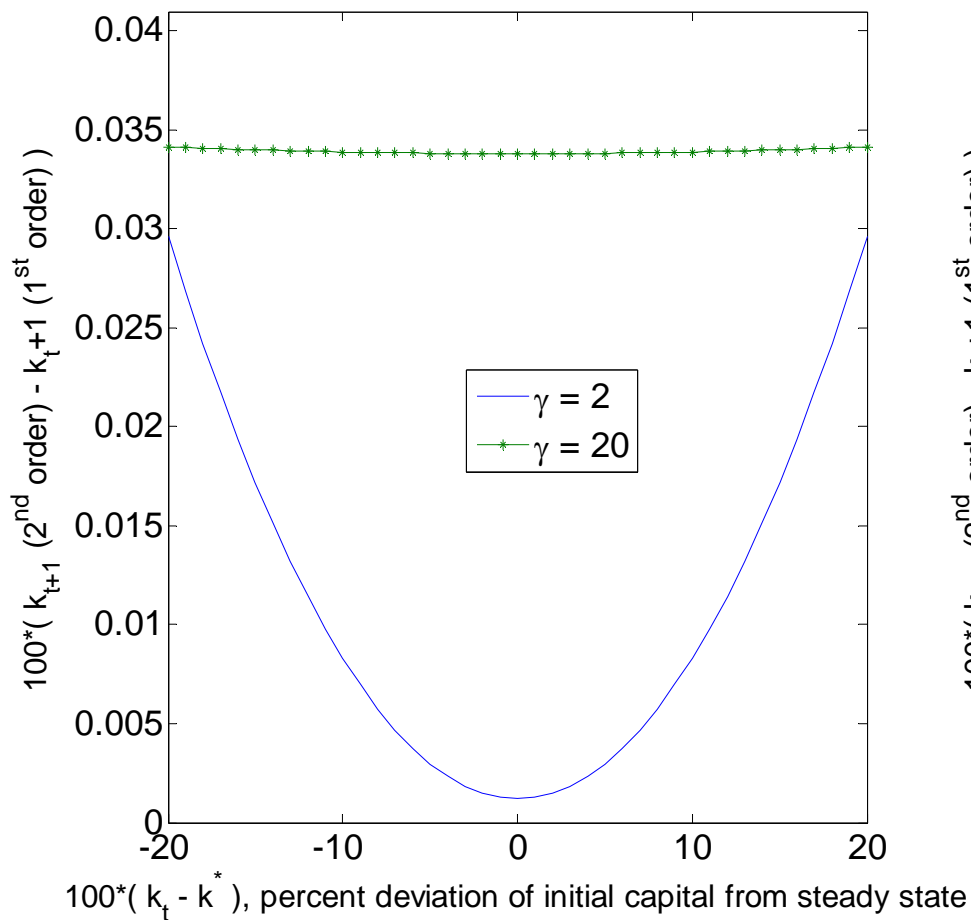
- Second order approximation:

$$\begin{aligned} \hat{g}(k_t, a_{t-1}, \varepsilon_t, \sigma) = & \overbrace{k^*}^{3.88} + \overbrace{g_k}^{0.98 (0.996)} (k_t - k^*) + \overbrace{g_a}^{0.06 (0.07)} a_t + \overbrace{g_\sigma}^0 \sigma \\ & + \frac{1}{2} \left[\overbrace{g_{kk}}^{0.014 (0.00017)} (k_t - k)^2 + \overbrace{g_{aa}}^{0.067 (0.079)} a_t^2 + \overbrace{g_{\sigma\sigma}}^{0.000024 (0.00068)} \sigma^2 \right] \\ & + \overbrace{g_{ka}}^{-0.035 (-0.028)} (k_t - k)a_t + \overbrace{g_{k\sigma}}^0 (k_t - k)\sigma + \overbrace{g_{a\sigma}}^0 a_t\sigma \end{aligned}$$

- Following is a graph that compares the policy rules implied by the first and second order perturbation.
- The graph itself corresponds to the baseline parameterization, and results are reported in parentheses for risk aversion equal to 20.

'If initial capital is 20 percent away from steady state, then capital choice differs by 0.03 (0.035) percent between the two approximations.'

'If shock is 6 standard deviations away from its mean, then capital choice differs by 0.14 (0.18) percent between the two approximations'



Number in parentheses at top correspond to $\gamma = 20$.

Conclusion

- For modest US-sized fluctuations and for aggregate quantities, it is reasonable to work with first order perturbations.
- First order perturbation: linearize (or, log-linearize) equilibrium conditions around non-stochastic steady state and solve the resulting system.
 - This approach assumes ‘certainty equivalence’. Ok, as a first order approximation.

List of endogenous variables determined at t

Solution by Linearization

- (log) Linearized Equilibrium Conditions:

$$E_t[\alpha_0 z_{t+1} + \alpha_1 z_t + \alpha_2 z_{t-1} + \beta_0 s_{t+1} + \beta_1 s_t] = 0$$

- Posit Linear Solution:

$$z_t = Az_{t-1} + Bs_t$$

$$s_t - Ps_{t-1} - \epsilon_t = 0.$$

Exogenous shocks

- To satisfy equil conditions, A and B must:

$$\alpha_0 A^2 + \alpha_1 A + \alpha_2 I = 0, \quad F = (\beta_0 + \alpha_0 B)P + [\beta_1 + (\alpha_0 A + \alpha_1)B] = 0$$

- If there is exactly one A with eigenvalues less than unity in absolute value, that's the solution. Otherwise, multiple solutions.
- Conditional on A , solve linear system for B .