# Perturbation Methods for the Numerical Analysis of DSGE Models: Lecture Notes

Stephanie Schmitt-Grohé and Martín Uribe[1]
Columbia University

January 2009
Incomplete

[1]E-mail: ss3501@columbia.edu.

# Contents

# Chapter 1

# Perturbation Methods: An Introduction

The equilibrium conditions of a wide variety of dynamic stochastic general equilibrium models can be written in the form of a nonlinear stochastic vector difference equation

$$E_t f(y_{t+1}, y_t, x_{t+1}, x_t) = 0, \tag{1.1}$$

where $E_t$ denotes the mathematical expectations operator conditional on information available at time $t$. The vector $x_t$ denotes predetermined (or state) variables and the vector $y_t$ denotes nonpredetermined (or control) variables. The initial value of the state vector $x_0$ is an initial condition for the economy. (Beyond the initial condition, the complete set of equilibrium conditions also includes a terminal condition, like a no-Ponzi game constraint. We omit such a constraint here because in the lectures we focus on approximating stationary solutions.) The state vector $x_t$ can be partitioned as $x_t = [x_t^1;\ x_t^2]'$. The vector $x_t^1$ consists of endogenous predetermined state variables and the vector $x_t^2$ of exogenous state variables. Specifically, we assume that $x_t^2$ follows the exogenous stochastic process given by

$$x_{t+1}^2 = \tilde{h}(x_t^2, \sigma) + \tilde{\eta}\sigma\epsilon_{t+1},$$

where both the vector $x_t^2$ and the innovation $\epsilon_t$ are of order $n_\epsilon \times 1$.[1] The vector $\epsilon_t$ is assumed to have a bounded support and to be independently and identically distributed, with mean zero and variance/covariance matrix

[1] It is straightforward to accommodate the case in which the size of the innovations vector $\epsilon_t$ is different from that of $x_t^2$.

$I$. The eigenvalues of the Jacobian of the function $\tilde{h}$ with respect to its first argument evaluated at the non-stochastic steady state are assumed to lie within the unit circle.

The solution to models belonging to the class given in equation (1.1) is of the form:

$$y_t = \hat{g}(x_t) \tag{1.2}$$

and

$$x_{t+1} = \hat{h}(x_t) + \eta\sigma\epsilon_{t+1}. \tag{1.3}$$

The vector $x_t$ of predetermined variables is of size $n_x \times 1$ and the vector $y_t$ of nonpredetermined variables is of size $n_y \times 1$. We define $n = n_x + n_y$. The function $f$ then maps $R^{n_y} \times R^{n_y} \times R^{n_x} \times R^{n_x}$ into $R^n$.

The matrix $\eta$ is of order $n_x \times n_\epsilon$ and is given by

$$\eta = \left[ \begin{array}{c} \emptyset \\ \tilde{\eta} \end{array} \right].$$

The shape of the functions $\hat{h}$ and $\hat{g}$ will in general depend on the amount of uncertainty in the economy. The key idea of perturbation methods is to interpret the solution to the model as a function of the state vector $x_t$ and of the parameter $\sigma$ scaling the amount of uncertainty in the economy, that is,

$$y_t = g(x_t, \sigma) \tag{1.4}$$

and

$$x_{t+1} = h(x_t, \sigma) + \eta\sigma\epsilon_{t+1}, \tag{1.5}$$

where the function $g$ maps $R^{n_x} \times R^+$ into $R^{n_y}$ and the function $h$ maps $R^{n_x} \times R^+$ into $R^{n_x}$.

Given this interpretation, a perturbation methods finds a *local* approximation of the functions $g$ and $h$. By a local approximation, we mean an approximation that is valid in the neighborhood of a particular point $(\bar{x}, \bar{\sigma})$.

Taking a Taylor series approximation of the functions $g$ and $h$ around the point $(x, \sigma) = (\bar{x}, \bar{\sigma})$ we have (for the moment to keep the notation simple, let's assume that nx=ny=1)

$$\begin{aligned} g(x,\sigma) &= g(\bar{x},\bar{\sigma}) + g_x(\bar{x},\bar{\sigma})(x-\bar{x}) + g_\sigma(\bar{x},\bar{\sigma})(\sigma-\bar{\sigma}) \\ &\quad + \frac{1}{2}g_{xx}(\bar{x},\bar{\sigma})(x-\bar{x})^2 + g_{x\sigma}(\bar{x},\bar{\sigma})(x-\bar{x})(\sigma-\bar{\sigma}) \\ &\quad + \frac{1}{2}g_{\sigma\sigma}(\bar{x},\bar{\sigma})(\sigma-\bar{\sigma})^2 + \ldots \end{aligned}$$

$$\begin{aligned} h(x,\sigma) &= h(\bar{x},\bar{\sigma}) + h_x(\bar{x},\bar{\sigma})(x-\bar{x}) + h_\sigma(\bar{x},\bar{\sigma})(\sigma-\bar{\sigma}) \\ &\quad + \frac{1}{2}h_{xx}(\bar{x},\bar{\sigma})(x-\bar{x})^2 \\ &\quad + h_{x\sigma}(\bar{x},\bar{\sigma})(x-\bar{x})(\sigma-\bar{\sigma}) \\ &\quad + \frac{1}{2}h_{\sigma\sigma}(\bar{x},\bar{\sigma})(\sigma-\bar{\sigma})^2 + \ldots, \end{aligned}$$

The unknowns of an $n^{th}$ order expansion are the n-th order derivatives of the functions $g$ and $h$ evaluated at the point $(\bar{x},\bar{\sigma})$.

To identify these derivatives, substitute the proposed solution given by equations (1.4) and (1.5) into equation (1.1), and define

$$\begin{aligned} F(x,\sigma) &\equiv E_t f(g(h(x,\sigma)+\eta\sigma\epsilon',\sigma), g(x,\sigma), h(x,\sigma)+\eta\sigma\epsilon', x) \quad (1.6) \\ &= 0. \end{aligned}$$

Here we are dropping time subscripts. We use a prime to indicate variables dated in period $t+1$.

Because $F(x,\sigma)$ must be equal to zero for any possible values of $x$ and $\sigma$, it must be the case that the derivatives of any order of $F$ must also be equal to zero. Formally,

$$F_{x^k\sigma^j}(x,\sigma) = 0 \quad \forall x,\sigma,j,k, \tag{1.7}$$

where $F_{x^k\sigma^j}(x,\sigma)$ denotes the derivative of $F$ with respect to $x$ taken $k$ times and with respect to $\sigma$ taken $j$ times.

As will become clear below, a particularly convenient point to approximate the functions $g$ and $h$ around is the non-stochastic steady state, $x_t = \bar{x}$ and $\sigma = 0$. We define the non-stochastic steady state as vectors $(\bar{x},\bar{y})$ such that

$$f(\bar{y},\bar{y},\bar{x},\bar{x}) = 0.$$

It is clear that $\bar{y} = g(\bar{x},0)$ and $\bar{x} = h(\bar{x},0)$. To see this, note that if $\sigma = 0$, then $E_t f = f$. The reason why the steady state is a particularly convenient point is that in most cases it is possible to solve for the steady state. With the steady state values in hand, one can then find the derivatives of the function $F$.

In principle, however, one can approximate the functions $g$ and $h$ around any arbitrary point $(x_t,\sigma)$ provided one can find the derivatives of $F(x,\sigma)$ evaluated at that point. For example, Dedola, Schmitt-Grohe, and Uribe (2002) approximates the functions $g$ and $h$ around a point $x_t \neq \bar{x}$ and $\sigma = 0$. This approach is useful in cases where one can find the exact deterministic solution of a model but needs to resort to approximation techniques to characterize the solution to the stochastic economy.

In what follows we will show how to use the derivatives $F_{x^k\sigma^j}(x,\sigma)$ to identify the derivatives of the functions $g$ and $h$. In chapter 2 we show how to find the first derivatives of $g$ and $h$ with respect to $x$ and $\sigma$. Finding the first order approximation involves solving a system of equations that is quadratic in the first derivatives of $h$ and $g$. We then show in chapter 3 how to find the second-order derivatives of the functions $h$ and $g$. This step involves solving systems of linear equations. In this sense, it is easier to find the second-order terms than it is to find the first order terms. In fact, given the n-th order derivatives of $h$ and $g$, the (n+1)-th order derivatives can be obtained by solving a linear system of equations. At least conceptually it is easy to envision how to compute, say a 5-th order approximation. In practice this has not been done for anything but very simple models with at most 2 or 3 state variables. The reason is that the system of linear equations that must be solved in such a case becomes quite large and the setup of such system can become challenging.

# Chapter 2

# Linear Perturbation Methods

In this lecture, we explain in detail how to solve for linear approximations to the policy function $h$ and $g$. In addition, we show how to use the solution to compute second moments and impulse response functions. Throughout, we will use a simple real business cycle economy to illustrate how to implement the method. We will review this model next.

## 2.1 Example economy: The neoclassical growth model

For simplicity, we will assume that labor is supplied inelastically.[1] Households are assumed to have the following preferences over a single consumption good, $c_t$,

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\gamma}}{1-\gamma} \qquad 0 < \beta < 1, \gamma > 0, \text{ and } \gamma \neq 1.$$

Here $\beta$ denotes the discount factor and $\gamma$ is a parameter measuring the inverse of intertemporal rate of substitution. The period-by-period budget constraint of the household is given by

$$A_t k_t^{\alpha} = c_t + k_{t+1} - (1-\delta)k_t, \tag{2.1}$$

[1]This simplification and others that we will introduce along the way mainly serve to keep the exposition short but are fairly inconsequential for the ease with which the computational method exposited here can be implemented.

where $k_t$ is the capital stock. In period $t$, the capital stock is predetermined. The variable $A_t$ denotes exogenous technological change and is assumed to evolve over time as

$$(A_{t+1} - 1) = \rho_A(A_t - 1) + \sigma\eta_A\epsilon_{t+1}$$

where $\epsilon_{t+1} \sim N(0,1)$, $\eta_A$ is the standard deviation of the innovation to the technology shock, and $0 \leq \rho_A < 1$ indicates the degree of serial correlation of the technology variable $A_t$. The household should also be subject to some kind of borrowing limit, which we ignore here, as we will focus on stationary solutions to the problem.

The Lagrangian of the household's optimization problem takes the form

$$\mathcal{L} = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{c_t^{1-\gamma}}{1-\gamma} + \lambda_t \left[ A_t k_t^{\alpha} - c_t - k_{t+1} + (1-\delta)k_t \right] \right\}.$$

The first-order optimality conditions are (2.1), and

$$c_t^{-\gamma} = \beta E_t c_{t+1}^{-\gamma} [\alpha A_{t+1} k_{t+1}^{\alpha-1} + 1 - \delta]. \tag{2.2}$$

The variable $k_t$ is an endogenous predetermined variable, so it belongs to the vector $x_t^1$ of endogenous predetermined state variables. The variable $A_t$ belongs to the vector $x_t^2$ of exogenous state variables. Consumption, $c_t$, being non-predetermined, is an element of the vector $y_t$.

To show that the equilibrium conditions of this model can be expressed in the form given in equation (1.1), let $y_t = c_t$ and $x_t = [x_t^1;\ x_t^2] = [k_t;\ A_t]'$. Then

$$E_t f(y_{t+1}, y_t, x_{t+1}, x_t) = E_t \begin{bmatrix} y_{1t}^{-\gamma} - \beta y_{1t+1}^{-\gamma} [\alpha x_{2t+1} x_{1t+1}^{\alpha-1} + 1 - \delta] \\ y_{1t} + x_{1t+1} - x_{2t} x_{1t}^{\alpha} - (1-\delta) x_{1t} \\ (x_{2t+1} - 1) - \rho_A (x_{2t} - 1) \end{bmatrix},$$

where $x_{it}$ and $y_{it}$ denote, respectively, the $i$-th element of the vectors $x_t$ and $y_t$.

## 2.2 Exercises

The following three exercises give examples of economies where the vector $x_t$ contains several lagged values of either endogenous or exogenous states, or of endogenous nonpredetermined variables.

1. Exogenous process with arbitrary AR structure: Suppose that the technology shock process instead of following an AR(1) would follow and AR(l) process with $l > 1$.[2] Show how to write the problem in the form given in equation (1.1). First consider the case $l = 2$, and then show the general case for a finite $l$.

2. Time-to-Build: Suppose that the production of capital is subject to gestation lags. Specifically, in order to build $s$ units of capital available in period $t + J$ the household has to invest $s/J$ units of goods for $J$ consecutive periods starting in period $t$. The evolution of the capital stocks is then given by

$$k_{t+J} = (1 - \delta)k_{t+J-1} + s_t. \tag{2.3}$$

   Investment in each period is equal to the sum of investments in each project initiated in the past $J-1$ periods including the current period:

$$i_t = J^{-1} \sum_{i=0}^{J-1} s_{t-i} \tag{2.4}$$

$$\text{given} \quad k_0,\ s_{-i}, \qquad i = 0, 1, \ldots, J-1,$$

   where $s_{t-i}$ denotes the number of investment projects initiated in period $t - i$ that will be productive in period $t - i + J$ .

   Show how to write the equilibrium conditions that is like the economy described in the text but for the assumption of time to build. Then show how the complete set of equilibrium conditions can be expressed in the form given in equation (1.1).

3. Habit Formation: Consider a preference specification featuring habit formation. Specifically, assume that instead of deriving utility from consumption, $c_t$, households derive utility from an object $h_t$ given by

$$h_t = c_t - \theta s_{t-1}$$

   where $s_{t-1}$ denotes the stock of habit in consumption, which is assumed to evolve over time according to the following law of motion

$$s_t = \rho s_{t-1} + (1 - \rho)c_t.$$

[2] An autoregressive process of order $l$, or AR($l$), is given by: $X_t = A_1 X_{t-1} + A_2 X_{t-2} + \ldots + A_l X_{t-l} + \sigma\epsilon_t$, where the matrices $A_i$ for $i = 1, 2, \ldots, l$ are of size $n \times n$ where $n$ is the length of the vector $X_t$.

The parameter $\rho \in [0, 1)$ measures the speed of adjustment of the stock of habit to variations in consumption. Assume that all aspects of the economy other than the preference specification, are as in the example economy described in the text. Show how to write the equilibrium conditions of the habit formation economy in the form given in equation (1.1).

## 2.3 Policy Function: the general case

We are looking for approximations to $g$ and $h$ around the point $(x, \sigma) = (\bar{x}, 0)$ of the form

$$g(x, \sigma) = g(\bar{x}, 0) + g_x(\bar{x}, 0)(x - \bar{x}) + g_\sigma(\bar{x}, 0)\sigma$$

$$h(x, \sigma) = h(\bar{x}, 0) + h_x(\bar{x}, 0)(x - \bar{x}) + h_\sigma(\bar{x}, 0)\sigma$$

As explained earlier,

$$g(\bar{x}, 0) = \bar{y}$$

and

$$h(\bar{x}, 0) = \bar{x}.$$

The remaining unknown coefficients of the first-order approximation to $g$ and $h$ are identified by using the fact that, by equation (1.7), it must be the case that:

$$F_\sigma(\bar{x}, 0) = 0.$$

and

$$F_x(\bar{x}, 0) = 0$$

To find those derivatives let's repeat equation (1.6)

$$\begin{aligned} F(x, \sigma) &\equiv E_t f(g(h(x, \sigma) + \eta\sigma\epsilon', \sigma), g(x, \sigma), h(x, \sigma) + \eta\sigma\epsilon', x) \\ &= 0. \end{aligned}$$

Taking derivative with respect to the scalar $\sigma$ we find:

$$\begin{aligned} F_\sigma(\bar{x}, 0) &= E_t f_{y'}[g_x(h_\sigma + \eta\epsilon') + g_\sigma] + f_y g_\sigma + f_{x'}(h_\sigma + \eta\epsilon') \\ &= f_{y'}[g_x h_\sigma + g_\sigma] + f_y g_\sigma + f_{x'} h_\sigma \end{aligned}$$

This is a system of $n$ equations. Then imposing

$$F_\sigma(\bar{x}, 0) = 0.$$

one can identify $g_\sigma$ and $h_\sigma$:

$$\left[f_{y'}g_x + f_{x'} \quad f_{y'} + f_y\right] \left[ \begin{array}{c} h_\sigma \\ g_\sigma \end{array} \right] = 0$$

This equation is linear and homogeneous in $g_\sigma$ and $h_\sigma$. Thus, if a unique solution exists, we have that

$$h_\sigma = 0.$$

and

$$g_\sigma = 0.$$

These two expressions represent an important theoretical result. They show that in general, up to first order, one need not correct the constant term of the approximation to the policy function for the size of the variance of the shocks.

This result implies that in a first-order approximation the expected values of $x_t$ and $y_t$ are equal to their non-stochastic steady-state values $\bar{x}$ and $\bar{y}$. In this sense, we can say that in a first-order approximation the certainty equivalence principle holds, that is, the policy function is independent of the variance-covariance matrix of $\epsilon_t$. This is an important limitation of first-order perturbation techniques. Because in many economic applications we are interested in finding the effect of uncertainty on the economy. For example, up to first-order the mean of the rate of return of all all assets must be same. Thus, first-order approximation techniques cannot be used to study risk premia. Another important question that can in general not be addressed with first-order perturbation techniques is how uncertainty affects welfare. This question is at the heart of the recent literature on optimal fiscal and monetary stabilization policy. Because in a first-order accurate solution the unconditional expectation of a variable is equal to the non-stochastic steady state, any two policies that give rise to the same steady state yield, up to first-order the same level of welfare.

To find $g_x$ and $h_x$ differentiate (1.6) with respect to $x$ to obtain the following system

$$F_x(\bar{x}, 0) \quad = \quad f_{y'}g_x h_x + f_y g_x + f_{x'}h_x + f_x$$

Note that the derivatives of $f$ evaluated at $(y', y, x', x) = (\bar{y}, \bar{y}, \bar{x}, \bar{x})$ are known. The above expression represents a system of $n \times n_x$ quadratic equations in the $n \times n_x$ unknowns given by the elements of $g_x$ and $h_x$. Imposing

$$F_x(\bar{x}, 0) = 0$$

the above expression can be written as:

$$[f_{x'} \quad f_{y'}] \begin{bmatrix} I \\ g_x \end{bmatrix} h_x = -[f_x \quad f_y] \begin{bmatrix} I \\ g_x \end{bmatrix}$$

Let $A = [f_{x'} \quad f_{y'}]$ and $B = -[f_x \quad f_y]$. Note that both $A$ and $B$ are known. We thus have the following system of $n \times n_x$ equations:

$$A \begin{bmatrix} I \\ g_x \end{bmatrix} h_x = B \begin{bmatrix} I \\ g_x \end{bmatrix} \tag{2.5}$$

Also, let $P$ the matrix of eigenvectors of the matrix $h_x$ such that

$$h_x P = P\Lambda,$$

where $\Lambda$ is diagonal. (The matlab command eig.m produces such a decomposition). Finally, let

$$Z \equiv \begin{bmatrix} I \\ g_x \end{bmatrix} P.$$

Then the above expression can be written as:

$$AZ\Lambda = BZ,$$

We can then map the above problem into a generalized eigenvalue problem. For given $n \times n$ matrices $A$ and $B$, there exists a matrix $V$ and a diagonal matrix $D$ such that (using the matlab command [V,D]=eig(B,A)):

$$A[V_1 \quad V_2] \begin{bmatrix} D_{11} & \emptyset \\ \emptyset & D_{22} \end{bmatrix} = B[V_1 \quad V_2]$$

Assume, without of loss of generality, that all the eigenvalues of $D_{11}$ have modulus less than unity. Then we have:

$$\Lambda = D_{11}$$

$$\begin{bmatrix} I \\ g_x \end{bmatrix} P = V_1 = \begin{bmatrix} V_{11} \\ V_{12} \end{bmatrix}$$

So that

$$\boxed{h_x = V_{11} D_{11} V_{11}^{-1}}$$

and

$$\boxed{g_x = V_{12} V_{11}^{-1}}$$

What regularity conditions are needed to use this solution algorithm? We need that the matrix $P$ is invertible. This will be the case as long as the the eigenvectors of $h_x$ are all linearly independent. The matlab program `gxhx.m` posted on the course webpage implements this generalized eigenvalue procedure to finding $h_x$ and $g_x$.

### 2.3.1 An alternative procedure: The Schur Decomposition Method

Rather than using the generalized eigenvalue decomposition one can use the Schur decomposition[3] to find the find $g_x$ and $h_x$. The matlab program `gx_hx.m` on our webpage implements the Schur Decomposition Method for finding $h_x$ and $g_x$.

Let $\hat{x}_t \equiv x_t - \bar{x}$, then postmultiplying the above system equation (2.5) by $\hat{x}_t$ we obtain:

$$A \begin{bmatrix} I \\ g_x \end{bmatrix} h_x \hat{x}_t = B \begin{bmatrix} I \\ g_x \end{bmatrix} \hat{x}_t$$

Consider for the moment, a perfect foresight equilibrium. In this case, $h_x \hat{x}_t = \hat{x}_{t+1}$.

$$A \begin{bmatrix} I \\ g_x \end{bmatrix} \hat{x}_{t+1} = B \begin{bmatrix} I \\ g_x \end{bmatrix} \hat{x}_t$$

We are interested in solutions in which

$$\lim_{t \to \infty} |\hat{x}_t| < \infty$$

We will use this limiting conditions to find the matrix $g_x$. In particular, we will use the Schur decomposition method.

To solve the above system, we use the generalized Schur decomposition of the matrices $A$ and $B$.[4] The generalized Schur decomposition of A and B is given by upper triangular matrices $a$ and $b$ and orthonormal matrices $q$ and $z$ satisfying:[5]

$$qAz = a$$

and

$$qBz = b.$$

[3]The exposition of the Schur decomposition method is adapted from Martin Uribe's manuscript 'Lectures in Open Economy Macroeconomics,' which is available online at his website: www.econ.duke.edu/∼uribe.

[4]More formal descriptions of the method can be found in Klein (2000).

[5]Recall that a matrix $a$ is said to be upper triangular if elements $a_{ij} = 0$ for $i > j$. A matrix $z$ is orthonormal if $z'z = zz' = I$.

Let

$$s_t \equiv z'[I; g_x]\hat{x}_t.$$

Then we have that

$$as_{t+1} = bs_t$$

Now partition $a$, $b$, $z$, and $s_t$ as

$$a = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix}, b = \begin{bmatrix} b_{11} & b_{12} \\ 0 & b_{22} \end{bmatrix}; z = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix}; s_t = \begin{bmatrix} s_t^1 \\ s_t^2 \end{bmatrix},$$

where $a_{22}$ and $b_{22}$ are of order $n_y \times n_y$, $z_{12}$ is of order $n_x \times n_y$, and $s_t^2$ is of order $n_y \times 1$. Then we have that

$$a_{22}s_{t+1}^2 = b_{22}s_t^2,$$

or

$$b_{22}^{-1}a_{22}s_{t+1}^2 = s_t^2.$$

Assume, without loss of generality, that the ratios $\text{abs}(a_{ii}/b_{ii})$ are decreasing in $i$. Suppose further that the number of ratios less than unity is exactly equal to the number of control variables, $n_y$, and that the number of ratios greater than one is equal to the number of state variables, $n_x$. By construction, the eigenvalues of $b_{22}^{-1}a_{22}$ are all less than unity in modulus.[6] Thus, the requirement $\lim_{j\to\infty}|s_{t+j}^2| < \infty$ is satisfied only if $s_t^2 = 0$. In turn, by the definition of $s_t^2$, this restriction implies that

$$(z_{12}' + z_{22}'g_x)\hat{x}_t = 0.$$

Because this condition has to hold for any value of the state vector, $\hat{x}_t$, it follows that it must be the case that

$$z_{12}' + z_{22}'g_x = 0.$$

Solving this expression for $g_x$ yields

$$g_x = -{z_{22}'}^{-1}z_{12}'.$$

The fact that $s_t^2 = 0$ also implies that

$$a_{11}s_{t+1}^1 = b_{11}s_t^1,$$

[6]Here we are applying a number of properties of upper triangular matrices. Namely, (a) The inverse of a nonsingular upper triangular matrix is upper triangular. (b) the product of two upper triangular matrices is upper triangular. (c) The eigenvalues of an upper triangular matrix are the elements of its main diagonal.

or

$$s^1_{t+1} = {a_{11}}^{-1} b_{11} s^1_t$$

Now

$$s^1_t = (z'_{11} + z'_{21} g_x) \hat{x}_t.$$

Replacing $g_x$, we have

$$s^1_t = [z'_{11} - z'_{21} {z'_{22}}^{-1} z'_{12}] \hat{x}_t.$$

Combining this expression with the equation describing the evolution of $s_t$ shown two lines above, we get

$$\hat{x}_{t+1} = [z'_{11} - z'_{21} {z'_{22}}^{-1} z'_{12}]^{-1} {a_{11}}^{-1} b_{11} [z'_{11} - z'_{21} {z'_{22}}^{-1} z'_{12}] \hat{x}_t;$$

so that

$$h_x = [z'_{11} - z'_{21} {z'_{22}}^{-1} z'_{12}]^{-1} {a_{11}}^{-1} b_{11} [z'_{11} - z'_{21} {z'_{22}}^{-1} z'_{12}].$$

We can simplify this expression for $h_x$ by using the following restrictions:

$$I = z'z = \begin{bmatrix} z'_{11} z_{11} + z'_{21} z_{21} & z'_{11} z_{12} + z'_{21} z_{22} \\ z'_{12} z_{11} + z'_{22} z_{21} & z'_{12} z_{12} + z'_{22} z_{22} \end{bmatrix}$$

to write:[7]

$$h_x = z_{11} a_{11}^{-1} b_{11} z_{11}^{-1}.$$

## 2.4 Local Existence and Uniqueness of Equilibrium

In the above discussion, we assumed that the number of eigenvalues of $D$ with modulus greater than unity is exactly equal to the number of control variables, $n_y$, and that the number of eigenvalues of $D$ with modulus less than one is equal to the number of state variables, $n_x$. In this case there is a unique local equilibrium. But not for every economy this is the case. Let's first consider the case that the number of eigenvalues of $D$ with modulus

[7] To obtain this simple expression for $h_x$, use element $(2,1)$ of $z'z$ to get $z'_{12} z_{11} = -z'_{22} z_{21}$. Premultiply by ${z'_{22}}^{-1}$ and post multiply by $z_{11}^{-1}$ to get ${z'_{22}}^{-1} z'_{12} = -z_{21} {z_{11}}^{-1}$. Use this expression to eliminate ${z'_{22}}^{-1} z'_{12}$ From the square bracket in the expression for $h_x$. Then this square bracket becomes $[z'_{11} + z'_{21} z_{21} {z_{11}}^{-1}]$. Now use element $(1,1)$ of $z'z$ to write $z'_{21} z_{21} = I - z'_{11} z_{11}$. Using this equation to eliminate $z'_{21} z_{21}$ from the expression in square brackets, we get $[z'_{11} + (I - z'_{11} z_{11}) {z_{11}}^{-1}]$, which is simply $z_{11}^{-1}$.

greater than unity is equal to $m < n_y$, which is less than the number of control variables. Then the requirement that we wish to study equilibria in which $lim_{j\to\infty} E_t|\hat{x}_{t+j}| < \infty$ will only yield $m$ restrictions, rather than $n_y$ restrictions. It follows that one can choose arbitrary initial values for $n_y - m$ elements of $y_0$ and the resulting first-order solution will still be expected to converge back to the steady state. In this case the equilibrium is indeterminate.

On the other hand, if the number of eigenvalues of $D$ with modulus greater than unity is greater than the number of control variables, $n_y$, then no local equilibrium exists. Let again $m$ denote the number of eigenvalues of $D$ greater than unity in modulus and assume that $m > n_y$. Then in order to ensure that $\lim E_t|\hat{x}_{t+j}| < \infty$ we must set $m$ elements of $[x_0 y_0]$ equal to zero. This implies that $m - n_y$ elements of $x_0$ must be functions of the remaining $n_x - (m - n_y)$ elements. But this can never be the case, because $x_0$ is a vector of predetermined or exogenous variables and therefore its elements can take arbitrary values. In this case, we say no local equilibrium exists.

## 2.5 Second Moments

Start with the equilibrium law of motion of the deviation of the state vector with respect to its steady-state value, which is given by

$$\hat{x}_{t+1} = h_x \hat{x}_t + \sigma\eta\epsilon_{t+1}, \tag{2.6}$$

**Covariance Matrix of $x_t$**

Let

$$\Sigma_x \equiv E\hat{x}_t\hat{x}_t'$$

denote the unconditional variance/covariance matrix of $\hat{x}_t$ and let

$$\Sigma_\epsilon \equiv \sigma^2\eta\eta'.$$

Then we have that

$$\Sigma_x = h_x\Sigma_x h_x' + \Sigma_\epsilon.$$

We will describe two numerical methods to compute $\Sigma_x$.

**Method 1**

One way to obtain $\Sigma_x$ is to make use of the following useful result. Let $A$, $B$, and $C$ be matrices whose dimensions are such that the product $ABC$

exists. Then

$$\text{vec}(ABC) = (C' \otimes A) \cdot \text{vec}(B),$$

where the vec operator transforms a matrix into a vector by stacking its columns, and the symbol $\otimes$ denotes the Kronecker product. Thus if the vec operator is applied to both sides of

$$\Sigma_x = h_x \Sigma_x h_x' + \Sigma_\epsilon,$$

the result is

$$\begin{aligned} \text{vec}(\Sigma_x) &= \text{vec}(h_x \Sigma_x h_x') + \text{vec}(\Sigma_\epsilon) \\ &= \mathcal{F}\ \text{vec}(\Sigma_x) + \text{vec}(\Sigma_\epsilon), \end{aligned}$$

where

$$\mathcal{F} = h_x \otimes h_x.$$

Solving the above expression for $\text{vec}(\Sigma_x)$ we obtain

$$\text{vec}(\Sigma_x) = (I - \mathcal{F})^{-1} \text{vec}(\Sigma_\epsilon)$$

provided that the inverse of $(I - \mathcal{F})$ exists. The eigenvalues of $\mathcal{F}$ are products of the eigenvalues of the matrix $h_x$. Because all eigenvalues of the matrix $h_x$ have by construction modulus less than one, it follows that all eigenvalues of $\mathcal{F}$ are less than one in modulus. This implies that $(I - \mathcal{F})$ is nonsingular and we can indeed solve for $\Sigma_x$. One possible drawback of this method is that one has to invert a matrix that has dimension $n_x^2 \times n_x^2$.

**Method 2**

The following iterative procedure, called doubling algorithm, may be faster than the one described above in cases in which the number of state variables $(n_x)$ is large.

$$\Sigma_{x,t+1} = h_{x,t} \Sigma_{x,t} h_{x,t}' + \Sigma_{\epsilon,t}$$

$$h_{x,t+1} = h_{x,t} h_{x,t}$$

$$\Sigma_{\epsilon,t+1} = h_{x,t} \Sigma_{\epsilon,t} h_{x,t}' + \Sigma_{\epsilon,t}$$

$$\Sigma_{x,0} = I$$

$$h_{x,0} = h_x$$

$$\Sigma_{\epsilon,0} = \Sigma_\epsilon$$

**Other second moments**

Once the covariance matrix of the state vector, $x_t$ has been computed, it is easy to find other second moments of interest. Consider for instance the covariance matrix $E\hat{x}_t\hat{x}'_{t-j}$ for $j > 0$. Let $\mu_t = \sigma\eta\epsilon_t$.

$$\begin{aligned} E\hat{x}_t\hat{x}'_{t-j} &= E[h_x^j\hat{x}_{t-j} + \sum_{k=0}^{j-1} h_x^k\mu_{t-k}]\hat{x}'_{t-j} \\ &= h_x^j E\hat{x}_{t-j}\hat{x}'_{t-j} \\ &= h_x^j\Sigma_x \end{aligned}$$

Similarly, consider the variance covariance matrix of linear combinations of the state vector $x_t$. For instance, the co-state, or control vector $y_t$ is given by $y_t = \bar{y} + g_x(x_t - \bar{x})$, which we can write as: $\hat{y}_t = g_x\hat{x}_t$. Then

$$\begin{aligned} E\hat{y}_t\hat{y}'_t &= Eg_x\hat{x}_t\hat{x}'_t g'_x \\ &= g_x[E\hat{x}_t\hat{x}'_t]g'_x \\ &= g_x\Sigma_x g'_x \end{aligned}$$

and, more generally,

$$\begin{aligned} E\hat{y}_t\hat{y}'_{t-j} &= g_x[E\hat{x}_t\hat{x}'_{t-j}]g'_x \\ &= g_x h_x^j\Sigma_x g'_x, \end{aligned}$$

for $j \geq 0$.

## 2.6 Impulse Response Functions

The impulse response to a variable, say $z_t$ in period $t + j$ to an impulse in period $t$ is defined as:

$$IR(z_{t+j}) \equiv E_t z_{t+j} - E_{t-1} z_{t+j}$$

The impulse response function traces the expected behavior of the system from period t on given information available in period t, relative to what was expected at time $t - 1$. Using the law of motion $E_t\hat{x}_{t+1} = h_x\hat{x}_t$ for the state vector, letting $x$ denote the innovation to the state vector in period 0, that is, $x = \eta\sigma\epsilon_0$, and applying the law of iterated expectations we get that the impulse response of the state vector in period $t$ is given by

$$IR(\hat{x}_t) \equiv E_0\hat{x}_t - E_{-1}\hat{x}_t = h_x^t[x_0 - E_{-1}x_0] = h_x^t[\eta\sigma\epsilon_0] = h_x^t x; \quad t \geq 0.$$

The response of the vector of controls $\hat{y}_t$ is given by

$$IR(\hat{y}_t) = g_x h_x^t x.$$

## 2.7 Matlab Code For Linear Perturbation Methods

Martin Uribe and I have written a suite of programs that are posted on `www.columbia.edu/~mu2166`. The program gxhx.m computes the matrices $g_x$ and $h_x$ using the generalized eigenvalue method. The program mom.m computes second moments. The program ir.m computes impulse response functions.

There are a number of other programs out there to compute the matrices $g_x$ and $h_x$. For example, Paul Klein of the University of Western Ontario has written Matlab programs that compute the matrices $h_x$ and $g_x$ given $A$, $B$, and $n_x$, using the Schur decomposition methods that we also discussed in the lecture notes (but not in class). To use this alternative method one needs to use a set of programs consisting of three files, solab.m, qzswitch.m, and qzdiv.m, which are available on Paul Klein's website, http://www.ssc.uwo.ca/economics/faculty/klein/. The programs qzswitch.m and qzdiv.m were created by Chris Sims of Princeton University. The files qzswitch.m and qzdiv.m are needed to arrive at a Schur decomposition in which the ratios abs($a_{ii}/b_{ii}$) are decreasing in $i$. The program solab.m, then uses the matrices $a$ and $b$ to find the matrices $g_x$ and $h_x$.

## 2.8 Exercises

The following exercises ask you to apply first-order perturbation methods to numerically analyze the real business cycle model.

1. How does depreciation affect the response of consumption to a positive technology shock? Plot the initial response of consumption to a one percent increase in $A_t$ as a function of $\delta$. Consider values of $\delta$ between 0 and 1. Give intuition for your findings. (Use the set of programs for the neoclassical growth model that are posted on our Second-Order Approximation website.)

2. How does the persistence of the technology shock affect the response of hours? Use the neoclassical growth model we studied earlier. However, now assume that labor supply is endogenous. In particular, the utility function now is given by

$$u(c,h) = \frac{[c(1-h_t)^{\xi}]^{1-\gamma} - 1}{1-\gamma}$$

and that the production function is Cobb Douglas in capital and hours. To calibrate this economy impose that in steady state households spend 20 percent of their time working ($h^{ss} = 0.2$) and that the labor share in value added is 70 percent $(1 - \alpha) = 0.7$. Assume that the rate of depreciation is 10 percent per year. Find the steady state of the economy and the parameter value for $\xi$. Then plot the initial response of hours to a one percent increase in the technology shock as a function of the serial correlation coefficient $\rho_A$. Provide intuition for your findings.

3. How does the value of the Frisch elasticity of labor supply affect (a) the volatility of hours relative to that of output; (b) the correlation between hours and output; (c) the correlation between hours and labor productivity; (d) the correlation between output and labor productivity;

# Chapter 3

# Second-Order Perturbation Methods

Most models used in modern macroeconomics are too complex to allow for exact solutions. For this reason, researchers have appealed to numerical approximation techniques. One popular and widely used approximation technique that we studied in the previous lecture is a first-order perturbation method delivering a linear approximation to the policy function. One reason for the popularity of first-order perturbation techniques is that they do not suffer from the 'curse of dimensionality.' That is, problems with a large number of state variables can be handled without much computational demands. Because models that are successful in accounting for many aspects of observed business cycles are bound to be large, this advantage of perturbation techniques is of particular importance for policy evaluation. However, first-order approximation techniques suffer from serious limitations when applied to welfare evaluation. The problem is that when welfare is evaluated using a first-order approximation to the equilibrium laws of motion of endogenous variables, some second- and higher-order terms of the equilibrium welfare function are omitted while others are included. Consequently, the resulting criterion is inaccurate to order two or higher. This inaccuracy may result in spurious welfare rankings. For instance, in a recent paper Jinill Kim and Sunghyun Kim show that in a simple two-agent economy, a welfare comparison based on an evaluation of the utility function using a linear approximation to the policy function may yield the erroneous result that welfare is higher under autarky than under full risk sharing. Similarly, Tesar (1995) reports that the welfare gains from risksharing a negative in some cases. Again this erroneous finding is due to the fact that she uses

first order approximations to the decision rules. If welfare were directly approximated to first order then, as we saw in the previous class, all policies that imply the same non-stochastic steady state will give rise to the same level of welfare up to first-order. So, for example, the welfare gains from risk sharing should be zero.

In general, to be able to measure welfare gains from different stabilization policies or from different risk sharing arrangements requires a correct second-order approximation of the equilibrium welfare function, which in turn requires a second-order approximation to the policy function. This is what we will study in this chapter. First, we derive an accurate second-order approximation to the solution of discrete-time rational expectations models whose equilibrium conditions can be expressed as in equation (1.1).

In today's lecture we will show that for any model belonging to this general class, the coefficients on the terms linear and quadratic in the state vector in a second-order expansion of the decision rule are independent of the volatility of the exogenous shocks. In other words, these coefficients must be the same in the stochastic and the deterministic versions of the model. Thus, up to second order, the presence of uncertainty affects only the constant term of the decision rules. But the fact that only the constant term is affected by the presence of uncertainty is by no means inconsequential. For it implies that up to second order the unconditional mean of endogenous variables can in general be significantly different from their non-stochastic steady state values. Thus, second-order approximation methods can in principle capture important effects of uncertainty on, for example, average rate of return differentials across assets with different risk characteristics or on the average level of consumer welfare. An additional advantage of higher-order perturbation methods is that like their first-order counterparts, they do not suffer from the curse of dimensionality. This is because given the first-order approximation to the policy function, finding the coefficients of a second-order approximation simply entails solving a system of linear equations.

## 3.1 Finding the second-order approximation

The second-order approximations to $g$ and $h$ around the point $(x, \sigma) = (\bar{x}, 0)$ are of the form

$$\begin{aligned}[g(x,\sigma)]^i &= [g(\bar{x},0)]^i + [g_x(\bar{x},0)]^i_a[(x-\bar{x})]_a + [g_\sigma(\bar{x},0)]^i[\sigma] \\ &\quad + \frac{1}{2}[g_{xx}(\bar{x},0)]^i_{ab}[(x-\bar{x})]_a[(x-\bar{x})]_b\end{aligned}$$

$$
\begin{aligned}
& & & +\frac{1}{2}[g_{x\sigma}(\bar{x},0)]^i_a[(x-\bar{x})]_a[\sigma] \\
& & & +\frac{1}{2}[g_{\sigma x}(\bar{x},0)]^i_a[(x-\bar{x})]_a[\sigma] \\
& & & +\frac{1}{2}[g_{\sigma\sigma}(\bar{x},0)]^i[\sigma][\sigma] \\
[h(x,\sigma)]^j &=& [h(\bar{x},0)]^j + [h_x(\bar{x},0)]^j_a[(x-\bar{x})]_a + [h_\sigma(\bar{x},0)]^j[\sigma] \\
& & & +\frac{1}{2}[h_{xx}(\bar{x},0)]^j_{ab}[(x-\bar{x})]_a[(x-\bar{x})]_b \\
& & & +\frac{1}{2}[h_{x\sigma}(\bar{x},0)]^j_a[(x-\bar{x})]_a[\sigma] \\
& & & +\frac{1}{2}[h_{\sigma x}(\bar{x},0)]^j_a[(x-\bar{x})]_a[\sigma] \\
& & & +\frac{1}{2}[h_{\sigma\sigma}(\bar{x},0)]^j[\sigma][\sigma],
\end{aligned}
$$

where $i = 1, \ldots, n_y$, $a, b = 1, \ldots, n_x$, and $j = 1, \ldots, n_x$.

Here we are using a notation suggested by Collard and Juillard (2001a). So, for example, $[f_{y'}]^i_\alpha$ is the $(i,\alpha)$ element of the derivative of $f$ with respect to $y'$. The derivative of $f$ with respect to $y'$ is an $n \times n_y$ matrix. Therefore, $[f_{y'}]^i_\alpha$ is the element of this matrix located at the intersection of the $i$-th row and $\alpha$-th column. Also, for example, $[f_{y'}]^i_\alpha[g_x]^\alpha_\beta[h_x]^\beta_j = \sum_{\alpha=1}^{n_y}\sum_{\beta=1}^{n_x}\frac{\partial f^i}{\partial y'^\alpha}\frac{\partial g^\alpha}{\partial x^\beta}\frac{\partial h^\beta}{\partial x^j}$.

The unknowns of this expansion are $[g_{xx}]^i_{ab}$, $[g_{x\sigma}]^i_a$, $[g_{\sigma x}]^i_a$, $[g_{\sigma\sigma}]^i$, $[h_{xx}]^j_{ab}$, $[h_{x\sigma}]^j_a$, $[h_{\sigma x}]^j_a$, $[h_{\sigma\sigma}]^j$, where we have omitted the argument $(\bar{x},0)$. These coefficients can be identified by taking the derivative of $F(x,\sigma)$ with respect to $x$ and $\sigma$ twice and evaluating them at $(x,\sigma) = (\bar{x},0)$. By the arguments provided earlier, these derivatives must be zero. Specifically, we use $F_{xx}(\bar{x},0)$ to identify $g_{xx}(\bar{x},0)$ and $h_{xx}(\bar{x},0)$. That is,[1]

$$
\begin{aligned}
[F_{xx}(\bar{x},0)]^i_{jk} &=& \left([f_{y'y'}]^i_{\alpha\gamma}[g_x]^\gamma_\delta[h_x]^\delta_k + [f_{y'y}]^i_{\alpha\gamma}[g_x]^\gamma_k + [f_{y'x'}]^i_{\alpha\delta}[h_x]^\delta_k + [f_{y'x}]^i_{\alpha k}\right)[g_x]^\alpha_\beta[h_x]^\beta_j \\
& & +[f_{y'}]^i_\alpha[g_{xx}]^\alpha_{\beta\delta}[h_x]^\delta_k[h_x]^\beta_j \\
& & +[f_{y'}]^i_\alpha[g_x]^\alpha_\beta[h_{xx}]^\beta_{jk} \\
& & +\left([f_{yy'}]^i_{\alpha\gamma}[g_x]^\gamma_\delta[h_x]^\delta_k + [f_{yy}]^i_{\alpha\gamma}[g_x]^\gamma_k + [f_{yx'}]^i_{\alpha\delta}[h_x]^\delta_k + [f_{yx}]^i_{\alpha k}\right)[g_x]^\alpha_j \\
& & +[f_y]^i_\alpha[g_{xx}]^\alpha_{jk}
\end{aligned}
$$

[1] At this point, an additional word about notation is in order. Take for example the expression $[f_{y'y'}]^i_{\alpha\gamma}$. Note that $f_{y'y'}$ is a three dimensional array with $n$ rows, $n_y$ columns, and $n_y$ pages. Then $[f_{y'y'}]^i_{\alpha\gamma}$ denotes the element of $f_{y'y'}$ located at the intersection of row $i$, column $\alpha$ and page $\gamma$.

$$
\begin{aligned}
&+\left([f_{x'y'}]^i_{\beta\gamma}[g_x]^\gamma_\delta[h_x]^\delta_k+[f_{x'y}]^i_{\beta\gamma}[g_x]^\gamma_k+[f_{x'x'}]^i_{\beta\delta}[h_x]^\delta_k+[f_{x'x}]^i_{\beta k}\right)[h_x]^\beta_j\\
&+[f_{x'}]^i_\beta[h_{xx}]^\beta_{jk}\\
&+[f_{xy'}]^i_{j\gamma}[g_x]^\gamma_\delta[h_x]^\delta_k+[f_{xy}]^i_{j\gamma}[g_x]^\gamma_k+[f_{xx'}]^i_{j\delta}[h_x]^\delta_k+[f_{xx}]^i_{jk}\\
=\;& 0;\quad i=1,\ldots n,\quad j,k,\beta,\delta=1,\ldots n_x;\quad \alpha,\gamma=1,\ldots n_y.
\end{aligned}
$$

Since we know the derivatives of $f$ as well as the first derivatives of $g$ and $h$ evaluated at $(y',y,x',x)=(\bar y,\bar y,\bar x,\bar x)$, it follows that the above expression represents a system of $n\times n_x\times n_x$ **linear** equations in the $n\times n_x\times n_x$ unknowns given by the elements of $g_{xx}$ and $h_{xx}$.

Similarly, $g_{\sigma\sigma}$ and $h_{\sigma\sigma}$ can be obtained by solving the linear system $F_{\sigma\sigma}(\bar x,0)=0$. More explicitly,

$$
\begin{aligned}
[F_{\sigma\sigma}(\bar x,0)]^i =\;& [f_{y'}]^i_\alpha[g_x]^\alpha_\beta[h_{\sigma\sigma}]^\beta\\
&+[f_{y'y'}]^i_{\alpha\gamma}[g_x]^\gamma_\delta[\eta]^\delta_\xi[g_x]^\alpha_\beta[\eta]^\beta_\phi[I]^\phi_\xi\\
&+[f_{y'x'}]^i_{\alpha\delta}[\eta]^\delta_\xi[g_x]^\alpha_\beta[\eta]^\beta_\phi[I]^\phi_\xi\\
&+[f_{y'}]^i_\alpha[g_{xx}]^\alpha_{\beta\delta}[\eta]^\delta_\xi[\eta]^\beta_\phi[I]^\phi_\xi\\
&+[f_{y'}]^i_\alpha[g_{\sigma\sigma}]^\alpha\\
&+[f_y]^i_\alpha[g_{\sigma\sigma}]^\alpha\\
&+[f_{x'}]^i_\beta[h_{\sigma\sigma}]^\beta\\
&+[f_{x'y'}]^i_{\beta\gamma}[g_x]^\gamma_\delta[\eta]^\delta_\xi[\eta]^\beta_\phi[I]^\phi_\xi\\
&+[f_{x'x'}]^i_{\beta\delta}[\eta]^\delta_\xi[\eta]^\beta_\phi[I]^\phi_\xi\\
=\;& 0;\quad i=1,\ldots,n;\quad \alpha,\gamma=1,\ldots,n_y;\quad \beta,\delta=1,\ldots,n_x;\quad \phi,\xi=1,\ldots,n_\epsilon.
\end{aligned}
\tag{3.1}
$$

This is a system of $n$ linear equations in the $n$ unknowns given by the elements of $g_{\sigma\sigma}$ and $h_{\sigma\sigma}$.

Finally, we show that the cross derivatives $g_{x\sigma}$ and $h_{x\sigma}$ are equal to zero when evaluated at $(\bar x,0)$. We write the system $F_{\sigma x}(\bar x,0)=0$ taking into account that all terms containing either $g_\sigma$ or $h_\sigma$ are zero at $(\bar x,0)$. Then we have,

$$
\begin{aligned}
[F_{\sigma x}(\bar x,0)]^i_j =\;& [f_{y'}]^i_\alpha[g_x]^\alpha_\beta[h_{\sigma x}]^\beta_j+[f_{y'}]^i_\alpha[g_{\sigma x}]^\alpha_\gamma[h_x]^\gamma_j+[f_y]^i_\alpha[g_{\sigma x}]^\alpha_j+[f_{x'}]^i_\beta[h_{\sigma x}]^\beta_j\\
=\;& 0;\quad i=1,\ldots n;\quad \alpha=1,\ldots,n_y;\quad \beta,\gamma,j=1,\ldots,n_x.
\end{aligned}
\tag{3.2}
$$

This is a system of $n\times n_x$ equations in the $n\times n_x$ unknowns given by the elements of $g_{\sigma x}$ and $h_{\sigma x}$. But clearly, the system is homogeneous in the unknowns. Thus, if a unique solution exists, it is given by

$$g_{\sigma x}=0$$

and

$$h_{\sigma x} = 0.$$

These equations represent an important theoretical result. They show that in general, up to second-order, the coefficients of the policy function on the terms that are linear in the state vector do not depend on the size of the variance of the underlying shocks.

This result implies that the second-order approximation to the policy function of a stochastic model belonging to the general class given in equation (1.1) differs from that of its non-stochastic counterpart only in a constant term given by $\frac{1}{2}g_{\sigma\sigma}\sigma^2$ for the control vector $y_t$ and by $\frac{1}{2}h_{\sigma\sigma}\sigma^2$ for the state vector $x_t$.

As we stressed above what is simple about finding the second order terms of the functions $g$ and $h$ that it only involves linear operations. This suggests that one can postulate the problem in matrix form. This is what is done, numerically, in the programs gxx_hxx.m and gss_hss.m.

## 3.2 Implementing Second-Order Methods

### 3.2.1 Computing the derivatives of $f$

Computing the derivatives of $f$, particulary, second derivatives, can be daunting task. It is almost impossible to do this by hand for any but the most simple of models. And most of us would get lost in algebra mistakes. So clearly one must resort to the computer for finding those derivates. Here one has two options. One is to try to find the numerical derivatives. But this is not the way we approach the problem in Schmitt-Grohe and Uribe (JEDC 2004). We instead make use of the capabilities of symbolic math software. The particular one we use is the MATLAB Toolbox Symbolic Math. Symbolic Math can handle analytical derivatives. We wrote programs, that compute the analytical derivatives of $f$ and evaluate them at the steady state. The program anal_deriv.m computes analyltical derivatives of $f$ and the program num_eval.m evaluates the analytical derivatives of $f$. You can check these programs out on the Second-Order Approximation website.

### 3.2.2 Which steady state to approximate around

In general, the second-order method exposited here can be used to compute welfare in the neighborhood of an arbitrary steady state.

However, in a monetary or fiscal policy stabilization problem, the steady state of the economy depends on the assumed policy. Suppose we wish

to evaluate the welfare consequences of alternative monetary stabilization policies with the objective to pick the best of those. In such an exercise one can, in principle, freely pick the steady state level of inflation around which the economy should be stabilized. For example, one could study the case in which inflation is stablized around, say 5 percent of inflation per year. But if, in the particular model studied the optimal inflation rate in the steady state is, say 0 percent per year, such a research strategy, may lead to the following problem. The benevolent policy maker may want to use stabilization policy to get the mean of inflation down, that is down from 5 percent to something closer to 0 percent. And the welfare gains from influencing the mean may be larger than stabilizing the economy around the deterministic steady state. For this reason, the accuracy of the approximation may be poor. Therefore, if one uses second-order approximation methods to find the best among a class of suboptimal policies, one should approximate the economy around the Ramsey steady state.

In other contexts, that is when one is not using the second-order approximation to find the best constrained policy, the perturbation method can be applied successfully to any arbitrary non-stochastic steady state. Say, if we want to characterize risk premia, then the accuracy of the second-order perturbation method should be independent of the Ramsey optimality of the point around which the solution is approximated.

### 3.2.3 How to construct second-order accurate time series

Suppose one wishes to simulate a time series path of the approximate solution for $x_t$ and $y_t$. One possibility of proceeding is to start with an initial value of the state, say $x_0$, and then to use the second-order approximation to the function $h$ iteratively to produce a time path. That is, iterate on: (Assume w.l.g. that the steady state value of $x$ is zero.)

$$x_1^i = h_x^i x0 + x_0' h_{xx}^i x_0 + \sigma \eta^i \epsilon_1.$$

In this way one can obtain time series for $x_t$ and $y_t$ conditional upon a particular value of $x_0$. This procedure is second-order accurate, but it introduces some higher-order terms into the expansion. These extra higher order-terms do not in general increase accuracy of the approximation as they do not correspoind to higer-order coefficients in a Taylor expansion. In practise they often lead to explosive time paths for $x_t$. To see what goes wrong let (this example is taken from Kim, Kim, Schaumburg, and Sims)

$$x_{t+1} = \rho x_t + \alpha x_t^2 + \epsilon_{t+1}; \quad 0 < \rho < 1$$

Note that this system has two steady states. One at $\bar{x} = 0$ and the other at $\bar{x} = (1-\rho)/\alpha$. Linearizing ths system around $x = \bar{x}$ yields:

$$\hat{x}_{t+1} = \rho\hat{x}_t + \epsilon_{t+1}$$

Given the assumpmtion that $\rho < 1$, this system is sstable in the neighborhood of $x = 0$. But the sysstem is unstable in the neighborhood of $x = (1-\rho)/\alpha$. Thus, once $x_t > (1-\rho)/\alpha$, the system will diverge.

Therefore, one should choose among the second-order accurate expansions the o ne that implies stability. A stable solution can be obtained by 'PRUNING' (a term coined by Chris Sims) out the the extraneous higher-order terms in each iteration by computing the projections of the second-order terms based on a first-order expansions. L

Let $x_t^f$ be the first-order iterate of $x_t$ given some $x_0$., that is , $x_0^f = x_0$, and let $x_t^s$ be the second-order iterate of $x_t$ given $x_0$ and set $x_0^s = 0$. Then define

$$x_{t+1}^f = h_x x_t^f + \sigma\eta\epsilon_{t+1}$$

and

$$x_{t+1}^{is} = h_x^i x_t^s + \frac{1}{2}x_t^{f\prime} h_{xx}^i x_t^f + \frac{1}{2} + h_{\sigma\sigma}^i\sigma^2$$

The program simu_2nd.m posted on our Second Order website constructs such time series.

### 3.2.4 How to compute unconditional first moments

First note that the variance of either $x_t$ or $y_t$ are correct to second-order accuracy when computed from the first-order approximation alone. (That is, you can use the program mom.m to find the second-oder accurate second-moments.) Then, suppose that the unconditional expectation of $x_t$ exists, we can approximate this from its law of motion (assume without loss of generality that $x^{SS} = 0$:

$$x_{t+1}^i \approx h_x^i x_t + \frac{1}{2}x_t' h_{xx}^i x_t + \frac{1}{2}h_{\sigma\sigma}^i\sigma^2 + \sigma\eta^i\epsilon_{t+1}$$

Note that eh scalar

$$x_t' h_{xx}^i x_t = x_t' \otimes x_t' vec(h_{xx}^i)$$

and that $Ex_t' \otimes x_t' = vec(\Sigma_x)'$. We then have that :

$$\begin{aligned} Ex_t' h_{xx}^i x_t &= Ex_t' \otimes x_t' vec(h_{xx}^i) \\ &= (vec(\Sigma_x))' vec(h_{xx}^i) \end{aligned}$$

It follows that

$$Ex_{t+1} = h_x Ex_t + \begin{bmatrix} vec(h^1_{xx})' \\ vec(h^2_{xx})' \\ \dots \\ vec(h^{n_x}_{xx})' \end{bmatrix} vec(\Sigma_x)/2 + \frac{1}{2} h_{\sigma\sigma}\sigma^2$$

$$Ex_t = (I - h_x)^{-1} \left( \begin{bmatrix} vec(h^1_{xx})' \\ vec(h^2_{xx})' \\ \dots \\ vec(h^{n_x}_{xx})' \end{bmatrix} vec(\Sigma_x)/2 + \frac{1}{2} h_{\sigma\sigma}\sigma^2 \right)$$

The program unconditional_means.m computes unconditioanl means.

## 3.3 Second-order accurate welfare approximations

Suppose we wish to evaluate the welfare consequences of alternative monetary policy arrangements. How should welfare be measured? Welfare is defined as the present dicounted value of lifetime utility. One could measure this either using the conditional expectation of lifetime utility or the unconditional expection:

### Conditional Welfare Measures

Let $V_t$ denote the time $t$ value of the present disounted value of lifetime utility. For example, in the case of the economy desribed in section 2.1 above we have:

$$V_t \equiv E_t \sum_{j=0}^{\infty} \beta^j U(c_{t+j})$$

What we are looking for is a second-order accurate approximation to $V_t$. At first instinct one may think that one way to obtain $V_t$ is to find a second-order accurate time series for $c_t$ and then just use the sample mean of say 1000 simulated paths for $\sum_{t=0}^{T} \beta^t U(c_t)$ for some rather large $T$. This way of proceeding is however unneccarily cumbersome and may introduce some errors stemming from the fact that one uses a sample path. Instead the way to proceed is to write $V_t$ recursively as:

$$V_t = U(c_t) + \beta E_t V_{t+1}$$

Then add this expression the vector valued function $f$ and make $V_t$ an element of the vector of endogenous non-predetermined variables $y_t$. In this

way, once one has computed second-order approximations to the function $g$, one has the second-order approximation of $V_t$ in hand. The idea is that we can express $V_t$ as some non-linear function of the state vector $x_t$ and of the scalar parameter $\sigma$, that is,

$$V_t = g^V(x_t, \sigma)$$

where, here the function $g^V$ is just one element of the function $g$ that we introduced earlier. It maps $R^{n_x} \times R^+$ into $R$. And by making $V_t$ one row of $f$, we can obtain the second-order accurate approximation to the function $g^V$.

Clearly, the level of welfare will depend on the initial state being considered, that is, on the value of the state vector $x_t$. Depending on the particular question one would like to address, one may consider a number of relevant initial values for $x_t$ and average over those to obtain an average value of conditional welfare levels.

### Welfare cost measures

Often in comparing alternative policies that are consistent with the same non-stochastic steady state equilibrium, one would like to put the welfare differences into perspective. Saying that the welfare of policy A is, say, 255.3, and that associated with policy B is, say 248.7, only tells us that policy A welfare dominates policy B. But often we are interested in asking how much better policy A is than policy B. One way of quantifying these welfare differences is to ask what percentage of the consumption stream associated with policy A, households are willing to give to be as well off, under policy A as under policy B. Let $c_t^A$ the contingent plan for consumption associated with policy A and $c_t^B$ the contingent plan for consumption associated with policy B. We can then implicitly define the welfare cost of adoptiong policy B rather than policy A as the value of $\lambda$ such that

$$V_t^B = E_t \sum_{t=0}^{\infty} \beta^t U(c_t^A(1-\lambda))$$

Suppose the utility function is of the form:

$$U(c) = \frac{c^{1-\gamma}}{1-\gamma}$$

with $0 < \gamma < 1$

So that

$$U(c(1-\lambda)) = (1-\lambda)^{1-\gamma}\left(\frac{c^{1-\gamma}}{1-\gamma}\right)$$

In this case we have:

$$V_t^B = (1-\lambda)^{1-\gamma}V_t^A$$

$$\ln(1-\lambda) = \frac{1}{1-\gamma}\left(\ln V_t^B - \ln V_t^A\right)$$

We also know the second-order approximation to say,

$$\ln V_t^A = \ln \bar{V} + g_x^{V^A}(x_t - \bar{x}) + \frac{1}{2}\left((x_t-\bar{x})' g_{xx}^{V^A}(x_t-\bar{x}) + g_{\sigma\sigma}^{V^A}\sigma^2\right)$$

and similarly:

$$\ln V_t^B = \ln \bar{V} + g_x^{V^B}(x_t - \bar{x}) + \frac{1}{2}\left((x_t-\bar{x})' g_{xx}^{V^B}(x_t-\bar{x}) + g_{\sigma\sigma}^{V^B}\sigma^2\right)$$

So that the approximate welfare cost is

$$\lambda \approx \frac{1}{1-\gamma}\left(\ln V_t^A - \ln V_t^B\right)$$

The welfare cost $\lambda \times 100$ indicates the percentage of the consumption stream associated with policy A that households are willing to give up to be as well off as under policy B. If $\lambda > 0$, it measures the welfare cost of policy B vis-à-vis policy A. If $\lambda < 0$, then in fact $\lambda$ indicates the percentage increase in the consumption stream required to make households as well off under policy A as under policy B.

## 3.4 Exercises

1. Consider the simple RBC model with inelastic labor supply studied earlier. Compute the second order accurate unconditional level of consumption, $E(C_t)$ under the assumption that technology is iid. Then change the information structure. Assume that you learn the time $t+1$ technology shock already in period $t$, that is, the technology shock is one period anticipated. Now compute again the unconditional expectation of consumption, $E(C_t)$. [I think consumption should be higher when the shock is one period anticipated.] How big do you need to make the variance of the technology shock so that the welfare difference is three percent of consumption. How big is the overall volatility of output in that case. [added June 3, 2011]

2. Compute the welfare cost of business cycles for the simple neoclassical growth model given in chapter 1. Set $\beta = 0.95$, $\gamma = 2$, and $\alpha = 0.3$. Assume that the depreciation rate is 10 perent per year. Use as a basis for the approximation the logarithm of the state vector and assume that $\ln A_{t+1} = \rho_A \ln A_t + \sigma\eta\epsilon_{t+1}$. Assume that $\rho_A = 0.95$. Set $\sigma = 1$ and set $\eta$ such that the standard deviaton of output is 1.72 percent. Find the welfare cost of business cycles for three different initial states of the economy, namely, the technology shock being 1 percent above, 1 percent below, and exactly equal to its non-stochastic steady state value. Assume that in all three states capital takes its steady state value. Compare conditional and unconditional welfare measures.

3. Using linear approximation techniques, Tesar (1995)[2] studies the welfare gains from international risk sharing. In some cases the welfare gain of moving from financial autarky to a situation of international risk sharing is reported to be negative. Recompute the welfare gains from risk-sharing using a second-order perturbation method.

4. Consider any DSGE model that displays equilibrium indeterminacy. For example, pick one of the 4 models discussed in Schmitt-Grohé (1997)[3] or a new-Keynesian model with a monetary policy specification that renders the equilibrium indeterminate, say a passive interest rate feedback rule. Use second-order methods to compute welfare. Explore for which parameterizations of the model chosen by you sunspots with a positive variance increase/decrease welfare.

5. Redo Boldrin, Christiano, and Fisher[4] to see if second-order approximations result in about the same risk premia as the numerical method used in that paper.

[2]Tesar, L. Evaluating the Gains from International Risksharing, *Carnegie-Rochester Conference Series on Public Policy*, 42, 1995, 95-143.

[3]Schmitt-Grohe, S., Comparing Four Models of Aggregate Fluctuations Due to Self-Fulfilling Expectations, *Journal of Economic Theory*, 72, January 1997, 96-147.

[4]Boldrin, M., L. Christiano, and J. Fisher, Habit Persistence, Asset Returns, and the Business Cycle, *American Economic Review*, 91, March 2001, 149-166.

# Chapter 4

# N-th Perturbation Methods

It is straightforward to apply the method described thus far to finding higher than second-order approximations to the policy function. For example, given the first- and second-order terms of the Taylor expansions of $h$ and $g$, the third-order terms can be identified by solving a linear system of equations. More generally, one can construct sequentially the nth-order approximation of the policy function by solving a linear system of equations whose (known) coefficients are the lower-order terms and the derivatives up to order $n$ of $f$ evaluated at $(y', y, x', x) = (\bar{y}, \bar{y}, \bar{x}, \bar{x})$.

To be completed.

# Chapter 5

# References


Benigno, Pierpaolo and Michael Woodford, "Inflation Stabilization and Welfare: The Case of a Distorted Steady State," NBER Working Paper No. 10838, October 2004 .

Blanchard, Olivier and Charles Kahn, "The solution of Linear Difference Models under Rational Expectations," *Econometrica 45,* July 1985, 1305-1311.

Collard, Fabrice, and Michel Juillard, "Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models," *Journal of Economic Dynamics and Control 25,* 2001, 979-999.

Collard, Fabrice, and Michel Juillard, "Perturbation Methods for Rational Expectations Models," Manuscript. Paris: CEPREMAP, February 2001.

Fleming, W. H., "Stochastic Control for Small Noise Intensities," *SIAM Journal of Control 9,* 1971, 473-517.

Jin, He-Hui and Kenneth L. Judd, "Perturbation Methods For General Dynamic Stochastic Models," manuscript Stanford University, 2002.

Judd, Kenneth, *Numerical Methods in Economics,* Cambridge, MA, MIT Press, 1998.

Kim, Jinill and Sunghyun Henry Kim, "Spurious Welfare Reversals in International Business Cycle Models," *Journal of International Economics 60,* 2003, 471-500.

Kim, J., Kim, S., Ernst Schaumburg, and Christopher A. Sims, "Calculating and Using Second-Order Accurate Solutions of Discrete-Time Dynamic Equilibrium Models," mimeo, Princeton University, June 19, 2003.

King, R.G., C.L. Plosser and S.T. Rebelo, "Production, Growth and Business Cycles I: The Basic Neoclassical Model," *Journal of Monetary Economics 21,* 1988, 191-232.

Klein, Paul, "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control 24,* 2000, 1405-1423.

Kydland, Finn and Edward Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica 50,* November 1982, 1345-1370.

Sims, Christopher, "Solving Linear Rational Expectations Models," *Computational Economics 20,* October 2002, 1-20.

Sims, Christopher, "Second Order Accurate Solution of Discrete Time Dynamic Equilibrium Models," Manuscript. Princeton: Princeton University, December 2000b.

Tesar, Linda L., "Evaluating the Gains from International Risksharing," *Carnegie-Rochester Conference Series on Public Policy 42,* 1995, 95-143.

Woodford, Michael, "Inflation Stabilization and Welfare," Manuscript. Princeton: Princeton University, June 1999.