Open Source Macroeconomics Laboratory Boot Camp
Perturbation Methods

Kerk Phillips
Brigham Young University

2017

1 Introduction

In this section we will explore in more detail the perturbation methods referenced in section 5.4 of the DSGE chapter. We will only consider a second order approximation of the policy function here, but approximations of yet higher order follow the same basic approach. There are many alterations of the standard perturbation method. Detailed discussions of perturbation methods can be found in chapters 13 – 15 of Judd (1998a), as well as in Collard and Juillard (2001), Schmitt-Grohe and Uribe (2004), and Heer and Maussner (2009a).

As noted previously, assuming that the policy functions are linear can be extremely useful in solving DSGE models. When second or higher order properties of the characterizing equations of the are important to the question being answered, however, linearization is undesirable. Linearization is essentially a first order Taylor series approximation about the steady state, as discussed in section 6 of the DSGE chapter. Such an approximation results in certainty equivalence, or the phenomenon of unconditional expectations of the endogenous variables being equal to their non-stochastic steady state values. This occurs because in a linear approximation, only the first moments of the shocks enter the linear equations. As these processes are assumed to be mean zero, these moments wash out when expectations are taken. Thus, the distribution of the shocks have no influence on the resultant policy equation solutions.

Applications where this can be troublesome include asset pricing models and welfare analysis. In asset pricing models the riskiness of an asset is directly related to the variance of the underlying shocks. Thus, failing to account for higher order characteristics of the model can invalidate the results. In welfare analysis where the utility functions have high curvature, failing to account for the second moment can similarly produce spurious results.

2 Pertubation Methods in General

To see how perturbation methods work consider the following simple example. Suppose we have a condition on a potentially nonlinear bivariate function:

$$F(x, u) = 0 \tag{1}$$

Assume u is an exogenously given variable, and x will be choson to satisfy (1). Denote the solution to this condition as $x(u)$ and assume that the value of $x(u_0)$ is known.

Taking the derivative of (1) with respect to u gives:

$$F_x\{x(u),u\}x_u(u) + F_u\{x(u),u\} = 0 \tag{2}$$

If we evaluate this at $u = u_0$ and solve for the first derivative of $x(u)$, we have:

$$x_u(u_0) = -\frac{F_u\{x(u_0),u_0\}}{F_x\{x(u_0),u_0\}} \tag{3}$$

Since $x(u_0)$ is known, as long as $F_x\{x(u_0),u_0\} \neq 0$ we can find the value for the first derivative. The first-order (linear) Taylor-series approximation of $x(u)$ will be:

$$x(u) = x(u_0) + x_u(u_0)(u - u_0) \tag{4}$$

To find the second-order terms we differentiate (2) again with respect to u.

$$\begin{aligned} &F_{xx}\{x(u),u\}x_u(u)x_u(u) + F_{xu}\{x(u),u\}x_u(u) \\ &+F_x\{x(u),u\}x_{uu}(u) + F_{xu}\{x(u),u\}x_u(u) \\ &+F_{uu}\{x(u),u\} = 0 \end{aligned} \tag{5}$$

Again evaluating at $u = u_0$ and solving this time for the second derivative of $x(u)$, we have:

$$x_{uu}(u_0) = -\frac{F_{xx}\{x(u_0),u_0\}[x_u(u_o)]^2 + 2F_{xu}\{x(u_0),u_0\}x_u(u_o) + F_{uu}}{F_x\{x(u_0),u_0\}} \tag{6}$$

Hence, the second-order (quadratic) Taylor-series approximation of $x(u)$ will be:

$$x(u) = x(u_0) + x_u(u_0)(u - u_0) + \tfrac{1}{2}x_{uu}(u_0)(u - u_0)^2 \tag{7}$$

Higher order terms can be obtained by successive differentiation of (5), setting $u = u_0$ and solving for the appropriate derivative. This will be a function of the various derivatives of $F(x,u)$ and the lower-order derivatives of $x(u)$ obtained from previous iterations.

3 Application to a Simple DGE Model

Consider a simple dynamic general equilibrium model with no stochastic shocks. An example of this is the model from Brock and Mirman (1972), which we have examined before. This model can be written in the following form.

$$\frac{1}{K_t^\alpha - K_{t+1}} = \beta E_t \left\{ \frac{\alpha K_{t+1}^{\alpha-1}}{K_{t+1}^\alpha - K_{t+2}} \right\}$$

We will rewrite this as follows.

$$\frac{1}{K_t^\alpha - K_{t+1}} - \beta \frac{\alpha K_{t+1}^{\alpha-1}}{K_{t+1}^\alpha - K_{t+2}} = 0 \tag{8}$$

We know the steady state is defined by $K_{t+2} = K_{t+1} = K_t = \bar{K}$. Substituting this into (8) gives:

$$\begin{aligned} \beta\alpha\bar{K}^{\alpha-1} &= 1 \\ \bar{K} &= (\beta\alpha)^{\frac{1}{1-\alpha}} \end{aligned} \tag{9}$$

In terms of our notatation from the previous section we have:

$$\begin{aligned} u &= K_t \\ x &= x(u) = K_{t+1} \\ y &= x(x) = K_{t+2} \\ F(y, x, u) = F(y(x(u)), x(u), u) &= \frac{1}{K_t^\alpha - K_{t+1}} - \beta \frac{\alpha K_{t+1}^{\alpha-1}}{K_{t+1}^\alpha - K_{t+2}} = 0 \end{aligned} \tag{10}$$

Take the derivative of (10) with respect to $u = K_t$:

$$\begin{aligned} F_u(y(x(u)), x(u), u) =& F_y(y(x(u)), x(u), \bar{u}) x_u(x(u)) x_u(u) \\ &+ F_x(y(x(u)), x(u), u) x_u(u) + F_u(y(x(u)), x(u), u) = 0 \end{aligned} \tag{11}$$

Evaluating (11) at $u = \bar{u} = \bar{K}$ and noting that $x(\bar{u}) = u$:

$$\begin{aligned} F_u(y(x(\bar{u})), x(\bar{u}), \bar{u}) =& F_y(y(x(\bar{u})), x(\bar{u}), \bar{u}) x_u(x(\bar{u})) x_u(\bar{u}) \\ &+ F_x(y(x(\bar{u})), x(\bar{u}), \bar{u}) x_u(\bar{u}) + F_u(y(x(\bar{u})), x(\bar{u}), \bar{u}) = 0 \\ F_u(\bar{u}, \bar{u}, \bar{u}) =& F_y(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^2 + F_x(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u}) + F_u(\bar{u}, \bar{u}, \bar{u}) = 0 \end{aligned}$$

Note that $F_y(\bar{u}, \bar{u}, \bar{u})$ is the same as F from the linearization notes. Similarly, $F_x(\bar{u}, \bar{u}, \bar{u})$ is G, and $F_u(\bar{u}, \bar{u}, \bar{u})$ is H. Also note that $x_u(\bar{u})$ is P. As in those notes the value of $x_u(\bar{u})$ comes from a solving a quadratic.

Now we take the derivative of (11) with respect to u and evalauate it at $u = \bar{u}$.

$$\begin{aligned} F_{uu}(\bar{u}, \bar{u}, \bar{u}) =& F_{yy}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^4 \\ &+ F_{yx}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^3 \\ &+ F_{yu}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^2 \\ &+ F_y(\bar{u}, \bar{u}, \bar{u})\ x_{uu}(\bar{u})\ x_u(\bar{u})^2 \\ &+ F_y(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})\ x_{uu}(\bar{u}) \\ &+ F_{yx}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^3 \\ &+ F_{xx}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^2 \\ &+ F_{xu}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u}) \\ &+ F_x(\bar{u}, \bar{u}, \bar{u})\ x_{uu}(\bar{u}) \\ &+ F_{yu}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u})^2 \\ &+ F_{xu}(\bar{u}, \bar{u}, \bar{u})\ x_u(\bar{u}) \\ &+ F_{uu}(\bar{u}, \bar{u}, \bar{u}) = 0 \end{aligned} \tag{12}$$

Supressing the function arguments for the sake of clarity we can rewrite (12) as below.

$$\begin{aligned} F_{uu} =& (F_{yy}\ x_u^4 + 2F_{yx}\ x_u^3 + 2F_{yu}\ x_u^2 + F_{xx}\ x_u^2 + 2F_{xu}\ x_u + F_{uu}) \\ &+ (F_y\ x_u^2 + F_y\ x_u + F_x) = 0 \end{aligned}$$

Note the F_{ij} are all second-derivatives evaluated at the steady state. Since x_u has already

been solved we can solve this for x_{uu}.

$$x_{uu} = \frac{F_{yy}\, x_u^4 + 2F_{yx}\, x_u^3 + 2F_{yu}\, x_u^2 + F_{xx}\, x_u^2 + 2F_{xu}\, x_u + F_{uu})}{(F_y\, x_u^2 + F_y\, x_u + F_x)} x_{uu} \tag{13}$$

The quadratic approximation to the policy function is given by:

$$\tilde{K}_{t+1} = x_u \tilde{K}_t + \frac{1}{2} x_{uu} \tilde{K}_t^2 \tag{14}$$

4 Pertubation Methods in Dynamic Stochastic Equilibrium Models

In this section we will explore in more detail the perturbation methods referenced in section 5.4 of the DSGE chapter. We will only consider a second order approximation of the policy function here, but approximations of yet higher order follow the same basic approach. There are many alterations of the standard perturbation method. Detailed discussions of perturbation methods can be found in chapters 13 – 15 of Judd (1998b), as well as in Collard and Juillard (2001), Schmitt-Grohe and Uribe (2004), Heer and Maussner (2009b) and Gomme and Klein (2011). Binning (2012) uses similar techniques to get cubic approximations.

Consider a non-linear system of dynamic equations. We can take natural logs or otherwise transform the equations to get:

$$E_t\{\Gamma(X_{t+1}, X_t, X_{t-1}, Y_{t+1}, Y_t, Z_{t+1}, Z_t)\} = 0 \tag{15}$$

X is a set of endogenous state variables, Y is a set of "jump", costate, or control variables, and Z is a list of exogenous state variables. The methods for obtaining a linear approximation of the policy function are well known from Uhlig (1999). Our task in this section is to obtain the quadratic terms from a second-order approximation of the same policy function.

We must recall that the exogenous state variables evolve according to a linear law of motion given in (16) with the long run value of $\bar{Z} = 0$.

$$Z_t = N Z_{t-1} + v\Omega\varepsilon_t; \varepsilon_t \sim (0, I_{n_Z}) \tag{16}$$

where v is a scalar, and Ω is a matrix that determines correlations of the elements in ε_t.

We are searching for the quadratic terms in the Taylor-series approximation of the policy function and jump variable functions which we will denote:

$$X_t = H(X_{t-1}, Z_t, v) \tag{17}$$

$$Y_t = G(X_{t-1}, Z_t, v) \tag{18}$$

For notational ease define the following.

$$A_t \equiv \begin{bmatrix} X_{t+1} & X_t & X_{t-1} & Y_{t+1} & Y_t & Z_{t+1} & Z_t \end{bmatrix}^T \tag{19}$$

$$S_t \equiv \begin{bmatrix} X_{t-1} & Z_t & v \end{bmatrix} \tag{20}$$

$$n_A \equiv 3n_X + 2n_Y + 2n_Z \tag{21}$$

$$n_s \equiv n_X + n_Z + 1 \tag{22}$$

4.1 Definitions and Preliminaries

The Taylor-series approximation of Γ with second-order terms for the variance is:

$$\begin{aligned} \Gamma(A_t) \doteq \Gamma(\bar{X}, ..., \bar{Z}) + \begin{bmatrix} \Gamma_1 & \dots & \Gamma_7 \end{bmatrix} \begin{bmatrix} \tilde{X}_{t+1} \\ \vdots \\ \tilde{Z}_t \end{bmatrix} \\ + \frac{1}{2} \left(I_{n_A} \otimes \begin{bmatrix} \tilde{X}_{t+1} & \vdots & \tilde{Z}_t \end{bmatrix} \right) \begin{bmatrix} \Gamma_{11} & \dots & \Gamma_{17} \\ \vdots & \ddots & \vdots \\ \Gamma_{71} & \dots & \Gamma_{77} \end{bmatrix} \begin{bmatrix} \tilde{X}_{t+1} \\ \vdots \\ \tilde{Z}_t \end{bmatrix} \end{aligned} \tag{23}$$

The Taylor-series approximation of H with second-order terms for the variance is:

$$\begin{aligned} H(X_{t-1}, Z_t, v) \doteq H(\bar{X}, \bar{Z}, \bar{v}) + \begin{bmatrix} H_X & H_Z & H_v \end{bmatrix} \begin{bmatrix} \tilde{X}_{t-1} \\ \tilde{Z}_t \\ \tilde{v} \end{bmatrix} \\ + \frac{1}{2} \left(I_{n_Y + n_X} \otimes \begin{bmatrix} \tilde{X}_{t-1}^T & \tilde{Z}_t^T & \tilde{v} \end{bmatrix} \right) \begin{bmatrix} H_{XX} & H_{XZ} & 0 \\ H_{ZX} & H_{ZZ} & 0 \\ 0 & 0 & H_{vv} \end{bmatrix} \begin{bmatrix} \tilde{X}_{t-1} \\ \tilde{Z}_t \\ \tilde{v} \end{bmatrix} \end{aligned} \tag{24}$$

H_X and H_Z terms are the P and Q matrices in Uhlig's notation. H_{XX}, H_{ZZ}, H_{ZX}, H_{XZ}^T and H_{vv} are all Magnus and Neudecker (1999) matrices of second-order coefficients. A similar setup is used for the approximation of the G function.

$$\begin{aligned} G^k(X_{t-1}, Z_t, v) \doteq G(\bar{X}, \bar{Z}, \bar{v}) + \begin{bmatrix} G_X & G_Z & G_v \end{bmatrix} \begin{bmatrix} \tilde{X}_{t-1} \\ \tilde{Z}_t \\ \tilde{v} \end{bmatrix} \\ + \frac{1}{2} \left(I_{n_Y + n_X} \otimes \begin{bmatrix} \tilde{X}_{t-1}^T & \tilde{Z}_t^T & \tilde{v} \end{bmatrix} \right) \begin{bmatrix} G_{XX} & G_{XZ} & 0 \\ G_{ZX} & G_{ZZ} & 0 \\ 0 & 0 & G_{vv} \end{bmatrix} \begin{bmatrix} \tilde{X}_{t-1} \\ \tilde{Z}_t \\ \tilde{v} \end{bmatrix} \end{aligned} \tag{25}$$

G_X and G_Z terms are the R and S matrices in Uhlig's notation.

We can substitute (17), (18) and (16) into (19) to get the following function:

$$A_t = F(S_t) = \begin{bmatrix} H\left(H(X_{t-1}, Z_t, v), NZ_t + v\Omega\varepsilon_{t+1}, v\right) \\ H(X_{t-1}, Z_t, v) \\ X_{t-1} \\ G\left(H(X_{t-1}, Z_t, v), NZ_t + v\Omega\varepsilon_{t+1}, v\right) \\ G(X_{t-1}, Z_t, v) \\ NZ_{t-1} + v\Omega\varepsilon_t \\ Z_t \end{bmatrix} \tag{26}$$

The Jacobian of this function is:

$$F_S(S_t) = \begin{bmatrix} H_X H_X & H_X H_Z + H_Z N & H_X H_v + H_Z \Omega \varepsilon_{t+1} + H_v \\ H_X & H_Z & H_v \\ 1 & 0 & 0 \\ G_X H_X & G_X H_Z + G_Z N & G_X H_v + G_Z \Omega \varepsilon_{t+1} + G_v \\ G_X & G_Z & G_v \\ 0 & N & \Omega \varepsilon_{t+1} \\ 0 & 1 & 0 \end{bmatrix} \tag{27}$$

The Magnus and Neudecker (1999) Hessian is:

$$
F_{SS}(S_t) = \begin{bmatrix}
H_{XX}H_XH_X + H_XH_{XX} & H_{XX}H_XH_Z + H_XH_{ZX} + H_{ZX}H_XN & H_{XX}H_xH_v + H_XH_{vX} + HZX\Omega\varepsilon_{t+1} + H_{vX} \\
H_{XX} & H_{ZX} & H_{vX} \\
0 & 0 & 0 \\
G_{XX}H_XH_X + G_XH_{XX} & G_{XX}H_XH_Z + G_XH_{ZX} + G_{ZX}H_XN & G_{XX}H_xH_v + G_XH_{vX} + GZX\Omega\varepsilon_{t+1} + G_{vX} \\
G_{XX} & G_{ZX} & G_{vX} \\
0 & 0 & 0 \\
0 & 0 & 0 \\
H_{XX}H_XH_Z + H_XH_{ZX} + H_{ZX}H_XN & H_{XZ}NH_Z + H_XH_{ZZ} + H_{ZZ}NN & H_{XZ}NH_v + H_XH_{vZ} + H_{ZZ}N\Omega\varepsilon_{t+1} \\
H_{XZ} & H_{ZZ} & H_{vZ} \\
0 & 0 & 0 \\
G_{XX}H_XH_Z + G_XH_{ZX} + G_{ZX}H_XN & G_{XZ}NH_Z + G_XH_{ZZ} + G_{ZZ}NN & G_{XZ}NH_v + G_XH_{vZ} + G_{ZZ}N\Omega\varepsilon_{t+1} \\
G_{XZ} & G_{ZZ} & G_{vZ} \\
0 & 0 & 0 \\
0 & 0 & 0 \\
H_{XX}H_xH_v + H_XH_{vX} + HZX\Omega\varepsilon_{t+1} + H_{vX} & H_{XZ}NH_v + H_XH_{vZ} + H_{ZZ}N\Omega\varepsilon_{t+1} & H_{Xv}H_v + H_XH_{vv} + H_{Zv}\Omega\varepsilon_{t+1} + H_{vv} \\
H_{Xv} & H_{Zv} & H_{vv} \\
0 & 0 & 0 \\
G_{XX}H_xH_v + G_XH_{vX} + GZX\Omega\varepsilon_{t+1} + G_{vX} & G_{XZ}NH_v + G_XH_{vZ} + G_{ZZ}N\Omega\varepsilon_{t+1} & G_{Xv}H_v + G_XH_{vv} + G_{Zv}\Omega\varepsilon_{t+1} + G_{vv} \\
G_{Xv} & G_{Zv} & G_{vv} \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
$$

We should recognize that the cross derivatives of the H and G functions for both X_{t-1} and Z_t with v are zero. That is $H_{Xv} = H_{Zv} = G_{Xv} = G_{Zv} = 0$. Also we note that for any function $f(x, y)$, $f_{xy} = f_{yx}$. Given these the matrix simplifies to:

$$
F_{SS}(S_t) = \begin{bmatrix}
H_{XX}H_XH_X + H_XH_{XX} & H_{XX}H_XH_Z + H_XH_{XZ} + H_{XZ}H_XN & H_{XX}H_xH_v + HXZ\Omega\varepsilon_{t+1} \\
H_{XX} & H_{XZ} & 0 \\
0 & 0 & 0 \\
G_{XX}H_XH_X + G_XH_{XX} & G_{XX}H_XH_Z + G_XH_{XZ} + G_{XZ}H_XN & G_{XX}H_xH_v + GXZ\Omega\varepsilon_{t+1} \\
G_{XX} & G_{XZ} & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
H_{XX}H_XH_Z + H_XH_{XZ} + H_{XZ}H_XN & H_{XZ}NH_Z + H_XH_{ZZ} + H_{ZZ}NN & H_{XZ}NH_v + H_{ZZ}N\Omega\varepsilon_{t+1} \\
H_{XZ} & H_{ZZ} & 0 \\
0 & 0 & 0 \\
G_{XX}H_XH_Z + G_XH_{XZ} + G_{XZ}H_XN & G_{XZ}NH_Z + G_XH_{ZZ} + G_{ZZ}NN & G_{XZ}NH_v + G_{ZZ}N\Omega\varepsilon_{t+1} \\
G_{XZ} & G_{ZZ} & 0 \\
0 & 0 & 0 \\
0 & 0 & 0 \\
H_{XX}H_xH_v + HXZ\Omega\varepsilon_{t+1} & H_{XZ}NH_v + H_{ZZ}N\Omega\varepsilon_{t+1} & H_XH_{vv} + H_{vv} \\
0 & 0 & H_{vv} \\
0 & 0 & 0 \\
G_{XX}H_xH_v + GXZ\Omega\varepsilon_{t+1} & G_{XZ}NH_v + G_{ZZ}N\Omega\varepsilon_{t+1} & G_XH_{vv} + G_{vv} \\
0 & 0 & G_{vv} \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \tag{28}
$$

Using (26) in (15) we get $\Delta(S_t) \equiv \Gamma(F(S_t)) = 0$. Magnus and Neudecker (1999) show that the chain-rule for this function with our setup for the organization of the Jacobian and Hessian matrices is as follows.

$$
\Delta_{SS} = (F_S \otimes I_{n_X+n_Y})^T \Gamma_{AA} F_S + (I_{n_S} \otimes \Gamma_A) F_{SS} \tag{29}
$$

4.2 Solving for Linear Terms

We can solve for the linear terms as we did in the linearization chapter. This will generate the coefficient matrices H_X, H_Z, G_X and G_Z. We do not get values for H_v and G_v, however.

Substituting the linear portions of (17), (18) and (16) into the linear portion of (19) gives:

$$\Gamma(A_t) \doteq \Gamma(\bar{X}, ..., \bar{Z}) + \begin{bmatrix} \Gamma_1 & \dots & \Gamma_7 \end{bmatrix} \begin{bmatrix} H_X H_S S_t + H_Z N + v\Omega\varepsilon_{t+1} + H_v\tilde{v} \\ H_S S_t \\ X_{t-1} \\ G_X H_S S_t + G_Z N + v\Omega\varepsilon_{t+1} + G_v\tilde{v} \\ G_S S_t \\ NZ_{t-1} + v\Omega\varepsilon_t \\ Z_t \end{bmatrix}$$

$$H_S \equiv \begin{bmatrix} H_X & H_Z & H_v \end{bmatrix}, \; G_S \equiv \begin{bmatrix} G_X & G_Z & G_v \end{bmatrix}$$

Noting that $\Gamma(\bar{X}, ..., \bar{Z}) = 0$, and multiplying out the $H_S S_t$ and $G_S S_t$ terms gives:

$$0 = \begin{bmatrix} \Gamma_1 & \dots & \Gamma_7 \end{bmatrix} \begin{bmatrix} H_X\tilde{X}_{t-1} + H_X H_Z\tilde{Z}_t + H_X H_v\tilde{v} + H_Z N + v\Omega\varepsilon_{t+1} + H_v\tilde{v} \\ H_X\tilde{X}_{t-1} + H_Z\tilde{Z}_t + H_v\tilde{v} \\ X_{t-1} \\ G_X H_X\tilde{X}_{t-1} + G_X H_Z\tilde{Z}_t + G_X H_v\tilde{v} + G_Z N + v\Omega\varepsilon_{t+1} + G_v\tilde{v} \\ G_X\tilde{X}_{t-1} + G_Z\tilde{Z}_t + G_v\tilde{v} \\ NZ_{t-1} + v\Omega\varepsilon_t \\ Z_t \end{bmatrix} \tag{30}$$

Taking the expectation of (30) and collecting the terms with $\tilde{v}$:

$$\left[\Gamma_1(H_X H_v + H_v) + \Gamma_2 H_v + \Gamma_4(G_X H_v + G_v) + \Gamma_4 G_v\right]\tilde{v}$$

As before the above term must be zero for all possible values of $\tilde{v}$. We note that all of the Γ_i matrices can be decomposed into equations that define Y in the first n_Y rows and

equations that define X in the final n_X rows. This means we can rewrite the condition as follows:

$$\begin{bmatrix}\Gamma_1^Y\\ \Gamma_1^X\end{bmatrix}(H_X H_v + H_v) + \begin{bmatrix}\Gamma_2^Y\\ \Gamma_2^X\end{bmatrix} H_v + \begin{bmatrix}\Gamma_4^Y\\ \Gamma_4^X\end{bmatrix}(G_X H_v + G_v) + \begin{bmatrix}\Gamma_5^Y\\ \Gamma_5^X\end{bmatrix} G_v = 0$$

Since the Y equations do not include X_{t+1} or Y_{t+1} terms, we know that $\Gamma_1^Y = \Gamma_4^Y = 0$. This gives:

$$\begin{aligned}\Gamma_2^Y H_v + \Gamma_5^Y G_v &= 0\\ \Gamma_1^X(H_X + I)H_v + \Gamma_2^X H_v + \Gamma_4^X(G_X H_v + G_v) + \Gamma_5^Y G_v &= 0\end{aligned}$$

Solving the first for G_v and substituting this into the second:

$$\Gamma_1^X(H_X + I)H_v + \Gamma_2^X H_v + \Gamma_4^X\left[G_X H_v + \left(-\frac{\Gamma_2^Y}{\Gamma_5^Y}H_v\right)\right] + \Gamma_5^Y\left(-\frac{\Gamma_2^Y}{\Gamma_5^Y}H_v\right) = 0$$

Solving this gives $H_v = G_v = 0$.

4.3 Solving for Quadratic Terms

With the first-order coeffiecients for the H and G functions known, we can use the expectation of (29) to solve for the second-order coefficients. We note that F_S is a function of the first-order coeffiecients as shown in (27). Similarly, we know that F_{SS} a function of both the first and second-order coefficients as shown in (28).

Before taking expectations, we need to multiply out the term $\Lambda \equiv (F_S \otimes I_{n_X+n_Y})^T \Gamma_{AA} F_S$.

Examine the F_S matrix and note that terms with ε_{t+1} appear only in the thrid column.

$$F_S = \begin{bmatrix} H_X H_X & H_X H_Z + H_Z N & H_X H_v + H_Z \Omega \varepsilon_{t+1} + H_v \\ H_X & H_Z & H_v \\ 1 & 0 & 0 \\ G_X H_X & G_X H_Z + G_Z N & G_X H_v + G_Z \Omega \varepsilon_{t+1} + G_v \\ G_X & G_Z & G_v \\ 0 & N & \Omega \varepsilon_{t+1} \\ 0 & 1 & 0 \end{bmatrix}$$

If we take the expectation of F_S the ε_{t+1} terms disappear.

$$E\{F_S\} = \begin{bmatrix} H_X H_X & H_X H_Z + H_Z N & H_X H_v + H_v \\ H_X & H_Z & H_v \\ 1 & 0 & 0 \\ G_X H_X & G_X H_Z + G_Z N & G_X H_v + G_v \\ G_X & G_Z & G_v \\ 0 & N & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Let's look at the $(3,3)$ block in Λ. Recall that H_v and G_v are zeros.

$$\Lambda(3,3) = \varepsilon_{t+1}^T (\Omega H_Z^T \Gamma_{11} H_Z \Omega + \Omega G_Z^T \Gamma_{44} G_Z \Omega) \varepsilon_{t+1}$$

This is the only term where the quadratic form of ε_{t+1} appears. In every other term it is either absent or appears as a linear term. Hence when expectations are taken the terms with ε_{t+1} disappear. This means we can use $E\{\Lambda\} = (E\{F_S\} \otimes I_{n_X + n_Y})^T \Gamma_{AA} E\{F_S\}$ and then replace the $(3,3)$ term, which will be zero.

To take expectations of $\Lambda(3,3)$ it is useful to know that if the elements of a column vector of random variables $\varepsilon \sim iid(0, I)$, then $E\{\varepsilon^T A \varepsilon\} = \text{tr}(A)$. To see this note that the matrix evaluates to $\sum_i \sum_j \varepsilon_i a_{ij} \varepsilon_j$. The expected value of each term is $a_{ij} = a_{ii}$ if $i = j$ and is zero otherwise. Hence the sum inclues only the diagonal elements of A, which is the trace of A.

So we replace the zero in the $(3,3)$ block with $\text{tr}(\Omega[H_Z^T\Gamma_{11}H_Z] + [G_Z^T\Gamma_{44}G_Z]\Omega)$.

Unfortunately, $I_{n_A} \otimes \Gamma_A$ is not a square matrix and therefore not invertable. However, we can solve for the second-order coefficients numerically. The coefficients we need to find are $\Theta = \{H_{XX}, H_{XZ}, H_{ZZ}, H_{vv}, G_{XX}, G_{XZ}, G_{ZZ}, G_{vv}\}$. $E\{F_S\}$, Γ_A and Γ_{AA} are known. We can therefore write a Δ_{SS} function as shown below and numerically solve for the values of Θ that set it equal to zero. We note that Δ_{SS} will return a matrix of size $n_S(n_X + n_Y) \times n_S$. This will be $n_X + n_Y$ blocks of symmetric $n_S \times n_S$ matrices.

$$\Delta_{SS}(\Theta) = (E\{\Lambda\} + (I_{n_S} \otimes \Gamma_A)E\{F_{SS}(\Theta)\} = 0$$
$$\Lambda = (F_S \otimes I_{n_X+n_Y})^T\Gamma_{AA}F_S$$

The symmetric blocks in the Δ_{SS} matrix will be denoted Δ_{SS}^i for $i \in \{1, 2, \ldots, n_X + n_Y\}$ and can be decomposed into nine parts.

$$\Delta_{SS}^i = \begin{bmatrix} \Delta_{XX}^i & \Delta_{XZ}^i & 0 \\ (\Delta_{XZ}^i)^T & \Delta_{ZZ}^i & 0 \\ 0 & 0 & \Delta_{vv}^i \end{bmatrix} \tag{31}$$

Hence we have $n_X^2 + n_Z^2 + n_X n_Z + 1$ unique values for each i, for a total of $(n_X^2 + n_Z^2 + n_X n_Z + 1)(n_X + n_Y)$. We have $(n_X^2 + n_X nZ + n_Z + 1)n_X$ terms in the H_{SS} coefficients and $(n_X^2 + n_X nZ + n_Z + 1)n_Y$ terms in the G_{SS} coefficients. Hence the $\Delta_{SS} = 0$ condition will exactly identify Θ.

5 Numerical Derivatives

Perturbation methods require taking a large number of derivatives. Oftimes these derivatives become prohibiitively complicated as the order of differentiation rises. The chances of making an error in calculating these derivatives is nontrivial. One way to aleviate this problem is to use a symbolic processor such as MAPLE, Mathematica, or SymPy to take these deriatives. Another method is to take derivatives numerically.

When taking numerical derivatives, the most commomn and intuitive method is to use is

"finite differences". We get the best asymoptotic convergence if we use "central" derivates; i.e. those where the function calls are centered around the evalutaion point, rather than forward or backward differencing. This can require more function calls, however.

For example, consider the simplest form of the derivative of the function $f(x)$ shown below in equation 32.

$$f_x(x) = \frac{f(x_0 + \varepsilon) - f(x_0)}{\varepsilon} \tag{32}$$

This is a forward differnce, whereas the central difference would be the formula given in 33.

$$f_x(x) = \frac{f(x_0 + \varepsilon) - f(x_0 - \varepsilon)}{2\varepsilon} \tag{33}$$

The number of function calls is the same in both these cases. However, we often know the value of $f(x_0)$ before we differentiate. This means that the forward differences requires only one *addtional* function call, while the central difference method requires two.

If we already have $f(x_0)$, then taking the two addtional function calls allows us to calculate the second derivative without any additional calls using the central differnce formula in 34

$$f_x(x) = \frac{\frac{f(x_0+\varepsilon)-f(x_0)}{\varepsilon} - \frac{f(x_0)-f(x_0-\varepsilon)}{\varepsilon}}{\varepsilon} = \frac{f(x_0 + \varepsilon) - 2f(x_0) + f(x_0 - \varepsilon)}{\varepsilon^2} \tag{34}$$

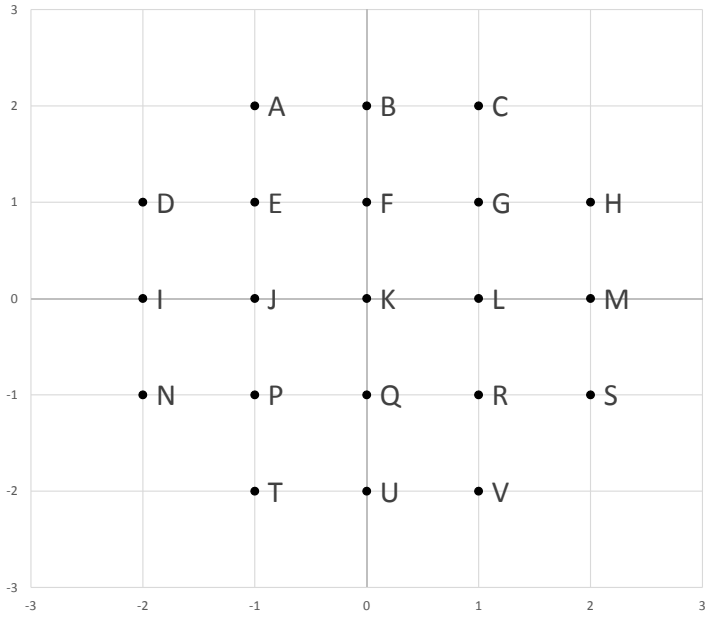
5.1 Bivarate Numerical Derivatives

Finite differece calculations for univariate functions are rather straightforward as long as one keeps centering in mind. Multivariate functions are bit harder because there are many variables, but the basic concepts are the same.

Consider the grid in Figure 1 which show the two inputs into a bivariate function, $f(x, y)$. The grid lines and points are laid out as multiples of ε, with the point K in the center corresponding to the evaluation point, (x_0, y_0).

The constant term is $f(\mathsf{K})$.

Figure 1: Points for Bivariate Numerical Differentiation



To find the first derivatives we evaluate our function at the points indicated.

$$f_x(x,y) = \frac{f(\mathsf{L}) - \mathsf{f}(\mathsf{J})}{2\varepsilon}$$
$$f_y(x,y) = \frac{f(\mathsf{F}) - \mathsf{f}(\mathsf{Q})}{2\varepsilon}$$

Second deriviatives are:

$$f_{xx}(x,y) = \frac{\frac{f(\mathsf{L})-\mathsf{f}(\mathsf{K})}{\varepsilon} - \frac{f(\mathsf{K})-\mathsf{f}(\mathsf{J})}{\varepsilon}}{\varepsilon} = \frac{f(\mathsf{L}) - 2\mathsf{f}(\mathsf{K}) + \mathsf{f}(\mathsf{J})}{\varepsilon^2}$$
$$f_{yy}(x,y) = \frac{\frac{f(\mathsf{F})-\mathsf{f}(\mathsf{K})}{\varepsilon} - \frac{f(\mathsf{K})-\mathsf{f}(\mathsf{Q})}{\varepsilon}}{\varepsilon} = \frac{f(\mathsf{F}) - 2\mathsf{f}(\mathsf{K}) + \mathsf{f}(\mathsf{Q})}{\varepsilon^2}$$
$$f_{yy}(x,y) = \frac{\frac{f(\mathsf{G})-\mathsf{f}(\mathsf{E})}{2\varepsilon} - \frac{f(\mathsf{R})-\mathsf{f}(\mathsf{P})}{2\varepsilon}}{2\varepsilon} = \frac{f(\mathsf{G}) - \mathsf{f}(\mathsf{E}) - \mathsf{f}(\mathsf{R}) + \mathsf{f}(\mathsf{P})}{4\varepsilon^2}$$

Third deriviatives are:

$$f_{xxx}(x,y) = \frac{\frac{f(\mathsf{M})-2\mathsf{f}(\mathsf{L})+\mathsf{f}(\mathsf{K})}{\varepsilon^2} - \frac{f(\mathsf{K})-2\mathsf{f}(\mathsf{J})+\mathsf{f}(\mathsf{I})}{\varepsilon^2}}{2\varepsilon} = \frac{f(\mathsf{M}) - 2\mathsf{f}(\mathsf{L}) + 2\mathsf{f}(\mathsf{J}) - \mathsf{f}(\mathsf{I})}{2\varepsilon^3}$$

$$f_{xxy}(x,y) = \frac{\frac{f(\mathsf{H})-\mathsf{f}(\mathsf{F})-\mathsf{f}(\mathsf{S})+\mathsf{f}(\mathsf{Q})}{4\varepsilon^2} - \frac{f(\mathsf{F})-\mathsf{f}(\mathsf{D})-\mathsf{f}(\mathsf{Q})+\mathsf{f}(\mathsf{N})}{4\varepsilon^2}}{2\varepsilon}$$

$$= \frac{f(\mathsf{H}) - 2\mathsf{f}(\mathsf{F}) - \mathsf{f}(\mathsf{S}) + \mathsf{f}(\mathsf{D}) + 2\mathsf{f}(\mathsf{Q}) - \mathsf{f}(\mathsf{N})}{8\varepsilon}$$

$$f_{xyy}(x,y) = \frac{\frac{f(\mathsf{A})-\mathsf{f}(\mathsf{J})-\mathsf{f}(\mathsf{C})+\mathsf{f}(\mathsf{L})}{4\varepsilon^2} - \frac{f(\mathsf{J})-\mathsf{f}(\mathsf{T})-\mathsf{f}(\mathsf{L})+\mathsf{f}(\mathsf{V})}{4\varepsilon^2}}{2\varepsilon}$$

$$= \frac{f(\mathsf{A}) - 2\mathsf{f}(\mathsf{J}) - \mathsf{f}(\mathsf{C}) + \mathsf{f}(\mathsf{T}) + 2\mathsf{f}(\mathsf{L}) - \mathsf{f}(\mathsf{V})}{8\varepsilon}$$

$$f_{yyy}(x,y) = \frac{\frac{f(\mathsf{B})-2\mathsf{f}(\mathsf{F})+\mathsf{f}(\mathsf{K})}{\varepsilon^2} - \frac{f(\mathsf{K})-2\mathsf{f}(\mathsf{Q})+\mathsf{f}(\mathsf{U})}{\varepsilon^2}}{2\varepsilon} = \frac{f(\mathsf{B}) - 2\mathsf{f}(\mathsf{F}) + 2\mathsf{f}(\mathsf{Q}) - \mathsf{f}(\mathsf{U})}{2\varepsilon^3}$$

Exercises

Exercise 1. Following the example in equations (5) and (6), find the formula for the cubic term, $x_{uuu}(u_0)$, as a function of the derivatives of the F function and the lower-order derrivatives of the x function, i.e. $x(u_0)$, $x_u(u_0)$ and $x_{uu}(u_0)$.

Exercise 2. Consider the following static general equlibrim model. Firms have a demand for labor curve given by $n^d = \left[\frac{(1-\alpha)z}{w}\right]^{\frac{1}{\alpha}} k$, where z is the level of technology, w is the wage rate, k is a fixed capital stock, and α is a capital share parameter from a Cobb-Douglas production function. Given this, the profits earned by the firm are $\pi = zk^{\alpha}(n^d)^{1-\alpha} - wn^d$. The supply of labor by households is $n^s = h - \frac{b}{w(1+b)}(wh + \pi - t)$, where h is the time endowment of the household, is t is a lump-sum tax, and b is a weight in utility on leisure versus consumption of goods. Assuming a unit measure of both households and firms, use the following parameter values $\alpha = .33$, $k = 5$, $z = 1$, $b = 2$, $t = .1$ and $h = 24$.

Find the market-clearing wage rate using `fsolve`. Find a first-order approxmation for wage as a function of k. Approximate about $k = 5$.

Find a second-order approximation also about $k = 5$.

Set up a grid on the space between $k = 1$ and $k = 15$. Use `fsolve` to find the equilibrium value of the wage at each point on the grid.

Plot the grid solution, the linear and quadratic approximations on the same graph.

Repeat the above exercise when the approximation point is $k = 10$.

Exercise 3. For the function $F(y, x) = (x^{.35} + .9x - y)^{-2.5} - .95(y^{.35} + .9y)^{-2.5} = 0$, simplify and then use perturbation methods to find the cubic approximation of $y = G(x)$ about the point $x_0 = 100$. In this case, $y_0 = G(x_0) = 10.5156$.

Write out the functional form of this polynomial with numbers for coefficients.

Set up a grid on the space between $x = 99$ and $x = 101$. Use `fsolve` to find the equilibrium value of y at each point on the grid.

Plot the difference between the linear, quadratic and cubic approximations and the grid solution on the same graph over the range $x \in (99, 101)$.

Exercise 4. For the Brock and Mirman (1972) model in section 3 with the default parameter values find the scalar values of H_X and H_{XX}. Use analytical derivatives.

Plot the the linear and quadratic approximations to policy function $K' = H(K)$. Compare this with the closed form solution from the notes.

Exercise 5. For the Brock and Mirman (1972) model with the default parameter values find the scalar values of H_X, H_Z, H_{XX}, H_{XZ}, H_{ZZ} and H_{vv}. Use numerical derivatives.

Plot the three-dimensional surface plot for the policy function $K' = H(K, z)$. Compare this with the closed form solution from the notes and the two approximations from the DSGE chapter Exercise 8 and the Linearization chapter Exercise 1.

Exercise 6. Repeat Exercise 5 above using the baseline model from section 6 in the linearization chapter. Generate figures similar to those you generated in the Linearization chapter in Exercise 7. Compare this plot with the one from that exercise.

References

Binning, Andrew, "Third-order approximation of dynamic models without the use of tensors," Technical Report, Norges Bank 2012.

Brock, William A. and Leonard J. Mirman, "Optimal economic growth and uncertainty: The discounted case," *Journal of Economic Theory*, June 1972, *4* (3), 479–513.

Collard, Fabrice and Michel Juillard, "A Higher-Order Taylor Expansion Approach to Simulation of Stochastic Forward-Looking Models with an Application to a Nonlinear Phillips Curve Model," *Computational Economics*, June 2001, *17* (2-3), 125–39.

Gomme, Paul and Paul Klein, "Second-order approximation of dynamic models without the use of tensors," *Journal of Economic Dynamics and Control*, 2011, *35* (4), 604–615.

Heer, Burkhard and Alfred Maussner, *Dynamic General Equilibrium Modeling: Computational Methods and Applications*, Springer,

_ and _ , *Dynamic General Equilibrium Modeling: Computational Methods and Applications*, second ed., Springer, 2009.

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

_ , *Numerical Methods in Economics*, MIT Press, 1998.

Magnus, Jan R. and Heinz Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics, 2nd ed.*, Wiley, 1999.

Schmitt-Grohe, Stephanie and Martin Uribe, "Solving Dynamic General Equilibrium Modes Using a Second-Order Approximation to the Policy Function," *Journal of Economic Dynamics and Control*, September 2004, *28* (10724), 755–775.

Uhlig, Harald, "A toolkit for analyzing nonlinear dynamic stochastic models easily," in Ramon Marimon, ed., *Computational Methods for the Study of Dynamic Economies*, Oxford University Press, 1999, pp. 30–61.