Notes on Linear Approximations

Michael Reiter, IHS, Vienna

2 Linear rational expectation models

Consider the scalar linear model

$$\mathrm{E}_t\, x_{t+1} = \alpha(x_t - u_t) \tag{1}$$

and assume that u_t is stochastic and stationary (for example, in the sense of constant mean and variance) and u_t is known at t.

2.1 Case 1: x_0 is given

Typical example: x is the capital stock of an economy.

Write (1) as

$$x_{t+1} = \alpha(x_t - u_t) + \epsilon_{t+1} \tag{2}$$

where ϵ_{t+1} is defined as the forecast error of x_{t+1}, namely $\epsilon_{t+1} \equiv x_{t+1} - \mathrm{E}_t\, x_{t+1}$. We assume in this case that ϵ_t is an exogenously given shock, for example a depreciation shock.

Obviously, the solution to (2) is unique, *conditional on the realization* of u_t and ϵ_t. If both u_t and ϵ_t are stationary, x_t is stationary if and only if $|\alpha| < 1$.

2.2 Case 2: x is a forward looking variable

Typical example: x is the price of an asset that has to be determined from an expected dividend stream.

Solve (1) for x_t and iterate forward:

$$\begin{aligned} x_t &= \frac{1}{\alpha} \operatorname{E}_t x_{t+1} + u_t \\ &= \frac{1}{\alpha} \operatorname{E}_t \left[\frac{1}{\alpha} \operatorname{E}_{t+1} x_{t+2} + u_{t+1} \right] + u_t \\ &= \ldots \\ &= \operatorname{E}_t \left[\alpha^{-n} x_{t+n} + \sum_{j=0}^{n-1} \alpha^{-j} u_{t+j} \right] \end{aligned} \tag{3}$$

- **Case 2a:** $|\alpha| > 1$

  If x_t is stationary, then $\alpha^{-n} x_{t+n} \to 0$ in some sense (for example, if x_t is stationary in the mean square sense, then $\lim_{n\to\infty} \alpha^{-n} x_{t+n} = 0$ in mean square). Letting $n \to \infty$ in (3), we obtain in this case that

  $$x_t = \operatorname{E}_t \sum_{i=0}^{\infty} \alpha^{-i} u_{t+i} \tag{4}$$

  This implies that (4) is the **unique stable solution** of (1).

  There are still many unstable solutions to (1). Denote the solution of (4) by x_t^*, and take **any** variable y that satisfies $y_t = \frac{1}{\alpha} \operatorname{E}_t y_{t+1}$. Then $z_t \equiv x_t^* + y_t$ satisfies (1)!

- **Case 2b:** $|\alpha| < 1$ From the logic of case 1, we see that we can choose any starting value x_0 and obtain a stable solution from (2). So the solution is indeterminate: there is a continuum of stable solutions, and we have to pick one of these solutions.

  In this case, the solution can even depend on exogenous variables that, from a point of view of the fundamentals, have nothing to do with the model (socalled "sunspots").

2.3 Conclusion:

In order to have a unique stable solution, we must have either

- x_0 given and $|\alpha| < 1$, or
- x_0 free and $|\alpha| > 1$

Generalization: in a system of linear expectations equations

$$\operatorname{E} x_{t+1} = A x_t + u_t \tag{5}$$

the matrix A should have as many eigenvalues smaller than 1 in absolute value as there are components of x_0 that are predetermined.

3 Linear vs. Nonlinear Solutions

3.1 Linear-quadratic optimization problems

Consider the problem

$$\min \frac{1}{2}\left[x_0^2 + \mathrm{E}_0(x_1 - a)^2\right] \tag{6}$$

subject to

$$x_1 = x_0 + \epsilon_1 \tag{7}$$

where

$$\mathrm{E}_0\,\epsilon_1 = 0 \tag{8}$$

Plugging (7) into (6) we get

$$\min \frac{1}{2}\left[x_0^2 + \mathrm{E}_0[(x_0 + \epsilon_1 - a)^2]\right] \tag{9}$$

The first order condition for (6) is

$$x_0 + \mathrm{E}_0[(x_0 + \epsilon_1 - a)] = 0 \tag{10}$$

Because of (8), (10) simplifies to

$$x_0 + (x_0 - a) = 0 \tag{11}$$

which gives the solution

$$x_0 = \frac{2}{a} \tag{12}$$

Using (12) in (9), we see that the optimum (minimum) value equals

$$\frac{1}{2}\left[(a/2)^2 + \mathrm{E}_0(a/2 + \epsilon_1 - a)^2\right] \quad = \quad \frac{1}{2}\left[a^2/4 + a^2/4 - \mathrm{E}_0(a\epsilon_1) + \mathrm{E}_0\,\epsilon_1^2\right] \quad = \quad \frac{a^2}{4} \;+\; \frac{1}{2}\sigma_\epsilon^2 \tag{13}$$

where σ_ϵ^2 is the variance of ϵ.

Conclusions from this example:

1. Uncertainty is harmful: the optimal value that one can achieve is worse (higher in a minimization problem) with higher uncertainty σ_ϵ^2.
2. The optimal policy, given in (13), is linear in the state variable x_0.
3. The optimal policy is not affected by σ_ϵ^2: this is called "certainty equivalence". The optimal x_0 is the same as if we knew x_1 for certain.

These conclusions hold up in a much more complicated models, as long as

- the objective function is quadratic in all variables
- the law of motion is linear in all variables.

3.2 Precautionary behavior

$$\max_{c_0} U(c_0) + \mathrm{E}_0\, U(c_1) \tag{14}$$

where

$$c_1 = M - c_0 + \epsilon_1 \tag{15}$$

The first order condition (Euler equation) is

$$U'(c_0) = \mathrm{E}_0\, U'(c_1) \tag{16}$$

Using $U(c) = log(c)$ and assuming

$$\epsilon_1 = \begin{cases} -\sigma & \text{with prob. } 0.5 \\ \sigma & \text{with prob. } 0.5 \end{cases} \tag{17}$$

this becomes

$$\begin{aligned} \frac{1}{c_0} &= \mathrm{E}_0 \frac{1}{M - c_0 + \epsilon_1} = \frac{1}{2}\left[\frac{1}{M - c_0 - \sigma} + \frac{1}{M - c_0 + \sigma}\right] \\ &= \frac{M - c_0}{(M - c_0)^2 - \sigma^2} \end{aligned} \tag{18}$$

Increasing σ makes the rhs of (18) increase; to match this, the lhs has to increase as well, this means, c_0 has to decrease. Higher uncertainty about the future lowers current consumption: this is called "precautionary saving". It is a consequence of the fact that the third derivative of the utility funciton (here the logarithm) is positive.

The linear-quadratic problem of Section 2.1 did not show any precautionary behavior, because the objective function is quadratic (third derivative is zero).

3.3 The effect of linearization

Let us now replace the first order condition (16) by its linearization. The "deterministic steady state" is given by $c_0 = c_1 = c^* = \frac{M}{2}$. Linearization around c^* gives

$$U'(c^*) + U''(c^*)(c_0 - c^*) = \mathrm{E}_0 \left[U'(c^*) + U''(c^*)(c_1 - c^*)\right] \tag{19}$$

which simplies to

$$U''(c^*)(c_0 - c^*) = U''(c^*)\, \mathrm{E}_0(c_1 - c^*) \tag{20}$$

The third derivative of the utility function, and therefore the precautionary saving effect, is thrown out by the linearization!

3.4 Log-linearization

If we do "log-linearization", rather than linearization in c, we get

$$U''(c^*) \left.\frac{\partial c}{\partial \log(c)}\right|_{c=c^*} (\log(c_0) - \log(c^*)) = U''(c^*) \left.\frac{\partial c}{\partial \log(c)}\right|_{c=c^*} \mathrm{E}_0(\log(c_1) - \log(c^*)) \tag{21}$$

This simplies to

$$U''(c^*)c^*(\log(c_0) - \log(c^*)) = U''(c^*)c^* \,\mathrm{E}_0(\log(c_1) - \log(c^*)) \tag{22}$$

(22) defines an equation for $\log(c_t)$, which is equal to the equation (21) for c_t, except for being multiplied by c^*. This means that the fluctuations in c_t implied by (21) are exactly the fluctuations in $log(c_t)$ implied by (22), multiplied by c^* (and the other way round, dividing by c^*). Therefore, we don't have worry about linearization vs. log-linearization at the time of solving the model. We can always go forth and back between the two by multipling or dividing the solution by the steady state values.

4 The Mechanics of Linearization

The whole economic model can be brought (cf. Section 3.3) into the form

$$F(x_t, x_{t-1}, \epsilon_t, \eta_t) = 0 \tag{23}$$

where x_t is the vector of all current variables, x_{t-1} is the vector of all lagged variables, ϵ_t is the vector of exogenous i.i.d. shocks, η_t is the vector of (endogenous) expectation errors.

4.1 The Deterministic Steady State

The steady state vector x^* satisfies (23) when all shocks and expectation errors are zero:

$$F(x^*, x^*, 0, 0) = 0 \tag{24}$$

4.2 Linearization around steady state

A first-order Taylor approximation about the steady state gives

$$\begin{aligned} F(x^*, x^*, 0, 0) + F_1(x^*, x^*, 0, 0)(x_t - x^*) + F_2(x^*, x^*, 0, 0)(x_{t-1} - x^*) \\ + F_3(x^*, x^*, 0, 0)\epsilon_t + F_4(x^*, x^*, 0, 0)\eta_t = 0 \end{aligned} \tag{25}$$

where F_i denotes the derivative of F w.r.t. to the i-th argument, and F is a vector-valued function. Because of (24), the first term in (25) drops out. We define $\tilde{x}_t$ as the deviation from steady state, $x_t - x^*$. Then we can rewrite (25) as

$$F_1(x^*, x^*, 0, 0)\tilde{x}_t + F_2(x^*, x^*, 0, 0)\tilde{x}_{t-1} + F_3(x^*, x^*, 0, 0)\epsilon_t + F_4(x^*, x^*, 0, 0)\eta_t = 0 \tag{26}$$

A solution consists of a linear system

$$\tilde{x}_t = A\tilde{x}_{t-1} + B\epsilon_t \tag{27}$$

that satisfies (26), for suitably chosen η_t.

4.3 How to Bring a Model Into the Canonical Form (23)

Equations with expectations

The expectation errors η_t are just a notational trick to get rid of the expectation operator. For example, we can write the household Euler equation

$$U_c(c_t, L_t) = \beta\,\mathrm{E}_t[(1 + r_{t+1})U_c(c_{t+1}, L_{t+1})] \tag{28}$$

as

$$U_c(c_t, L_t) = \beta[(1 + r_{t+1})U_c(c_{t+1}, L_{t+1}) - \eta_{t+1}] \tag{29}$$

if we define

$$\eta_{t+1} = [(1 + r_{t+1})U_c(c_{t+1}, L_{t+1})] - \mathrm{E}_t[(1 + r_{t+1})U_c(c_{t+1}, L_{t+1})] \tag{30}$$

Since (29) has to hold for all t, we can lower the time index, and write

$$U_c(c_{t-1}, L_{t-1}) - \beta[(1 + r_t)U_c(c_t, L_t) - \eta_t] = 0 \tag{31}$$

which is of the form (23). This means, it is one of the equations in (23).

Models with longer lags or leads

If we have an equation

$$y_t = f(y_{t-1}, y_{t-2}) \tag{32}$$

we can define the auxiliary variable

$$z_t = y_{t-1} \tag{33}$$

Then (37) becomes

$$y_t = f(y_{t-1}, z_{t-1}) \tag{34}$$

(33) and (34) are both of the form (23).

If we have an equation

$$y_t = \mathrm{E}_t\, f(y_{t+1}, y_{t+2}) \tag{35}$$

we cannot define $z_t = y_{t+1}$. This would be a severe mistake, because y_{t+1} becomes known only in period $t + 1$, and z_t has a time index t. But we can define

$$z_t = \mathrm{E}_t\, y_{t+1} \tag{36}$$

Then we do a linearization and use (36):

$$
\begin{aligned}
y_t &= \mathrm{E}_t\, f(y_{t+1}, y_{t+2}) \\
&\approx \mathrm{E}_t \left[f(y^*, y^*) + f_1(y^*, y^*)(y_{t+1} - y^*) + f_2(y^*, y^*)(y_{t+2} - y^*) \right] \\
&= f(y^*, y^*) + f_1(y^*, y^*)(\mathrm{E}_t\, y_{t+1} - y^*) + f_2(y^*, y^*)(\mathrm{E}_t\, y_{t+2} - y^*) \\
&= f(y^*, y^*) + f_1(y^*, y^*)(\mathrm{E}_t\, y_{t+1} - y^*) + f_2(y^*, y^*)(\mathrm{E}_t\, \mathrm{E}_{t+1}\, y_{t+2} - y^*) \\
&= f(y^*, y^*) + f_1(y^*, y^*)(\mathrm{E}_t\, y_{t+1} - y^*) + f_2(y^*, y^*)(\mathrm{E}_t\, z_{t+1} - y^*) \qquad (37)
\end{aligned}
$$

which is again of the form (23). The fourth line in (37) uses the law of iterated expectations.

Notice that we have not brought the original, nonlinear model into the form of Equation (23), but only the linearized version of it.

5 Limits to linear approximation

From the discussion in Section 2, it is clear that linear approximations are not appropriate if one wants to study the effects of uncertainty on behavior, such as precautionary saving. These effects are important in models where uncertainty is large, such as household models with unemployment.

Furthermore, linearization cannot be used in the following cases:

- Inequality constraints, such as irreversibilities (for example, that a firm can increase, but not decrease the capital stock) cannot be captured in a linearized model.
- More generally, asymmetries are thrown out by linearization. For example, if it costs a firm 1000 dollars to buy a machine, but the firm gets back only 500 if it sells the same machine, this cannot be handled in a linearized model.
- Discrete decision problems, such as the decision of quitting or staying in a job, don't have an Euler equation (first order condition). One has to compare the value of doing one thing (quitting) to the value of doing the other thing. This needs a value function approach, and cannot be solved (in a straightforward way) by linearization.
- Models where the first order conditions are not sufficient for a solution (non-convex optimization problems) need special care. We will talk about this when we do dynamic programming.
- For problems of asset choice (for example, holding stocks versus holding bonds), uncertainty is essential: it is the uncertainty that makes the assets different. This cannot be handled by linearization; however, one can use higher-order approximations (**?**)).
- Standard packages (including Dynare) cannot be used if, within one period, different pieces of information arrive sequentially, and decisions are based on different information sets, as for example in **?**). One can still use linear approximatons, as explained in their original paper.

5.1 Deriving Euler equations

Consider the model

$$\max_{u_0,u_1,\ldots} \mathrm{E}_0 \sum_{t=0}^{\infty} \beta^t R(x_t, u_t, z_t) \tag{38a}$$

subject to

$$x_{t+1} = u_t + \xi(z_{t+1}), \qquad \mathrm{E}_t\, \xi(z_{t+1}) = 0 \tag{38b}$$

$$G(x_t, u_t, z_t) \geq 0 \tag{38c}$$

$$x_0 \text{ given} \tag{38d}$$

and the transversality condition

$$\mathrm{E}_0 \lim_{t\to\infty} \beta^t R_x(x_t, u_t, z_t) \cdot x_t = 0 \tag{38e}$$

Here, x_t is a vector of endogenous state variables, u_t is a vector of control variables, and z_t is a vector of exogenous random variables which follows a Markovian transition law that need not be specified here. The random vector can appear in the objective function as well as in the dynamic equation, with the requirement that $\mathrm{E}_t\, \xi(z_{t+1}) = 0$. In other words, the dynamic equation (38b) is written such that the control u_t is the expected value of next period's state vector. It is usually possible to rewrite a model in this form.

The Bellman equation is

$$V(x_t, z_t) = \max_{u_t \text{ s.t. } G(x_t,u_t,z_t)\geq 0} \left\{R(x_t, u_t, z_t) + \beta\, \mathrm{E}_t\, V\left(u + \xi(z_{t+1}), z_{t+1}\right)\right\} \tag{39}$$

The maximization at the rhs of (39) can be solved by the Lagrangian

$$\mathcal{L} = R(x_t, u_t, z_t) + \lambda_t' G(x_t, u_t, z_t) + \beta\, \mathrm{E}_t\, V\left\{u_t + \xi(z_{t+1}), z_{t+1}\right\} \tag{40}$$

where λ_t is a vector of Lagrange multipliers, one multiplier per equation in (38c). The FOC is

$$R_u(t) + \lambda_t G_u(t) + \beta\, \mathrm{E}_t\, V_x(t+1) = 0 \tag{41}$$

Here and in the following, subscripts denote partial derivatives. Differentiating (39) w.r.t. x we get

$$\begin{aligned} V_x(t) &= R_x(t) + \left(R_u(t) + \beta\, \mathrm{E}_t\, V_x(t+1)\right) \frac{\partial u_t}{\partial x_t} \\ &= R_x(t) - \lambda_t' G_u(t) \frac{\partial u_t}{\partial x_t} \end{aligned}$$

Since $\lambda(x_t, z_t)' G(x_t, u(x_t), z_t) = 0$ for all (x_t, z_t), we can differentiate this identity and get

$$\lambda_t' \left(G_x(t) + G_u(t) \frac{\partial u_t}{\partial x_t} \right) = 0$$

which gives

$$V_x(t) = R_x(t) + \lambda_t' G_x(t)$$

Inserting this into (41) gives the Euler equation

$$
\begin{aligned}
0 = R_u\left(x_t, u_t, z_t\right) &+ \sum_{i=1}^{q} G_u^i\left(x_t, u_t, z_t\right) \lambda_t^i(x_t, z_t) \\
&+ \beta\, \mathrm{E}_t \left[R_x\left(x_{t+1}, u_{t+1}, z_{t+1}\right) + \sum_{i=1}^{q} G_x^i\left(x_{t+1}, u_{t+1}, z_{t+1}\right) \lambda_{t+1}^i\left(x_{t+1}, z_{t+1}\right) \right]
\end{aligned}
\tag{42}
$$