DYNAMIC PROGRAMMING

FOR DUMMIES

Parts I & II

Gonçalo L. Fonseca

Contents:

Part II

(4) Stochastic Dynamic Programming
- (a) Some Basics
- (b) Some Examples:
  - (i) Consumption with Uncertainty
  - (ii) Asset Prices
  - (iii) Search Models

(5) Mathematical Appendices
- (a) Principle of Optimality
- (b) Existence of Bellman's Equation

Texts

There are actually not many books on dynamic programming methods in economics. The following are standard references:

Stokey, N.L. and Lucas, R.E. (1989) *Recursive Methods in Economic Dynamics*. (Harvard University Press)

Sargent, T.J. (1987) *Dynamic Macroeconomic Theory* (Harvard University Press)

Lars Ljungqvist and Sargent, T.J. (2000) *Recursive Macroeconomic Theory* (MIT Press)

Stokey and Lucas does more the formal "mathy" part of it, with few worked-through applications. In contrast, Sargent's books have a lot of relatively simple and interesting applications to macroeconomic theory. Here, we draw a lot of our examples from Sargent (1987).

(1) Some Basic Intuition in Finite Horizons

(A) Optimal Control vs. Dynamic Programming

The method of dynamic programming is analagous, but different from optimal control in that optimal control uses continuous time while dynamic programming uses discrete time. Recall the general set-up of an optimal control model (we take the Cass-Koopmans growth model as an example):

$$\max \int u(c(t))e^{-rt}dt$$

s.t.

$$dk/dt = f(k(t)) - c(t) - nk$$

plus initial and transversality conditions of some sort. Thus, we are trying to find paths for control and state variables that maximize a continuous, discounted stream of utility over time. We find this path by setting up the current-value Hamiltonian (dropping time subscripts):

$$H = u(c) + \lambda(f(k) - c - nk)$$

where λ is the co-state variable. Solving this using the first order conditions for a Hamiltonian which, after some math, yields us a set of differential equations for dk/t and dc/dt and, when, solved, yield us the optimal paths $\{c(t)^*, k(t)^*\}$.

In "dynamic programming", we have a different story. Here, we use *discrete* time rather than continuous time. So, instead of maximizing the integral of a continuous utility over time, we maximize the sum of discrete utilities over time. Also, our constraint is no longer a differential equation (as in optimal control) but rather a *difference* equation. So our problem looks something like:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t)$$

s.t.

$$k_{t+1} = f(k_t) - c_t$$

or some version thereof. β^t is the discrete time discount factor (discrete time analogue of e^{-rt} in continuous time). How do we solve this? We can't use Hamiltonians here, our workhorse for the optimal control problem. We must use something else - namely, the *value* function.

(B) The Finite Case: Value Functions and the Euler Equation

Let us follow the basic intuition of a 12-year-old: imagine a path over time that starts at $t = 0$ and ends at $t = T$. Let us divide time up into T units of time, where a particular unit of time is denoted "t".

Let us concentrate on this single unit of time (t) and ignore all the other time periods - so that, for the moment, we are thinking of a static problem. We begin with a *given* amount of capital, k_t. Now, we wish to give this capital stock k_t some "value", which we write $V(k_t)$. But what sort of "value" ought we to assign to it? One possibility is the amount of utility we get out of that capital. However, as capital yields no utility directly, the value of the capital stock must be considered differently - perhaps, we may suspect, as the maximum amount of utility we obtain from that capital stock *once* we've converted it to output and consumed it. We know output is derived from capital, so $f(k_t)$ is the output we obtain from k_t. So, we would like to write something like $V(k_t) = \max u(c_t)$ s.t. $c \leq f(k_t)$, i.e. the value of the capital stock k_t at t is the maximum utility we obtain from consuming the output made with that capital. If we assume preferences are monotonic, etc., then *all* output is consumed and thus the constraint holds with equality so that $V(k_t) = u(f(k_t))$.

Note that so far, there is "no future" and thus this is a static problem. However, suppose now there is a future, i.e. some t+1. Suppose that capital tomorrow (k_{t+1}) must be constructed (or rather saved) from output today ($f(k_t)$). In this case, consumption is restricted to the *unsaved* output, i.e. $c_t \leq f(k_t) - k_{t+1}$. Assuming this holds with equality, then:

$$V(k_t) = \max u(c_t) = \max u(f(k_t) - k_{t+1})$$

This is still the same problem in essence: we seek to choose the consumption level such that we obtain the maximum amount of utility possible from that intial stock of capital k_t. Recall that k_t is given and thus, by extension, $f(k_t)$ is given. However, k_{t+1}, the capital we build for tomorrow, is not exogenously given but it is an item of choice. Note also that we can increase c_t by decreasing k_{t+1}. Thus, our "control" over c_t translates into "control" over k_{t+1}. So, by this simple manipulation, we've switched our "control" variable from c_t to k_{t+1}. One should easily see that these controls are in fact the same: regardless of which control we decide to employ to maximize utility, we should obtain the same result.

However, if there is no use for capital in the future, then optimally speaking, $k_{t+1} = 0$. Thus, the value of capital k_t becomes $V(k_t) = u(f(k_t))$, i.e. we consume all output. This akin to a "final period" solution. But *now* suppose that future capital k_{t+1} has *some* use. Of what use is it? Well, we already have a term for "usefulness" of present capital - we called it $V(k_t)$. So, tentatively, the "usefulness" of future capital we can simply call $V(k_{t+1})$. So our original value function ought to change to something like:

$$V(k_t) = \max \{u(c_t) + V(k_{t+1})\}$$

as we now have utility gains from future capital. So, the gain of "present" capital is not only how much utility we get from converting some of it to output and consuming that but also the utility we get tomorrow from *not* consuming all of present capital now but rather saving some for the future period.

But we've gotten into a little tangle. $V(k_t)$ is the value, at the present time t, of current capital, whereas $V(k_{t+1})$ is the value, at the present time t, of future capital. Thus, subscripting the value functions by t, we have $V_t(k_t)$ and $V_t(k_{t+1})$. But these value functions ought *not* to be comparable as they are considering two wholly different things (present capital versus future capital). Thus, to make the value functions comparable, we should make $V_t(k_{t+1}) = \beta V_{t+1}(k_{t+1})$. Thus, the present value of future capital is the discounted future value of future capital, where $V_{t+1}(k_{t+1})$ is thus analogous, logically, to $V_t(k_t)$ - albeit iterated one period ahead. Thus, our value function ought properly to be:

$$V_t(k_t) = \max \{u(c_t) + \beta V_{t+1}(k_{t+1})\} \qquad \text{(VF)}$$

Wonderful, isn't it? What is double wonderful is that we already have a term for k_{t+1}, namely the difference equation itself, i.e. $k_{t+1} = f(k_t) - c_t$. If we wanted, we could plug that in to obtain:

$$V_t(k_t) = \max \{u(c_t) + \beta V_{t+1}[f(k_t) - c_t]\}$$

and, abracadabra, that's the value of present capital. (alternatively, we could have used inverted the transition function and used k_{t+1} as our control).

How do we maximize the value of present capital? If differentiable, we can use the simple Lagrangian method: obtain the first order condition by taking the derivative of this term with respect to the control (in this case, c_t). Thus, the first order condition is:

$$\frac{dV_t(k_t)}{dc_t} = \frac{dU}{dc_t} + \beta\left[\frac{dV_{t+1}}{dk_{t+1}} \cdot \frac{dk_{t+1}}{dc_t}\right] = 0 \qquad \text{(FOC)}$$

a fairly obvious result. As we ostensibly know the functional form of U(.), we can easily get dU/dc_t; and we know $dk_{t+1}/dc_t = -1$, from the difference equation, then we can write the FOC as:

$$dU/dc_t = \beta[dV_{t+1}/dk_{t+1}]$$

Thus, this claims that we will adjust consumption today until the marginal utility of present consumption equals its marginal disutility - the latter being measured by the (discounted) utility foregone tomorrow by reducing future capital (i.e. savings) by a unit.

But we are not quite finished as we still don't know what dV/dk_{t+1} is. (not knowing the value function is one of the major problems in all applications of dynamic programming). However, the following trick can be used - the "Benveniste-Scheinkman" (B-S) condition. Namely, take the derivative of $V_t(k_t)$ with respect to the state, k_t. Then:

$$\frac{dV_t(k_t)}{dk_t} = \beta\left[\frac{dV_{t+1}}{dk_{t+1}} \cdot \frac{dk_{t+1}}{dk_t}\right] \qquad \text{(BS)}$$

Now, $dk_{t+1}/dk_t = f'(k_t)$, i.e. the marginal product of capital. Thus:

$$dV_t/dk_t = \beta(dV_{t+1}/dk_{t+1})\cdot f'(k_t)$$

Now, recall that the FOC claimed that $dU/dc_t = \beta(dV_{t+1}/dk_{t+1})$. Thus, plugging in:

$$dV_t/dk_t = (dU/dc_t)\cdot f'(k_t)$$

iterating everything one period forward:

$$dV_{t+1}/dk_{t+1} = (dU/dc_{t+1})\cdot f'(k_{t+1})$$

et voila! We have an expression for dV_{t+1}/dk_{t+1}, which we can plug into our FOC, which will yield:

$$dU/dc_t = \beta(dU/dc_{t+1})\cdot f'(k_{t+1})$$

or, recognizing that $k_{t+1} = f(k_t) - k_t$, then:

$$dU/dc_t = \beta(dU/dc_{t+1})\cdot f'(f(k_t) + k_t)$$

which one ought to recognize as the discrete version of the "Euler Equation", so familiar in dynamic optimization and macroeconomics.

To see the Euler Equation more clearly, perhaps we should take a more familiar example. Consider, for simplicity, an intertemporal "consumption-savings" model which can be expressed as:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_t)$$

s.t.
$x_{t+1} = R(x_t - c_t)$
x_0 given

where c_t is consumption, which is deducted from "non-human" wealth or assets, x_t. Whatever is left over is saved into assets x_{t+1} for the future, which receive a constant return R = (1+r). Thus, consumption (c_t) is the control and x_t is the state. The value function at t is:

$$V_t(x_t) = \max \{u(c_t) + \beta V_{t+1}(x_{t+1})\} \qquad \text{(VF)}$$

taking the first order condition:

$$\frac{dV_t(x_t)}{dc_t} = \frac{dU}{dc_t} + \beta\left[\frac{dV_{t+1}}{dx_{t+1}} \cdot \frac{dx_{t+1}}{dc_t}\right] = 0 \qquad \text{(FOC)}$$

or, as $x_{t+1} = R(x_t - c_t)$, then $dx_{t+1}/dc_t = -R$, so the FOC becomes:

$$dU/dc_t = \beta R[dV_{t+1}/dx_{t+1}]$$

As we do not know what dV/dx_{t+1} is, then we need the Benveniste-Scheinkman (BS) condition. Taking the derivative of $V_t(x_t)$ with respect to the state, x_t:

$$\frac{dV_t(x_t)}{dx_t} = \beta\left[\frac{dV_{t+1}}{dx_{t+1}} \cdot \frac{dx_{t+1}}{dx_t}\right] \qquad \text{(BS)}$$

or, as $dx_{t+1}/dx_t = R$:

$$dV_t/dk_t = \beta R(dV_{t+1}/dx_{t+1})$$

Now, as the FOC claimed that $dU/dc_t = \beta R(dV_{t+1}/dx_{t+1})$, then this becomes:

$$dV_t/dx_t = dU/dc_t$$

or iterating one period forward:

$$dV_{t+1}/dx_{t+1} = dU/dc_{t+1}$$

and plugging this back into the FOC, we obtain:

$$dU/dc_t = \beta R(dU/dc_{t+1})$$

a *far* more familiar Euler Equation.

How does this Euler Equation help us? If we gave it an explicit functional form, we could say something about the relationship between c_{t+1} and c_t - which might be informative. After all, the Euler Equation is a difference equation itself - which could be solved for on optimal path $\{c_t^*\}$ *if* we had an initial condition on c_0. But we do not have this initial condition. Our sole initial condition is x_0 - and x does not enter this equation explicitly. The solution to the optimal path must be derived with some additional machinery.

One way would be if we could derive a set of "policy functions". A "policy function" $h_t(.)$ relates the control variable at period t with the contemporaneous state variable at period t, i.e. a function $c_t^* = h_t(x_t)$. In the finite horizon case, policy functions are different for different periods. The advantage of getting a policy function is that *then* we could plug these into our Euler Equation instead of the controls. Thus, if we could find $c_t^* = h_t(x_t)$ and $c_{t+1}^* = h_{t+1}(x_{t+1})$, then our Euler becomes:

$$dU/d(h_t(x_t)) = \beta R(dU/d(h_{t+1}(x_{t+1}))$$

As we have an initial x_0, we this becomes a solvable difference equation. However, we rarely, if ever, can derive the policy function so easily from FOC. However, in the finite horizon case, another method of solution is available - namely, that of *backward recursion*.

(C) The Recursive Solution

How does one go about solving for $\{c_t^*\}$ from the Euler Equation? The simplest procedure, in a finite horizon, is backwards recursion. Here is where the magic of the "recursive" structure really comes in. Firstly, the term recursive means that we have a system of equations which "fit" into one another sequentially. For instance, consider the two equations, $y = f(x)$ and $z = f(y)$ - then starting with x, we get y from the first equation, plug into the second equation and then obtain z. That's a recursive system.

Well, the value function derived earlier is merely a method of solution for a recursive system. The general idea can be thought of this way. Consider the Solowian model again. Suppose you have a finite time path, broken up into T time periods (beginning at 0 ending at T) which are of length t each. Now, look at the *final* time period, T, which we begin with a *given* amount of capital, k_T. The value function is:

$$V_T(k_T) = \max_{c_T} \{u(c_T)\} \qquad \text{s.t. } f(k_T) - c_T \geq 0$$

where the term $\beta V_{T+1}(k_{T+1})$ is omitted the time period T+1 does not exist - thus no future value for any capital created today. This is akin to our earlier one-period static problem which is simply solved: c_T^*, the optimal consumption in the final period will be exactly the entire output, i.e. $c_T^* = f(k_T)$, the output produced from the capital stock we began the final period T with. Thus, $V_T(k_T) = \max \{u(c_T)\} = u(c_T^*) = u(f(k_T))$

But now let's go back a step. Let's go the "next-to-last" period, i.e. period T-1. Here we have a value function that looks like:

$$V_{T-1}(k_{T-1}) = \max_{c_{T-1}} \{u(c_{T-1}) + \beta V_T(k_T)\} \quad \text{s.t. } k_T = f(k_{T-1}) - c_{T-1}$$

since T, the final period, is a "future" which exists. We could proceed to solve it - but recall that we have *already* solved $V_T(k_T) = u(c_T^*)$. We could just plug that back in so that the value function becomes:

$$V_{T-1}(k_{T-1}) = \max_{c_{T-1}} \{u(c_{T-1}) + \beta u(c_T^*)\}$$

where now we are facing a one-period problem: finding the c_{T-1} that maximizes utility in period T-1 while at the same time providing the capital *necessary* to achieve c_T^* in the final period. So, c_T^* enters here as a "constraint" upon the values c_{T-1} can take. Thus, maximizing this one-period problem, we obtain some solution $c_{T-1}*$. Plugging this in, we have:

$$V_{T-1}(k_{T-1}) = u(c_{T-1}^*) + \beta u(c_T^*)$$

Now, let us turn to the *previous* period, i.e. T-2. Here we face the value function:

$$V_{T-2}(k_{T-2}) = \max_{c_{T-2}} \{u(c_{T-2}) + \beta V_{T-1}(k_{T-1})\} \qquad \text{s.t. } k_{T-1} = f(k_{T-2}) - c_{T-2}$$

here $V_{T-1}(k_{T-1})$ is now the value of our "future" capital next period - which we have already solved before. Thus, plugging that in, we obtain:

$$V(k_{T-2}) = \max \{u(c_{T-2}) + \beta[u(c_{T-1}^*) + \beta u(c_T^*)]\}$$

or simply:

$$V(k_{T-2}) = \max \{u(c_{T-2}) + \beta u(c_{T-1}^*) + \beta^2 u(c_T^*)\}$$

so, once again, we have a one-period problem - with c_{T-1}^* and c_T^* already given and known and thus acting as constraints on the values of c_{T-2} we can choose. So we take FOCs, etc. and find the optimal $c_{T-2}*$ as simple static optimization. And so we continue on iterating backwards. We go to period T-3 and use c_{T-2}^*, c_{T-1}^* and c_T^* as the constraints to that maximization problem, etc. We do this continuously until we get back to the beginning (time t = 0). Thus, it is a backwardly "recursive" system. We start from the end and solve backwards by this normal fashion. When we have iterated backwards completely via this recursive method through all the value functions, we effectively generated an entire series of optimal values:

$$\{c_1^*, c_2^*, \ldots\ldots, c_{T-2}^*, c_{T-1}^*, c_T^*\}$$

which, when plotted, is quite simply the "optimal path".

We could solve an entire intertemporal optimizing problem of max $\sum^T_{t=1} \beta^t u(c_t)$ s.t. $k_{t+1} = f(k_t) - c_t$ via this recursive method. Unlike optimal control, we did not need to go looking for a specific dynamic equation to integrate and get a solution. How do we know it is the same solution? Consider the original intertemporal problem again. If we had set it up as a Lagrangian, we would obtain something like:

$$L = \sum_{t=1}^T \beta^t u(c_t) + \lambda_1(k_2 - f(k_1) - c_1) + \lambda_2(k_3 - f(k_2) - c_2) + \ldots.\ \lambda_T(k_{T+1} - f(k_T) - c_T)$$

where final capital stock k_{T+1} must conform to some prespecified transversality facet. We obtain, from this, a set of first order conditions which yield a path $\{c_t^*\}_{t=1}^{\infty}$. To see equvialence (if you are not yet convinced), recall that at any period t, our value function was:

$$V_t(k_t) = \max_{c_t} \{u(c_t) + \beta V_{t+1}(k_{t+1})\} \qquad \text{s.t. } k_{t+1} = f(k_t) - c_t$$

where we have noted underneath the max operator that we are *merely* choosing c_t. Now, as $V_{t+1}(k_{t+1}) = \max\{u(c_{t+1}) + \beta V_{t+2}(k_{t+2})\}$ where we are merely choosing c_{t+1}, then plugging this into our original problem:

$$V_t(k_t) = \max_{c_t} \{u(c_t) + \beta(\max_{c_{t+1}}\{u(c_{t+1}) + \beta V_{t+2}(k_{t+2})\})\}$$

or, combining max operators:

$$V_t(k_t) = \max_{\{c_t,\, c_{t+1}\}} \{u(c_t) + \beta u(c_{t+1}) + \beta^2 V_{t+2}(k_{t+2})\} \qquad \text{s.t. } k_{t+2} = f(k_{t+1}) - c_{t+1}, \quad k_{t+1} = f(k_t) - c_t$$

thus, we are now choosing two variables, c_t and c_{t+1} - subject to two constraints. Iterating again and again until the final period T, we obtain:

$$V_t(k_t) = \max_{\{c_t,\, c_{t+1},\, ...,\, c_T\}} \{\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(c_\tau)\} \qquad \text{s.t. } k_{\tau+1} = f(k_\tau) - c_\tau \text{ for all } t \leq \tau \leq T$$

thus, we are trying to find an "entire" path $\{c_\tau\}_{\tau=t}^{T}$ with a single value function - appropriately iterated. If we consider k_t as our initial k, then $V_t(k_t)$ is merely the maximized value of the earlier intertemporal optimization problem's objective function. The constraint works implicitly in a similar manner as before as the $\{c_t\}_{t=1}^{T}$ can only take values restricted by the difference equation, $k_{t+1} = f(k_t) - c_t$. Thus, if we wished, recognizing that $c_\tau = f(k_\tau) - k_{\tau+1}$, then we could rewrite this entirely with k_{t+1} as our controls so:

$$V_t(k_t) = \max_{\{k_{t+1},\, k_{t+2},\, ...,\, k_T\}} \{\textstyle\sum_{\tau=t}^{T} \beta^{\tau-t} u(f(k_\tau) - k_{\tau+1})\}$$

and the system is explicit and complete. The solution is identical as to the one obtained by the maximization of the original system via Lagrangian methods. Going forward, or going backwards, we obtain in the end the essential result that:

$$V_0(k_0) = \max_{\{k_1,\, k_2,\, ...,\, k_T\}} \{\textstyle\sum_{t=1}^{T} \beta^{t} u(f(k_t) - k_{t+1})\}$$

to solve the entire system - as k_0 is given.

All this can be illustrated as in Figure 1. We have here five time periods from t_0 to t_5. We have an initial state (k_0) and a final state (k_5). Going forward from k_0, we can choose either k_{11} or k_{12}, depending on which we think is "better" by the value function. Suppose we choose k_{12}. Then at period t_1, we begin at k_{12} and must now choose between going to k_{22} or k_{23} (k_{21} being no longer possible). If we believe k_{22} is better, we go in that direction. At time t_2, starting at k_{22}, we are now forced to go to k_{32} as that is the only option. At t_3, starting at k_{32}, we can choose either k_{42} or k_{43}, and we choose to go to k_{42}. Finally, at k_{42} at t_4, we are forced to go to k_T as that is our endpoint constraint. Thus, our series of multi-stage decisions have plotted out a path from k_0 to k_T via $k_0 \rightarrow k_{12} \rightarrow k_{22} \rightarrow k_{32} \rightarrow k_{42} \rightarrow k_T$ and this is the "optimal" state path, even though we could have gone through many others.

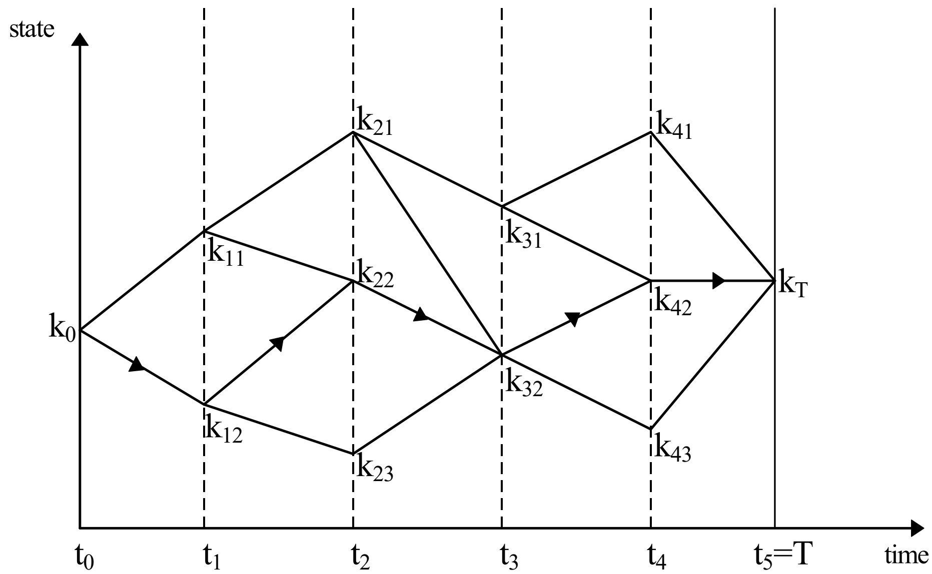

Figure 1 - Optimal Path via Multi-Stage Dynamic Programming

Backwards recursion works effectively the opposite way. Suppose we are at the next-to-last period t_4 and we are restricted to go to k_T in the final time period. We have three choices: k_{41}, k_{42} and k_{43}. We choose k_{42} as the "best". Consequently, going back one period to t_3, we must now choose between k_{31} and k_{32} *subject* to the constraint that k_{42} will be our the state in the next period t_4. Suppose we now choose k_{32} as being "better" than k_{31}. Then going backwards one more period to t_2, we now must choose between k_{21}, k_{22} and k_{23} *subject* to the constraint that we are at k_{32} in the next period. But notice that this knocks k_{21} out of contention as an option. Thus, we are forced to choose between k_{22} and k_{23}. If we take k_{23}, then going back a period to t_1, we must now choose between k_{11} and k_{12} subject to the constraint that k_{23} is the next period's state. Again, this knocks k_{11} out of consideration, and we must choose k_{12}. Finally, in the initial time period, we choose the path that takes us from k_0 at t_0 to k_{12} in time period t_1. Thus, by backwards recursion we trace out the same optimal path, $k_0 \rightarrow k_{12} \rightarrow k_{22} \rightarrow k_{32} \rightarrow k_{42} \rightarrow k_T$.

It might be worthwhile considering a specific exercise. Here, we shall consider two: an optimal consumption-savings decision problem and an optimal investment model.

(i) *Example No.1 - Consumption-Savings Decisions*

Suppose we have the simple consumption-savings problem for three periods:

$$\max \; u(c_1) + \beta u(c_2) + \beta^2 u(c_3)$$

s.t.

$$x_{t+1} = R(x_t - c_t)$$
$$x_1 \text{ given}$$
$$x_4 \geq 0.$$

For simplicity, let $u(c_i) = \ln c_i$. Then, let us consider each of the value functions individually. In the last period, we maximize period 3 consumption subject to the constraint that $R[x_3 - c_3] \geq 0$. As we do not want any assets in period 4, we face the value function:

$$V_3(x_3) = \max_{c3}\{u(c_3)\} = \max_{c3}\{\ln c_3\} \qquad \text{s.t. } R[x_3 - c_3] \geq 0$$

which, deriving FOCs, yields at the optimum $c_3^* = x_3$. Thus, plugging this back in, the value function in period 3 is:

$$V_3(x_3) = \ln x_3$$

In period 2, we face the value function:

$$V_2(x_2) = \max_{c2}\{u(c_2) + \beta V_3(x_3)\} \qquad \text{s.t. } x_3 = R[x_2 - c_2]$$

so the first order condition yields:

$$dV_2/dc_2 = 1/c_2 - \beta(dV_3/dx_3)(dx_3/dc_2) = 0 \qquad (FOC_2)$$

Now, by the transition function, $dx_3/dc_2 = -R$, thus the FOC becomes:

$$1/c_2 = R\beta(dV_3/dx_3)$$

The only thing that remains is dV_3/dx_3. As this V_3 must *already* be optimal, we know from before that $V_3(x_3) = \ln x_3$ or simply, by the transition function, $V_3 = \ln [R(x_2 - c_2)]$, so $dV_3/dx_3 = 1/[R(x_2 - c_2)]$, so the FOC becomes

$$1/c_2 = R\beta/[R(x_2 - c_2)] = \beta/(x_2 - c_2)$$

or:

$$c_2[1 + 1/\beta] = x_2/\beta$$

thus:

$$c_2^* = x_2/(1+\beta) \qquad (PF_2)$$

which is our solution. This is what is called the "policy function" for period 2, i.e. $c_2^* = h_2(x_2) = x_2/(1+\beta)$ - thus, given x_2, we can find c_2^* via this function. Notice that this also implies that $x_3 = R[x_2 - c_2^*] = Rx_2\beta/(1+\beta)$. Thus, the value function at time 2 can be rewritten:

$$V_2(x_2) = \ln (x_2/(1+\beta)) + \beta \ln [Rx_2\beta/(1+\beta)]$$

or simply:

$$V_2(x_2) = \ln x_2 - \ln (1+\beta) + \beta[\ln R + \ln x_2 - \ln \beta - \ln (1+\beta)]$$

Now, the Benveniste-Scheinkman condition states that $dU/dc_2^* = dV/dx_2$. So, as we know, $dU/dc_2^* = 1+\beta/x_2$. Does this hold? Well, note that $dV/dx_2 = 1/x_2 + \beta/x_2 = (1+\beta)/x_2$. So it does.

In period 1, we face the value function:

$$V_1(x_1) = \max_{c1}\{u(c_1) + \beta V_2(x_2)\} \qquad \text{s.t. } x_2 = R[x_1 - c_1]$$

So the first order condition yields:

$$dV_1/dc_1 = 1/c_1 - \beta(dV_2/dx_2)(dx_2/dc_1) = 0.$$

By the transition function, $dx_2/dc_1 = -R$ again, so:

$$1/c_1 = \beta R(dV_2/dx_2) \qquad (FOC_1)$$

We now face the problem of dV_2/dx_2. The Benveniste-Scheinkman condition claims that:

$$dV_1/dx_1 = \beta(dV_2/dx_2)(dx_2/dx_1) = \beta R(dV_2/dx_2) \qquad (BS)$$

thus, plugging the FOC into this and iterating forward by one period:

$$dV_2/dx_2 = 1/c_2^*$$

(we have to remember that the c_2^* is optimal). Now, we derived earlier the policy function $c_2^* = x_2/(1+\beta)$, thus plugging this into the equation:

$$dV_2/dx_2 = (1+\beta)/x_2$$

and then plugging this into the FOC:

$$1/c_1 = \beta R(1+\beta)/x_2$$

Or, as $x_2 = R[x_1 - c_1]$, then:

$$1/c_1 = \beta R(1+\beta)/[R(x_1 - c_1)] = \beta(1+\beta)/(x_1 - c_1)$$

or simply:

$$c_1^* = x_1/(1 + \beta + \beta^2) \qquad (PF_1)$$

which is our "policy function" for period 1, i.e. $c_1^* = h_1(x_1)$. Note that the *form* of the policy function for period 1 (PF_1) is *different* from the policy function for period 2. Thus, our value function at period 1 is:

$$V_1(x_1) = \ln c_1^* + \beta V_2(x_2)$$

Now, $c_1^* = x_1/(1+\beta+\beta^2)$ implies that:

$$V_1(x_1) = \ln x_1 - \ln(1+\beta+\beta^2) + \beta V_2(x_2)$$

But recall that $V_2(x_2) = \ln(x_2/(1+\beta)) + \beta \ln[Rx_2\beta/(1+\beta)]$. Thus, the value function becomes:

$$V_1(x_1) = \ln x_1 - \ln(1+\beta+\beta^2) + \beta\{\ln(x_2/(1+\beta)) + \beta \ln[Rx_2\beta/(1+\beta)]\}$$

But, recall that $x_2 = R[x_1 - c_1^*]$, so, substituting for c_1^*, then $x_2 = Rx_1\beta(1+\beta)/[1 + \beta(1+\beta)])$. Thus:

$$V_1(x_1) = \ln x_1 - \ln(1+\beta+\beta^2)$$
$$+ \beta(1+\beta) \ln[Rx_1\beta(1+\beta)/(1 + \beta(1+\beta))] + \beta[\beta \ln(R\beta/(1 +\beta)) - \ln(1+\beta)].$$

or:

$$V_1(x_1) = (1+\beta+\beta^2) \ln x_1 - \ln(1+\beta+\beta^2) + \ln[R\beta(1+\beta)/(1 + \beta(1+\beta))]$$
$$+ \beta[\beta \ln(R\beta/(1 +\beta)) - \ln(1+\beta)].$$

which is a very, very ugly term. However, it might be useful to verify if the Benveniste-Scheinkman condition holds. We can see from the ugly term above that $dV_1/dx_1 = (1 + \beta + \beta^2)/x_1$. But, we also know that $dU/dc_1^* = (1 + \beta + \beta^2)/x_1$. Thus, indeed, $dU/dc_1^* = dV/dx_1$, i.e. the BS condition holds.

Thus, note that we have obtained three different value functions $V_1(x_1)$, $V_2(x_2)$ and $V_3(x_3)$ - *all* of which have different forms. We have also obtained three different policy functions,

$c_1^* = h_1(x_1) = x_1/(1+\beta + \beta^2)$, $c_2^* = h_2(x_2) = x_2/(1+\beta)$ and $c_3^* = h_3(x_3) = x_3$ - all of which have different forms. Nonetheless, note that the solution to our intertemporal problem is equal to the solution to our dynamic programming method via value functions.

(ii) *Example No.2 - Investment with Adjustment Costs*

Firm faces investment decision in a finite horizon. There is no depreceiation, so $k_{t+1} = k_t + y_t$, where y_t is investment. Current profits are defined as $\pi_t[k_t, y_t] = A_t k_t - B/2y_t^2$ - thus, note, we have "adjustment costs" in the form of profits declining profits by some amount proportional to the square of investment. A_t can be thought of as the average productivity of capital. Firm's maximization problem is:

$$\max \sum_{t=0}^{T} \beta^t \{\pi_t\}$$

$$\text{s.t.} \quad k_{t+1} = k_t + y_t$$
$$k_0 \text{ fixed.}$$

The dynamic programming problem can be written with y_t as control and k_t as state. As we have finite time, the value function at time t is:

$$V_t(k_t) = \max_{y_t} \{\pi_t[k_t, y_t] + \beta V_{t+1}(k_{t+1})\}$$

or:

$$V_t(k_t) = \max_{y_t} \{A_t k_t - B/2y_t^2 + \beta V_{t+1}(k_t + y_t)\}$$

We can obtain the Euler Equation for time $t < T$ via the FOC and Benveniste-Scheinkman. Namely, taking the first order condition:

$$d\pi_t/dy_t + \beta(dV_{t+1}/dk_{t+1})(dk_{t+1}/y_t) = 0$$

or, as $d\pi_t/dy_t = -By_t$ and $dk_{t+1}/y_t = 1$:

$$By_t = \beta(dV_{t+1}/dk_{t+1}) \qquad \text{(FOC)}$$

Our next step is to derive the Benveniste-Scheinkman condition. Differentiating:

$$dV_t/dk_t = d\pi_t/dk_t + \beta(dV_{t+1}/dk_{t+1})(dk_{t+1}/dk_t)$$

or, as $dk_{t+1}/dk_t = 1$ and as $d\pi_t/dk_t = d(Ak_t - B/2y_t^2)/dk_t = A_t$, then:

$$dV_t/dk_t = A_t + \beta(dV_{t+1}/dk_{t+1}) \qquad \text{(BS)}$$

Now, plugging in the FOC for the second term:

$$dV_t/dk_t = A_t + By_t$$

iterating one period forward:

$$dV_{t+1}/dk_{t+1} = A_{t+1} + By_{t+1}$$

and plugging this back into the FOC:

$$By_t = \beta(A_{t+1} + By_{t+1})$$

or simply:

$$y_t = \beta(A_{t+1} + By_{t+1})/B \qquad \text{(EE)}$$

Which is our "Euler equation".

The issue now comes around to deriving the optimal control path. To do this, we must iterate backwards recursively. We can do this either by iterating on the Euler Equation or on the Value Function. Let us begin with the Euler. Starting at T, then, the optimal investment is $y_T^* = 0$ because it is costly and yields no future benefit. At T-1, we have:

$$y_{T-1}^* = \beta(A_T + By_T^*)/B = \beta A_T/B$$

because $y_T^* = 0$. At T-2, we have:

$$y_{T-2}^* = \beta(A_{T-1} + By_{T-1}^*)/B = \beta(A_{T-1} + B(\beta A_t/B))/B$$

or simply:

$$y_{T-2}^* = \beta(A_{T-1} + \beta A_T)/B$$

and so on. Thus, it is apparent that, for any t:

$$y_t^* = \beta A_{t+1} + \beta^2 A_{t+2} + ... + \beta^{T-t} A_T)/\beta$$

is the optimal investment at any time t < T.

Let us now do the same exercise of backward recursion on the value function. At T:

$$V_T(k_T) = A_T k_T$$

as $y_T^* = 0$. Then, at T-1:

$$V_{T-1}(k_{T-1}) = \max_{y_{T-1}} \{\pi_{T-1}[k_{T-1}, y_{T-1}] + \beta V_T(k_T)\} \qquad \text{s.t. } k_T = k_{T-1} + y_{T-1}$$

or:

$$V_{T-1}(k_{T-1}) = \max_{y_{T-1}} \{A_{T-1}k_{T-1} + B/2y_{T-1}^2 + \beta V_T(k_{T-1} + y_{T-1})\}$$

From the previous result, $V_T(k_{T-1} + y_{T-1}) = A_T(k_{T-1} + y_{T-1})$, so:

$$V_{T-1}(k_{T-1}) = \max_{y_{T-1}} \{A_{T-1}k_{T-1} + B/2y_{T-1}^2 + \beta A_T(k_{T-1} + y_{T-1})\}$$

Thus, maximizing, we obtain the first order condition:

$$dV_{T-1}/dy_{T-1} = - By_{T-1} + \beta A_T = 0$$

or:

$$y_{T-1}^* = \beta A_T/B$$

exactly the same solution we obtained earlier. Thus, completing, the value function by plugging in the optimal value y_{T-1}^*, we obtain:

$$V_{T-1}(k_{T-1}) = A_{T-1}k_{T-1} - B/2(\beta A_T/B)^2 + \beta A_T(k_{T-1} + \beta A_T/\beta)$$

Going to period T-2, we have the value function:

$$V_{T-2}(k_{T-2}) = \max_{y_{T-2}} \{\pi_{T-2}[k_{T-2}, y_{T-2}] + \beta V_{T-1}(k_{T-1})\} \qquad \text{s.t. } k_{T-1} = k_{T-2} + y_{T-2}$$

or substituting in for our previous result on $V_{T-1}(k_{T-1})$:

$$V_{T-2}(k_{T-2}) = \max_{y_{T-2}} \{A_{T-2}k_{T-2} + B/2(y_{T-2})^2 + \beta [A_{T-1}k_{T-1} - B/2(\beta A_T/B)^2 + \beta A_T(k_{T-1} + \beta A_T/B)]\}$$

substituting for k_{T-1}:

$$V_{T-2}(k_{T-2}) = \max_{y_{T-2}} \{A_{T-2}k_{T-2} + B/2(y_{T-2})^2 + \beta[A_{T-1}(k_{T-2} + y_{T-2}) - B/2(\beta A_T/B)^2 + \beta A_T(k_{T-2} + y_{T-2} + \beta A_T/B)]\}$$

Taking FOC for y_{T-2}:

$$dV_{T-2}/dy_{T-2} = -By_{T-2} + \beta(A_{T-1} + \beta A_T) = 0$$

or simply:

$$y_{T-2}^* = (\beta A_{T-1} + \beta^2 A_T)/B$$

which, again, is identical to before.

Thus, in this simple, finite-dimensional world, backwards iteration the Euler Equation and backwards iteration on value functions yields an identical solution for a quite simple optimal investment path, $\{y_t\}^T_{t=0}$.

(iii) *Habit Formation*

Consider the following finite-horizon consumption-savings problem with "habit-formation":

$$\max \sum_{t=1}^{T} \beta^{t-1} u(c_t, c_{t-1})$$

s.t.

$$x_{t+1} = R(x_t - c_t)$$

x_0 given

where $du/dc_t = U^1(c_t, c_{t-1}) > 0$ and $du/dc_{t-1} = U^2(c_t, c_{t-2}) < 0$.

where the lagged term c_{t-1} entering the utility function captures the phenomenon of "habit formation". U^1 and U^2 denote the first and second partial derivatives of the utility function $u(c_t, c_{t-1})$. The first issue is to tackle states and controls: in this case, c_{t-1} and x_t are the states as these are the things which we enter period t with, whereas c_t is the control (alternatively, we could take x_{t+1} as control). Then, the value function at any time t is:

$$V_t(x_t, c_{t-1}) = \max_{c_t} \{u(c_t, c_{t-1}) + \beta V_{t+1}(x_{t+1}, c_t)\} \quad \text{s.t. } x_{t+1} = R(x_t - c_t)$$

The first-order condition for a maximum is:

$$dV_t/dc_t = U^1(c_t, c_{t-1}) + \beta(dV_{t+1}/dx_{t+1})(dx_{t+1}/dc_t) + \beta(dV_{t+1}/dc_t) = 0$$

or, as $dx_{t+1}/dc_t = -R$, then:

$$U^1(c_t, c_{t-1}) = \beta R(dV_{t+1}/dx_{t+1}) - \beta(dV_{t+1}/dc_t) \qquad \text{(FOC)}$$

Now, we have two state variables and two V-terms to get rid of - thus we will need two Benveniste-Scheinkman conditions. Let us take the dV_t/dx_t case first:

$$dV_t/dx_t = \beta(dV_{t+1}/dx_{t+1})(dx_{t+1}/dx_t) = \beta R(dV_{t+1}/dx_{t+1}) \qquad \text{(BS1)}$$

Now, for the second case dV_t/dc_{t-1}:

$$dV_t/dc_{t-1} = U^2[c_t, c_{t-1}] \qquad \text{(BS2)}$$

Now, plugging in (BS1) into the (FOC), we obtain:

$$U^1(c_t, c_{t-1}) = (dV_t/dx_t) - \beta(dV_{t+1}/dc_t)$$

then, iterating one-period forward:

$$U^1(c_{t+1}, c_t) = (dV_{t+1}/dx_{t+1}) - \beta(dV_{t+2}/dc_{t+1})$$

then, iterating (BS2) *two* periods forward so $dV_{t+2}/dc_{t+1} = U^2(c_{t+2}, c_{t+1})$ and plugging in:

$$U^1(c_{t+1}, c_t) = (dV_{t+1}/dx_{t+1}) - \beta U^2(c_{t+2}, c_{t+1})$$

multiplying through by βR and rearranging:

$$\beta R(dV_{t+1}/dx_{t+1}) = \beta R[U^1(c_{t+1}, c_t) + \beta U^2(c_{t+2}, c_{t+1})]$$

thus, equating this with the (FOC):

$$U^1(c_t, c_{t-1}) + \beta(dV_{t+1}/dc_t) = \beta R[U^1(c_{t+1}, c_t) + \beta U^2(c_{t+2}, c_{t+1})]$$

finally, plugging in (BS2) iterated a period ahead:

$$U^1(c_t, c_{t-1}) + \beta U^2[c_{t+1}, c_t] = \beta R[U^1(c_{t+1}, c_t) + \beta U^2(c_{t+2}, c_{t+1})]$$

and rearranging:

$$U^1(c_t, c_{t-1}) = \beta R U^1(c_{t+1}, c_t) + \beta^2 R U^2(c_{t+2}, c_{t+1}) - \beta U^2[c_{t+1}, c_t] \qquad \text{(EE)}$$

which is our Euler Equation. We could attempt a solution by backward iteration of this recursive system, but we shall avoid doing so. This merely serves to illustrate how to find Euler Equations when there are multiple state variables.

There are four main lessons to be drawn from the recursive method with a finite time horizon:

(1) the optimal solution to the finite-horizon intertemporal optimization problem is identical to the set of solutions obtained recursively from maximizing the one-period problem represented by the Value Function.

(2) in some cases, we can obtain the optimal solution path by backwards recursion on the Euler Equation or by backwards recursion on the Value Function - as in our investment case;

(3), if we cannot get an explicit optimal solution path, at least we can obtain an explicit optimal policy function, as in the consumption case where we could express optimal control variable in any time period t c_t* as some function of the concurrent state x_t, i.e. $c_t^* = h_t(x_t)$ - as in our consumption-savings case.

(4) in principle, the *form* of the value function and/or the policy function at any time t are *different* from the value function and/or policy function at any other time, i.e. $h_t \neq h_{t+1}$ and $V_t \neq V_{t+1}$.

(2) The Infinite Case: Bellman's Equation

(A) Some Basic Intuition

A question emerges immediately: what happens when we have an infinite horizon, i.e. when there is no "final" time period T? How do we solve that "recursively" when there is nowhere to recurse from? The workhorse in this case is the *Bellman's Equation* - which is simply the value function we saw earlier, only now applied indiscriminately between any two time periods. Recall the value function for finite time was:

$$V_t(k_t) = \max \{u(c_t) + \beta V_{t+1}(k_{t+1})\} \qquad \text{(VF)}$$

Then, letting prime (′) denote "future", Bellman's Equation is merely:

$$V(k) = \max \{u(c) + \beta V(k')\} \qquad \text{(BE)}$$

where note the *very important point* that the V for V(k) and the V for V(k′) are identical - i.e. the value function is time-invariant.

What does this have to do with "solving" for the optimal path $\{c_t^*\}_{t=0}^{\infty}$? Well, that's the whole purpose of Stokey and Lucas's (1989) dense discussion from Chapters 3 and 4, namely to prove that the "solution" to the intertemporal problem is equivalent to "solving" Bellmans' equation and that a unique solution exists. What Stokey and Lucas demonstrate is that, under certain conditions, the function V(k) that solves $V(k) = \max \{u(c) + \beta V(k')\}$ is *exactly* the same function V(k) that solves $V(k_0) = \max \sum_{t=0}^{\infty} \beta^t u(c_t)$ s.t. $k_{t+1} = f(k_t) - c_t$ for a given initial k_0. This is a striking assertion but we already saw this equivalence in the finite horizon case. This is Bellman's (1957) "Principle of Optimality".

What is more striking is that in the infinite horizon case, the value function form is time-invariant. We can think of it as follows. Consider our original VF for the finite horizon case. This was:

$$V_t(k_t) = \max \{u(c_t) + \beta V_{t+1}(k_{t+1})\} \qquad \text{(VF)}$$

where we are choosing a $c_t \in C_t$ (the set of admissable values for c_t). If C_t is time-invariant or can be made so (so $C_t = C_{t+1} = C$), then $c_t \in C$. This is not that restrictive as usually C = R or some other rather conventional term. More restrictive is the assumption that U and f are time-invariant as well - but these are conditions we must impose. As a result, t is arbitrary for the *arguments*, so that we can rewrite VF as:

$$V_t(k) = \max \{u(c) + \beta V_{t+1}(k')\}$$

where k′ is one period ahead of k. The only thing that is not time-independent, then, are the value functions themselves. However, recall that in our finite-horizon case, the pattern of recursion of one value function into another was $V_T(k) \rightarrow V_{T-1}(k) \rightarrow V_{T-2}(k) \rightarrow ...... \rightarrow V_1(k) \rightarrow V_0(k)$. What Stokey and Lucas assert is that, in infinite time, then from *any* initial V_0, then as $t \rightarrow \infty$, we $V_t \rightarrow V$, a *unique* value function. This is what permits us to write our Bellman's equation as:

$$V(k) = \max \{u(c) + \beta V(k')\}$$

where, note again, the V for V(k) is the same as the V for V(k′).

The intuition of why we should obtain a single V for all time periods in the infinite case should be obvious. In a finite horizon case, this should not be possible because $V_t(k_t)$ and $V_{t+1}(k_{t+1})$ have *different* futures - the former has (T - t) time periods ahead of it, whereas the latter has (T-t-1) time periods ahead of it. The "doomsday" date T effectively makes the difference. But if we have no doomsday, if we have an infinite horizon, then both $V_t(k_t)$ and $V_{t+1}(k_{t+1})$ have the same number of future time periods ahead of them - namely, infinity. Thus, if they face the same "future", in a sense, they ought to have the same "value" function V(.). Thus, V(k) and V(k′) are different only in the arguments (k and k′) and *not* in the functional form (V(.)).

(B) Why does Bellman's Equation Exist?

Under what conditions is this Bellman's Equation, this value function with unique V(.) for any period, possible? Let us consider this more carefully and in a more general case. Let us take the now-familiar general notation of Kamien and Schwartz (1991). Suppose we have state variable x_t and control variable u_t (they can be vectors). Suppose we have "return function" $F(x_t, u_t, t)$ and a "transition function" $x_{t+1} = g(x_t, u_t, t)$. Thus, we seek to maximize the following:

$$\max \sum_{t=0}^{\infty} F(x_t, u_t, t)$$

$$\text{s.t.} \quad x_{t+1} = g(x_t, u_t, t)$$
$$x_0 \text{ given.}$$

Thus, in the previous consumption-savings example, the utility function was our "return" function and the asset-accumulation equation was our "transition" function with consumption as control and asset as state. In the investment example, profit was our return function and capital-accumulation our transition function with investment as control and capital as state.

Assumption (1): the functions are time-autonomous, i.e. the return function and the transition function are independent of time, thus we can rewrite them as $F(x_t, u_t)$ and $x_{t+1} = g(x_t, u_t)$.

This was true for U and *f* in the Solowian model, the consumption model and the investment model. They may not be so in general cases, but to streamline our problems, this must be assumed. At this point, this is not too damning - but when we consider stochastic problems, it causes some difficulties.

Assumption (2): $F(x_t, u_t)$ is concave and bounded.

Assumption (3): the set $\Phi = \{x_{t+1}, x_t, u_t \mid x_{t+1} \leq g(x_t, u_t)\}$ is convex and compact for admissable values of u_t.

These assumptions yield the following results: namely, we can specify a functional equation:

$$V(x) = \max_u \{F(x, u) + \beta V(x')\} \qquad \text{s.t. } x' = g(x, u), \; x_0 \text{ given.}$$

and it will be the case that:

(1) (Existence and Uniqueness) the functional equation has a *unique* concave solution, V(x).

(2) (Recursive Limit) this solution V(x) is approached in the limit as $t \rightarrow \infty$ by iterations on the time-specific value function $V_t(x) = \max_u \{F(x, u) + \beta V_{t+1}(x')\}$.

(3) (Principle of Optimality) For $x = x_0$, the V(.) that solves the functional equation is the same as the V(.) that solves the intertemporal problem, i.e.

$$V(x_0) = \max \sum_{t-0}^{\infty} \beta^t F(x_t, u_t)$$
$$\text{s.t. } x_{t+1} = g(x_t, u_t)$$
$$x_0 \text{ given}$$

(4) (Policy Function) there is a unique and time-invariant "policy function" $u = h(x) = \text{argmax } [F(x, u) + \beta V(x')]$.

(5) (Benveniste-Scheinkman) off corners, the limiting value function V is differentiable with:

$$dV(x)/dx = dF(x, h(x))/dx + \beta[dV/dg]\cdot[dg(x, h(x))/dx]$$

The conditions stated are sufficient (but not necessary) conditions to yield the propositions (1)-(5) we have here. More general cases can be found. Economically, the main point is that the solution to the Bellman's equation, V(x), is the same as that of the intertemporal optimization problem (Result 3) and that we obtain a unique "policy function" u = h(x) which we can then use to map out the "optimal path" (Result (4)). Mathematically, the main point is that a unique, time-invariant V(.) exists (Result 1) and that it is the limit of a recursive sequence of time-dependent value functions (Result 2). Result (5), the Benveniste-Scheinkman condition, is, as we have already seen, highly useful for applications.

The way this mathematical set of results ((1) and (2)) is proved is via a contraction mapping. Namely, as $V_t(x)$ is defined for particular time and value functions enter recursively into each other, we obtain a "transition function" from one value function to another which, as is demonstrated in Stokey and Lucas (1992: p.79, Th. 4.6), yields a contraction mapping and thus a fixed point (see Appendix I). More specifically, we had, in the time-dependent case:

$$V_t(x) = \max_u \{F(x, u) + \beta V_{t+1}(x')\} \qquad \text{(VF)}$$

where $u \in R^n$ (as we may have n controls) and $x' \leq g(x, u)$, $x \in X$ and $\beta \in (0, 1)$. We can define the operator T as $V_t = TV_{t+1}$. In this case:

$$TV_{t+1}(x) = \max_u \{F(x, u) + \beta V_{t+1}(x')\}$$

but note that the two value functions, TV_{t+1} and V_{t+1}, are still different from each other. If $F(x, u)$ is real valued, continuous, concave and bounded and the set $\{x', x, u \mid x' \leq g(x, u), u \in R^n\}$ is convex and compact, and $V_{t+1} \in C(X)$ (where C(X), the set of continuous bounded functions on X, is a complete normed vector space equipped with a sup norm), then, $T: C(X) \rightarrow C(X)$ is mapping a continuous bounded function to a continuous bounded function. It is easy to note that T is monotone and satisfies discounting, i.e. if $V_{t+1}(x) \geq W(x)$ for all $x \in X$, then:

$$TV_{t+1}(x) = \max_u \{F(x, u) + \beta V_{t+1}(x')\} \geq \max_u \{F(x, u) + \beta W(x')\} = TW(x)$$

thus monotonicity is fulfilled. Similarly, for any scalar α:

$$T(V_{t+1}+\alpha)(x) = \max_u \{F(x, u) + \beta[V_{t+1}(x') + \alpha]\}$$

$$= \max_u \{F(x, u) + \beta V_{t+1}(x') + \beta\alpha\} = TV_{t+1}(x) + \beta\alpha$$

thus discounting is fulfilled. By Blackwell's Sufficiency Criteria, a fixed point exists, i.e. there is a unique $V_{t+1}^* \in C(X)$ such that $TV_{t+1}^* = V_{t+1}^*$. This fixed-point is approached by iterations $V_{t+1}^k = T^k(V^0)$ where V^0 is some initial bounded, continuous function - which, in this space, implies uniform convergence of the functions V_{t+1}^k. (also, if C(X) is restricted to concave functions, then if TV_{t+1} and V_{t+1} are concave functions and so is V_{t+1}^*).

What does this imply? Recall that the fixed point claims that $V_{t+1}^* = TV_{t+1}^*$. But, we also know that $TV_{t+1} = V_t$ by our definition of T. Thus, $V_{t+1}^* = V_t^*$ - which will be true for all t. Thus, dropping the asterisk, we have a *unique* value function $V = V_t = V_{t+1}$ for all t, which implies:

$$V(x) = \max_u \{F(x, u) + \beta V(x')\}$$

which is our Bellman's equation. The equivalence of the solution V(x) of this Bellman's equation to the solution to the intertemporal problem - Richard Bellman's "Principle of Optimality" - we shall pass over in silence and refer to Appendix A. We already saw the equivalence for the finite case; the infinite case is really a mere extension of that analogy. The existence of a unique policy function u = h(x) = argmax V(x) is coincidental to the proof of a unique V.

(C) Euler Once Again

We actually already know how to employ Bellman's equation to solve intertemporal problems - for the general solution method has already been outlined earlier when we obtained the Euler Equation from the time-dependent value function. This is not a general solution, nor even a solution properly speaking, but it drops out of the Bellman's Equation for many of the problems we deal with in economics - and, even if we cannot derive an optimal path for controls and states (the ultimate objective), we get some idea of their properties.

The method of obtaining the Euler from the value functions has already been outlined earlier. Now, we do it once again, but this time for the Bellman's and only in general form. Consider the consumption-savings problem outlined earlier, which yields the Bellman's Equation:

$$V(x) = \max\{u(c) + \beta V(x')\} \qquad \text{(BE)}$$

where $x' = R[x - c]$.

Steps:

(1) plug the difference equation $x' = R[x - c]$ in instead of x'.

$$V(x) = \max\{u(c) + \beta V(R[x - c])\} \qquad \text{(BE)}$$

(rules of thumb: in principle, either c or x′ is the control; x is always the state; whenever you can, manipulate the equations you have so that your control variables appear in the V(x) equation as often as possible).

(2) get the first order condition (FOC) by taking the derivative with respect to c (or whatever the control variable is) and then setting to zero. (- if there is more than one control variables at one time, then you must take a FOC for each control)

$$dV/dc = Uc - \beta(dV/dx')(dx'/dc) = U_C - \beta R(dV/dx') = 0 \qquad \text{(FOC)}$$

(3) get the Benveniste-Scheinkman (BS) condition by taking the derivative with respect to the state variable x (but don't set this to zero!) (- if there is more than one state variable at one time, then you must take a BS condition for each state)

$$dV/dx = \beta(dV/dx')(dx'/dx) = \beta R(dV/dx') \qquad \text{(BS)}$$

(4) Plug the FOC into the BS and get an expression for dV/dx

$$dV/dx = U_c$$

(5) Iterate the FOC one period forward (i.e. prime everything once).

$$dV/dx' = U_{C'}$$

(6) Plug the iterated term dV/dx′ into the FOC to obtain the Euler Equation (EE).

$$Uc = \beta R U_{C'} \qquad \text{(EE)}$$

and that's all there is to it!

Actually, that's far easier said than done. It is not only difficult to get the Euler Equation in many cases, but in most cases the Euler is not enough. The objective is to draw out the optimal path - usually by finding the optimal policy function, c = h(x), and this involves often little more than educated guesses. An optimal policy function allows us to possibly obtain the optimal path $\{c_t, k_t\}_{t=0}^{\infty}$ analytically, but, far more often, a computer and an algorithm is absolutely necessary.

In general cases where V(x) = max [F(x, u) + βV(x′)] s.t. x′ = g(x, u), the steps to obtaining an Euler Equation are essentially identical.

(1) The first thing is to identify the states and controls; a usual rule of thumb is to look at the transition function, see which are states and controls. As noted, there are two ways of writing Bellman's - with x′ as controls or with u as controls, so:

$$V(x) = \max [F(x, u) + \beta V(g(x, u))] \qquad \text{(BE with u)}$$

is a BE which uses u as control, whereas, if you can invery so that $u = g^{-1}(x, x')$

$$V(x) = \max [F(x, g^{-1}(x, x') + \beta V(x')] \qquad \text{(BE with x')}$$

is a BE which uses x′ as control. As noted, always make your controls appear "everywhere".

(don't confuse states with controls! If you get confused, look at the transition equation and see which variables are iterated forwards (or backwards), then, take the "present" version of that term as your state; thus, examining x′ = g(x, u), we see immediately that x′ is iterated forward, so the present version of it, x, is the state).

(2) get the first order condition (FOC) by taking the derivative with respect to controls and then setting to zero.

- this is standard differentiation. If you wrote your Bellman's well so that the control variables appear everywhere they can, this should be a cakewalk. In general:

$$dV/du = dF/du + \beta[(dV/dx')\cdot(dx'/du)] = 0 \qquad \text{(FOC)}$$

where dV/dx' is, of course, $dV/dg(.)$ and dx'/du is merely $dg(.)/du$.

(3) get the Benveniste-Scheinkman (BS) condition by taking the derivative with respect to the state variables x

- the Benveniste-Scheinkman condition is supposed to be done *after* you find the "optimal" controls as functions of the states, $u = h(x)$, which makes differentiation look harder. But it actually isn't because a LOT of terms will disappear by the "envelope theorem". To illustrate this, differentiate the entire Bellman's with respect to x (allowing for the fact that u may be a function of x):

$$dV/dx = dF/dx + (dF/du)\cdot(du/dx) + \beta[(dV/dx')\cdot(dx'/dx) + (dV/dx')\cdot(dx'/du)\cdot(du/dx)]$$

where du/dx is the derivative of $u = h(x)$. Lots of terms, lots of chain rules, right? Well, usually the "envelope theorem" allows you to set *all* terms $\{du/dx\}$ to zero! So this becomes much reduced to:

$$dV/dx = dF/dx + \beta[(dV/dx')\cdot(dx'/dx)] \qquad \text{(BS)}$$

Much better! In many problems, *often* (but not always), F(.) is independent of x, so $dF/dx = 0$. If you're this lucky, we have an even simpler Benveniste-Scheinkman condition.

(4) Plug the FOC into the BS condition:

- specifically, you usually can connect $\beta(dV/dx']$ in the two equations. (you may also have to invert the rightmost side of the FOC, so $-(dF/du)(du/dx') = \beta[dV/dx']$ would be the expression to be inserted). Plug that into BS:

$$dV/dx = dF/dx - (dF/du)(du/dx')\cdot(dx'/dx)$$

or, recalling that $x' = g(x, u)$, then we rewrite this for the next step:

$$dV/dx = dF/dx - (dF/du)(du/dg(x,u)\cdot(dg(x, u)/dx)$$

(5) Iterate this one period forward (i.e. prime everything once).

$$dV/dx' = dF/dx' - (dF/du')(du'/dg(x',u')\cdot(dg(x', u')/dx')$$

It looks ugly here - but in actual problems it looks *much* simpler.

(6) Plug the iterated term back into FOC to obtain the Euler Equation (EE).

As x' = g(.) by the transition function, then dV/dx' = dV/dg(.). So substitute the entire BS2 for the the term dV/dg(.) in the original FOC.

$$dF/du = - \beta(dx'/du)[dF/dx' - (dF/du')(du'/dg(x',u')\cdot(dg(x', u')/dx')] \qquad \text{(EE)}$$

So, note, there are *no* "V" terms inside the equation - so we need to know nothing about the form of the value function. This is the Euler Equation - very ugly in its present form, but guaranteed in concrete applications to be *much* simpler.

And that's all folks! - at least for the Euler...The solutions, as noted earlier, are more complicated.

(D) Setting up the Bellman's: Some examples

Here go a few examples of how to set up Bellman's equations:

(i) Labor Supply/Household Production model

A worker's instantanteous utility depends on the amount of market-produced commodities consumed c_{1t} and home-produced goods (e.g. leisure) c_{2t}. In order to acquire market produced goods, worker must allocate some time l_{1t} to market activities that earn a salary w_t. Market wage evolves according to $w_{t+1} = h(w_t)$. Quantity of home-produced goods depends on stock of expertise at beginning of period, which we label a_t, via a home production function $f(.)$. We assume this stock depreciates at rate δ and can be increased by allocated time to non-market activities. l^ is the maximimum amount of labor time available. $f(.)$ and u(.) are bounded and continuous and $a_0 > 0$ is given.

The problem to be maximized is:

$$\max \sum_{t=0}^{\infty} \beta^t u(c_{1t}, c_{2t}) \qquad 0 < \beta < 1$$

$$\begin{aligned} \text{s.t.} \quad & c_{1t} \leq w_t l_{1t} \\ & c_{2t} \leq f(a_t) \\ & a_{t+1} \leq (1-\delta)a_t + l_{2t} \\ & l_{1t} + l_{2t} \leq l\hat{} \\ & w_{t+1} = h(w_t) \\ & a_0 > 0 \text{ given.} \end{aligned}$$

The state variables are (a, w) and the controls are (c_1, c_2, l_1, l_2). Thus, the Bellman's equation is:

$$V(a, w) = \max_{c_1, c_2, l_1, l_2} \{u(c_1, c_2) + \beta V(a', w')\}$$

subject to $c_1 \leq w_t l_1$, $c_2 \leq f(a)$, $a' \leq (1\text{-}\delta)a + l_2$, $l_1 + l_2 \leq l^\wedge$, $w' = h(w)$. The solution follows standard means (FOC for each control, BS for each state, etc.).

(ii) Investment with Adjustment Costs

Firm maximizes present value of cash flow with future earnings discounted at β. Income at t is given by sales revenue, $p_t q_t$, where p_t is price and q_t is quantity produced. Firm takes prices as given and prices evolve according to law of motion $p_{t+1} = f(p_t)$. Total production depends on the amounts of capital, k_t, labor n_t and on the square of the difference between current ratio of sales to investment and the previous-period's ratio. Assume that wage rate w is constant. Capital depreciates at rate δ.

The firm's problem is:

$$\max \sum_{t=0}^{\infty} \beta^t [p_t q_t - w n_t] \qquad 0 < \beta < 1$$

$$\text{s.t. } q_t + x_t \leq g[k_t, n_t, (q_t/x_t - q_{t-1}/x_{t-1})^2]$$
$$k_{t+1} \leq (1\text{-}\delta)k_t + x_t \qquad 0 < \delta < 1$$
$$p_{t+1} = f(p_t)$$
$$q_{-1}/x_{-1} \text{ given.}$$
$$k_0 > 0 \text{ given.}$$

g(.) is bounded and increasing in the first two arguments and decreasing in the third.

The state variables are k, p and the *ratio* q/x which we label z (thus z = q/x). Note that as w is constant, it is not a state variable - even though it affects returns. Controls are sales q, investment x and employment n. The Bellman's equation, then, is:

$$V(k, p, z) = \max_{q, x, n} \{pq - wn + \beta V(k', p', z')\}$$

subject to $q + x \leq g[k, n, q^2/(x - z)^2]$, $z' = q/x$, $k' = (1\text{-}\delta)k + x$ and $p' = f(p)$ - where the last three are the laws of motion for the state variables. Notice then, in this problem, the tricky use of the ratio q/x to define a *new* state variable and a *new* law of motion. Such tricks are commonly-used short-cuts to solving dynamic programming problems.

(3) Solving Bellman's Equation for Policy Functions

In dynamic programming, "solving" things means something specific. Recall our general problem:

$$V(x_0) = \max \sum_{t=0}^{\infty} \beta^t F(x_t, u_t)$$

$$\text{s.t.} \quad x_{t+1} = g(x_t, u_t)$$
$$x_0 \text{ given}$$

Now, our previous result, we know the solution to this problem is equivalent to the solution to the Bellman's Equation:

$$V(x) = \max_u \{F(x, u) + \beta V(x')\}$$

where $x' = g(x, u)$. We still have not satisfied ourselves that the control path $\{u_t\}^{\infty}_{t=0}$ is obtainable from the Bellman's. But it is. Namely, what we wish to obtain is a time-invariant "policy" function $u = h(x)$ so that, given x, we can choose a specific u. Where is this? Well, if V(x) truly solves the Bellman's, then the optimal control variables $u = h(x)$ ought to be the variables that maximize it, i.e. $u = h(x) = \text{argmax}\{F(x, u) + \beta V(x')\}$. Thus:

$$V(x) = F(x, h(x)) + \beta V(g(x, h(x)))$$

where we have plugged in $g(x, h(x))$ for x' and $h(x)$ for u. Thus, V(x) and h(x) are solved "jointly". Thus, if we can specify this policy function, we are home because given $u_t = h(x_t)$ and given $x' = g(x, u)$, we can obtain the optimal path for the controls $\{u_t\}^{\infty}_{t=0}$ and the states $\{x_t\}^{\infty}_{t=0}$.

How do we get this policy function from the Bellman's? You simply can't. Or, at least not generally. Specifically, there are only two recognized methods of going about it (a third method - "Howard's Improvement Algorithm" is omitted but extensively discussed in Sargent (1997)):

Method 1: Computer: iterate value function in a recursive fashion until we obtain the path.

Method 2: Guess-And-Verify. There are two types of "Guess and Verify" methods:

(a) Given the Euler Equation, guess the form of the policy function $u = h(x)$ (controls as functions of states) and then verify that this indeed holds true.

(b) Guess a specific form for the value function, verify that it solves the Bellman's equation, i.e. that $V = \max_u \{F(x, u) + \beta V(x')\}$ and then derive the form of $u = h(x)$.

Method 1 is very complicated and rarely works to our satisfaction. You need to set up an algorithm and simulate it computatonally and then only maybe a solution will be found in a decent amount of time.

Method 2 involves some involved steps - but note the following warning: t*he guess-and-verify method has proven to work in only two cases:*

(a) when the return function F(x, u) is quadratic and the constraints g(x, u) are linear.

(b) the return function F(x, u) is logarithmic and the constraints g(x, u) are Cobb-Douglas.

Any other form of specification of return functions and constraints will usually *not* work by the "guess and verify method" - in which case, one really has to go turn on that computer, write up some iterative algorithm and pray.

[Now, ladies and gentlemen of the jury, you may begin to realize why so many macroeconomists use specific functional forms for utility (usually quadratic or logarithmic) and production functions (usually linear or Cobb-Douglas). These are the only forms which yield tractable analytical solutions. Applying any other form requires some heavy-duty computational work. Apparently, in macroeconomics, it is practicality and not truth that rules.]

(A) Guess and Verify Method: The Idea

As noted earlier, there are two guess-and-verify methods. We shall give the basic idea of each using the consumption problem, so $V(x) = \max_c\{u(c) + \beta V(x')\}$ where $x' = R[x - c]$. Thus, our controls are c and our states are x:

(i) Guessing the policy function u = h(x) (in our case, c = h(x)) should only be done if you get as far as the Euler Equation. The steps are than as follows:

(1) Guess a shape c* = h(x)
(2) If you got an Euler from your Bellman's, plug h(x) in the Euler Equation so dU/dc = Rβ(dU/dc') becomes:

$$(dU/dh(x)) = R\beta(dU/dh(x'))$$

(3) use the transition equation x' = R[x - c] and plug that in for x':

$$(dU/dh(x)) = R\beta(dU/dh(R[x- c])$$

(4) As c shows up again here, then plug the policy function *again* for c* = h(x):

$$(dU/dh(x)) = R\beta(dU/dh(R[x- h(x)])$$

(5) Rearrange the equation so that you obtain c* as a function of x in the shape you originally proposed.

(ii) Guessing the Value Function form V(x).

Recall that the value function is always a function of the states *only*. A common example is the "optimal linear regulator" problem, when we guess a linear function for V, e.g.

$$V = E + Fx$$

where E and F are vectors of parameters and x is the vector of parameters. So, again for our particular case:

(1) Guess the form of V(x) (e.g. V = E + Fx)

(2) Equate this with the Bellman's Equation:

$$E + Fx = \max\{u(c) + \beta V(x')\}$$

(3) We have V(x′) inside the Bellman's, so plug in the guessed form again:

$$E + Fx = \max\{u(c) + ß(E + Fx')\}$$

(4) Derive the first-order conditions:

$$d(E + Fx)/dc = ... = 0$$

(5) Solve for optimal c* (recall c* = h(x))

(6) Plug c^* back into step (3) equation:

$$E + Fx = u(c^*) + ß(E + Fx')$$

(7) Simplify (but *do not* try to solve for c^* as a function of x, just do a lot factoring wherever you can.)

(8) If this works, you *should* have a coefficient attached to the x terms. These should be your term F. The rest is your term E. Thus you guessed "correctly".

(9) Plug back E and F into the equation for $c^* = h(x)$ and simplify. The resulting $c^* = h(x)$ is your solution.

Does this all sounds fishy and obscure? Well, the best way to deal with it is to take an example of each.

(B) Guess and Verify Method: Two Examples

(i) *One Example: Log Return and Cobb-Douglas Transition*

We work through a simple example of the Solowian model. Let us take log-preferences and Cobb-Douglas production function, so $u(c_t) = \ln c_t$ and $f(k_t) = Ak_t^{\alpha}$. So we now have:

$$\max \sum_{t=0}^{\infty} \beta^t \ln c_t$$

s.t.

$$k_{t+1} = Ak_t^{\alpha} - c_t$$

which yields a Bellman's Equation:

$$V(k) = \max_{c} \{\ln c + \beta V(k')\} \qquad \text{s.t. } k' = Ak^{\alpha} - c$$

where, please note again, that as we are in infinite time, the time subscripts have been dropped and "prime" denotes future. Now, remember that we can choose either c or k' as the control variable - and we opt for k′. Thus, we must plug in our transition function for c, i.e. $c = Ak^{\alpha} - k'$, so the Bellman's becomes:

$$V(k) = \max_{k'} \{\ln[Ak^{\alpha} - k'] + \beta V(k')\}$$

So, now, the "solution" is really a sequence of k'*, or k' = h(k) (controls as functions of states) (to express for u, just run it through the transition function). Let us use the guess-and-verify method for the value function and denote the steps concurrently with those outlined earlier. Thus, (Step 1), we guess the value function has the form:

$$V(k) = E + F(\ln k)$$

So equating (Step 2) this to the Bellman's:

$$E + F(\ln k) = \max\{\ln[Ak^{\alpha} - k'] + \beta V(k')\}$$

and realizing that V(k') = E + F(ln k'), then plugging that in (Step 3):

$$V(k') = E + F(\ln k) = \max\{\ln[Ak^{\alpha} - k'] + \beta[E + F(\ln k')]\}$$

Deriving FOC with respect to k' (Step 4):

$$dV/dk' = -1/[Ak^{\alpha} - k']) + \beta F/k' = 0$$

or simply:

$$k' = \beta F(Ak^{\alpha} - k')$$

factoring out the terms k':

$$(1 + \beta F)k' = \beta F(Ak^{\alpha})$$

so dividing through by $(1+\beta F)$ we obtain:

$$k' = \beta F(Ak^{\alpha})/(1+\beta F)$$

That was Step 5. In effect, we have solved for $k'^* = h(k)$ - although we still have unknown coefficients E and F in there which we must find. We must also verify that the form originally proposed for the value function holds true. Thus, now, in step 6, we must plug this back into the Bellman's, (so the max drops out):

$$E + F(\ln k) = \ln[Ak^{\alpha} - \beta FAk^{\alpha}/(1+\beta F)] + \beta[E + F(\ln (\beta FAk^{\alpha}/(1+\beta F))]$$

Forget about solving this for anything! All one needs to do is rearrange it, e.g. put the E's and F's together, do some log derivations. Working through ugly algebra, one will obtain the following (which is step 7):

$$E + F(\ln k) = [\ln A(1-\alpha\beta) + (\beta\alpha/(1-\beta\alpha))(\ln A\beta\alpha)]/[1-\beta] + [\alpha/(1-\alpha\beta)](\ln k)$$

so that there are no F or E terms on the right-hand-side and we have factored for ln k. This then implies that:

$$E = [\ln A(1-\alpha\beta) + (\beta\alpha/(1-\beta\alpha))(\ln A\beta\alpha)]/[1-\beta]$$

$$F = [\alpha/(1-\alpha\beta)]$$

which proves that our guess was right: the value function *does* take the form $E + F(\ln k)$. This is step 8. So, having identified E and F, we plug these back into the equation we had in Step 5, $k' = \beta F(Ak^{\alpha})/(1+\beta F)$, to obtain:

$$k' = \beta\, [\alpha/(1-\alpha\beta)](Ak^{\alpha})/(1+\beta[\alpha/(1-\alpha\beta)])$$

then (Step 9) simplifying this, we obtain:

$$k' = A\beta\alpha k^{\alpha}$$

et voila! - this is our explicit policy function, expressing the control as a function of the state ($k'^* = h(k)$) thus we can put it through the transition function and obtain the optimal u* path for an initial k_0.

With all the manipulating of these equations algebraically, we can see immediately that with more complicated functions which are not log/Cobb-Douglas combinations, we will have tremendous amounts of trouble using the guess-and-verify method - if we can at all. So, for those, simulation and computational algorithms are really the only feasible way of doing them.

(ii) *The Other Example: Quadratic Return and Linear Transition*

We said that the quadratic/linear combination was also feasible. The example for this is derived from Sargent and is called the "Optimal Linear Regulator Problem". Here we have quadratic return function in the form $F(x, u) = x'Rx + u'Qu$ (note: the primes in this term denote transposes of vectors and not "future" vectors) and linear transition function, $x' = g(x, u) = Ax + Bu$. Or:

$$\max \sum_{t=0}^{\infty} [x_t'Rx_t + u_t'Qu_t]$$

s.t.

$$x_{t+1} = Ax_t + Bu_t$$

Note that we are *not* discounting, so β doesn't appear in the objective. A, B, R and Q are matrices of coefficients. Of course, $x'Rx$ is a quadratic function whereas Ax is a linear function (note: x and u are vectors). The Bellman's equation for this case is, generally:

$$V(x) = \max\{F(x, u) + V(x')\}$$

as there is no discounting - but Sargent demonstrates that the Bellman's conditions are still fulfilled. Let us take u as our control. Thus, $x' = g(x, u)$ is plugged in place of x'. Thus, using our terms:

$$V(x) = \max\{x'Rx + u'Qu + V(Ax + Bu)\}$$

So let's go through the guess-and-verify method for the value function. In Step 1, Sargent guesses the value function is a quadratic function, i.e.$V(x) = x'Px$ - which is a quadratic function. So, going through Step 2:

$$V(x) = x'Px = \max\{x'Rx + u'Qu + V(Ax + Bu)\}$$

But now we have $V(Ax + Bu)$. So, if indeed $V(x) = x'Px$, then $V(Ax + Bu)$ must be $(Ax + Bu)'P(Ax + Bu)$, so as in Step 3:

$$V(x) = x'Px = \max\{x'Rx + u'Qu + (Ax + Bu)'P(Ax + Bu)\}$$

We must now go to Step 4 - taking first order conditions. Here it gets a bit ugly - check a linear algebra or calculus book if you want to go through derivatives of matrices. We obtain, as a result:

$$dV/du = Qu + B'PBu + B'PAx = 0$$

or:

$$(Q + B'PB)u = - B'PAx$$

which, solving for u:

$$u = - (Q + B'PB)^{-1}B'PAx$$

(assuming it's invertible). So calling $F = (Q + B'PB)^{-1}B'PA$, we get:

$$u = - Fx$$

which is our step 5. So, going back to the Bellman's (Step 6), we plug this in for u (and the max drops out because it's already "optimal"):

$$V(x) = x'Rx + (-Fx)'Q(Fx) + (Ax + B(-Fx))'P(Ax + B(-Fx))\}$$

so there are no u's on the right hand side. So, substituting our terms for F and rearranging, you should get:

$$V(x) = x'Px = x'[R + A'PA - A'PB(Q + B'PB)^{-1}B'PA]x$$

which is another quadratic form. But we got Ps on the right hand side and on the left hand side. So we are not finished with steps 8 and 9. We must make P a function only of A, B, R and Q and not a function of itself. Sargent suggests several rather complicated ways of doing this - partly by reducing this to a "Riccatti Equation" - which we don't have the patience to go through. Hopefully, the point is made: P can be found and the form is right.

PART II

(4) Stochastic Dynamic Programming

Stochastic Dynamic Programming is a more important issue. In general, the Bellman's was written:

$$V(x) = \max\{F(x, u) + \beta V(x')\} \qquad \text{(BE)}$$

where $x' = g(x, u)$. But x' is in the future. What if we don't know what the actual result of x' is? Or rather, what if there are stochastic shocks to the future. Consider the following standard Solowian stochastic growth example:

$$\max E_0[\textstyle\sum^{\infty}_{t=0} \beta^t u(c_t)]$$

s.t.

$$k_{t+1} = f(k_t) - c_t + \varepsilon_{t+1}$$

$$\varepsilon_{t+1} \sim N(0, \sigma^2)$$

where there have been three modifications. Firstly, we are no longer maximizing a utility stream but rather maximizing an "expected" utility stream. This is because, as we see in the constraint, k_{t+1} is no longer a certain variable - there is a shock ε_{t+1} at time t+1 which may make the resulting k_{t+1} bigger or smaller than the k_{t+1} we expected when we chose c_t. We have assumed that ε_{t+1} is normally distributed with zero mean and constant variance. Note that we can assume any structure for the shock. It can be white noise (i.e. independently and identically distributed with zero mean and constant variance) or can follow a random walk, $\varepsilon_{t+1} = \varepsilon_t + u_t$ where u_t is white noise - whatever we want.

The issue is that when previously we chose the control c_t, we always assumed we knew what the result that followed (the subsequent state k_{t+1}) and thus the next c_{t+1} could be chosen. But now we are no longer certain what the subsequent state is so we are no longer certain what choosing a particular control will do in terms of affecting the future. However, we can still *solve* it, but the solution will be different.

(A) Some Basics

In general, the stochastic problem can be stated as:

$$\max E_0[\textstyle\sum^{\infty}_{t=0} \beta^t F(u_t, x_t)] \qquad 0 < \beta < 1$$

s.t.

$$x_{t+1} = g(x_t, u_t, \varepsilon_{t+1})$$
$$\text{Prob}\{\varepsilon_t \leq e\} = F(e) \text{ for all } t$$
$$x_0 \text{ given}$$

where F(e) is a cumulative probability distribution (not to be confused with the return function!). Thus, the Bellman's equation is now written as:

$$V(x) = \max\{F(x, u) + \beta E[V(x')| x]\} \qquad \text{(SBE)}$$

where $E[V(x')|x]$ is the conditional expectation of the future value $V(x')$, i.e. the expected future value function $V(x')$ *given* that we are in x at present. Thus:

$$E[V[g(x, u, \varepsilon)] \mid x] = \int V[g(x, u, \varepsilon)]dF(\varepsilon)$$

The proof that a unique V exists to justify our use of a Bellman's in the stochastic case is provided, in an analagous manner, to the certainty case by Lucas and Stokey (1989: Ch. 9).

The practical method of solution remains effectively as in the certainty case. Namely, given BE, we take FOC:

$$dV/du = dF/du + \beta E[(dV/dx')(dx'/du) \mid x] = 0 \qquad \text{(FOC)}$$

where, note, we maintain the conditional expectation. The Benveniste-Scheinkman condition, evaluated at the internal optimum $u^* = h(x)$, can be taken with respect to the state variables:

$$dV/dx = dF/dx + \beta E[(dV/dx')(dx'/dx) \mid x] = 0 \qquad \text{(BS)}$$

which we combine with the FOC to obtain a Euler Equation:

$$dF/du + \beta E[(dF[x', u']/dx)(dg(.)/du) \mid x] = 0 \qquad \text{(EE)}$$

of this sort.

Stokey and Lucas's notation is different from Sargent's. They seem obsessed with expressing the conditional expectation term in Markov transition matrix form. In this case, they like to write stochastic Bellman's equations as:

$$V(x, z) = \max_{u}\{F(x, u, z) + \beta \int_Z V(x') Q(z, dz')\} \qquad \text{(SBE)}$$

s.t. $x' = g(x, u, z)$ is their representation of the stochastic Bellman's equation (SBE). Note that the random variable z is a *state* variable and enters F(.) and g(.). Consequently, it is denoted in the initial list of states in the value function V(x, z).

A word or two ought to be spent on the term $Q(z, dz')$. This is merely a probability measure. Consider the measurable space $(S, \Im)$ where S is the set of states and $\Im$ is the set of "events" (i.e.set of subsets of S) and it is presumed that $\Im$ forms a σ-algebra on S. Thus, an element $A \in \Im$ is also a subset of S (a set of states). When a measurable space is equipped with a "measure", i.e. a function μ that assigns to every event (i.e. $A \in \Im$) a particular real number, then we have what is called a "measure space", $(S, \Im, \mu)$. The only measure we care about in this context is the probability measure, i.e. the one which assigns "probabilities" (numbers) to events. Thus, $\mu(A)$ is the "probability of A happening". Or, alternatively, μ is merely a "distribution."

A "random variable", as the old saying goes, is neither random nor a variable but rather it is a "measurable function" $f: S \rightarrow R$. A "measurable" function is the measure-theoretic analog of a "continuous" function. Recall that we say a function $f: X \rightarrow Y$ is "continuous" if for any open set B in Y, the inverse $f^{-1}(B)$ is itself an open set in X. Similarly, a function $f: X \rightarrow Y$ is "measurable" if for any set B in the σ-algebra of Y, the inverse $f^{-1}(B)$ is itself in the σ-algebra of X. Thus, a "random variable" $f: S \rightarrow R$ is measurable takes sets in the "Borel σ-algebra" of R and the inverse of that is in the σ-algebra of S.

Passing over a lot of intermediate material, and merely write down the "integral" of a measurable function as:

$$\int f\, d\mu = \sup\{\Sigma_i(\inf_{s \in A_i} f(s)\}\mu(A_i)\}$$

where the supremum is taken over all finite partition of S into sets $A_i \in \Im$. If the supremum does not exist, the integral has value $+\infty$. If f is measurable function to R and is measurable with respect to the Borel σ-algebra of R, then this integral is the "Lebesgue integral".

With this in hand, then we can define:

$$E(f) = \int_S f(s)\, \mu(ds) = \int_S f\, d\mu(S)$$

as the expected value of the random variable f. If you don't like the notation, think of $X(\omega)$ as a random variable and P as a probability measure and define:

$$E(X) = \int_\Omega X(\omega)\, dP(\omega)$$

as the expected value of X. If you still don't like that, then write:

$$E(X) = \int_{-\infty}^{+\infty} x\, dF_x(x)$$

where F(x) is a cumulative probability distribution. All these representations are effectively analagous. It does not really matter. The only thing worth mentioning are "conditional expectations". These are defined as follows:

$$E(f \mid A) = \int_A f\, d\mu_A$$

Now, let us return to that odd character $Q(z, dz')$. This is a Markov process, i.e. a "probability distribution" which assigns particular probabilities to being in any state in the next period conditional on being in state z in this period, thus Q(a, A) is the probability of being in a state in the set A given that one is already at a. It is a well-defined measure. Thus, if f is a random variable, then a one period-conditional expectation:

$$E(f \mid z) = \int_Z f(z') Q(z, dz')$$

is simply the expected value of f in the next time period *given* that we are in z now and z′ can only take values in Z. Thus, when Stokey and Lucas define the stochastic Bellman's equation as:

$$V(x, z) = \max_{u} \{F(x, u, z) + \beta \int_Z V(x') Q(z, dz')\} \quad \text{(SBE)}$$

the entire term to the right is merely $E[V(x') \mid z]$, the expected future x given that the present stochastic shock was z. Note that it is not, like Sargent's, conditional on x. This is because it is assumed in this notation that stochastic shocks z are distributed independently of x - or, rather, conditional only on the previous shock, z, whereas I suppose Sargent might allow a more general sort of transition. $Q(z, dz')$, thus, is merely the equivalent of the transition function for the random state variable.

(B) Some Examples

(i) *Consumption with Uncertainty*

Consider the following standard stochastic consumption-savings model:

$$\max E_0[\sum_{t=0}^{\infty} \beta^t U(c_t)]$$

$$\text{s.t.}$$
$$x_{t+1} = R_t(x_t - c_t)$$
$$x_0 \text{ given}$$

where R_t is a random variable governed by a one-period (or "first order") Markov process where $\text{prob}\{R_t \leq R^* \mid R_{t-1} = R\} = F(R^*, R)$. Thus, the stochastic Bellman's equation is now:

$$V(x_t, R_{t-1}) = \max \{U(c_t) + \beta E[V(x_{t+1}, R_t)]\}$$

where the state is now the current state, x_t and the previous random shock, R_{t-1}. We keep track of time subcripts in this case because of the multiple periods involved. We would like to define the control, in this case, as $u_t = x_t - c_t$ so that now $c_t = x_t - u_t$ and the transition function is now $x_{t+1} = R_t u_t$. Thus, plugging all this in:

$$V(x_t, R_{t-1}) = \max_{u_t} \{U(x_t - u_t) + \beta E_t[V(R_t u_t, R_t)]\}$$

Thus, with u_t as our control, the FOC is:

$$dV/du_t = (dU/dc_t)(dc_t/du_t) + \beta E_t[(dV/dx_{t+1})(dx_{t+1}/du_t)] = 0$$

as $dc_t/du_t = -1$ and $dx_{t+1}/du_t = R_t$, then this becomes:

$$(dU/dc_t) = \beta R_t E_t[(dV/dx_{t+1})] \qquad \text{(FOC)}$$

Now, we have two states x_t and R_{t-1}, however we only apply the Benveniste-Scheinkman condition to the non-stochastic term x_t. Thus:

$$dV/dx_t = (dU/dc_t)(dc_t/dx_t)$$

or, as $dc/dx_t = 1$, then:

$$dV/dx_t = dU/dc_t \qquad \text{(BS)}$$

Thus, iterating this forward by one period:

$$dV/dx_{t+1} = dU/dc_{t+1}$$

and plugging into the FOC:

$$dU/dc_t = \beta R_t E_t[dU/dc_{t+1}] \qquad \text{(EE)}$$

which is our simple stochastic Euler Equation. If we had specified this a little further, we could guess a policy function $u_t^* = \phi(x_t, R_{t-1})$ and then converted via $u_t^* = x_t - c_t^*$ the consumption policy function, $c_t^* = h(x_t, R_{t-1})$ which we could plug in and verify.

(ii) *Asset Prices*

Consider the following version of the famous Lucas (1978) "asset pricing" model. The set up is as follows: agents own "securities" or "trees" which yield a dividend per period of d_t. It is proposed that the dividend stream $\{d_t\}_{t=0}^{\infty}$ is stochastic, i.e. a Markov process where $\text{prob}\{d_{t+1} \leq d' \mid d_t = d\} = F(d', d)$. Let s_t be the number of securities owned in period t which agents can buy and sell at price p_t. In fact, we can posit that in any period t, an agent sells his entire stock of securities, s_t and buys up a new stock of securities, s_{t+1}. Thus, a consumer thus gains "income" from the dividends yielded by amount of securities he has ($d_t s_t$) and the selling of these securities ($p_t s_t$) and he uses this income to purchase commodities (c_t, which he consumes and whose price is 1) and a new set of securities ($p_t s_{t+1}$). Prices move according to a law of motion $p_t = g(d_t)$, so that prices are related to dividends. Thus, the problem is specified as:

$$\max E_0[\sum^{\infty}_{t=0} \beta^t U(c_t)]$$

s.t.

$c_t + p_t s_{t+1} \leq (p_t + d_t)s_t$

$p_t = g(d_t)$

d_t is governed by first order Markov process $F(d', d)$

s_0 given

Thus, we must now set up the Bellman's equation. We must first specify the control variable: this can be either c (consumption) *or* s′ (future securities). The states are more tricky. We really have three of them - stock of securities entered in the period s, the divided d and the price p. Thus, in general, we can think of the Bellman's equation as:

$$V(s, p, d) = \max_c \{U(c) + \beta\int V(s', p', d') dF(d', d)\}$$

thus, future earnings are random. However, we shall make this more precise by recognizing that, by non-satiation, that $c + ps' = (p+d)s$. Thus, $c = (p + d)s - ps'$. As $p = g(d)$, then $c = (g(d)+d)s - g(d)s'$. Thus, letting $s[g(d) + d]$ be the "state" and s′ the "control", then we have the stochastic Bellman's equation in the form:

$$V(s[g(d) + d]) = \max_{s'} \{U[(g(d) + d)s - g(d)s'] + \beta \int V(s'[g(d') + d'])dF(d', d)\} \qquad \text{(SBE)}$$

For shorthand, we shall occasionally use $z = s[g(d) + d]$. Now, the first order condition for a maximum is then:

$$dV/ds' = (dU/dc)(dc/ds') + \beta\int[(dV/dz')(dz'/ds')]dF(d', d) = 0$$

where, as $dc/ds' = - g(d)$ and $dz'/ds' = (g(d') + d')$, the FOC becomes:

$$(dU/dc)g(d) = \beta\int[(dV/dz')(g(d') + d')]dF(d', d) \qquad \text{(FOC)}$$

We must now derive the Benveniste-Scheinkman condition. Namely:

$$dV/dz = (dU/dc)(dc/dz)$$

or, as $dc/dz = 1$, then:

$$dV/dz = dU/dc \qquad \text{(BS)}$$

iterating one period forward: $dV/dz' = dU/dc'$, we can plug this into the FOC:

$$(dU/dc)g(d) = \beta\int[(dU/dc')(g(d') + d')]dF(d', d)$$

or, splitting up the terms under the integral:

$$(dU/dc)g(d) = \beta\int[(dU/dc')g(d')]dF(d', d) + \beta\int[(dU/dc')d']dF(d', d)$$

defining $w(d) = (dU/dc)g(d)$, then:

$$w(d) = \beta\int w(d')\, dF(d', d) + \beta\int[(dU/dc')d']dF(d', d)$$

Lucas imposes the equilibrium condition that $s' = s = 1$, i.e. there is one security per "person". This implies that in equilibrium, $c = [g(d)- d]s - g(d)s' = d$. Thus, $dU/dc = dU/dd$ in equilibrium. Conseuently:

$$w(d) = \beta\int w(d')\, dF(d', d) + \beta\int[(dU/dd')d']dF(d', d)$$

The process now involves around finding $w(d)$. Because $w(d) = (dU/dc)g(d)$ and dU/dc is known, then once $w(d)$ is determined, we can immediately compute $g(d) = w(d)/(dU/dc)$. Thus, we must solve this equation for $w(d)$. Lucas (1978) does this with standard arguments on the boundedness and concavity of u(.), etc. to prove there is a unique solution to the Bellman's $w(d)$. As Sargent suggests, the resulting solution $w(d)$ implies $g(d)$ as a "fixed point". But recall that $g(d)$ is merely a price function. This might be seen as a "rational expectations" equilibrium as the "fixed point of this mapping froma perceived pricing functions to actual pricing functions" (Sargent, 1987: p.100). We suggest to examine Sargent (1987: Ch. 3) for further examples of asset pricing models.

(iii) *Search Models*

Consider the following search model for jobs. Suppose an unemployed worker looking for a job gets a wage offer w drawn independently from a distribution $F(W) = \text{prob}\ \{w \leq W)$ with $F(0) = 0$, $F(B) = 1$ for some $B < \infty$. She can accept the offer, in which case she receives w per period forever or she can reject the offer - in which case she receives an unemployment compensation payment c for this period and a chance for another wage draw from the same distribution in the next period. Let y_t be the worker's income in a particular period t - which can either be, as we saw, c or w. The worker seeks to maximize:

$$\max E_0[\sum^{\infty}_{t=0} \beta^t U(y_t)]$$

For the sake of simplicity, let $U(y_t) = y_t$. Thus, if the agent accepts a job, her gain is the entire intinite discounted utility stream, i.e. $w/(1-\beta)$. If she rejects, she gets c and the chance of another draw from the distribution $F(W)$. Thus, state variable is the wage offer w, and the choice set in any period is the binary one of "accept" or "reject" a wage offer. In his appendix, Sargent proves that a Bellman's exists for this problem. Consequently, the Bellman's equation can be written:

$$V(w) = \max\ \{w/(1-\beta), c + \beta\int V(w')dF(w')\} \qquad \text{(BE)}$$

where, note, $\int V(w')dF(w')$ denotes the expected wage offer next period - suitably discounted by a single period. Is there a solution V? Sargent (1987: A.8) shows there is by checking that an operator T can be defined so that $TV(w) = \max \{w/(1-\beta), c + \beta \int V(w)\, dF(w)\}$ fulfills the Blackwell's sufficiency critieria and thus is a contraction mapping with a fixed point V(.).

Having gotten that far, we now want to guess a form of solution to the Bellman's equation. Obviously, this problem seems to call for a reservation wage - w_r - and this is indeed what we guess. Namely, we propose that the solution to the value function will be such that the agent's strategy is to determine a reservation w_r such that if $w \geq w_r$, she will acccept, whereas if $0 \leq w < w_r$, she will reject. The reservation wage, then, is determined by the unemployment benefit payment c and the expectated future wage offer, i.e. $w_r/(1-\beta) = c + \beta\int V(w')dF(w')$. us, we guess that the Bellman's equation has a solution of the form:

$$V(w) = \begin{cases} w_r/(1-\beta) = c + \beta\int V(w')dF(w') & \text{if } 0 < w < w_r \\ w/(1-\beta) & \text{if } w \geq w_r \end{cases}$$

Note: this is a guess. To see if it works, we must equate it. Now, to get rid of the interior value function $V(w')$ in the Bellman's, we must impose our structure upon the future draw w' as well. Thus, examine $V(w')$. We should expect $V(w') = w_r/(1-\beta)$ if $w' < w_r$ and $V(w') = w'/(1-\beta)$ if $w' \geq w_r$. Consequently, we can propose that if w_r is indeed the reservation wage, then the following holds true:

$$w_r/(1-\beta) = c + \beta\int V(w')dF(w')$$

or splitting integrals into two areas - from 0 to w_r (where $V(w') = w_r/(1-\beta)$) and from w_r to ∞ (where $V(w') = w'/(1-\beta)$):

$$w_r/(1-\beta) = c + \beta\int_0^{wr} w_r/(1-\beta)\, dF(w') + \beta\int_{wr}^{\infty} w'/(1-\beta)\, dF(w')$$

Splitting the left hand side into two integrals over the same range:

$$\int_0^{wr} w_r/(1-\beta)\, dF(w') + \int_{wr}^{\infty} w_r/(1-\beta)\, dF(w') = c + \beta\int_0^{wr} w_r/(1-\beta)\, dF(w') + \beta\int_{wr}^{\infty} w'/(1-\beta)dF(w')$$

and then factoring out $w_r/(1-\beta)$ from the interiors of the integrals (as w_r is constant):

$$w_r/(1-\beta)\int_0^{wr} dF(w') + w_r/(1-\beta) \int_{wr}^{\infty} dF(w') = c + \beta w_r/(1-\beta) \int_0^{wr} dF(w') + \beta\int_{wr}^{\infty} w'/(1-\beta)\, dF(w')$$

and bringing an integral on the right hand side to the left and from the left to the right so that therms with integrals over equivalent ranges are together:

$$(1-\beta)w_r/(1-\beta)\int_0^{wr} dF(w') = c + \beta\int_{wr}^{\infty} w'/(1-\beta)\, dF(w') + w_r/(1-\beta) \int_{wr}^{\infty} dF(w')$$

or:

$$w_r \int_0^{w_r} dF(w') = c + \int_{w_r}^{\infty} (\beta w' - w_r)/(1-\beta)\, dF(w')$$

or adding $w_r \int_{w_r}^{\infty} dF(w')$ to both sides:

$$w_r = c + \int_{w_r}^{\infty} (\beta w' - w_r)/(1-\beta)\, dF(w') + w_r \int_{w_r}^{\infty} dF(w')$$

or simply:

$$w_r - c = [\beta/(1-\beta)] \int_{w_r}^{\infty} (w' - w_r)\, dF(w')$$

where the left hand side is the marginal cost of rejecting an offer and searching one more time when an offer w_r is on hand while the right hand side is the expected marginal benefit of searching one more time. Sargent then defines the following function:

$$h(w) = [\beta/(1-\beta)] \int_{w_r}^{\infty} (w' - w_r)\, dF(w')$$

which, as he shows, can determined w_r. How? As w_r is determined by the intersection of w-c curve and the h(w) curve. When w - c = h(r), then we have found the reservation wage.

For more details, consult Sargent (1987, Ch.2).

(5) Mathematical Appendices

(A) Principle of Optimality

Richard Bellman's "Principle of Optimality" asserts that a solution to a standard intertemproal programming problem is equivalent to the solution obtained from Bellman's equation. To this end, we need to change our notation a bit to conform to Stokey and Lucas (1989: Ch. 4).

Let x be a state variable and y a control variable.
X = set of admissable values of state variable x
$\Gamma: X \rightarrow X$ = correspondence describing feasible values for state variable next period, i.e. $y \in \Gamma(x)$ is an admissable state variable in the future.
$A = \{(x, y) \in X \times X \mid y \in \Gamma(x)\}$ = graph of Γ
$F: A \rightarrow R$ = return function, $F(x, y)$.
$0 < \beta < 1$ = given discount factor
$\prod(x_0) = (\{x_t\}_{t=0}^{\infty} \mid x_{t+1} \in \Gamma(x_t)\}$ is the set of feasible plans starting from x_0.

The principle of optimality then states the following. Namely, for $x = x_0$, the V(.) that solves the functional equation:

$$V(x) = \max_y \{F(x, y) + \beta V(y)\} \text{ s.t. } y \in \Gamma(x), \text{ for all } x \in X \qquad \text{(BE)}$$

is the same as the V(.) that solves the intertemporal problem, i.e.

$$V(x_0) = \max \sum_{t-0}^{\infty} \beta^t F(x_t, x_{t+1}) \qquad \text{(IP)}$$
$$\text{s.t. } x_{t+1} \in \Gamma(x_t)$$
$$x_0 \in X \text{ given}$$

where, note, $y_t = x_{t+1}$. To this end we need the following sets of assumptions:

Assumptions

(A.1) Return functions are time autonomous, $F(x_t, x_{t+1})$ and not $F(x_t, x_{t+1}, t)$.

(A.2) $\Gamma(x)$ is non-empty for all $x \in X$.

(A.3) For all $x_0 \in X$ and $\{x\} \in \prod(x_0)$, $\lim_{n\rightarrow\infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1})$ exists (although it may be $+\infty$ or $-\infty$).

The second assumption (A.2) guarantees that $\prod(x_0)$ is non-empty. (A.3) guarantees that any feasible plan can be valued by using the objective function (F(.)) and the discount rated β. The following are sufficient conditions for (A.3):

Sufficiency 1: F is bounded above/below and $0 < \beta < 1$.

Sufficiency 2: for each $x_0 \in x$ and $\{x\} \in \prod(x_0)$, there exists $\theta \in (0, 1/\beta)$ and $0 < c < \infty$ such that $F(x_t, x_{t+1}) \leq c\theta^t$ for all t.

If (A.1)-(A.3) hold, then objective function in (IP) is well-defined, i.e. $\sup_{\{x\} \in \prod(x0)} \sum\beta^t F(x_t, x_{t+1})$ exists for any x_0. We can define the supremum function $V: X \to R\cup\{\infty\}$ as:

$$V(x_0) = \sup_{\{x\} \in \prod(x0)}\sum\beta^t F(x_t, x_{t+1})$$

Is V unique? The following assumptions are sufficient:

(S.1) $|V(x_0)| < \infty$, then $V(x_0) \geq u(\{x\})$ for all $\{x\} \in \prod(x_0)$ and for any $\varepsilon > 0$, there is a $\{x\} \in \prod(x_0)$ such that $V(x_0) \leq u(\{x\}) + \varepsilon$.

(S.2) if $V(x_0) = +\infty$, then there is a sequence $\{x^k\} \in \prod(x_0)$ such that $\lim_{k\to\infty} u(\{x^k\}) = +\infty$.
(S.3) if $V(x_0) = -\infty$, then $u(\{x\}) = -\infty$ for all $\{x\} \in \prod(x_0)$.

Theorem: (BE → IP)

Proof: Let $V(\cdot)$ satisfy the intertemporal programming problem. Suppose $\beta > 0$. Suppose $V(x_0)$ is finite. Then, $V^*(x_0) \geq u(x)$ for all $x \in \prod(x_0)$

(B) Existence of Bellman's Equation

Stokey and Lucas (1989: Ch. 4) go through the proof of the existence of a Bellman's equation in the case of bounded, continuous functions. We here follow their Exercise 4.4 (Stokey and Lucas, 1989: p.82-3) to illustrate this for bounded functions case where the state space X is finite or countable.

Preliminaries:

Let $X = \{x_1, x_2, ...\}$ be a finite or countable set;
let the correspondence $\Gamma: X \to X$ be non-empty and finite-valued;
let $A = \{(x, y) \in X \times X \mid y \in \Gamma(x)\}$ (i.e. graph of Γ);
let $F: A \to R$ be a bounded function;
let $0 < \beta < 1$;
let B(X) be the set of bounded functions $f: X \to R$ with sup norm, i.e. $\|f\| = \sup_{x\in X} |f(x)|$
Define the operator T as $(Tf)(x) = \max_{y\in\Gamma(x)} [F(x, y) + \beta f(y)]$.

Question (A):

(A) Show the following:

(i) T: $B(X) \rightarrow B(X)$

(ii) T has a unique fixed point $v \in B(x)$

(iii) $||T^n v_0 - v|| \leq \beta^n ||v_0 - v||$, n = 0, 1, 2, ... holds for all $v_0 \in B(X)$

(iv) Optimal policy correspondence $G:X \rightarrow X$ defined as $G(x) = \{y \in \Gamma(x) \mid v(x) = F(x, y) + \beta v(y)\}$ is non-empty.

Answers to (A):

We first need the following two lemmas:

Lemma 1: B(X) the set of bounded functions on a non-empty space X equipped with a sup norm is a normed vector space.

Proof: Let B(X) be the set of bounded functions on a non-empty set X (note: we do not require it to be compact or a topological space). Let us equip it with the sup norm, so that:

$$||f||_\infty = \sup_{x \in X} |f(x)|$$

As B(X) is the set of bounded functions, then there is an $M > 0$ such that $|f(x)| \leq M$ for all $f \in B(X)$. That it is a norm is easy to see:

(i) $||f|| = \sup_t |f(t)| \geq 0$ with equality iff $f(t) = 0$ for all $t \in [a, b]$;

(ii) $||\alpha f|| = \sup_t |\alpha f(t)| = |\alpha| \cdot [\sup_t |f(t)|] = |\alpha| \cdot ||f||$

(iii) $\sup_t |f(t)| + \sup_t |g(t)| \geq \sup_t |f(t) + g(t)|$ by the properties of the sup operator, thus $||f|| + ||g|| \geq ||f + g||$.

Thus, B(X) with sup norm is a normed vector space.♣

Lemma 2: B(X) equipped with a sup norm is a complete normed vector space (i.e. a Banach space).

Proof: We already know from Lemma 1 that B(X) is a normed vector space, so all that remains is completeness. Thus, we must prove than any Cauchy sequence $\{f_n\}$ in B(X) converges to a point in B(X). Let us propose f as the candidate for such a point. Now, as $\{f_n\}$ is a Cauchy sequence, then for any $\varepsilon > 0$, there is an N such that $\rho(f_n, f_m) < \varepsilon$ for all $n, m \geq N$. For our metric, we take the sup norm, i.e. $\rho(f_n, f_m) = ||f_n - f_m|| = \sup_{x \in X} |f_n(x) - f_m(x)|$. Thus, for any $\varepsilon/2 > 0$, there is an N such that $\sup_{x \in X} |f_n(x) - f_m(x)| < \varepsilon/2$ for all $n, m \geq N$.

Now, f: X $\rightarrow$ R is a real-valued function so that $f_n(x)$ is a real number for any x $\in$ X. Let us fix x′ such that we obtain $\{f_n(x')\}$ as a sequence of real numbers. Obviously, it must be that, by definition of the supremum and using the Cauchy criteria for the original sequence$\{f_n\}$:

$$|f_n(x') - f_m(x')| \leq \sup_{x \in X} |f_n(x) - f_m(x)| < \varepsilon/2$$

for all n, m $\geq$ N. This implies that $\{f_n(x')\}$ is itself a Cauchy sequence in R. Since R is complete, then $f_n(x') \rightarrow f(x') \in$ R. Now, by the triangular inequality:

$$|f_n(x') - f(x')| \leq |f_n(x') - f_m(x')| + |f_m(x') - f(x')|$$

Consequently, by our earlier inequality:

$$|f_n(x') - f(x')| \leq \sup |f_n(x) - f_m(x)| + |f_m(x') - f(x')| < \varepsilon/2 + |f_m(x') - f(x')|$$

But, as we know, like our earlier $\{f_n(x')\}$, we can use the same Cauchy sequence in B(X) argument to show that $\{f_m(x')\}$ is merely a sequence of numbers in R which converges to $f(x')$, i.e. for a given x $\in$ X and $\varepsilon/2 > 0$, there is an N such that for all m > N, $|f_m(x') - f(x')| < \varepsilon/2$. Thus $\varepsilon/2 + |f_m(x') - f(x')| < \varepsilon/2 + \varepsilon/2 = \varepsilon$ for all m > N. Consequently $|f_n(x') - f(x')| < \varepsilon$ for all n $\geq$ N. As this true for all x′ $\in$ X, it will hold true for the x that yields the sup, i.e. $\sup_{x \in X}|f_n(x) - f(x)| < \varepsilon$ for all n $\geq$ N. As $\sup_{x \in X}|f_n(x) - f(x)| = \|f_n - f\|$, then it must be that $\|f_n - f\| < \varepsilon$ for all n $\geq$ N. Thus, $f_n \rightarrow f$, i.e. $\{f_n\}$ converges to a point f.

The only point that remains is to show that f is itself a bounded function, i.e. $f \in$ B(X). Since $\|f_n - f\| = \sup_{x \in X} |f_n(x) - f(x)| < \varepsilon$ implies that $|f_n(x) - f(x)| < \varepsilon$ for all x, then $f \in (f_n(x) - \varepsilon, f_n(x) + \varepsilon)$ for all x $\in$ X. As all f_n are bounded and ε is some number, then f is bounded, i.e. $f \in$ B(X). Thus, any Cauchy sequence $\{f_n\}$ in B(X) converges to a point in B(X). Thus the space B(X) is complete and thus, a Banach space.♣

Let us now turn to the questions at hand:

(i) T: B(X) $\rightarrow$ B(X).

Proof: Recall that operator T is defined as $(Tf)(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta f(y)]$. Now, from Lemma 1, we know that B(X) is a normed vector space. $f(.)$ is a bounded function on X, thus $f \in$ B(X). F(., .) is a bounded function on X × X, thus F(x, .) is a bounded function on X, or F(x, .) $\in$ B(X). As B(X) is a vector space, this implies that for any $\beta \in (0, 1)$, we know that F(x, .) + $\beta f(.) \in$ B(X). Now, as $\Gamma(x) \subset X$ is a finite-valued and non-empty set and F(x, .) + $\beta f(.)$ is a bounded function over X and thus is bounded over $\Gamma(x)$ as well, then a supremum exists, i.e. there is a $\sup_{y \in \Gamma(x)}[F(x, y) + \beta f(y)]$. (Why? Define g = [F(x, .) + $\beta f(.)$] which is a bounded function over $\Gamma(x)$. Suppose there is no supremum. Then, there is an element y $\in \Gamma(X)$ such that g(y) > m for any m $\in$ R. This contradicts the notion that g is bounded. Thus, a supremum must exist). Furthermore, as $\Gamma(x)$ is closed, then this supremum is a well-defined maximum, i.e. $\max_{y\in\Gamma(x)}[F(x, y) + \beta f(y)]$ exists. Define $T(f, x) = \max_{y\in\Gamma(x)}[F(x, y) + \beta f(y)]$ as, obviously, the maximum itself is a function of the function $f \in B(X)$ and variable $x \in X$ chosen. Fixing f, then $T(f, .)$ is a function over X alone, i.e. $T(f, .): X \rightarrow R$. As $F(. , .) + \beta f(.)$ is bounded over $X \times X$, then $F(., y^*) + \beta f(y^*)$ is bounded over X (where y* is the argmax of the function). Thus, as $T(f, .) = F(., y^*) + \beta(y^*)$ then $T(f, .)$ is itself a bounded function on X, i.e. $T(f, .) \in B(X)$. As $(Tf): X \rightarrow R$ is merely another term for $T(f, .)$, then $Tf \in B(X)$. Thus, as $f \in B(X)$, and $Tf \in B(X)$, then $T: B(X) \rightarrow B(X)$.♣

(ii) T has a unique fixed point $v \in B(x)$

Proof: The simplest way to proceed would be to apply Blackwell's Sufficiency Criteria for a contraction mapping (Stokey and Lucas, 1992: p.54, Th.3.3). This states, succinctly, that if the operator $T: B(X) \rightarrow B(X)$ satisfies monotonicity and discounting criteria, then T is a contraction mapping with modulus β.

(i) Monotonicity. T is monotonic if for any $f, g \in B(X)$ where $f(x) \leq g(x)$ for all $x \in X$, then $(Tf)(x) \leq (Tg)(x)$ for all $x \in X$. To see this, consider $f, g \in B(X)$ where $f(x) \leq g(x)$ for all $y \in X$. Pick an arbitrary $x' \in X$, then $(Tf)(x') = \sup_{y\in\Gamma(x')}[F(x', y) + \beta f(y)]$ and $(Tg)(x') = \sup_{y\in\Gamma(x')}[F(x', y) + \beta g(y)]$. Now, by the properties of the sup operator:

$$(Tf)(x') - (Tg)(x') = \sup_{y\in\Gamma(x')}[F(x', y) + \beta f(y)] - \sup_{y\in\Gamma(x')}[F(x', y) + \beta g(y)]$$

$$\leq \sup_{y\in\Gamma(x')}[F(x', y) + \beta f(y) - F(x', y) - \beta g(y)]$$

or simply:

$$(Tf)(x') - (Tg)(x') \leq \beta \sup_{y\in\Gamma(x')}[f(y) - g(y)] \leq 0$$

as, by hypothesis, $\beta \in (0, 1)$ and $f(x) \leq g(x)$ for all $x \in X$ and $y \in \Gamma(x) \subset X$ implies $f(y) \leq g(y)$ for all $y \in X$. Thus $(Tf)(x') \leq (Tg)(x')$ for all $y \in X$. As x' was chosen arbitrarily, then it is true for any $x \in X$ that if $f(x) \leq g(x)$ for all $x \in X$, then $(Tf)(x) \leq (Tg)(x)$ for all $x \in X$. Thus, T is monotonic.

(ii) Discounting: If discounting holds, then for all $f \in B(X)$ and $x \in X$ and any $\alpha \geq 0$ there exists some $\beta \in (0, 1)$ such that $[T(f + \alpha)](x) \leq (Tf)(x) + \beta\alpha$, where $(f+a)(x) = f(x) + a$. In our case:

$$(Tf)(x) = \max_{y\in\Gamma(x)}[F(x, y) + \beta f(y)]$$

$$= \max_{y\in\Gamma(x)}[F(x, y) + \beta f(y)] + \beta\alpha - \beta\alpha$$

$$= \max_{y\in\Gamma(x)}[F(x, y) + \beta f(y) + \beta\alpha] - \beta\alpha$$

$= \max_{y \in \Gamma(x)}[F(x, y) + \beta[f(y) + \alpha]] - \beta\alpha$

$= [T(f + \alpha)](x) - \beta\alpha$

Thus:

$(Tf)(x) + \beta\alpha = [T(f + \alpha)](x)$

thus discounting is satisfied.

Consequently, by Blackwell's Sufficiency Criteria, T is a contraction mapping with modulus β. Now, Lemma 2 proves that B(X) is a *complete* normed vector space. Consequently, by the Contraction Mapping Theorem, T has exactly one fixed point in B(X), i.e. there is a $f \in B(X)$ such that $Tf = f$. Or, using the desired Stokey-Lucas notation, there is a $v \in B(x)$ such that $Tv = v$. Thus, (ii) is satisfied.♣

(iii) $||T^n v_0 - v|| \leq \beta^n ||v_0 - v||$, n = 0, 1, 2, ... holds for all $v_0 \in B(X)$

This is merely a restatement of the familiar contraction mapping theorem (Stokey and Lucas, 1992: p.50, Th. 3.2). Namely, as T is a contraction mapping with modulus β, over a complete metric space (B(X)) then for any $v_0 \in B(X)$, it must be that $||Tv_0 - Tv|| \leq \beta\, ||v_0 - v||$. As $v = Tv$ (fixed point), then $||Tv_0 - v|| \leq \beta\, ||v_0 - v||$. Thus iterating n times, then, $||T^n v_0 - v|| \leq \beta^n ||v_0 - v||$. As this is true for any n = 0, 1, 2. ...., then we obtain:

$||T^n v_0 - v|| \leq \beta^n ||v_0 - v||$, n = 0, 1, 2, ... holds for all $v_0 \in B(X)$

with no further ado.♣

(iv) Optimal policy correspondence $G: X \rightarrow X$ defined as $G(x) = \{y \in \Gamma(x) \mid v(x) = F(x, y) + \beta v(y)\}$ is non-empty.

Proof: Let v be a fixed point of T. Then $(Tv)(x) = \max_{y \in \Gamma(x)} [F(x, y) + \beta v(y)] = v(x)$, which exists and is attained. As Γ(x) is non-empty, then for any $x \in X$, there is a $y \in \Gamma(x)$ such that $v(x) = \max_{y \in \Gamma(x)}[F(x, y) + \beta v(y)]$. Consequently, $G(x) = \{y \in \Gamma(x) \mid v(x) = F(x, y)|$, is non-empty.♣

Question (B):

Let H be the set of functions $h: X \rightarrow X$ such that $h(x) \in \Gamma(x)$, all $x \in X$, For any $h \in H$, define the operator T_h on B(X) by $(T_h f)(x) + F[x, h(x)] + \beta f[h(x)]$. Show that for any $h \in H$, $T_h: B(X) \rightarrow B(X)$ and T_h as a unique fixed point $w \in B(X)$.

Answer to (B):

Proof: Since $h(x) \in \Gamma(x)$, then, by the definition of T, $(Tf)(x) \geq (T_h f)(x)$ for each $x \in X$. As shown earlier, (Tf) is bounded for all $f \in B(X)$. Thus, it must also be that $(T_h f) \in B(X)$. By simple extension of the argument in 4.4.a.(i), then T_h itself is bounded, i.e. $T_h \in B(X)$, so that $T_h:B(X) \to B(X)$. Finally, by a similar argument to the one outlined earlier in 4.4.a.(ii), we can show that T_h satisfies Blackwell's Sufficiency Criteria for a contraction mapping with modulus β in a complete metric space B(X). Consequently, T_h has a unique fixed point, i.e. there is a $w \in B(X)$ such that $T_h w = w$.♣

Question (C):

Let $h_0 \in H$ be given and consider the following algorithm. Given h_n, let w_n be the unique fixed point of T_{hn}. Given w_n, choose h_{n+1} so that $h_{n+1} \in \text{argmax}_{y \in F(X)}[F(x, y) + \beta w_n(y)]$. Show that the sequence of functions $\{w_n\}$ converges to v, the unique fixed point of T. [Hint: Show that $w_0 \leq Tw_0 \leq w_1 \leq Tw_1 \leq ....$]

Answer to (C):

Proof: Given the hint, we want to prove that $w_{n+1} \geq Tw_n \geq w_n$ for all n. Since w_n is a fixed point of T_{hn}, then $w_n = T_{hn}w_n$, or, for any $x \in X$, $w_n(x) = F[x, h_n(x)] + \beta w_n[h_n(x)]$. As $h_{n+1}(x) = \text{argmax}_{y \in \Gamma(x)}[F(x, y) + \beta w_n(y)]$, then $T_{hn+1}w_n(x) = F(x, h_{n+1}(x)) + \beta w_n[h_{n+1}(x)] = \max_{y \in \Gamma(x)}[F(x, y) + \beta w_n(y)] = Tw_n(x)$. Thus, $(T_{hn+1}w_n)(x) = Tw_n(x) \geq w_n(x)$ for all $x \in X$ and all n, i.e. $T_{hn+1}w_n = Tw_n \geq w_n$ for all n.

Now, we can see that T_{hn+1} is monotonic, i.e. we can show that if $f(x) \leq g(x)$ for all x X, then $(T_{hn+1}f)(x) = F(x, h_{n+1}(x) + \beta f[h_{n+1}(x)] \leq F[x, h_{n+1}(x)] + \beta g[h_{n+1}(x)] = (T_{hn+1}g)(x)$ in an analagous manner to the earlier proof. Consequently, by monotonicity, $T^2_{hn+1}w_n \geq T_{hn+1}w_n \geq w_n$. Or, in general, $T^k_{hn+1}w_n \geq T^{k-1}_{hn+1}w_n \geq .... T_{hn+1}w_n = Tw_n \geq w_n$. However, as we know that T_{hn+1} is a contraction mapping, then as $k \to \infty$, $T^k_{hn+1} \to w_{n+1}$, where w_{n+1} is a fixed point, i.e. $w_{n+1} = T_{hn+1}w_{n+1}$. Thus, $w_{n+1} \geq Tw_n \geq w_n$ for all n.

Now, we want to proceed to the limit. We have uncovered a non-decreasing, monotonic sequence, .. $w_{n+1} \geq Tw_n \geq w_n$...that is bounded above by v, the unique fixed point. Thus, as $n \to \infty$, $w_n \to v = Tv$.♣