Overview of Dynamic Programming

Russell Cooper

May 5, 2013

1 An Asset Pricing Example

How much would you pay for a fruit tree? Should depend upon:

- the fruit it bears
- how long it will live
- how you discount the future
- the price of the tree in the future

1.1 Some Basic Analytics

To formalize this, let q^T be the value today of a tree which will live for T years. If the tree lives forever, then its value is denoted q. Further, suppose that the fruit (dividend) of the tree in with t periods until the end of the horizon is denoted d_t, for $t = 1, 2, ...T$. This is just notation but be careful with the distinction between the current period, the number of periods left in the problem, a subscript, and T, a superscript, the horizon.

Suppose that the tree will live for one period and that your utility comes from eating the fruit of the tree in a linear fashion. So clearly,

$$q^1 = d_1. \tag{1}$$

That is the value of the tree is simply the flow of dividends which you receive from the tree.

Now suppose that the tree lives for two periods, $t = 1, 2$ so that $T = 2$. Further, suppose that you will hold tree for both periods so that

$$q^2 = d_2 + \beta d_1. \tag{2}$$

Here d_2 is the value of the dividend with two periods to go, which is the first period that you own the two period tree. The value of the tree is the discounted present value of the dividends over the two periods. Here the value is discounted using $\beta < 1$. In computing this value, we assumed you would hold the tree for two periods.

Instead, suppose that you buy the tree, collect the dividend in the first period and then sell it, before the second period dividend. What is the value of the tree under this strategy?

The value of the tree when you sell it after one period is q_1. So the value to you of this strategy is given by:

$$q^{2s} = d_2 + \beta q^1. \tag{3}$$

Using (1), $q^1 = d_1$ so that the values given in (2) and (3) are the same.

This should be intuitive. Today you anticipate that the market value of the tree will be the value of the dividends in the last period. Since the market value and your value are the same, the price you would pay for the tree should be the same regardless of whether you plan to hold the tree or sell it.

Expressed yet another way, in (3), the market value of the tree tomorrow completely summarizes the future value of the dividends from the tree. We will build on that intuition.

Now suppose the tree lives for three periods, $t = 1, 2, 3$ so $T = 3$. Further, suppose you will hold tree for all periods so

$$q^3 = d_3 + \beta d_2 + \beta^2 d_1 = \sum_{s=0,T-1} \beta^s d_{T-s}. \tag{4}$$

That is, the value of the tree is the discounted present value of the dividends over the three periods. In computing this value, we assumed you would hold the tree for three periods.

With the three period example in mind, lets consider the other options from selling the tree. Suppose you buy the tree, collect dividends this period and then sell it at the start of the next period. Your value would be given by

$$q^{3s} = d_3 + \beta q^2. \tag{5}$$

Note that while this is a three period example, there are only two terms in (5). Somehow, the three period problem is summarized by two terms. The first term is the dividend flow in the first period, d_3. The second term is the price you will get from selling the tree in the second period, q^2. Since the tree lives a total of three periods, the value q^2 is the value of the two-period tree.

If you use (2) to substitute for q^2 in (5), you see that $q^3 = q^{3s}$. Thus the value of the tree can be stated it two ways:

- it is the discounted present value of dividends, as in (4)
- it is the value of dividends plus the value of selling the tree next period, as in (5).

You should be able to make a similar argument to show that q^3 is the value of a tree which is held for two periods and then sold at the start of the final period, before receiving dividends d_3.

Looking at this problem if the horizon is infinite,

$$q_t = \sum_{\tau=0}^{\text{inf}} \beta^\tau d_{t+\tau}. \tag{6}$$

The value of the tree in period t is the discounted present value of the dividends from that point onwards. For this problem, we have to index the value of the tree by time as the problem is not stationary though the horizon is infinite. Here d_t is the dividend in period t as there is no final period to count back from.

If the problem is stationary so that the distribution of dividends is the same every year, then $d_t = d$ for all t as there is no uncertainty in this problem. In this case, $q_t = q$ as the time period, per se, contains no information about asset values. Thus

$$q = \sum_{\tau=0} \beta^\tau d = \frac{d}{1-\beta}. \tag{7}$$

1.2 Adding Uncertainty

Til now, we have treated dividends as known with certainty. Suppose that instead dividends are drawn each period from a given distribution.[1] There are a couple of ways to think about the valuation of the asset relative to the realization of dividends.

First, suppose the asset is bought prior to the realization of the stochastic dividend. In this case, the value of the tree for the one period problem is simply

[1] So the problem is stationary in that the distribution of dividends is fixed.

$$q^1 = Ed \tag{8}$$

where Ed is the expected value of the dividend stream under the known distribution of dividends. So, for example, if dividends take on two values, $d \in \{d_L, d_H\}$ with probabilities (π_L, π_H), then $Ed = \pi_L * d_L + \pi_H * d_H$.

With this same timing, if the horizon is infinite and the draws of the dividend are iid, then the value is given by

$$q = \sum_{\tau=0} \beta^\tau Ed = \frac{Ed}{1-\beta}. \tag{9}$$

Second, suppose that we value the asset after the realization of the dividend. Since the current value of d is known when the asset is valued, let $q(d)$ be the value of the asset when the current realization of the dividend is d. Further let d' be the future value of the dividend. Then,

$$q(d) = d + \beta E_{d'|d} q(d') \tag{10}$$

is the value of the asset when the current dividend is $d \in D$.

Here the notation $E_{d'|d}$ means that the take the expectation of the future value of dividends, d' conditional on knowing the current value of d this period. Returning to our earlier example of d taking on two values, let π_i, j be the probability that $d' = d_j$ tomorrow given that $d = d_i$ today for $i, j = L, H$. Further let π be the transition matrix for dividends:

Table 1: Transition Matrix

	H	L
H	π_{HH}	π_{HL}
L	π_{LH}	π_{LL}

Note: the rows are the states tomorrow and the columns are the states today.

Equation (10) has the same interpretation as (5): the value of the asset is the current dividend plus its (expected) value the next period. As discussed earlier, assuming that you sell the asset after one period is without loss of generality since the resale price next period will fully reflect the expected discounted present value of dividends.

For this case, an educated guess along with verification of that guess is insightful. Suppose that dividends follow an autoregressive process, taking on continuous rather than discrete values. So represent dividends through: $d' = (1-\rho)\mu + \rho d + \varepsilon$ where ε is iid with mean zero. The mean of d is then μ and ρ is the serial correlation.

Guess that $q(d) = \chi_0 + \chi_1 d$ where (χ_0, χ_1) are unknown parameters. Substitute this guess into (10) and you have

$$q(d) = \chi_0 + \chi_1 d = d + \beta[\chi_0 + \chi_1 E_{d'|d} d'] \tag{11}$$

for all $d \in D$. As d follows an AR(1) process with mean μ, $E_{d'|d} d' = (1-\rho)\mu + \rho d$. where ρ is the autoregressive parameter. We then have

$$\chi_0 + \chi_1 d = \beta\chi_0 + \beta\chi_1(1-\rho)\mu + d[1 + \beta\chi_1\rho] \tag{12}$$

for all d. For this to hold for all d, the constant terms on the left must equal those on the right and the terms multiplying d on the left must equal those on the right. This implies:

$$\chi_0 = \beta\chi_0 + \beta\chi_1(1-\rho)\mu \tag{13}$$

and

$$\chi_1 = [1 + \chi_1\rho] \tag{14}$$

so that $\chi_1 = \frac{1}{1-\beta\rho}$. As long as $\beta\rho < 1$, $\chi_1 > 0$. Solve for $\chi_0 = \frac{\beta\chi_1(1-\rho)\mu}{1-\beta}$ using this value of χ_1. Note that $\beta < 1$ is usually assumed in these problems. Here, if $\beta = 1$ so that there is no discounting, then there will be no solution as long as $\mu > 0$.

This verifies the guess. That is we have found the (χ_0, χ_1) parameters to solve the equation for $q(d)$.

In fact, this guess works even if dividends are integrated, $d' = d\epsilon'$ where the mean of the innovation is 1. In this case, it is useful to start from (10) and divide the problem through by d. Let $\tilde{q}(d) = \frac{q(d)}{d}$ so that

$$\tilde{q}(d) = 1 + \beta E_{\epsilon'|d}\tilde{q}(d') \tag{15}$$

for all $d \in D$. Guess that $\tilde{q}(d) = \kappa$ for all d. Then clearly $\kappa = \frac{1}{1-\beta}$ satisfies the functional equation.

Instead of valuing the dividend at its flow value, we can also evaluate the utility of that flow, $u(d)$. Here $u(\cdot)$ would be strictly increasing and concave. Then the valuation becomes

$$q(d) = u(d) + \beta E_{d'|d} q(d') \tag{16}$$

is the value of the asset when the current dividend is $d \in D$.

The form of (16) is something we will encounter many times in dynamic programming problems. In general, we will find that the value of utility or profits or, most generally, payoffs have a current component, as in $u(d)$, and a second component, the discounted expected value, as in $\beta E_{d'|d} q(d')$.

1.3 Solving this problem

How could you solve (15)? Here by solve we mean find the function $q(d)$ that satisfies (15) for all d.

One approach was demonstrated above: guess and verify. When this works, it is great. But normally problems are too complex to generally use this approach.

Another follows from the fact that the unknown function in (15) enters linearly. So you can use linear algebra to solve this system. For example, if d takes on the two values (d_L, d_H), then there are two equations and two unknowns, $(q(d_L), q(d_H))$ to solve

$$q(d_L) = u(d_L) + \beta E_{d'|d_L} q(d') \tag{17}$$

and

$$q(d_H) = u(d_H) + \beta E_{d'|d_H} q(d'). \tag{18}$$

To take conditional expectations here, use the Π matrix introduced earlier.

A final technique is **value function iteration**. The idea is to start from a guess on the $q(d)$ function and then use (16) to update (iterate) the guess. If the process converges (there are theorems on this convergence), then we will have found a solution to (16). Let $i = 0, 1, 2, ...$ index the iteration where $i = 0$ is the initial guess. Then given guess i we obtain guess $i + 1$ by using (16)

$$q^{i+1}(d) = u(d) + \beta E_{d'|d} q^i(d') \tag{19}$$

where the superscript on the $q(d)$ function notates the iteration.

1.4 Estimation

Let Θ be the vector of parameters for this problem. Thus the solution to (10) depends on Θ, which we denote as $q_\Theta(d)$.

Further, suppose the dividend follows a first order Markov process. Specifically, assume d takes values in a finite set D, where $D = (d_1, d_2, ..., d_I)$. Further let $\pi_{ij} = P(d_{t+1} = d_j | d_t = d_i)$ where $P(x|y)$ is a probability of x given y. The transition matrix, Π, is an IxI matrix of transition probabilities with the restriction that $1 > \pi_{ij} > 0$ and $\sum_j \pi_{ij} = 1$ for all i.

For this application, Θ would include the discount factor, β, the utility function $u(d)$ and the transition matrix, Π, which characterizes the distribution of $d'|d$.

If you have data on the dividends for T periods, d_t for $t = 1, 2, 3....$, then you can estimate Π. In the simple case where $d_t \in D$, then you can simply count the transitions from state i to j and divide by the number of occurrences of state i to obtain an estimate of π_{ij}. Alternatively, if you can represent d_t as an AR(1) process, you can use the methods of Tauchen, discussed in Adda-Cooper drawing on Tauchen's *Economic Letters* paper, to generate a discrete representation of this process.

So suppose that you can estimate Π from data on dividends. What else can you estimate from this? Well since this is an exogenous process, neither variations in β nor in $u(\cdot)$ will have any influence on realized dividends. Thus nothing more can be estimated from observing dividends.

What if you also observed asset prices? In an equilibrium model with representative agents where assets are priced by fundamentals, then the state contingent asset price is the same as the valuation of the representative agents. Thus you are observing $q(d_t)$. Since the model predicts that $q(d_t)$ depends on Θ, you can try to infer Θ from the observed prices.

To see how this could be achieved, assume that only β is unknown. Now we will exploit the fact (which we can demonstrate by simulation) that the $q_\Theta(d)$ solving (10) varies with β.

There are many ways to estimate β from asset prices. Here is one way called *simulated method of moments*. Look at the average of the prices in the data, denoted μ^d. Here the superscript refers to the data.

Using the model, for each value of β you could solve (10) to obtain $q_\Theta(d)$. Then you can simulate a sequence of d realizations from the Markov process or even use those from the data. This will give you a time series of simulated values of the asset prices. Then take the mean, denoted $\mu^s(\beta)$. Since $q_\Theta(d)$ depends on Θ, then the mean of the prices, which is just one of many moments will depend on β. In fact, the mean is monotonically increasing in β.

Suppose that there exists a β^* such that $\mu^d = \mu^s(\beta^*)$. So at this value of β the simulated and actual moment are exactly equal. If such a value of β exists and is unique, then this is the estimate of β for this model.

We will discuss later in the course the properties of this type of estimator. We will also talk in length about how to estimate multiple parameters etc.

Another approach, which fits into the approach called *indirect inference*, calculates moments from a reduced-form regression. Though the reduced form does not directly come from the structure of the model, it can be informative about the parameters of the model.

To illustrate, suppose that we estimate an AR(1) representation of the observed asset prices, $q_t, t = 1, 2, 3, ..T$. Let ρ^D be the estimated serial correlation of the process using the data.

Solving the model for a given parameter β (we retain the assumption that other parameters are known), we can create a simulated time series and run the **same** AR(1) that we ran on the actual data. Let $\rho(\beta)$ denote the estimated serial correlation in asset prices from the simulated data. This will generally depend on β. We can then find an estimate of β from solving: $min_\beta(\rho^D - \rho(\beta))^2$. This is a minimum distance estimator.

2 Cake Eating Problem

Suppose that you have a cake of size W_1. At each point of time, $t = 1, 2, 3, ...T$ you can consume some of the cake and thus save the remainder. Let c_t be your consumption in period t and let $u(c_t)$ represent the flow of utility from this consumption. The utility function is not indexed by time: preferences are stationary. Assume $u(\cdot)$ is real-valued, differentiable, strictly increasing and strictly concave. Assume that $\lim_{c\to 0} u'(c) \to \infty$. Represent lifetime utility by

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

where $0 \leq \beta \leq 1$ and β is called the **discount factor**.

For now, assume that the cake does not depreciate (melt) or grow. Hence, the evolution of the cake over time is governed by:

$$W_{t+1} = W_t - c_t \tag{20}$$

for $t = 1, 2, ..T$. How would you find the optimal path of consumption, $\{c_t\}_1^T$?[2]

[2]Throughout, the notation $\{x_t\}_1^T$ is used to define the sequence $(x_1, x_2, ...x_T)$ for some variable x.

2.1 Direct Attack

One approach is to solve the constrained optimization problem directly. This is called the **sequence problem**. Consider the problem of:

$$\max_{\{c_t\}_1^T, \{W_t\}_2^{T+1}} \sum_{t=1}^{T} \beta^{(t-1)} u(c_t) \tag{21}$$

subject to the transition equation (20), which holds for $t = 1, 2, 3, ...T$. Also, there are non-negativity constraints on consumption and the cake given by: $c_t \geq 0$ and $W_t \geq 0$. For this problem, W_1 is given.

Alternatively, the flow constraints imposed by (20) for each t could be combined yielding:

$$\sum_{t=1}^{T} c_t + W_{T+1} = W_1. \tag{22}$$

The non-negativity constraints are simpler: $c_t \geq 0$ for $t = 1, 2, ..T$ and $W_{T+1} \geq 0$. For now, we will work with the single resource constraint. This is a well-behaved problem as the objective is concave and continuous and the constraint set is compact. So there is a solution to this problem.

Letting λ be the multiplier on (22), the first order conditions are given by:

$$\beta^{t-1} u'(c_t) = \lambda$$

for $t = 1, 2, ..., T$ and

$$\lambda = \phi$$

where ϕ is the multiplier on the non-negativity constraint on W_{T+1}. The non-negativity constraints on $c_t \geq 0$ are ignored as we assumed that the marginal utility of consumption becomes infinite as consumption approaches zero within any period.

Combining equations, we obtain an expression that links consumption across any two periods:

$$u'(c_t) = \beta u'(c_{t+1}). \tag{23}$$

This is a necessary condition for optimality for **any** t: if it was violated, the agent could do better by adjusting c_t and c_{t+1}. Frequently, (23) is referred to as an **Euler equation**.

To understand this condition, suppose that you have a proposed (candidate) solution for this problem given by $\{c_t^*\}_1^T, \{W_t^*\}_2^{T+1}$. Essentially, the Euler equation says that the marginal utility cost of reducing consumption by ε in period t equals the marginal utility gain from consuming the extra ε of cake in the next period, which is discounted by β. If the Euler equation holds, then it is impossible to increase utility by moving consumption across adjacent periods given a candidate solution.

It should be clear though that this condition may not be sufficient: it does not cover deviations that last more than one period. For example, could utility be increased by reducing consumption by ε in period t saving the "cake" for two periods and then increasing consumption in period t+2? Clearly this is not covered by a single Euler equation. However, by combining the Euler equation that hold across period t and $t+1$ with that which holds for periods $t+1$ and $t+2$, we can see that such a deviation will not increase utility. This is simply because the combination of Euler equations implies:

$$u'(c_t) = \beta^2 u'(c_{t+2})$$

so that the two-period deviation from the candidate solution will not increase utility.

As long as the problem is finite, the fact that the Euler equation holds across all adjacent periods implies that any finite deviations from a candidate solution that satisfies the Euler equations will not increase utility.

Is this enough? Not quite. Imagine a candidate solution that satisfies all of the Euler equations but has the property that $W_T > 0$ so that there is cake left over. This is clearly an inefficient plan: having the Euler equations holding is necessary but not sufficient. Hence the optimal solution will satisfy the Euler equation for each period and the agent will consume the entire cake!

Formally, this involves showing the non-negativity constraint on W_{T+1} must bind. In fact, this constraint is binding in the above solution: $\lambda = \phi > 0$. This non-negativity constraint serves two important purposes. First, in the absence of a constraint that $W_{T+1} \geq 0$, the agent would clearly want to set $W_{T+1} = -\infty$ and thus die with outstanding obligations. This is clearly not feasible. Second, the fact that the constraint is binding in the optimal solution guarantees that cake is not being thrown away after period T.

So, in effect, the problem is pinned down by an initial condition (W_1 is given) and by a terminal condition ($W_{T+1} = 0$). The set of $(T-1)$ Euler equations and (22) then determine the time path of consumption.

Let the solution to this problem be denoted by $V_T(W_1)$ where T is the horizon of the problem and W_1 is the initial size of the cake. $V_T(W_1)$ represents the maximal utility flow from a T period problem given a size W_1 cake. From now on, we call this a **value function**. This is completely analogous to the indirect utility functions expressed for the household and the firm.

As in those problems, a slight increase in the size of the cake leads to an increase in lifetime utility equal to the marginal utility in any period. That is,

$$V_T'(W_1) = \lambda = \beta^{t-1}u'(c_t), t = 1, 2, ...T.$$

It doesn't matter when the extra cake is eaten given that the consumer is acting optimally. This is analogous to the point raised above about the effect on utility of an increase in income in the consumer choice problem with multiple goods.

2.2 Dynamic Programming Approach

Suppose that we change the above problem slightly: we add a period 0 and give an initial cake of size W_0. One approach to determining the optimal solution of this augmented problem is to go back to the sequence problem and resolve it using this longer horizon and new constraint. But, having done all of the hard work with the T period problem, it would be nice not to have to do it again!

2.2.1 Finite Horizon Problem

The dynamic programming approach provides a means of doing so. It essentially converts a (arbitrary) T period problem into a 2 period problem with the appropriate rewriting of the objective function. In doing so, it uses the value function obtained from solving a shorter horizon problem.

So, when we consider adding a period 0 to our original problem, we can take advantage of the information provided in $V_T(W_1)$, the solution of the T period problem given W_1 from (21). Given W_0, consider the problem of

$$\max_{c_0} u(c_0) + \beta V_T(W_1) \tag{24}$$

where

$$W_1 = W_0 - c_0; W_0 \text{ given.}$$

In this formulation, the choice of consumption in period 0 determines the size of the cake that will be available starting in period 1, W_1. So instead of choosing a sequence of consumption levels, we are just choosing c_0. Once c_0 and thus W_1 are determined, the value of the problem from then on is given by $V_T(W_1)$. This function completely summarizes optimal behavior from period 1 onwards. For the purposes of the dynamic programming problem, it doesn't matter how the cake will be consumed after the initial period. All that is important is that the agent will be acting optimally and thus generating utility given by $V_T(W_1)$. This is the **principle of optimality**, due to Richard Bellman, at work. With this knowledge, an optimal decision can be made regarding consumption in period 0.

Note that the first order condition (assuming that $V_T(W_1)$ is differentiable) is given by:

$$u'(c_0) = \beta V_T'(W_1)$$

so that the marginal gain from reducing consumption a little in period 0 is summarized by the derivative of the value function. As noted in the earlier discussion of the T period sequence problem,

$$V_T'(W_1) = u'(c_1) = \beta^t u'(c_{t+1})$$

for $t = 1, 2, ...T - 1$. Using these two conditions together yields

$$u'(c_t) = \beta u'(c_{t+1}),$$

for $t = 0, 1, 2, ...T - 1$, a familiar necessary condition for an optimal solution.

Since the Euler conditions for the other periods underlie the creation of the value function, one might suspect that the solution to the $T + 1$ problem using this dynamic programming approach is identical to that from using the sequence approach.[3] This is clearly true for this problem: the set of first order conditions for the two problems are identical and thus, given the strict concavity of the $u(c)$ functions, the solutions will be identical as well.

The apparent ease of this approach though is a bit misleading. We were able to make the problem look simple by pretending that we actually knew $V_T(W_1)$. Of course, we had to solve for this either by tackling a sequence problem directly or by building it recursively starting from an initial single period problem.

[3] By the sequence approach, we mean solving the problem using the direct approach outlined in the previous section.

On this latter approach, we could start with the single period problem implying $V_1(W_1)$. We could then solve (24) to build $V_2(W_1)$. Given this function, we could move to a solution of the $T+3$ problem and proceed iteratively, using (24) to build $V_T(W_1)$ for any T.

2.2.2 Example

We illustrate the construction of the value function in a specific example. Assume $u(c) = ln(c)$. Suppose that $T = 1$. Then $V_1(W_1) = \ln(W_1)$.

For $T = 2$, the first order condition from (21) is

$$1/c_1 = \beta/c_2$$

and the resource constraint is

$$W_1 = c_1 + c_2.$$

Working with these two conditions:

$$c_1 = W_1/(1+\beta) \text{ and } c_2 = \beta W_1/(1+\beta).$$

From this, we can solve for the value of the 2-period problem:

$$V_2(W_1) = \ln(c_1) + \beta\ln(c_2) = A_2 + B_2\ln(W_1) \tag{25}$$

where A_2 and B_2 are constants associated with the two period problem. These constants are given by:

$$A_2 = \ln(1/(1+\beta)) + \beta\ln(\beta/(1+\beta)) \qquad B_2 = (1+\beta)$$

Importantly, (25) does not include the *max* operator as we are substituting the optimal decisions in the construction of the value function, $V_2(W_1)$.

Using this function, the $T = 3$ problem can then be written as:

$$V_3(W_1) = \max_{W_2} \ln(W_1 - W_2) + \beta V_2(W_2)$$

where the choice variable is the state in the subsequent period. The first order condition is:

$$\frac{1}{c_1} = \beta V_2'(W_2).$$

Using (25) evaluated at a cake of size W_2, we can solve for $V_2'(W_2)$ implying:

$$\frac{1}{c_1} = \beta \frac{B_2}{W_2} = \frac{\beta}{c_2}.$$

Here c_2 the consumption level is the second period of the three-period problem and thus is the same as the level of consumption in the first period of the two-period problem. Further, we know from the 2-period problem that

$$1/c_2 = \beta/c_3.$$

This plus the resource constraint allows us to construct the solution of the 3-period problem:

$$c_1 = W_1/(1+\beta+\beta^2), \quad c_2 = \beta W_1/(1+\beta+\beta^2), \quad c_3 = \beta^2 W_1/(1+\beta+\beta^2).$$

Substituting into $V_3(W_1)$ yields

$$V_3(W_1) = A_3 + B_3 \ln(W_1)$$

where

$$A_3 = \ln(1/(1+\beta+\beta^2)) + \beta \ln(\beta/(1+\beta+\beta^2)) + \beta^2 \ln(\beta^2/(1+\beta+\beta^2)), \qquad B_3 = (1+\beta+\beta^2)$$

This solution can be verified from a direct attack on the 3 period problem using (21) and (22).

3 Extensions

3.1 Infinite Horizon

Suppose that we consider the above problem and allow the horizon to go to infinity. As before, one can consider solving the infinite horizon sequence problem given by:

$$\max_{\{c_t\}_1^\infty, \{W_t\}_2^\infty} \sum_{t=1}^{\infty} \beta^t u(c_t)$$

along with the transition equation of

$$W_{t+1} = W_t - c_t$$

for t=1,2,......

Specifying this as a dynamic programming problem,

$$V(W) = \max_{c \in [0,W]} u(c) + \beta V(W - c)$$

for all W. Here $u(c)$ is again the utility from consuming c units in the current period. $V(W)$ is the value of the infinite horizon problem starting with a cake of size W. So in the given period, the agent chooses current consumption and thus reduces the size of the cake to $W' = W - c$, as in the transition equation. We use variables with primes to denote future values. The value of starting the next period with a cake of that size is then given by $V(W - c)$ which is discounted at rate $\beta < 1$.

For this problem, the **state variable** is the size of the cake (W) that is given at the start of any period. The state completely summarizes all information from the past that is needed for the forward looking optimization problem. The **control variable** is the variable that is being chosen. In this case, it is the level of consumption in the current period, c. Note that c lies in a compact set. The dependence of the state tomorrow on the state today and the control today, given by

$$W' = W - c$$

is called the **transition equation**.

Alternatively, we can specify the problem so that instead of choosing today's consumption we choose tomorrow's state.

$$V(W) = \max_{W' \in [0,W]} u(W - W') + \beta V(W') \tag{26}$$

for all W. Either specification yields the same result. But choosing tomorrow's state often makes the algebra a bit easier so we will work with (26).

This expression is known as a **functional equation** and is often called a Bellman equation after Richard Bellman, one of the originators of dynamic programming. Note that the unknown in the Bellman equation is the value function itself: the idea is to find a function $V(W)$ that satisfies this condition for all W. Unlike the finite horizon problem, there is no terminal period to use to derive the value function. In effect, the fixed point restriction of having $V(W)$ on both sides of (26) will provide us with a means of solving the functional equation.

Note too that time itself does not enter into Bellman's equation: we can express all relations without an indication of time. This is the essence of **stationarity**.[4] In fact, we will ultimately use the stationarity of the problem to make arguments about the existence of a value function satisfying the functional equation.

A final very important property of this problem is that all information about the past that bears on current and future decisions is summarized by W, the size of the cake at the start of the period. Whether the cake is of this size because we initially had a large cake and ate a lot or a small cake and were frugal is not relevant. All that matters is that we have a cake of a given size. This property partly reflects the fact that the preferences of the agent do not depend on past consumption. But, in fact, if this was the case, we could amend the problem to allow this possibility.

The next part of this chapter addresses the question of whether there exists a value function that satisfies (26). For now, we assume that a solution exists and explore its properties.

The first order condition for the optimization problem in (26) can be written as

$$u'(c) = \beta V'(W').$$

This looks simple but what is the derivative of the value function? This seems particularly hard to answer since we do not know $V(W)$. However, we take use the fact that $V(W)$ satisfies (26) for all W to calculate V'. Assuming that this value function is differentiable,

$$V'(W) = u'(c),$$

a result we have seen before. Since this holds for all W, it will hold in the following period yielding:

$$V'(W') = u'(c').$$

Substitution leads to the familar Euler equation:

$$u'(c) = \beta u'(c').$$

The solution to the cake eating problem will satisfy this necessary condition for all W.

[4] As you may already know, stationarity is vital in econometrics as well. Thus making assumptions of stationarity in economic theory have a natural counterpart in empirical studies. In some cases, we will have to modify optimization problems to ensure stationarity.

The link from the level of consumption and next period's cake (the controls from the different formulations) to the size of the cake (the state) is given by the **policy function:**

$$c = \phi(W), \quad W' = \varphi(W) \equiv W - \phi(W).$$

Using these in the Euler equation reduces the problem to these policy functions alone:

$$u'(\phi(W)) = \beta u'(\phi(W - \phi(W)))$$

for all W.

These policy functions are very important in applied research since they provide the mapping from the state to actions. When elements of the state as well as the action are observable, then these policy functions will provide the foundation for estimation of the underlying parameters.

3.1.1 An Example

In general, actually finding closed form solutions for the value function and the resulting policy functions is not possible. In those cases, we try to characterize certain properties of the solution and, for some exercises, we solve these problems numerically.

However, as suggested by the analysis of the finite horizon examples, there are some versions of the problem we can solve completely. Suppose then, as above, that $u(c) = ln(c)$. Given the results for the T-period problem, we might conjecture that the solution to the functional equation takes the form of:

$$V(W) = A + B\ln(W)$$

for all W. With this guess we have reduced the dimensionality of the unknown function $V(W)$ to two parameters, A and B. But can we find values for A and B such that $V(W)$ will satisfy the functional equation?

Taking this guess as given and using the special preferences, the functional equation becomes:

$$A + B\ln(W) = \max_{W'} \ln(W - W') + \beta(A + B\ln(W')) \tag{27}$$

for all W. After some algebra, the first-order condition implies:

$$W' = \varphi(W) = \frac{\beta B}{(1+\beta B)} W.$$

Using this in (27) implies:

$$A + B\ln(W) = \ln\frac{W}{(1+\beta B)} + \beta(A + B\ln(\frac{\beta BW}{(1+\beta B)}))$$

for all W. Collecting terms into a constant and terms that multiply $ln(W)$ and then imposing the requirement that the functional equation must hold for all W, we find that

$$B = 1/(1-\beta)$$

is required for a solution. Given this, there is a complicated expression that can be used to find A. To be clear then we have indeed guessed a solution to the functional equation. We know that because we can solve for (A, B) such that the functional equation holds for all W using the optimal consumption and savings decision rules.

With this solution, we know that

$$c = W(1-\beta), W' = \beta W.$$

Evidently, the optimal policy is to save a constant fraction of the cake and eat the remaining fraction.

Interestingly, the solution to B could be guessed from the solution to the T-horizon problem, section 2.2.2, where

$$B_T = \sum_{t=1}^{T} \beta^{t-1}.$$

Evidently, $B = lim_{T\to\infty} B_T$. In fact, we will be exploiting the theme that the value function which solves the infinite horizon problem is related to the limit of the finite solutions in much of our numerical analysis.

3.2 Taste Shocks

To allow for variations in tastes, suppose that utility over consumption is given by:

$$\varepsilon u(c)$$

where ε is a random variable whose properties we will describe below. The function $u(c)$ is again assumed to be strictly increasing and strictly concave. Otherwise, the problem is the original cake eating problem with an initial cake of size W.

In problems with stochastic elements, it is critical to be precise about the timing of events. Does the optimizing agent know the current shocks when making a decision? For this analysis, assume that the agent knows the value of the taste shock when making current decisions but does not know future values. Thus the agent must use expectations of future values of ε when deciding how much cake to eat today: it may be optimal to consume less today (save more) in anticipation of a high realization of ε in the future.

For simplicity, assume that the taste shock takes on only two values: $\varepsilon \in \{\varepsilon_h, \varepsilon_l\}$ with $\varepsilon_h > \varepsilon_l > 0$. Further, we assume that the taste shock follows a first -order Markov process which means that the probability a particular realization of ε occurs in the current period depends **only** the value of ε attained in the previous period.[5] For notation, let π_{ij} denote the probability that the value of ε goes from state i in the current period to state j in the next period. For example, π_{lh} is defined from:

$$\pi_{lh} \equiv \text{ Prob}(\varepsilon' = \varepsilon_h | \varepsilon = \varepsilon_l)$$

where ε' refers to the future value of ε. Clearly $\pi_{ih} + \pi_{il} = 1$ for $i = h, l$. Let Π be a $2x2$ matrix with a typical element π_{ij} which summarizes the information about the probability of moving across states. This matrix is naturally called a **transition matrix**.

Given this notation and structure, we can turn to the cake eating problem. It is critical to carefully define the state of the system for the optimizing agent. In the nonstochastic problem, the state was simply the size of the cake. This provided all the information the agent needed to make a choice. When taste shocks are introduced, the agent needs to take this into account as well. In fact, the taste shocks provide information about current payoffs and, through the Π matrix, are informative about the future value of the taste shock as well.[6]

Formally, the Bellman equation is:

$$V(W, \varepsilon) = \max_{W'} \varepsilon u(W - W') + \beta E_{\varepsilon'|\varepsilon} V(W', \varepsilon')$$

[5]The evolution can also depend on the control of the previous period. Note too that by appropriate rewriting of the state space, richer specifications of uncertainty can be encompassed.

[6]This is a point that we return to below in our discussion of the capital accumulation problem.

for all (W, ε) where $W' = W - c$ as usual. Note that the conditional expectation is denoted here by $E_{\varepsilon'|\varepsilon} V(W', \varepsilon')$ which, given Π, is something we can compute.[7]

The first order condition for this problem is given by:

$$\varepsilon u'(W - W') = \beta E_{\varepsilon'|\varepsilon} V_1(W', \varepsilon')$$

for all (W, ε). Using the functional equation to solve for the marginal value of cake, we find:

$$\varepsilon u'(W - W') = \beta E_{\varepsilon'|\varepsilon}[\varepsilon' u'(W' - W'')] \tag{28}$$

which, of course, is the stochastic Euler equation for this problem.

The optimal policy function is given by

$$W' = \varphi(W, \varepsilon)$$

The Euler equation can be rewritten in these terms as:

$$\varepsilon u'(W - \varphi(W, \varepsilon)) = \beta E_{\varepsilon'|\varepsilon}[\varepsilon' u'(\varphi(W, \varepsilon) - \varphi(\varphi(W, \varepsilon), \varepsilon')))]$$

The properties of this policy function can then be deduced from this condition. Clearly both ε' and c' depend on the realized value of ε' so that the expectation on the right side of (28) cannot be split into two separate pieces.

3.3 A Stochastic Discrete Cake Eating Example

Suppose you have a cake which must, by assumption, be eaten in one period. Perhaps we should think of this as the wine drinking problem recognizing that once a good bottle of wine is opened, it should be consumed! The value of eating a cake of size W today is $\varepsilon u(W)$ where $u(\cdot)$ is increasing and concave and ε is a shock to tastes.

If the cake is not eaten, then tomorrow it is of size $W' = \rho W$. In the next period, the agent decides to eat the cake or not.

The problem is then an example of a dynamic, stochastic discrete choice problem. This is an example of a family of problems called **optimal stopping problems**. The common element in all of these problems is the emphasis on timing of a single event: when to eat the cake; when to take a job; when to stop school, when to stop revising a chapter, etc. In fact, for many of these problems, these choices are not once in a lifetime

[7]Throughout we denote the conditional expectation of ε' given ε as $E_{\varepsilon'|\varepsilon}$.

events and so we will be looking at problems even richer than the optimal stopping variety.

Let $V^E(W,\varepsilon)$ and $V^N(W,\varepsilon)$ be the value of eating the size W cake now (E) and waiting (N) respectively given the current taste shock, $\varepsilon \in \{\varepsilon_h, \varepsilon_l\}$. Then,

$$V^E(W,\varepsilon) = \varepsilon u(W) \tag{29}$$

and

$$V^N(W,\varepsilon) = \beta E_{\varepsilon'|\varepsilon} V(\rho W, \varepsilon'). \tag{30}$$

where

$$V(W,\varepsilon) = \max(V^E(W,\varepsilon), V^N(W,\varepsilon)) \tag{31}$$

for all (W,ε).

To understand these terms, $\varepsilon u(W)$ is the direct utility flow from eating the cake. Once the cake is eaten the problem has ended. So $V^E(W,\varepsilon)$ is just a one-period return. If the agent waits, then there is no cake consumption in the current period and next period the cake is of size (ρW). As tastes are stochastic, the agent choosing to wait must take expectations of the future taste shock, ε'. The agent has an option next period of eating the cake or waiting further. Hence the value of having the cake in any state is given by $V(W,\varepsilon)$, which is the value attained by maximizing over the two options of eating or waiting. The cost of delaying the choice is determined by the discount factor β while the gains to delay are associated with the growth of the cake, parameterized by ρ. Further, the realized value of ε will surely influence the relative value of consuming the cake immediately.

As a special case, suppose that $\rho = 1$ and ε is an iid random variable. Then there will generally exist a critical value of the taste shock, $\varepsilon^*(W)$, such that the value of eating the cake and waiting are equal. In the iid case, $V^N(W,\varepsilon))$ is independent of ε. But, the value of eating the cake is monotonically increasing in ε. So generally (unless the domain of the shock is restricted), there will exist a value of the shock such that: $\varepsilon u(W) = V^N(W,\varepsilon))$. Moreover, since the cake is ultimately eaten, $V^N(W,\varepsilon)$ is proportional to $u(W)$. Thus $\varepsilon^*(W)$ is independent of W.

As another special case, assume that ε takes on two values, i.e. $\varepsilon \in \{\varepsilon_h, \varepsilon_l\}$ and follows a first-order Markov process with a transition matrix, π. In this case, there is no gain from delay when $\varepsilon = \varepsilon_h$. If the agent delays, then utility in the next period will have to be lower due to discounting and, with probability π_{hl}, the taste shock will switch from low to high. So, waiting to eat the cake in the future will not be desirable. Hence,

$$V(W, \varepsilon_h) = V^E(W, \varepsilon_h) = \varepsilon_h u(W)$$

for all W.

In the low ε state, matters are more complex. If β and ρ are sufficiently close to 1 then there is not a large cost to delay. Further, if π_{lh} is sufficiently close to 1, then it is likely that tastes will switch from low to high. Thus it will be optimal not to eat the cake in state (W, ε_l).

4 Growth Model

To begin our exploration of applications of dynamic programming problems in macroeconomics, a natural starting point is the stochastic growth model. This framework has been used for understanding fluctuations in the aggregate economy. We begin with the non-stochastic model to get some basic concepts straight and then enrich the model to include shocks and other relevant features in other classes.[8]

4.1 Non-Stochastic Growth Model

We study the planner's problem. The economy has a representative household which lives forever and a technology to produce output from capital and labor inputs.

The household is endowed with *one* unit of leisure each period and supplies this inelastically to a production process. The household consumes an amount c_t each period which it evaluates using a utility function, $u(c_t)$. Assume that $u(\cdot)$ is strictly increasing and strictly concave. The household's lifetime utility is given by

$$\sum_{1}^{\infty} \beta^{t-1} u(c_t) \tag{32}$$

The technology produces output (y) from capital (k), given its inelastically supplied labor services. Let $y = f(k)$ be the production function. Assume that $f(k)$ is strictly increasing and strictly concave.

[8]This presentation is taken from Adda, J. and Cooper, R., **Dynamic Programming: Theory and Macroeconomic Applications**

There is a resource constraint that decomposes output into consumption and investment (i_t):

$$y_t = c_t + i_t.$$

The capital stock accumulates according to:

$$k_{t+1} = k_t(1-\delta) + i_t$$

where $\delta \in (0,1)$ is the rate of physical depreciation. So, the capital input into the production process is accumulated from forgone consumption.

The planner's problem is to determine an optimal path for capital by splitting output between these two competing uses. For now, our focus is on solving for this allocation as the solution of a dynamic optimization problem. Decentralization is not considered.

To study this problem, we use the dynamic programming approach and consider the following functional equation:

$$V(k) = \max_{k'} u(f(k) + (1-\delta)k - k') + \beta V(k') \tag{33}$$

for all k. Here the state variable is the stock of capital at the start of the period and the control variable is the capital stock for the next period.[9]

With $f(k)$ strictly concave, there will exist a maximal level of capital achievable by this economy given by $\bar{k}$ where

$$\bar{k} = (1-\delta)\bar{k} + f(\bar{k}).$$

This provides a bound on the capital stock for this economy and thus guarantees that our objective function, $u(c)$, is bounded on the set of feasible consumption levels, $[0, f(\bar{k}) + (1-\delta)\bar{k}]$. We assume that both $u(c)$ and $f(k)$ are continuous and real-valued so there exists a $V(k)$ that solves (33).

The first-order condition is given by:

$$u'(c) = \beta V'(k'). \tag{34}$$

Of course, we don't know $V(k)$ directly so that we need to use (33) to determine $V'(k)$. As (33) holds for all $k \in [0, \bar{k}]$, we can take a derivative and obtain:

$$V'(k) = u'(c)(f'(k) + (1-\delta)).$$

[9]Equivalently, we could have specified the problem with k as the state, c as the control and then used a transition equation of: $k' = f(k) + (1-\delta)k - c$.

Updating this one period and inserting this into the first-order condition implies:

$$u'(c) = \beta u'(c')(f'(k') + (1 - \delta)).$$

This is an Euler condition that is not unlike the one we encountered in the cake eating problem. Here the left side is the cost of reducing consumption by ε today. The right side is then the increase in utility in the next period from the extra capital created by investment of the ε. As in the cake eating structure, if the Euler equation holds then no single period deviations will increase utility of the household. As with that problem, this is a necessary but not a sufficient condition for optimality.[10]

You can show that $V(k)$ is strictly concave. Consequently, from (34), k' must be increasing in k. To see why, suppose that current capital increases but future capital falls. Then current consumption will certainly increase so that the left side of (34) decreases. Yet with k' falling and $V(k)$ strictly concave, the right side of (34) increases. This is a contradiction.

4.2 An Example

Suppose that $u(c) = ln(c), f(k) = k^\alpha$ and $\delta = 1$. With this special structure, we can actually solve this model. Guess that the value function is given by:

$$V(k) = A + B \ln k$$

for all k. If this guess is correct, then we must be able to show that it satisfies (33). If it does, then the first-order condition, (34), can be written:

$$\frac{1}{c} = \frac{\beta B}{k'}.$$

Using the resource constraint ($k^\alpha = c + k'$),

$$\beta B(k^\alpha - k') = k'$$

or

$$k' = (\frac{\beta B}{1 + \beta B})k^\alpha. \tag{35}$$

[10] As noted in the discussion of the cake eating problem, this is but one form of a deviation from a proposed optimal path. Deviations for a finite number of periods also do not increase utility if (4.1) holds. In addition, a transversality condition must be imposed to rule out deviations over an infinite number of period.

So, if our guess on $V(k)$ is correct, this is the policy function.

Given this policy function, we can now verify whether or not our guess on $V(k)$ satisfies the functional equation, (33). Substitution of (35) into (33) yields

$$A + B\ln k = \ln[(\frac{1}{1+\beta B})k^{\alpha}] + \beta[A + B\ln((\frac{\beta B}{1+\beta B})k^{\alpha})] \tag{36}$$

for all k. Here we use $c = y - k'$ so that

$$c = (\frac{1}{1+\beta B})k^{\alpha}.$$

Grouping constant terms implies:

$$A = \ln(\frac{1}{1+\beta B}) + \beta[A + B\ln(\frac{\beta B}{1+\beta B})]$$

and grouping terms that multiply ln k,

$$B = \alpha + \beta B\alpha.$$

Hence $B = \frac{\alpha}{1-\beta\alpha}$. Using this, A can be determined. Thus, we have found the solution to the functional equation.

As for the policy functions, using B, we find

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

and

$$c = (1-\beta\alpha)k^{\alpha}.$$

It is important to understand how this type of argument works. We started with a guess of the value function. Using this guess, we derived a policy function. Substituting this policy function into the functional equation gave us an expression, (36), that depends only on the current state, k. As this expression must hold for all k, we grouped terms and solved for the unknown coefficients of the proposed value function.

4.2.1 Computational Approach

Check out Matlab code, grow.m, on the course page.

4.3 Stochastic Growth

To make the model stochastic we add shocks. The most common shock is in technology: $y = Af(k)$ where A is a serially correlated productivity shock. You can then solve the model to determine policy functions in state (A, k). You might also consider shocks to preferences and to the process creating capital from investment.