ECON 202A : Week 2

Paula Beltran*

October 16, 2020

In the previous session, we learned how to find the optimal sequence of choice variables over time using the recursive formulation. To do so, we assumed that V is differentiable and existence of steady state. In these notes, we discuss these results and take a closer look at the dynamics in the neoclassical growth models and how to solve dynamic programming problems with constraints.

So far, we have considered the deterministic case. That is, there is no uncertainty. We introduce the concept of Markov Chains, which is the first step towards introducing uncertainty in our models.

1 Uniqueness and Existence of Steady State in Neoclassical Model

First, we prove exercise 4.5 in Stokey and Lucas to check uniqueness and existence of *steady state* in neoclassical model[1]. Second, we observe how the economy *converges* to the unique steady state from a given initial state. Third, we consider the *general framework* of dynamic programming with generic state and control variables. During this, we will see how the envelop theorem can help us construct a well-defined system of equations which characterizes the optimal solution of the planner. In this section we prove some properties of the policy function prior to discussing the steady state and dynamic properties of the deterministic neoclassical growth model. We will show that this model is characterized by a unique steady state capital stock and that if the economy begins away from that steady state level, it will monotonically converge to the steady state over time.

Recall that we have assumed the following for the production function $f(k)$ in the neoclassical growth model:

*I thank the former TAs for this course: Sntiago Justel, Emmanouil Chatzikonstantinou, Andrés Schneider, Adrien des Enffans d'Avernas, Roberto Fattal, Matt Luzzetti, Kyle Herkenhoff and Jinwook Hur. These notes are **heavily** based on their notes. All errors are my own. Email me at pabeltran90@gmail.com for any errors, corrections, or suggestions.

[1]Although there exist two steady states in the neoclassical growth model (the other one is a trivial case where $k_0 = 0$), I will refer to the model as having a unique steady state and this unique steady state being that which is globally stable.

1. $f(0) = 0$
2. $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{k \to \infty} f'(k) = 0$
3. $f'(k) > 0, \forall k \in \mathbb{R}_+$
4. $f(\cdot)$ is strictly concave

Assumption 4 is derived from the fact that the original production function $F(\cdot, \cdot)$ was assumed to be strictly quasi-concave and satisfy constant returns to scale. Exercise 4.8 in Stokey and Lucas demonstrates that these assumptions imply that $F(\cdot, \cdot)$ is strictly concave, implying that $f(\cdot)$ is strictly concave.

We often define the maximum sustainable capital stock as that capital stock that could be maintained if consumption is set to zero. Denoting the maximum sustainable capital stock by k^*, this definition implies that k^* is the fixed point satisfying the following mapping: $k^* = f(k^*)$. Given the assumptions on the production function, this fixed point is unique. It is useful to truncate the state space (when using numerical methods).

Since $f(k_t)$ represents the maximum level of output than can be achieved with a capital stock equal to k_t, it must be that $g(k_t) \leq f(k_t), \forall k_t \in \mathbb{R}_+$. That is, the production function bounds the policy function from above. This is true because it would never be optimal to forego consumption completely and save all output.

We now look to formally establish two additional properties of the policy function: it is *strictly increasing* and *its slope at any point is bounded above* by the slope of the production function. This question is the subject of Exercise 4.5 in Stokey and Lucas.

The point of the proof below is to establish that if f has certain properties that guarantee the existence of a unique maximum capital stock, which is the amount of capital an economy can sustain if no one consumes and everyone saves, then the policy function inherits those properties and there is a unique steady state capital stock.

Exercise 4.5 (Stokey and Lucas) *Assume that $U, f,$ and V are strictly increasing, strictly concave, and once continuously differentiable, and that $0 < g(x) < f(x)$, for all x. Use the first order condition: $U'(f(x) - g(x)) - \beta V'(g(x)) = 0$, to show that $g(\cdot)$ is strictly increasing and has slope less than the slope of $f(\cdot)$. That is: $0 < g(\tilde{x}) - g(x) < f(\tilde{x}) - f(x)$, if $\tilde{x} > x$.*

Proof: We first show that $g(\cdot)$ is strictly increasing. To derive a contradiction, suppose that $g(\tilde{x}) \leq g(x)$ for some $\tilde{x} > x$. Using the first-order condition, we have the following:

$$\begin{aligned}
\beta V'(g(\tilde{x})) &= U'(f(\tilde{x}) - g(\tilde{x})) \\
&< U'(f(x) - g(\tilde{x})) \quad \text{(by } U \text{ strictly concave and } f \text{ strictly increasing)} \\
&\leq U'(f(x) - g(x)) \quad \text{(by } U \text{ strictly concave and } g(x) \geq g(\tilde{x})) \\
&= \beta V'(g(x)).
\end{aligned}$$

However, this contradicts the assumption that V is strictly concave. Thus, g must be strictly increasing.

Next, we prove that g's slope must be everywhere bounded above by f's slope. Take any $\tilde{x} > x$. Since V is strictly concave and g is strictly increasing:

$$\begin{aligned} & V'(g(\tilde{x})) < V'(g(x)) \\ \iff\ & U'(f(\tilde{x}) - g(\tilde{x})) < U'(f(x) - g(x)) \\ \iff\ & f(\tilde{x}) - g(\tilde{x}) > f(x) - g(x) \\ \iff\ & f(\tilde{x}) - f(x) > g(\tilde{x}) - g(x) > 0. \end{aligned}$$

Therefore, g is strictly increasing and has a slope bounded above by the slope of f. We have thus established the additional assumptions on g used in class. Yet, we have seen in lecture that these assumptions did not seem to prove that a unique steady state existed in the neoclassical growth model. To see that a unique steady state exists we will have to invoke the strict concavity of f implied by the assumptions that f exhibits constant returns to scale and is strictly quasi-concave.

2 Steady State

In the deterministic neoclassical growth model, there is a unique value of capital stock where if the economy reaches the level, it optimally stays at that level indefinitely. This means that the steady state capital stock, denoted by $\bar{k}$, must be a fixed point of the policy function, or $\bar{k} = g(\bar{k})$. By the resource constraint of the model, this means that the consumption level of this economy also remains constant in all periods once the economy reaches there.

Note that this steady state is globally stable, ch. 6 from Stockey and Lucas deals with this fact formally, but just to give you a sketch of the argument, let us focus on the standard problem that we described last week. In general $k_{t+1} = g(k_t)$ describes the policy function, let me linearize this solution around the steady state k^*.

$$k_{t+1} \approx k^* + g'(k^*)(k_t - k^*)$$

$$\underbrace{k_{t+1} - k^*}_{x_{t+1}} \approx g'(k^*) \underbrace{(k_t - k^*)}_{x_t}$$

We can express the solution as a linear difference equation $x_{t+1} = g'(k^*)x_t$, the solution for this equation is $x_t = [g'(k^*)]^t x_0$. It is easy to check that if $|g'(k^*)| < 1$ then $\lim_{t\to\infty} x_t = 0$, which is equivalent to $\lim_{t\to\infty} k_t = k^*$, thus the solution is *locally* stable. So a condition for stability is $|g'(k^*)| < 1$.[2]

This stability argument means that, even if the initial capital stock is not equal to the steady state level, the economy will converge to the steady state value over time (as long

[2] The problem with this argument is that, in general, we don't know g or if g is differentiable. That is why one can apply similar logic to the Euler equation (SLP, section 6.4).

as $k_0 > 0$). In this section, we see how we can solve for the steady state from the Euler equations, without conducting value function iteration procedure.

Let's see the Bellman equation that we dealt with last week:

$$V(k) = \max_{k'} \Big[U(f(k) - k') + \beta V(k') \Big]. \tag{1}$$

As we did it last week, taking the first-order condition of 1 with respect to k',

$$U'(f(k) - k') = \beta V'(k'). \tag{2}$$

However, this equation contains the unknown function V. Also, the differentiability of V is also unknown. Fortunately, the next theorem shows the differentiability of V and how can we compute this derivative.

Theorem 2.1. Envelope Theorem *or Benveniste and Scheinkman* *If V is* concave *and W is* concave and differentiable *with $W(x_0) = V(x_0)$ and $W(x) \leq V(x)$ $\forall x$ in a neighborhood of x_0, then V itself is differentiable at x_0 and $V_i(x_0) = W_i(x_0)$.*[3]

As in lecture, define $W(k) = U(f(k) - g(k_0)) + \beta V(g(k_0))$. Since U is assumed to be concave and $V(g(k_0))$ is just a constant, W is concave. Furthermore, W is differentiable since both U and f are differentiable by assumption. The remainder to be checked is whether V is concave. Theorem 4.8 in Stokey and Lucas establishes the conditions under which V is concave. Taking these facts as given, we can apply Envelope Theorem to find the derivative of the value function in our model. The envelope condition yields

$$V'(k) = U'(f(k) - k')f'(k).$$

Advance this condition for one period. Then,

$$V'(k') = U'(f(k') - k'')f'(k'). \tag{3}$$

Now, plug (2) into (3). Then, we finally obtain the Euler Equation as follows:

$$U'(f(k) - k') = \beta U'(f(k') - k'')f'(k').$$

Note that this equation is the same as we obtained from the sequential version of this problem.

Now, we can compute the steady state level of capital stock using this Euler equation. In the steady state, we have $k = k' = k'' = \bar{k}$. Therefore, evaluating the Euler equation at the steady state we have:

$$\beta f'(\bar{k}) = 1. \tag{4}$$

By strict concavity of f, there exists a unique solution to this equation and thus a unique steady state value. One can see that, from what we have done so far, the steady state of

[3] Here, $V_i(x_0)$ means the partial derivative of V with respect to its ith argument. Naturally, we are allowing the possibility that x_0 is a vector in principle.

this economy can be characterized even without solving for the value function and the policy function. To examine whether this solution of steady state is consistent with the one from the computed policy function, let's assume $f(k) = k^\alpha$ as we did last week. Then (4) becomes:

$$\begin{aligned} \beta\alpha\bar{k}^{\alpha-1} &= 1 \\ \bar{k} &= (\alpha\beta)^{\frac{1}{1-\alpha}}. \end{aligned}$$

Recall the policy function that we computed last week: $g(k) = \alpha\beta k^\alpha$. Then, at the steady state,

$$\begin{aligned} \bar{k} &= g(\bar{k}) \\ &= \alpha\beta\bar{k}^\alpha \\ &= (\alpha\beta)^{\frac{1}{1-\alpha}}. \end{aligned}$$

Therefore, we obtain the same level of steady state capital stock. Finally, let us check if this steady state is stable, that is we need to show that $g(k^*) < 1$. In this case, $g'(k) = \alpha^2\beta k^{\alpha-1}$, at the steady state $g'(k^*) = \alpha^2\beta k^{*\alpha-1} = \alpha^2\beta(\alpha\beta)^{-1} = \alpha < 1$. Hence, steady state is stable.

3 Dynamics

The deterministic neoclassical growth model exhibits rather simple dynamics given the unique globally stable steady state. By dynamics, we mean the process by which the capital stock moves from its initial point $k_0 > 0$ as time proceeds.[4] The dynamics for this model are characterized in the following points:

1. For all $0 < k_0 < \bar{k}$, $k_{t+1} > k_t$ and k_t converges to $\bar{k}$ monotonically from below.

2. For all $k_0 > \bar{k} > 0$, $k_{t+1} < k_t$ and k_t converges to $\bar{k}$ monotonically from above.

The monotonic convergence can be easily seen at Figure 1. For any strictly positive initial capital stock k_0 and tracking its optimal evolution using the policy function and 45 degree line, we can observe the capital stock converges to the steady state level monotonically.

[4]Note that the trivial steady state, where $k_0 = 0$ is given and $k_t = 0$ for all t. This solution is unstable since $\lim_{k\to 0} g'(k) = \infty > 1$

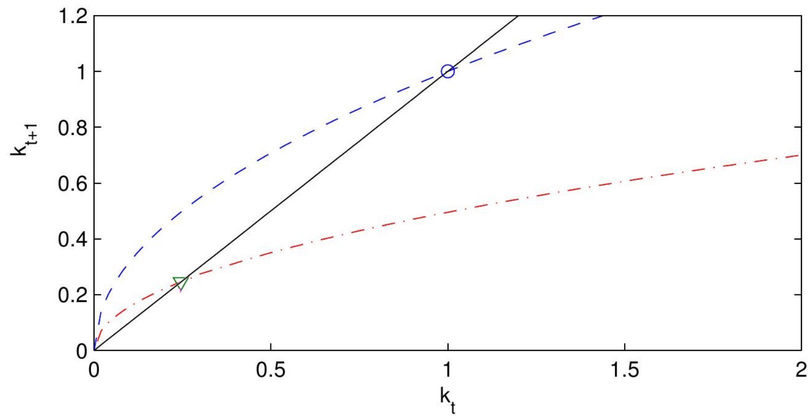

Figure 1: Steady State and Dynamics in the Standard Neoclassical Growth Model

4 Solving Dynamic Programming Problems with Constraints

In this section we introduce a technique to derive the Euler equations of a model using the *Lagrangian*. In the simplest neoclassical model, we did not need to construct Lagrangian since it was possible to substitute the constraint into the objective function. However, in more complex models with many variables and many markets, this may be impossible to do so. Without transition and if time permits, I will also define the *sequential competitive equilibrium*. In order to solve problems in a recursive form, we need to get rid of the unknown value function in the necessary first order conditions. To do this, we need to apply the following Envelope Theorem for the case of constrained maximization.[5]

Theorem 4.1. ***Envelope Theorem 2.0*** *Consider the following maximization problem:*

$$V(y) = \max_x f(x,y) \tag{5}$$

subject to

$$g(x,y) \geq 0.$$

The Lagrangian of this problem is given by:

$$L(x,y,\lambda) = f(x,y) + \lambda g(x,y)$$

Then: $V'(y) = \frac{\partial L}{\partial y}$

4.1 Envelope theorem and neoclassical problem

Back to the neoclassical problem, consider now the following generalized version of dynamic programming problem.

$$V(k) = \max_{c,k'} \Big[U(c) + \beta V(k') \Big] \tag{6}$$

[5]We prove this theorem in the appendix

subject to

$$g(c, k, k') \geq 0.$$

Suppose g is differentiable and concave in each of its arguments. We can derive the Euler equation using the following steps:

1. Define the Lagrangian.
2. Take the first-order conditions with respect to all control variables and the state variables.
3. Apply the Envelope theorem to the Lagrangian

Following the steps above, we can solve the problem (6).

Step 1: Define the Lagrangian:

$$L(c, k, k') = U(c) + \beta V(k') + \lambda g(c, k, k')$$

Step 2: Take first-order conditions.

$$c: \quad \frac{\partial L}{\partial c} = U'(c) + \lambda g_1(c, k, k') = 0 \tag{7}$$

$$k': \quad \frac{\partial L}{\partial k'} = \beta V'(k') + \lambda g_3(c, k, k') = 0 \tag{8}$$

By rearranging (7), we can solve for the Lagrange multiplier with respect to the other variables:

$$\lambda = -\frac{U'(c)}{g_1(c, k, k')}. \tag{9}$$

Plugging (9) into (8),

$$\beta V'(k') = \frac{U'(c)}{g_1(c, k, k')} g_3(c, k, k'). \tag{10}$$

Step 3:

The envelope theorem in this model, therefore, yields:

$$\begin{aligned} V'(k) &= \frac{\partial L}{\partial k} = \lambda g_2(c, k, k') \\ &= -\frac{U'(c)}{g_1(c, k, k')} g_2(c, k, k') \end{aligned} \tag{11}$$

Advancing (11) by one period,

$$V'(k') = -\frac{U'(c')}{g_1(c', k', k'')} g_2(c', k', k'')$$

Combine this with (10). Then, we obtain the following Euler equation:

$$-\beta \frac{U'(c')}{g_1(c', k', k'')} g_2(c', k', k'') = \frac{U'(c)}{g_1(c, k, k')} g_3(c, k, k') \tag{12}$$

In the simple neoclassical world, $g(c,k,k') = f(k) - c - k'$ and thus $g_1(c,k,k') = -1$, $g_2(c,k,k') = f'(k)$, and $g_3(c,k,k') = -1$. Therefore, (12) becomes,

$$U'(c) = \beta U'(c')f'(k'). \tag{13}$$

Combining (13) with the resource constraint $c = f(k) - k'$ yields the familiar expression

$$U'(f(k) - k') = \beta U'(f(k') - k'')f'(k').$$

4.2 Exercise

<Midterm 2007, Question 2>

A social planner solves the following problem:

$$\max_{\{c_t, m_t, i_t, n_t, h_t, k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\log c_t + n_t A \log(1 - h_t)\right]$$

subject to

$$\begin{aligned} & c_t + i_t + m_t \leq h_t k_t^{\theta} n_t^{1-\theta} \\ & m_t = \alpha(n_t - n_{t-1})^2 \\ & k_{t+1} = (1-\delta)k_t + i_t \\ & n_{-1} \text{ and } k_0 \text{ given} \end{aligned}$$

The variables c_t, i_t, m_t, h_t, n_t, k_t are time t values of consumption, investment, moving costs, length of a work shift, employment rate, and capital stock, respectively. The interpretation is that the planner chooses both employment (number of people), n, and hours worked per person, h. If $\alpha > 0$, it is costly to hire or fire workers.

A. Write down the Bellman's equation associated with this planner's problem. (15 points)

Solution:[6]

$$V(k, n_{-1}) = \max_{c,n,h,k'} \Big[\log c + nA\log(1-h) + \beta V(k', n) \Big]$$

subject to

$$c + k' + \alpha(n - n_{-1})^2 = hk^{\theta}n^{1-\theta} + (1-\delta)k \tag{14}$$

B. Derive a set of first order conditions and envelope conditions that characterize a solution to this problem. Derive the intertemporal Euler equation(s). (10 points)

Solution: First, we define the Lagrangian and take the first order conditions.

$$\mathbf{L} = \log c + nA\log(1-h) + \beta V(k', n) + \lambda\Big[hk^{\theta}n^{1-\theta} + (1-\delta)k - c - k' - \alpha(n - n_{-1})^2\Big]$$

[6] The solution for this question is not unique. That is, any ways of eliminating variables are correct. For example, one can eliminate c as well by substituting the resource constraint into c in the return function.

$$
\begin{aligned}
[c] &\quad : \quad && \frac{1}{c} = \lambda && (15)\\
[n] &\quad : \quad && 0 = A\log(1-h) + \beta V_2(k', n) + \lambda\Big[(1-\theta)hk^\theta n^{-\theta} - 2\alpha(n - n_{-1})\Big] && (16)\\
[h] &\quad : \quad && \frac{nA}{1-h} = \lambda k^\theta n^{1-\theta} && (17)\\
[k'] &\quad : \quad && \lambda = \beta V_1(k', n) && (18)
\end{aligned}
$$

We can eliminate λ by substituting (15) into (16)$\sim$(18).

$$
\begin{aligned}
[n] &\quad : \quad && 0 = A\log(1-h) + \beta V_2(k', n) + \frac{1}{c}\Big[(1-\theta)hk^\theta n^{-\theta} - 2\alpha(n - n_{-1})\Big] && (19)\\
[h] &\quad : \quad && \frac{nA}{1-h} = \frac{k^\theta n^{1-\theta}}{c} && (20)\\
[k'] &\quad : \quad && \frac{1}{c} = \beta V_1(k', n) && (21)
\end{aligned}
$$

Now we derive the envelope conditions,

$$
\begin{aligned}
V_1(k, n_{-1}) = \frac{\partial \mathbf{L}}{\partial k} &= \lambda\Big[\theta h k^{\theta-1} n^{1-\theta} + 1 - \delta\Big]\\
&= \frac{1}{c}\Big[\theta h k^{\theta-1} n^{1-\theta} + 1 - \delta\Big], && (22)\\
V_2(k, n_{-1}) = \frac{\partial \mathbf{L}}{\partial n_{-1}} &= \lambda \cdot 2\alpha(n - n_{-1})\\
&= \frac{2\alpha(n - n_{-1})}{c}, && (23)
\end{aligned}
$$

and advance (22) and (23) by one period. Then we can derive two intertemporal Euler equations: one by combining (19) and (23) and the other by combining (21) and (22). By reviving the time subscripts, we get:

$$
\begin{aligned}
0 &= A\log(1-h_t) + \beta\frac{2\alpha(n_{t+1} - n_t)}{c_{t+1}} + \frac{1}{c_t}\Big[(1-\theta)h_t k_t^\theta n_t^{-\theta} - 2\alpha(n_t - n_{t-1})\Big] && (24)\\
\frac{1}{c_t} &= \beta\frac{1}{c_{t+1}}\Big[\theta h_{t+1} k_{t+1}^{\theta-1} n_{t+1}^{1-\theta} + 1 - \delta\Big] && (25)
\end{aligned}
$$

C. Let $\alpha = 0$. Show that h_t is equal to a constant (that it is not affected by the state of the economy). (10 points)

Solution: Let's rewrite the equations that characterize the optimum. There are 2 intertemporal Euler equations (24) and (25), one intratemporal first-order condition (17), and the resource constraint (14).

Assume $\alpha = 0$, then (24) is simplified as

$$0 = A\log(1-h_t) + \frac{1}{c_t}(1-\theta)h_t k_t^\theta n_t^{-\theta}.$$

Dividing (20) both sides by n_t yields

$$\frac{A}{1-h_t} = \frac{k_t^\theta n_t^{-\theta}}{c_t}.$$

Plugging these together,

$$\begin{aligned} 0 &= A\log(1-h_t) + (1-\theta)h_t \frac{A}{1-h_t}, \\ 0 &= \log(1-h_t) + (1-\theta)\frac{h_t}{1-h_t}. \end{aligned} \tag{26}$$

You can observe that (26) consists only of h_t and parameters. Therefore, optimal h_t is a constant when $\alpha = 0$.

D. Characterize the steady state in this model. (added)

5 Stochastic Growth: Markov Chains

So far, we determined how to solve a sequence and dynamic programming problem (Bellman equation, Neoclassical problem, and Envelope Theorem). As we want to be able to tackle aggregate uncertainty, we start by defining the concept of *Markov Chains*, a key and easy way to include randomness, and show you how a first-order auto regressive (AR(1)) process can be transformed into a first-order Markov chain. Later in class, we will relate our workhorse model with Markov chains and discuss the concept of Business Cycles (short/medium term economic fluctuations).

Stochastic Process A stochastic process (for the discrete time case) is a sequence of random variables $\{z_t : t \in T\}$, where T is the index set or a collection of time periods. (This is a very informal definition, and if we want to be more precise we should talk about the probability space for the stochastic process, as well as the corresponding probability measure). Usually, we are interested in stochastic processes that preserve the recursive nature of the macroeconomic problems. Markov processes have this property and as such are the building blocks of stochastic dynamic programs.

Markov Process A stochastic process is a (first order) Markov process if

$$\Pr[z_{t+1}|z_t, z_{t-1}, z_{t-2}, \ldots] = \Pr[z_{t+1}|z_t].$$

This says that z_t contains all the past relevant information that affects the realization of z_{t+1}. In other words, the entire history of past shocks does not matter. In this sense the process is "memoryless". The above condition is called **Markov property**.

Markov Chain A Markov chain is a mathematical object that characterizes a Markov process when the state space is finite, i.e. when there is a finite number of possible realizations of z_t at each t (later on we will generalize for the case of continuum of states). Formally, a Markov chain is a triplet $\{z, P, \pi_0\}$, where

(a) $z = \{z_1, z_2, \ldots, z_n\} \in \mathbb{R}^n$ is a n-dimensional vector of possible values of z_t in each t

(b) P is an $n \times n$ transition matrix

$$P = \{P_{ij}\} = \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{pmatrix}$$

which records the probabilities of moving from one state to another state in one period. The rows of P are associated with the current states and the columns are associated with tomorrow's states. The following properties apply to the transition matrix

(i) $P_{ij} = \Pr[z_{t+1} = z_j | z_t = z_i]$. This describes the probability of going to "state" j from i.

(ii) $P_{ij} \geq 0$ for all i, j,

(iii) $\sum_{j=1}^{n} P_{ij} = 1$ for all i.

(c) $\pi_0 = \{\pi_0(1), \ldots, \pi_0(n)\}$ is a $n \times 1$ vector recording the initial distribution, satisfying

(i) $\pi_0(i) = \Pr[z_0 = z_i]$,

(ii) $\pi_0(i) \geq 0$ for all i,

(iii) $\sum_{j=1}^{n} \pi_0(i) = 1$ for all i.

Having these definitions, you can see for example that the probability of moving from state i to state j in k periods is

$$\Pr[z_{t+k} = z_j | z_t = z_i] = \sum_{h=1}^{n} P_{ih}^{k-1} P_{hj} = P_{ij}^{k},$$

where P_{ij}^k is the (i, j) element of P^k. You should first try to understand this formula when $k = 2$.

Note that P defines the conditional distribution of z, meaning, given some state $z_0 = z_i$, $P_{i,\cdot}$ describes the probability distribution given that the current state is z_i. In reality, we rarely see these conditional distributions, we only see the unconditional one (wealth distribution, firm size distribution, etc). With a Markov Chain we can easily obtain the unconditional distribution.

Definition: Unconditional Distribution: Remember that z_t can assume n possible values at each t. Let π_t be the unconditional probability of state i being realized in period t, and let π_t be a $n \times 1$ vector recording the probabilities of all possible states in period t, that is

$$\pi_t = \begin{pmatrix} \pi_t(1) \\ \pi_t(2) \\ \vdots \\ \pi_t(n) \end{pmatrix}.$$

The unconditional distribution π_1 will, of course, depend on the initial distribution and on the transition probabilities of states between period $t = 0$ and period $t = 1$ as

$$\pi_1^\intercal = \pi_0^\intercal P,$$

or more explicitly,

$$\begin{aligned}
\begin{pmatrix} \pi_1(1) \\ \pi_1(2) \\ \vdots \\ \pi_1(n) \end{pmatrix}^\intercal &= \begin{pmatrix} \pi_0(1) \\ \pi_0(2) \\ \vdots \\ \pi_0(n) \end{pmatrix}^\intercal \begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1n} \\ P_{21} & P_{22} & \cdots & P_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nn} \end{pmatrix} \\
&= \begin{pmatrix} \pi_0(1)P_{11} + \ldots + \pi_0(n)P_{n1} \\ \pi_0(1)P_{12} + \ldots + \pi_0(n)P_{n2} \\ \cdots \\ \pi_0(1)P_{1n} + \ldots + \pi_0(n)P_{nn} \end{pmatrix}.
\end{aligned}$$

Generalizing we have

$$\pi_{t+1}^\intercal = \pi_t^\intercal P = \pi_0^\intercal P^{t+1}.$$

The vector π_t of unconditional probabilities satisfies

(i) $\pi_t(i) = \Pr[z_t = z_i]$,

(ii) $\pi_t(i) \geq 0$ for all i and all t,

(iii) $\sum_{i=1}^n \pi_t(i) = 1$ for all t.

5.1 Stationary Distributions of Markov Chains

The unconditional probabilities derived in the previous section are time-dependent. As you already know, we are usually interested in the properties of stationary equilibria in macro models. This is one of the motivations for studying **stationary distributions** of Markov processes. For the sake of time, I will briefly address the issues of (1) existence and (2) uniqueness.

5.1.1 Existence

A distribution is stationary if it is time-invariant, that is, if

$$\pi_{t+1} = \pi_t = \pi.$$

This must imply that a stationary distribution π satisfies

$$\pi^\intercal = \pi^\intercal P.$$

Note that π is simply a fixed point of the operator $P : X^{n^2} \to X^{n^2}$ where X^{n^2} denotes the space of $n \times n$-dimensional probability matrices. The above equation can be rewritten as

$$(I - P^\intercal)\pi = 0.$$

This equation has a non-zero solution only if the matrix $(I - P^\intercal)$ is singular, i.e. its determinant is zero. This means that the rows (or columns) of $(I - P^\intercal)$ are linearly dependent. Therefore, you can determine at most $n-1$ elements of π. The remaining element is pinned down using $\sum_{i=1}^{n} \pi(i) = 1$. You can also notice that the invariant distribution is simply the eigenvector of the linear function $P^\intercal$ associated with the unit eigenvalue.

Because of the natural properties of a Markov Chain (all probabilities are non negative and they add up to one), it always exists at least one stationary distribution.

5.1.2 Uniqueness

Definition 5.1. Let π be the unique vector that satisfies $(I - P^\intercal)\pi = 0$. If for all initial distributions π_0 it is true that $\lim_{t\to\infty} {P^\intercal}^t \pi_0 = \pi$ we say that the Markov chain is asymptotically stationary (or ergodic) with a unique invariant distribution π.

Definition 5.2. Every Markov chain can be divided into transitory (or transient) states and ergodic sets. A state is transient if there is a positive probability of leaving and never returning to that state. An ergodic set E is the smallest subset of the state space for which $\Pr[z_{t+1} \in E | z_t \in E] = 1$, that is, the system never leaves E once it reaches E.

Two theorems show that under certain conditions the invariant distribution is unique.[7]

Theorem 5.1. *Let P be a transition matrix such that $P_{ij} > 0$ for all i, j. Then P has a unique stationary distribution, and the process is asymptotically stationary.*

Theorem 5.2. *Let P be a transition matrix such that $[P^t]_{ij} > 0$ for all i, j for some value of $t \geq 1$. Then P has a unique stationary distribution, and the process is asymptotically stationary.*

Both theorems indicate that as long as there is a positive probability of moving to any other state in one (or t) steps, then the stationary distribution is unique.

5.2 Exercises

For each of the examples below, determine the ergodic set(s) and the invariant distribution(s) π.

1.

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$$

[7] See Sargent and Ljunqvist, Ch.2

2.

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0.25 & 0.75 \end{pmatrix}$$

3.

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

4.

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

5. i.i.d. process where $\Pr[z_{t+1} = z_j | z_t = z_i] = \Pr[z_{t+1} = z_j]$

$$P = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix}$$

6. serially correlated process

$$P = \begin{pmatrix} \frac{\epsilon}{2} & 1-\frac{\epsilon}{2} \\ 1-\frac{\epsilon}{2} & \frac{\epsilon}{2} \end{pmatrix}$$

7.

$$P = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0 & 0.2 & 0.8 \\ 0 & 0.7 & 0.3 \end{pmatrix}$$

8.

$$P = \begin{pmatrix} 0.8 & 0.2 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0 & 0.4 & 0.6 \end{pmatrix}$$

9.

$$P = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

10.

$$P = \begin{pmatrix} 0.5 & 0.5 \\ 0 & 0.1 \end{pmatrix}$$

11. Consider two states: alive and dead with the probability of staying alive in a given period being p. Compute the life expectancy.

12. According to the NBER, between 1854 and 2001 the US economy experienced about 31 economic cycles. The duration of a typical cycle was 56 months, with the economic contraction lasting on average 17 months and the economic expansion lasting on average 38 months. Assume the US economy can be in three different states each month: boom (z_1), recession (z_2) and depression (z_3), according to the following transition matrix:

$$P = \begin{pmatrix} 0.8 & 0.2 & 0 \\ 0.5 & 0.4 & 0.1 \\ 0.3 & 0.5 & 0.2 \end{pmatrix}.$$

That is, if the economy is in a boom in the current month, there is a good chance of remaining in a boom next month. However, there is a 20 percent probability of entering a recession and no chance of entering a severe depression in just one month. On the other hand, if the economy is in recession, it could remain in the recession for one more month, it could go to a depression, but it could also recover. Finally, if the economy is in the middle of a depression, it could go to a less severe situation and eventually recover. The probability of staying in depression is relatively low (only 20 percent in this case) because depressions are rare events. What percentage of time is the US experiencing a boom or a recession?

6 Appendix A: Envelope Theorem for Constrained Maximization

Here we provide a proof for Theorem 4.1.

Proof. Let $x^*(y)$ and $\lambda^*(y)$ be the corresponding policy function and Lagrange multiplier associated with this problem. For simplicity, I will assume that these functions are differentiable. Then, the Kuhn-Tucker conditions of this problem are:

$$\frac{\partial L}{\partial x} = f_1(x^*(y), y) + \lambda^*(y) g_1(x^*(y), y) = 0 \tag{27}$$

$$\lambda^*(y) g(x^*(y), y) = 0 \tag{28}$$

Given this, we can write down:

$$V(y) = f(x^*(y), y) + \lambda^*(y) g(x^*(y), y)$$

Differentiating both sides we have:

$$V'(y) = f_1(x^*(y), y)\frac{dx^*}{dy} + f_2(x^*(y), y) + \frac{d\lambda^*}{dy} g(x^*(y), y) + \lambda^*(y)\left(g_1(x^*(y), y)\frac{dx^*}{dy} + g_2(x^*(y), y)\right)$$

Combining terms

$$V'(y) = (f_1(x^*(y),y)+\lambda^*(y)g_1(x^*(y),y))\frac{dx^*}{dy}+f_2(x^*(y),y)+\frac{d\lambda^*}{dy}g(x^*(y),y)+\lambda^*(y)g_2(x^*(y),y)$$

Notice that using (27), we can rewrite the last expression as:

$$V'(y) = f_2(x^*(y),y) + \frac{d\lambda^*}{dy}g(x^*(y),y) + \lambda^*(y)g_2(x^*(y),y) \tag{29}$$

Finally, by definition:

$$\frac{d\lambda^*}{dy}g(x^*(y),y) = \left(\lim_{\epsilon\to 0}\frac{\lambda(y+\epsilon)-\lambda(y)}{\epsilon}\right)g(x^*(y),y)$$

If the restriction g does not bind at a given y, then $\lambda(y+\epsilon) = \lambda(y) = 0$ for a ϵ small enough. If the restriction is binding $g(x^*(y),y) = 0$, so in any case:

$$\frac{d\lambda^*}{dy}g(x^*(y),y) = \lim_{\epsilon\to 0}\frac{0}{\epsilon} = 0$$

So, we can write (29) as

$$V'(y) = f_2(x^*(y),y) + \lambda^*(y)g_2(x^*(y),y) = \frac{\partial L}{\partial y}. \tag{30}$$

□

This proof is standard, although it assumes differentiability of policy function and Lagrange multiplier.[8]

7 Appendix C

7.1 AR(1) to Markov Chain

Suppose productivity (z_t) evolves according to the following AR(1) process

$$z_{t+1} = \rho z_t + \epsilon_{t+1}.$$

where $\epsilon_{t+1} \overset{i.i.d}{\sim} N(0,\sigma_\epsilon^2)$ and $\rho \in (0,1)$. From this process, we want to calibrate a two-state Markov chain of the form $z_t \in \{-\sigma,\sigma\}$, where the transition matrix for z is given by

$$P = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}$$

with $q \in [0,1)$. By calibrating the Markov chain we mean that we need to find values for σ and q as functions of parameters of the AR(1) process, such as the resulting invariant distribution of the Markov chain is consistent with the unconditional mean, unconditional variance, and first-order autocovariance of the AR(1) process.

[8] Another way of proving it is using the Benveniste and Scheinkman theorem . Then, we need to assume that g is differentiable and concave. Finally, you just need to define $W(x,y)$ as the Lagrangian and the theorem gives you exactly the result

7.1.1 Final (2006) Question 3

We are told that z_t follows a Markov chain with $z_{t+1} \in \{-\sigma, \sigma\}$, $\sigma > 0$ and transition matrix

$$P = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}.$$

(a) Suppose you are told that the unconditional variance of z is equal to a and the first order autocovariance of z is equal to b. Express calibrated values for σ and q in terms of a and b.

Solution: In the previous section we found that the unconditional variance of the Markov chain is given by σ^2. In order to obtain a calibrated value for this parameter as a function of the parameters of the AR(1) process, we must set the unconditional variances of the two processes equal to each other, implying $\sigma^2 = a$ thus $\sigma = \sqrt{a}$.

The first-order autocovariance resulting from this same Markov chain is given by $\gamma_1 = \sigma^2(2q-1)$. Thus, in order to pin down q, we must equate b with the first-order autocovariance implied by the Markov chain, i.e. $q = \frac{b}{2a} + \frac{1}{2}$.