ECON 6130: Dynamic Programming

Mathieu Taschereau-Dumouchel

Dynamic Programming

Through this section, we will be interested in problems of the form

$$v(x) = \max_{y \in \Gamma(x)} \{F(x, y) + \beta v(y)\}$$

where

- x is the set of state variables
- y is the set of controls
- F is the period return function
- Γ is the constraint set

For the neoclassical growth model

- x corresponds to k
- y corresponds to k'
- $F(k, k') = U(f(k) - k')$
- $\Gamma(k) = \{k' \in \mathbb{R} : 0 \leq k' \leq f(k)\}$

Dynamic Programming

Define operator T:

$$(Tv)(x) \equiv \max_{y \in \Gamma(x)} \{F(x,y) + \beta v(y)\}$$

T takes a function v as input and spits out a new function Tv

Using this notation, a solution v^* to our original functional equation is a *fixed point* of the operator T:

$$v^* = Tv^*$$

Questions:

1. Under what conditions does T have a fixed point v^*?
2. Under what conditions is v^* unique?
3. Under what conditions does the sequence $\{v_n\}_{n=0}^{\infty}$ defined recursively by $v_{n+1} = Tv_n$ and v_0 is a guess converges to v^*.

Dynamic Programming

Define operator T:

$$(Tv)(x) \equiv \max_{y \in \Gamma(x)} \{F(x, y) + \beta v(y)\}$$

T takes a function v as input and spits out a new function Tv

Using this notation, a solution v^* to our original functional equation is a *fixed point* of the operator T:

$$v^* = Tv^*$$

Questions:

1. Under what conditions does T have a fixed point v^*?
2. Under what conditions is v^* unique?
3. Under what conditions does the sequence $\{v_n\}_{n=0}^{\infty}$ defined recursively by $v_{n+1} = Tv_n$ and v_0 is a guess converges to v^*.

Answer: Contraction mapping theorem

Metric space

Definition 1

A metric space is a set S and a function, called distance, $d : S \times S \to \mathbb{R}$ such that for all $x, y, z \in S$

1. $d(x, y) \geq 0$
2. $d(x, y) = 0$ *if and only if* $x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, z) \leq d(x, y) + d(y, z)$

Definition 2

A sequence $\{x_n\}_{n=0}^{\infty}$ with $x_n \in S$ for all n is said to converge to $x \in S$ if for every $\epsilon > 0$ there exists a $N_\epsilon \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n \geq N_\epsilon$. In this case we write $\lim_{n\to\infty} x_n = x$.

Metric space

Definition 3

A sequence $\{x_n\}_{n=0}^{\infty}$ *with* $x_n \in S$ *for all* n *is said to be a Cauchy sequence if for every* $\epsilon > 0$ *there exists a* $N_\epsilon \in \mathbb{N}$ *such that* $d(x_n, x_m) < \epsilon$ *for all* $n, m \geq N_\epsilon$*.*

Definition 4

A metric space (S,d) is complete if every Cauchy sequence $\{x_n\}_{n=0}^{\infty}$ *with* $x_n \in S$ *for all* n *converges to some* $\underline{x \in S}$*.*

Example: Lex $X \subseteq \mathbb{R}^l$ and $S = C(X)$ be the set of all continuous and bounded functions $f : X \to \mathbb{R}$. Define the distance $d : C(X) \times C(X) \to \mathbb{R}$ as $d(f, g) = \sup_{x \in X} |f(x) - g(x)|$. This distance is called the sup-norm. Then (S, d) is a complete metric space. (The proof is in SLP)

Contraction mapping theorem

Definition 5

Let (S, d) be a metric space and $T : S \to S$. The function T is a contraction mapping if there exists a number $\beta \in (0, 1)$ satisfying

$$d(Tx, Ty) \leq \beta d(x, y) \text{ for all } x, y \in S$$

β is called the modulus of the contraction.

Theorem 1 (Contraction Mapping Theorem)

Let (S, d) be a complete metric space and suppose that $T : S \to S$ is a contraction mapping with modulus β. Then

1. *the operator T has exactly one fixed point $v^* \in S$*
2. *for any $v_0 \in S$ and any $n \in \mathbb{N}$ we have*

$$d(T^n v_0, v^*) \leq \beta^n d(v_0, v^*)$$

Proof of the first part of CMT (lemma)

Lemma 1

Let (S,d) be a metric space and $T : S \to S$. If T is a contraction mapping, then T is continuous.

Proof.

We need to show: for all $s_0 \in S$ and all $\epsilon > 0$ there exists a $\delta(\epsilon, s_0)$ such that if $s \in S$ and $d(s, s_0) < \delta(\epsilon, s_0)$, then $d(Ts, Ts_0) < \epsilon$. Fix arbitrary $s_0 \in S$ and $\epsilon > 0$ and pick $\delta(\epsilon, s_0) = \epsilon$. Then

$$d(Ts, Ts_0) \leq \beta d(s, s_0) < \beta\delta(\epsilon, s_0).$$

□

Proof of contraction mapping theorem (part 1)

Proof of the first part of CMT:
Start with an arbitrary $v_0 \in S$ an consider the sequence $v_n = T^n v_0$. Our candidate for a fixed point is $v^* = \lim_{n\to\infty} v_n$.
Step 1: Show that $v_n \to v^* \in S$.
Since T is a contraction:

$$\begin{aligned} d(v_{n+1}, v_n) &= d(Tv_n, Tv_{n-1}) \leq \beta d(v_n, v_{n-1}) \\ &\leq \beta d(Tv_{n-1}, Tv_{n-2}) \leq \beta^2 d(v_{n-1}, v_{n-2}) \\ &\leq \cdots \leq \beta^n d(v_1, v_0) \end{aligned}$$

Proof of contraction mapping theorem (part 1)

We now use the triangle inequality. For any $m > n$:

$$\begin{aligned}
d(v_m, v_n) &\le d(v_m, v_{m-1}) + d(v_{m-1}, v_n) \\
&\le d(v_m, v_{m-1}) + d(v_{m-1}, v_{m-2}) + \dots d(v_{n+1}, v_n) \\
&\le \beta^{m-1} d(v_1, v_0) + \beta^{m-2} d(v_1, v_0) + \dots \beta^n d(v_1, v_0) \\
&= \beta^n(\beta^{m-n-1} + \cdots + \beta + 1) d(v_1, v_0) \\
&\le \frac{\beta^n}{1-\beta} d(v_1, v_0)
\end{aligned}$$

Therefore, the sequence $\{v_n\}_{n=0}^{\infty}$ is a Cauchy sequence. Since (S, d) is a complete metric space, $\{v_n\}_{n=0}^{\infty}$ converges in S. We have shown that

$$v_n \to v^* \in S$$

Proof of contraction mapping theorem (part 1)

Step 2: We now establish that v^* is a fixed point of T:

$$Tv^* = T(\lim_{n\to\infty} v_n) = \lim_{n\to\infty} T(v_n) = \lim_{n\to\infty} v_{n+1} = v^*$$

Step 3: We now prove that the fixed point is unique. Suppose there is another $\hat{v} \in S$ such that $\hat{v} = T\hat{v}$ and $\hat{v} \neq v^*$. Then there exists $a > 0$ such that $d(\hat{v}, v^*) = a$. But then

$$0 < a = d(\hat{v}, v^*) = d(T\hat{v}, Tv^*) \leq \beta d(\hat{v}, v^*) = \beta a$$

which is a contradiction.

Proof of contraction mapping theorem (part 2)

We proceed by induction. For $n = 0$, the claim holds. Now suppose that

$$d(T^k v_0, v^*) \leq \beta^k d(v_0, v^*)$$

We need to show that

$$d(T^{k+1} v_0, v^*) \leq \beta^{k+1} d(v_0, v^*)$$

But

$$d(T^{k+1} v_0, v^*) = d(T(T^k v_0), Tv^*) \leq \beta d(T^k v_0, v^*) \leq \beta^{k+1} d(v_0, v^*)$$

which complete the proof of the contraction mapping theorem. $\square$

Blackwell's theorem

The CMT is extremely powerful. However, it is sometimes hard to show that an operator is a contraction.

Theorem 2 (Blackwell)

Let $X \subseteq \mathbb{R}^L$ and $B(X)$ be the space of bounded functions $f : X \to \mathbb{R}$ with the distance being the sup-norm. Let $T : B(X) \to B(X)$ be an operator satisfying:

1. *Monotonicity: If $f, g \in B(X)$ are such that $f(x) \leq g(x)$ for all $x \in X$, then $(Tf)(x) \leq (Tg)(x)$ for all $x \in X$*
2. *Discounting: Let the function $f + a$, for $f \in B(X)$ and $a \in \mathbb{R}_+$ be defined by $(f + a)(x) = f(x) + a$. There exists $\beta \in (0, 1)$ such that for all $f \in B(X), a \geq 0$ and all $x \in X$*
$$[T(f + a)](x) \leq [Tf](x) + \beta a$$

then T is a contraction mapping with modulus β.

Blackwell's theorem

Proof.

If $f(x) \leq g(x)$ for all $x \in X$ we write $f \leq g$. For any $f, g \in B(X)$, $f \leq g + d(f,g)$, where d is the sup-norm. The monotonicity and discounting imply that

$$Tf \leq T\left(g + d(f,g)\right) \leq Tg + \beta d(f,g)$$

Reversing the roles of f and g gives, by the same logic,

$$Tg \leq Tf + \beta d(f,g)$$

Combining these inequalities, we find $d(Tf, Tg) \leq \beta d(f,g)$ so T is a contraction. □

Application to the neoclassical growth model

Can these theorems help with the growth model?

- Metric space $(B[0,\infty), d)$ the space of bounded function with d being the sup-norm.
- Define an operator

$$(Tv)(k) = \max_{0 \leq k' \leq f(k)} \{U(f(k) - k') + \beta v(k')\}$$

- Verify that T maps $B[0,\infty)$ into itself: Take v to be bounded, since U is bounded by assumption, then Tv is also bounded.

Application to the neoclassical growth model

- Monotonicity: Suppose $v \le w$. Let $g_v(k)$ denote an optimal policy (need not be unique) corresponding to v. Then for all $k \in [0, \infty)$

$$\begin{aligned} Tv(k) &= U(f(k) - g_v(k)) + \beta v(g_v(k)) \\ &\le U(f(k) - g_v(k)) + \beta w(g_v(k)) \\ &\le \max_{0 \le k' \le f(k)} \{U(f(k) - k') + \beta w(k')\} \\ &= Tw(k) \end{aligned}$$

- Discounting:

$$\begin{aligned} T(v + a)(k) &= \max_{0 \le k' \le f(k)} \{U(f(k) - k') + \beta(v(k') + a)\} \\ &= \max_{0 \le k' \le f(k)} \{U(f(k) - k') + \beta(v(k')\} + \beta a \\ &= Tv(k) + \beta a \end{aligned}$$

Application to the neoclassical growth model

We have shown that the neoclassical model with bounded utility satisfies Blackwell's conditions and is therefore a contraction mapping with modulus β. Hence there is a unique fixed point to the functional equation that can be computed from any starting guess v_0 by repeated application of the operator T.

Theorem of the maximum - Preliminaries

We're interested in problem of the form

$$h(x) = \max_{y \in \Gamma(x)} f(x, y)$$

Define

$$G(x) = \{y \in \Gamma(x) : f(x, y) = h(x)\}$$

Intuitively, what is $G(x)$?

Question: What can we say about the properties of h and G?

Definition 6

Let X, Y be arbitrary sets. A correspondence $\Gamma : X \to Y$ maps each element $x \in X$ into a subset $\Gamma(x)$ of Y.

Theorem of the maximum - Preliminaries

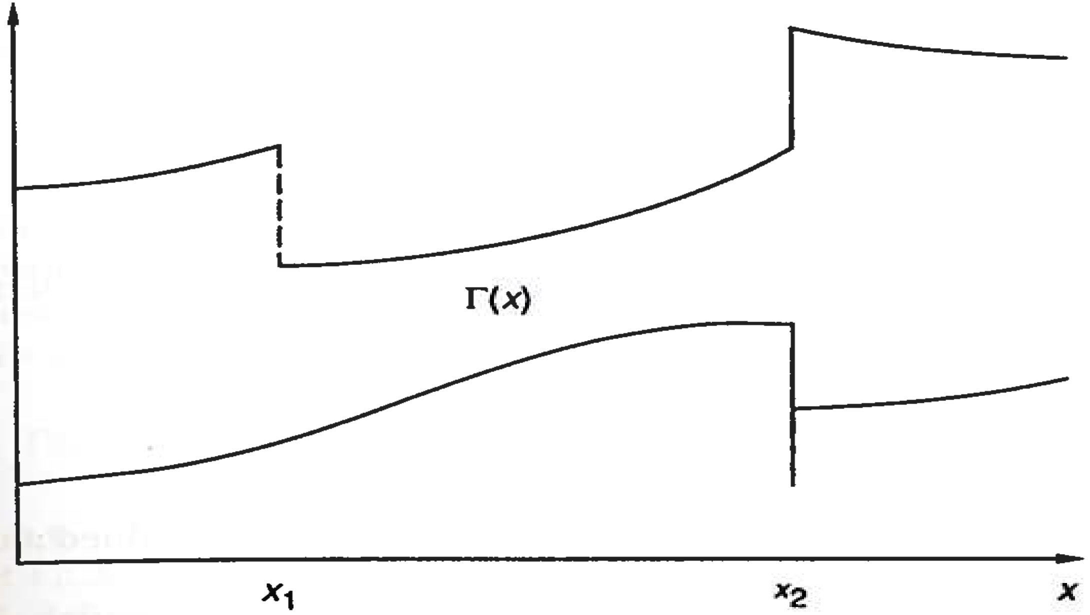
Definition 7

A correspondence $\Gamma : X \to Y$ is lower-hemicontinuous at a point x if $\Gamma(x) \neq \emptyset$ and if for every $y \in \Gamma(x)$ and every sequence $\{x_n\}$ in X converging to $x \in X$ there exists $N \geq 1$ and a sequence $\{y_n\} \in Y$ converging to y such that $y_n \in \Gamma(x_n)$ for all $n \geq N$.

Definition 8

A compact-valued correspondence $\Gamma : X \to Y$ is upper-hemicontinuous at a point x if $\Gamma(x) \neq \emptyset$ and if for all sequences $\{x_n\}$ in X converging to $x \in X$ and all sequences $\{y_n\}$ in Y such that $y_n \in \Gamma(x_n)$ for all n, there exists a convergent subsequence of $\{y_n\}$ that converges to some $y \in \Gamma(x)$.

Note: a single-valued correspondence (i.e. a function) that is upper-hemicontinuous is continuous.



Γ(x)
x1
x2
x

Theorem of the maximum

Definition 9

A correspondence $\Gamma : X \to Y$ is continuous if it is both upper-hemicontinuous and lower-hemicontinuous.

$$h(x) = \max_{y \in \Gamma(x)} f(x, y)$$

$$G(x) = \{y \in \Gamma(x) : f(x, y) = h(x)\}$$

Theorem 3 (Theorem of the maximum)

Let $X \subseteq \mathbb{R}^L$ an $Y \subseteq \mathbb{R}^M$, let $f : X \times Y \to \mathbb{R}$ be a continuous function, and let $\Gamma : X \to Y$ be a compact-valued and continuous correspondence. Then $h : X \to \mathbb{R}$ is continuous and $G : X \to Y$ is nonempty, compact-valued and upper-hemicontinuous.

The proof is in SLP.

Application to the neoclassical growth model

$$(Tv)(k) = \max_{0 \le k' \le f(k)} \{U(f(k) - k') + \beta v(k')\}$$

- $x = k, y = k', X = Y = \mathbb{R}_+$
- $f(x, y) = U(f(x) - y) + \beta v(y)$
- $\Gamma : X \to Y$ is given by $\Gamma(x) = \{y \in \mathbb{R}_+ | 0 \le y \le f(x)\}$

Suppose that v is continuous, then the theorem of the maximum implies that $Tv(\cdot)$ is a continuous function and that optimal policy $g(\cdot)$ is an uhc correspondence. If $g(\cdot)$ is a function, then it is continuous.

Principle of optimality

Functional equation (FE)

$$v(x) = \sup_{y \in \Gamma(x)} \{F(x, y) + \beta v(y)\}$$

has a unique solution v^* which is approached from any initial guess v^0.

Sequential problem (SP)

$$w(x_0) = \sup_{\{x_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1})$$

subject to

$$x_{t+1} \in \Gamma(x_t)$$
$$x_0 \in X \text{ given}$$

Questions:

1. When do $v = w$?
2. When is $\{x_{t+1}\}_{t=0}^{\infty}$ the same as $y = g(x)$?

Principle of optimality - Preliminaries

Define some notation

- Let X be the set of possible values that the state x can take
- Correspondence $\Gamma : X \to X$ describes the feasible set of next period's state y, given that today's state is x
- Graph of Γ, A is defined as

$$A = \{(x, y) \in X \times X : y \in \Gamma(x)\}$$

- Period return function $F : A \to \mathbb{R}$
- Fundamentals of the analysis are (X, F, β, Γ). For neoclassical growth model F and β describe preferences and X, Γ describe technology.
- Any sequence of states $\{x_t\}_{t=0}^{\infty}$ is a plan
- For a given x_0, the set of feasible plans $\Pi(x_0)$ is $\Pi(x_0) = \{\{x_t\}_{t=1}^{\infty} : x_{t+1} \in \Gamma(x_t)\}$

Principle of optimality - Preliminaries

We need some assumptions for the Principle of Optimality

Assumption 1 (1)

$\Gamma(x)$ *is nonempty for all* $x \in X$

Assumption 2 (2)

For all initial x_0 *and all feasible plans* $\bar{x} \in \Pi(x_0)$

$$\lim_{n \to \infty} \sum_{t=0}^{n} \beta^t F(x_t, x_{t+1})$$

exists (although it may be $+\infty$ *or* $-\infty$*)*

Principle of optimality

Theorem 4 (Principle of optimality)

Suppose that (X, Γ, F, β) *satisfy the two previous assumptions. Then*

1. *the function* w *satisfies the functional equation (FE)*
2. *if for all* $x_0 \in X$ *and all* $x \in \Pi(x_0)$ *a solution* v *to the functional equation (FE) satisfies*

$$\lim_{n \to \infty} \beta^n v(x_n) = 0$$

   then $v = w$.

In words

- Supremum function from SP solves the functional equation
- Result 2 is key. It states a condition under which a solution to FE is a solution to SP which is what we are looking for.

Principle of optimality

Equivalence of policies:

Theorem 5 (Principle of optimality)

Suppose that (X, Γ, F, β) *satisfy the two previous assumptions.*

1. *Let* $\bar{x} \in \Pi(x_0)$ *be a feasible plan that attains the supremum in SP. Then for all* $t \geq 0$
$$w(\bar{x}_t) = F(\bar{x}_t, \bar{x}_{t+1}) + \beta w(\bar{x}_{t+1})$$
2. *Let* $\hat{x} \in \Pi(x_0)$ *be a feasible plan satisfying, for all* $t \geq 0$
$$w(\hat{x}_t) = F(\hat{x}_t, \hat{x}_{t+1}) + \beta w(\hat{x}_{t+1})$$
and
$$\lim_{t \to \infty} \sup \beta^t w(\hat{x}_t) \leq 0$$
then $\hat{x}$ *attains the supremum in SP for* x_0.

Dynamic Programming with Bounded Returns

Functional equation:

$$v(x) = \sup_{y \in \Gamma(x)} \{F(x,y) + \beta v(y)\}$$

with associated operator $T : C(X) \to C(X)$

$$(Tv)(x) = \max_{y \in \Gamma(x)} \{F(x,y) + \beta v(y)\}$$

We will make a number of stronger assumptions on (X, F, β, Γ) to be able to characterize v and g where:

$$g(x) = \{y \in \Gamma(x) : v(x) = F(x,y) + \beta v(y)\}$$

is the policy correspondence associated with v.

DP with Bounded Returns - Uniqueness of solution

Assumption 3 (3)

X is a convex subset of $\mathbb{R}^L$ and the correspondence $\Gamma : X \to X$ is nonempty, compact-valued and continuous.

Assumption 4 (4)

The function $F : A \to \mathbb{R}$ is continuous and bounded, and $\beta \in (0, 1)$.

Note that these Assumptions imply Assumptions 1 and 2.

Theorem 6

Under Assumptions 3 and 4 the operator T maps $C(X)$ into itself. T has a unique fixed point v and for all $v_0 \in C(X)$

$$(T^n v_0, v) \leq \beta^n d(v_0, v)$$

Furthermore, the policy correspondence g is compact-valued and upper-hemicontinuous.

DP with Bounded Returns - Monotonicity of value function

Assumption 5 (5)

For fixed y, $F(\cdot, y)$ is strictly increasing in each of its L components.

Assumption 6 (6)

Γ *is monotone in the sense that* $x \leq x'$ *implies* $\Gamma(x) \subseteq \Gamma(x')$.

Theorem 7

Under Assumptions 3 to 6 the unique fixed point v of T is strictly increasing.

DP with BR - Strict concavity of v and unique policy

Assumption 7 (7)

F is strictly concave: for all $(x, y), (x', y') \in A$ and $\theta \in (0, 1)$

$$F[\theta(x, y) + (1 - \theta)(x', y')] \geq \theta F(x, y) + (1 - \theta)F(x', y')$$

Assumption 8 (8)

Γ is convex in the sense that for $\theta \in [0, 1]$, $x, x' \in X$, $y \in \Gamma(x)$, $y' \in \Gamma(x')$ then

$$\theta y + (1 - \theta)y' \in \Gamma(\theta x + (1 - \theta)x')$$

Theorem 8

Under Assumption 3-4 and 7-8 the unique fixed point v is strictly concave and the optimal policy g is a single-valued continuous function.

DP with BR - Differentiability of value function

Assumption 9 (9)

F is continuously differentiable.

Theorem 9 (Benveniste-Scheinkman or Envelope Theorem)

Under assumption 3-4 and 7-9 if $x_0 \in int(X)$ and $g(x_0) \in int(\Gamma(x_0))$, then the unique fixed point v is continuously differentiable at x_0 with

$$\frac{\partial v(x_0)}{\partial x_0} = \frac{\partial F(x_0, g(x_0))}{\partial x_0}$$

All the proofs are in SLP.

Solving Bellman equations with Benveniste-Scheinkman

We have the functional equation

$$v(k) = \max_{0 \leq k' \leq f(k)} U(f(k) - k') + \beta v(k')$$

Taking the FOC with respect to k' gives:

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

Then with Benveniste-Scheinkman

$$v'(k) = U'(f(k) - g(k))f'(k)$$

and hence

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

which is the Euler equation we found earlier.

Stochastic growth model - Markov process

Most of what we've done works in a stochastic environment as long as we can summarize the state of the world in a simple way.

Here we specify a specific structure to uncertainty that makes our models tractable: discrete time, discrete state, time homogeneous Markov processes.

- Let
$$\pi(j|i) = \text{prob}(s_{t+1} = j|s_t = i)$$
Conditional probabilities of s_{t+1} only depend on realization of s_t not s_{t-1} or other past realization.
- *Time homogeneity* means that π is not indexed by time

Stochastic growth model - Markov process

Given that $s_{t+1} \in S$ and $s_t \in S$ and S is a finite set, the distribution $\pi(\cdot, \cdot)$ is an $N \times N$-matrix of the form

$$\pi = \begin{pmatrix} \pi_{11} & \dots & \pi_{1j} & \dots & \pi_{1N} \\ \vdots & & \vdots & & \vdots \\ \pi_{i1} & \dots & \pi_{ij} & \dots & \pi_{iN} \\ \vdots & & \vdots & & \vdots \\ \pi_{N1} & \dots & \pi_{Nj} & \dots & \pi_{NN} \end{pmatrix}$$

- Generic element: $\pi_{ij} = \pi(j|i) = \text{prob}(s_{t+1} = j | s_t = i)$.
- Since $\pi_{ij} \geq 0$ and $\sum_j \pi_{ij} = 1$ for all i, matrix π is called a *stochastic matrix*

Stochastic growth model - Markov process

Dynamics of the probability distribution

- Suppose probability distribution over states today is given by the N-dimensional column vector $P_t = (p_t^1, \dots, p_t^N)^T$ with $\sum_i p_t^i = 1$.
- Probability of being in state j tomorrow is

$$p_{t+1}^j = \sum_i \pi_{ij} p_t^i$$

or, in compact form

$$P_{t+1} = \pi^T P_t$$

Stochastic growth model - Markov process

Stationary distribution

- A *stationary* distribution Π of the Markov chain π is

$$\Pi = \pi^T \Pi$$

- A Markov process π has at least one stationary Π: the eigenvector (normalized to 1) associated with the eigenvalue $\lambda = 1$ of π^T.
- If only one such eigenvalue exists, then unique stationary distribution. If more than one unit eigenvalue, then there are multiple stationary distributions.
- If s_t is a Markov chain, we have

$$\pi(s^{t+1}) = \pi(s_{t+1}|s_t) \times \pi(s_t|s_{t-1}) \times \ldots \pi(s_1|s_0) \times \Pi(s_0)$$

Stochastic growth model - Markov process

Suppose

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

for some $p \in (0,1)$. Unique invariant distribution is $\Pi(s) = 1/2$ for both s.

Suppose

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

then any distribution over the two states is an invariant distribution.

Stochastic growth model

- Technology
$$y_t = e^{z_t} F(k_t, n_t)$$
where z_t is a technology shock that has unconditional mean 0 and follows a N-state Markov chain with state space $Z = \{z_1, z_2, \dots, z_N\}$ and transition matrix $\pi = (\pi_{ij})$. Let Π denote stationary distribution.
- Evolution of capital stock $k_{t+1} = (1-\delta)k_t + i_t$
- Resource constraint $y_t = c_t + i_t$
- Preferences
$$E_0 \sum_{t=0}^{\infty} \beta^t U(c_t) = \sum_{t=0}^{\infty} \sum_{z^t \in Z^t} \beta^t \pi(z^t) U(c_t(z^t))$$
- Endowment: initial capital k_0 and one unit of time.
- Information: z_t is publicly observable. $z_0 \sim \Pi$.

Stochastic growth model

We can use our new cool tools to solve this model.

- State variables (k, z)
- Control variable k'
- Bellman equation

$$v(k,z) = \max_{k'} \left\{ U(e^z F(k,1) + (1-\delta)k - k') + \beta \sum_{z'} \pi(z'|z) v(k', z') \right\}$$

subject to:

$$0 \leq k' \leq e^z F(k,1) + (1-\delta)k$$

Stochastic growth model

An important part of output fluctuations is coming from labor.

- Add labor-leisure choice: $U(c_t, 1 - n_t)$
- New Bellman equation

$$v(k,z) = \max_{k',n}\{U(e^z F(k,n) + (1-\delta)k - k', 1-n) + \beta \sum_{z'} \pi(z'|z) v(k',z')\}$$

subject to:

$$0 \leq k' \leq e^z F(k,n) + (1-\delta)k, 0 \leq n \leq 1$$

- This is the benchmark model of modern business cycle research. See: Cooley and Prescott: Economic Growth and Business Cycles, in Frontiers of Business Cycle Research, edited by Thomas F. Cooley (1995).

Stochastic growth model

Solving the model

- *Intra*temporal optimality condition

$$e^z F_n(k,n) = \frac{U_l(c,1-n)}{U_c(c,1-n)}$$

- *Inter*temporal optimality condition

$$U_c(c,1-n) = \beta \sum_{z'} \pi(z'|z) v'(k',z')$$

- Envelope condition

$$v'(k,z) = (e^z F_k(k,n) + 1 - \delta) U_c(c,1-n)$$

Combining:

$$U_c(c,1-n) = \beta \sum_{z'} \pi(z'|z)(e^{z'} F_k(k',n') + 1 - \delta) U_c(c',1-n')$$

Calibration

Purpose: choose (or estimate) parameters of the model so that it can be used for quantitative analysis of real world and counterfactual analysis.

Idea of calibration

1. Choose a set of empirical facts that the model should match
2. Choose parameters so that equilibrium of model matches the facts

Note: fact that model fits these facts can not be used as claim of success. Evaluation of success has to be on other dimensions.

We will calibrate a simple version of the deterministic neoclassical model with population and technology growth.

Calibration

- Functional forms

$$U(c) = \frac{c^{1-\sigma} - 1}{1-\sigma}$$
$$F(K, N) = K^{\alpha} \left((1+g)^t N\right)^{1-\alpha}$$

- Parameters: Technology (α, δ, g), Demographics n, Preferences (β, σ)
- Empirical targets: Choose parameters such that balanced growth path (BGP) of model matches long-run average facts for the U.S. economy.
- Need to decide on period length. Take period to be one year.

Main facts about long-run growth

Kaldor (1959) popularized the following six stylized facts concerning long run economic growth

1. Output per capita, Y/N, grows at a constant rate
2. The capital to labor ratio, K/N, grows at constant rate
3. The interest rate, R, is fairly constant
4. The output to capital ratio, Y/K, is fairly constant
5. The share of value added going to labor and capital are fairly constant
6. There are wide dispersion in Y_i/N_i across countries

Calibration

Parameters directly taken from long run averages in the data

- Population growth rate in model is n, in data $n = 1.1\%$
- Growth rate of per capita GDP in model is g, in data $g = 1.8\%$

Exploiting BCG relationships

$$w_t = (1-\alpha)K_t^{\alpha} N_t^{-\alpha} \left((1+g)^t\right)^{1-\alpha}$$

$$\frac{w_t N_t}{Y_t} = 1 - \alpha$$

In the U.S. the labor share of income has averaged about $2/3$, so $\alpha = 1/3$.

Calibration

To calibrate the depreciation rate δ start with the resource constraint at the BGP (remember that $\tilde{x}_t = x_t/(1+g)^t$ and $x_t = X_t/(1+n)^t$)

$$\tilde{c} + (1-n)(1+g)\tilde{k} = F(\tilde{k},1) + (1-\delta)\tilde{k}$$
$$\tilde{c} + [(1-n)(1+g) - (1-\delta)]\tilde{k} = F(\tilde{k},1)$$

In the BGP, investment is given by

$$\tilde{i} = [(1+n)(1+g) - (1-\delta)]\tilde{k}$$
$$\frac{I/Y}{K/Y} = \frac{I}{K} = \frac{\tilde{i}}{\tilde{k}} = (1+n)(1+g) - (1-\delta)$$

In the data, $I/Y \approx 0.2$ and $K/Y \approx 3$, using our previous parameters, we find $\delta \approx 4\%$.

Calibration

We need to pick parameters for the utility function. From the Euler equation with CRRA utility function:

$$(1+n)(1+g)(\tilde{c}_t)^{-\sigma} = (1+r_{t+1}-\delta)\tilde{\beta}(\tilde{c}_{t+1})^{-\sigma}$$

In the BGP

$$\begin{aligned}(1+n)(1+g) &= (1+r-\delta)\beta(1+g)^{1-\sigma} \\ \beta(1+g)^{-\sigma} &= \frac{1+n}{1+r-\delta}\end{aligned}$$

We need to find r. The rental rate of capital is:

$$r_{t+1} = \alpha K_t^{\alpha-1}\left[(1+g)^t N_t\right]^{1-\alpha} = \alpha \frac{Y_t}{K_t}$$

with $K/Y \approx 3$ and $\alpha \approx 1/3$ we find $r \approx 0.11$.

Calibration

Plugging back these values in the FOC:

$$\beta(1.018)^{-\sigma} = 0.944$$

Note that without growth ($g = 0$) this relationship pins down β but doesn't inform us about σ. With growth, the typical approach is to pick σ from information outside the model.

One can estimate σ by taking the log of

$$(1+n)(1+g)(\tilde{c}_t)^{-\sigma} = (1 + r_{t+1} - \delta)\tilde{\beta}(\tilde{c}_{t+1})^{-\sigma}$$

and do the estimation using consumption data:

- with macro data (Hall 1982): $\frac{1}{\sigma} = 0.1$
- with micro data (Attanasio et al, 1993, 1995) $\frac{1}{\sigma} \in [0.3, 0.8]$
- We pick $\sigma = 1$.

Calibration

Summarizing the parameters:

Param.	Value	Target
g	1.8%	g in data
n	1.1%	n in data
α	0.33	labor share
δ	4%	$\frac{I/Y}{K/Y}$
σ	1	Outside evidence
β	0.961	K/Y

How does the model fare on other moments?
We will come back to the growth model (in continuous time) later.