Stochastic Problems

• References: SLP chapters 9, 10, 11; L&S chapters 2 and 6

1 Examples

1.1 Neoclassical Growth Model with Stochastic Technology

• Production function

\[ y = Af(k) \]

where \( A \) is random

• Let \( A(s^t) \) be productivity in history \( s^t \)

  – We could have \( A \) be a function of \( s_t \) only, i.e. \( A(s_t) \) but allocations will still be functions of histories

• Sequence problem

\[
V^*(k_0, A_0) = \max_{c(s^t), k(s^t)} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \Pr(s^t) u(c(s^t)) \\
\text{s.t.} \\
k(s^{t+1}) = A(s^t) f(k(s^t)) + (1 - \delta) k(s^t) - c(s^t) \\
k(s^t) \geq 0 \\
c(s^t) \geq 0 \\
k_0, A_0 \text{ given}
\]

1.2 A Model of Job Search

• Each period a worker begins by receiving a wage offer \( w \in [0, \bar{w}] \)

• If he accepts the offer, he gets paid \( w \) for that period

• If he rejects the offer he gets unemployment benefits \( b \) instead
• (The worker cannot borrow or save, so he consumes either $w$ or $b$ respectively)

• If the worker wasn’t working the previous period, then $w$ is drawn iid from a distribution $G$
  
  \[- G \text{ could have mass at } w = 0, \text{ which can be interpreted as not finding a job} \]

• If the worker was working the previous period, then
  
  \[- \text{ with iid probability } (1 - \lambda) \text{ he gets the same wage offer as the previous period (interpretation: he keeps his job)} \]
  
  \[- \text{ with iid probability } \lambda \text{ he draws a new } w \text{ from } G \text{ (interpretation: he goes back to the unemployment pool)} \]

• The exogenous state variables are given by:
  
  \[- w: \text{ the wage realization in case the worker draws from } G \]
  
  \[- \theta \in \{0, 1\}: \text{ an indicator of whether the worker keeps his previous job in case he had one} \]

• Endogenous state variables:
  
  \[- J \in \{0, 1\}: \text{ whether the worker was employed yesterday} \]
  
  \[- z \in [0, \bar{w}]: \text{ wage that you enter the period with (only relevant if } J = 1) \]

• Decision: does the worker accept the current offer? $a \in \{0, 1\}$

• Graph:
• Sequence problem:

\[
V^* (w, \theta, J, z) = \max_{a(s^t), J(s^t), c(s^t), z(s^t)} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \Pr (s^t) u (c(s^t))
\]

s.t.

\[
c (s^t) = \begin{cases} 
  z (s^t) & \text{if } J (s^t) = 1, \theta (s^t) = 1, a (s^t) = 1 \\
  w (s^t) & \text{if } J (s^t) = 0 \text{ or } \theta (s^t) = 0 \text{ and } a (s^t) = 1 \\
  b & \text{if } a (s^t) = 0 
\end{cases}
\]

\[
J (s^{t+1}) = a (s^t) \\
z (s^{t+1}) = c (s^t) \\
w_0, y_0, J_0, z_0 \text{ given}
\]
1.3 Consumption-Savings Under Uncertainty

• Household gets income $y(s^t)$ in history $s^t$
  – Again, we could simplify this to $y(s_t)$
• Can borrow (up to a limit) and save at the risk-free interest rate $R$
• Does not have access to insurance (i.e. there are no Arrow-Debreu securities he can buy)
• Sequence problem

$$V^*(A_0, y_0) = \max_{A(s^t), c(s^t)} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \Pr(s^t) u(c(s^t))$$

s.t.

$$A(s^{t+1}) = RA(s^t) + y(s^t) - c(s^t)$$

$$A(s^{t+1}) \geq -B$$

$A_0$ and $y_0$ given

2 Markov Processes

• Stochastic process: sequence of random vectors $x_t \in X$
  – Example:

$$x_t = \{A_t, k_t\}$$

(capital and productivity)
  – (s could be composed of exogenous variables, endogenous variables or a combination thereof)
• Markov property:

$$\Pr(x_{t+1}|x_t, x_{t-1}, \ldots, x_{t-k}) = \Pr(x_{t+1}|x_t)$$

for all $k$
• The state today is a sufficient statistic for forecasting the state tomorrow
  – This is not so restrictive as it seems
• To avoid going too much into measure theory, we are going to focus mostly on cases where $X$ is a finite set
  – In those cases, a Markov process can also be called a Markov chain
  – We’ll briefly mention how some of the results generalize to a continuous state space
2.1 Markov chains

- Discrete state space

- \( x \in \{x_1, \ldots, x_N\} \)

- The evolution of \( x \) is governed by a matrix \( P \) sometimes called transition matrix, Markov matrix or stochastic matrix

**Definition 1.** A stochastic matrix is an \( N \times N \) matrix \( P \) such that:

\[
\sum_{j=1}^{N} P_{ij} = 1 \quad \forall i
\]

- \( P_{i,j} = \) Probability that tomorrow you’ll be in state \( j \) given that today you are in state \( i \)

- Suppose we represent the probability distribution over states in period \( t \) by \( \pi_t \) where \( \pi_t \) is a row vector such that

\[
\sum_{i=1}^{N} \pi_{ti} = 1
\]

then \( \pi P \) represents the probability distribution over states in period \( t + 1 \)

- Example 1:

\[
\pi = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}
\]

(i.e. there are three possible states and we start at \( x = x_2 \) for sure)

\[
P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ 0.3 & 0.2 & 0.5 \\ p_{31} & p_{32} & p_{33} \end{pmatrix}
\]

so

\[
\pi P = \begin{pmatrix} 0.3 & 0.2 & 0.5 \end{pmatrix}
\]

so the second row of \( P \) tells us the probability distributions over states tomorrow given the state today

- Example 2:

\[
\pi = \begin{pmatrix} 0.5 & 0.5 & 0 \end{pmatrix}
\]

\[
P = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.2 & 0.5 \\ p_{31} & p_{32} & p_{33} \end{pmatrix}
\]
\[
\pi P = \begin{pmatrix}
0.4 & 0.3 & 0.3
\end{pmatrix}
\]
so if in period \(t\) there’s a fifty-fifty chance of being in states \(x_1\) or \(x_2\), then \(\pi P\) gives us the probabilities of being in the three states in period \(t+1\)

- Notation note:
  - Sometimes people use the opposite row/column convention for transition matrices:
  - \(P_{i,j}\) = Probability that tomorrow you’ll be in state \(i\) given that today you are in state \(j\)
  - For this you should write \(\pi\) as a column vector and \(\pi' = P\pi\)
  - Of course, this is all equivalent

2.2 The Markov Assumption

- Suppose I have the following process:
  - \(x \in X = \{\text{Rain, Sun}\}\)
  - \(\Pr \{x_{t+1} = \text{Rain} | x_t, x_{t-1}\}\)
    \[
    \begin{cases}
    0.1 & \text{if } x_t = x_{t-1} = \text{Rain} \\
    0.6 & \text{if } x_t = \text{Rain} \text{ and } x_{t-1} = \text{Sun} \\
    0.3 & \text{if } x_t = \text{Sun}
    \end{cases}
    \]
  - This seems not to satisfy the Markov assumption because the weather today is not a sufficient statistic to forecast the weather tomorrow
  - But a simple transformation restores that Markov property:
    - \(X = \{\text{NewRain, OldRain, Sun}\}\)
    - Transition Matrix
      \[
P = \begin{pmatrix}
0 & 0.6 & 0.4 \\
0 & 0.1 & 0.9 \\
0.3 & 0 & 0.7
\end{pmatrix}
\]
- Many history-dependent processes can be rewritten this way as Markov Processes
- For now, we’ll use the Markov assumption to have a justification for saying
  \[
  \Pr(x'|x)
  \]
  and knowing that we can ignore the history of \(x\)
- Later we’ll ask questions about Markov processes themselves:
What do they look like in the long run?
Under what conditions do they converge?
(What do we mean by converge?)

3 Recursive Setup

• Often, there is more than one way to set things up

• Typical sequence:
  – Stuff happens
  – Make decisions
  – Stuff happens
  – Make decisions
  – ...

• Do I compute the value function at the time I’m about to make a decision or when something is about to happen?

• What do I define as the state variable?

• The results on the equivalence of the sequence problem and the functional equation extend to the stochastic case, s.t. some technical caveats
  – See SLP Theorems 9.2 and 9.4

3.1 Neoclassical Growth Model with Stochastic Technology

• State variables:
  – capital stock (endogenous)
  – productivity (exogenous)

• Bellman equation:

\[
V(k, A) = \max_{c,k'} u(c) + \beta \mathbb{E}(V(k', A') | A)
\]

\text{s.t.}

\[
k' = Af(k) + (1 - \delta)k - c
\]
\[
k' \geq 0
\]
\[
c \geq 0
\]
or simply

\[ V(k, A) = \max_{k' \in [0, Af(k) + (1 - \delta)k]} u(Af(k) + (1 - \delta)k - k') + \beta \mathbb{E}(V(k', A') | A) \]

where

\[ \mathbb{E}(V(k', A') | A) = \sum_{A'} \text{Pr}(A'|A) V(k', A') \]

for a discrete space for \( A \)

- FOC for \( c \):
  \[ u'(c) - \lambda = 0 \]

- FOC for \( k' \):
  \[ \beta \sum_{A'} \text{Pr}(A'|A) V_K(k', A') - \lambda = 0 \]

- Envelope condition:
  \[ V_K(k, A) = \lambda [Af'(k) + (1 - \delta)] \]

  and therefore
  \[ V_K(k', A') = \lambda' [A'f'(k') + (1 - \delta)] \]

- Putting these together
  \[ u'(c(k, A)) = \beta \sum_{A'} \text{Pr}(A'|A) u'(c(k', A')) [A'f'(k') + (1 - \delta)] \]

which is the Euler equation that we encountered at the beginning of the class:

\[ u'(c_t) = \beta \mathbb{E}[Ru'(c_{t+1})] \]

where

\[ R = A'f'(k') + (1 - \delta) \]

is the stochastic return of a unit of capital

- The solution defines a system of stochastic difference equations in \( k, c \) and \( A \):

\[ u'(c_t) = \beta \mathbb{E}[(A_{t+1}f'(k_{t+1}) + (1 - \delta))u'(c_{t+1})] \]

\[ k_{t+1} = A_t f(k_t) + (1 - \delta) - c_t \]

\[ A_{t+1} \sim \text{exogenous Markov process} \]

- This is a Markov process for the vector \( k, c, A \)
Unfortunately, unless we restrict $k$ and $c$ to grids, this Markov process lives in a continuous state space even if $A$ lives in a discrete state space.

Questions:

- How does this system behave over time?
- Does it tend to a steady state? In what sense?
- Are there values that it will never reach / reach repeatedly?

3.2 Job search model

- A not-so-useful way to do it:

- State variables:
  - Exogenous: $w$: Wage drawn from $G$ this period
  - Exogenous: $\theta$: whether the worker keeps his job this period (if he had it)
  - Exogenous: $J$: the worker’s incoming job status
  - Exogenous: $z$: The worker’s incoming wage (if employed)

- Bellman equation:

$$
V(w, \theta, J, z) = \max_{a, J', z', c} u(c) + \beta \mathbb{E}[V(w', \theta', J, z)]
$$

s.t.

$$
c = \begin{cases}
  z & \text{if } J = 1, \theta = 1, a = 1 \\
  w & \text{if } J = 0 \text{ or } \theta = 0 \text{ and } a = 1 \\
  b & \text{if } a = 0 
\end{cases}
$$

$$
J' = a \\
z' = c
$$

where

$$
w \sim G
$$

$$
\Pr(\theta = 1) = \lambda
$$

- Alternatively, notice that at the point the worker gets a wage offer, that’s the only thing that matters to him.

- Things that don’t matter:
– Whether the wage was offered by his old job or a new one
– Whether he is drawing from $G$ because he got fired or because he was unemployed

• Define $x$ as today’s job offer

$$V(x) = \max_a u(xa + b(1-a)) + \beta \mathbb{E}(V(x')|x,a)$$

where

$$x'|x,a = \begin{cases} x & \text{with probability } 1 - \lambda \text{ if } a = 1 \\ \sim G & \text{with probability } \lambda \text{ if } a = 1 \\ \sim G & \text{if } a = 0 \end{cases}$$

• Note: if we didn’t assume that new wage offers were iid, then $x$ would not be a sufficient state variable

• Alternatively, compute the value function for a worker who has a job (but might get fired or quit) and a worker who doesn’t have a job separately

– Let $z$ be the incoming wage for a worker who has a job
– $V^J$ is the value of entering with a job that pays $z$
– $V^U$ is the value of entering the period unemployed

$$V^J(z) = \lambda V^U + (1 - \lambda) \max_a \left[ a \left( z + \beta V^J(z) \right) + (1 - a) \left( b + \beta V^U \right) \right]$$

$$V^U = \max_{a(w)} \mathbb{E} \left[ a(w) \left( w + \beta V^J(w) \right) + (1 - a(w)) \left( b + \beta V^U \right) \right]$$

$$= \max_{a(w)} \int \left[ a(w) \left( w + \beta V^J(w) \right) + (1 - a(w)) \left( b + \beta V^U \right) \right] dG(w)$$

• Let’s work with this last formulation

• Unemployed workers have a reservation wage $\bar{w}$ which satisfies

$$\bar{w} + \beta V^J(\bar{w}) = b + \beta V^U$$

• Workers never quit

– If they chose to take the job when unemployed, this means that the wage $z$ they currently have satisfied $z \geq \bar{w}$, which is the same condition for not quitting

– Hence

$$V^J(z) = \lambda V^U + (1 - \lambda) \left( z + \beta V^J(z) \right)$$  \hspace{1cm} (1)
• Simplify $V^U$ to

$$V^U = \int \left[ \max \left\{ w + \beta V^J(w), b + \beta V^U \right\} \right] dG(w)$$  \hspace{1cm} (2)

• Special case $\lambda = 0$ (you won’t get fired)

• Simplifies to

$$V^J(z) = z + \beta V^J(z)$$

$$V^J(z) = \frac{z}{1 - \beta}$$

• Replace in (2):

$$V^U = \int \left[ \max \left\{ \frac{w}{1 - \beta}, b + \beta V^U \right\} \right] dG(w)$$

$$= G(\bar{w}) \left[ b + \beta V^U \right] + \int_{\bar{w}}^{\infty} \frac{w}{1 - \beta} dG(w)$$

$$= \frac{1}{1 - \beta} \left[ G(\bar{w}) b + \int_{\bar{w}}^{\infty} \frac{w}{1 - \beta} dG(w) \right]$$

with

$$\frac{\bar{w}}{1 - \beta} = b + \beta V^U$$

**Exercise 1.**

1. Show that the reservation wage increases if $G$ becomes more dispersed in a SODS sense

2. Show that $\bar{w}$ is decreasing in $\lambda$

### 3.3 Consumption-Savings

• State variables:

  - $A$: Level of assets
  - $y$: today’s income

$$V(A, y) = \max_{c,A'} u(c) + \beta \sum_{y'} \Pr(y'|y) V(A', y')$$  \hspace{1cm} (3)

s.t.  \hspace{1cm} $A' \leq y - c + RA$

\hspace{1cm} $A' \geq -b$
• Alternatively, let \( x \equiv RA + y \) (“cash on hand”):

\[
V(x, y) = \max_{c, x'(y')} u(c) + \beta \sum_{y'} \Pr(y'|y)V(x'(y'), y')
\]

s.t. \( x'(y') \leq R |x - c| + y' \)

\( c \leq x + b \)

• This relabeling is especially useful if \( y \) is iid: reduce problem to a single state variable

  – because with \( x, y \) as the state variables, \( y \) only matters through \( \Pr(y'|y) \)

• FOC:

\[
u'(c) - \sum_{y'} \lambda(y') R - \mu = 0
\]

\[
\beta \Pr(y'|y) \frac{\partial V(x'(y'), y')}{\partial x} - \lambda(y') = 0
\]

\[
\Rightarrow u'(c) \geq \beta R \sum_{y'} \Pr(y'|y) \frac{\partial V(x'(y'), y')}{\partial x}
\]

with equality if \( c < x + b \)

• Envelope condition

\[
\frac{\partial V(x, y)}{\partial x} = u'(c(x, y))
\]

• Euler equation:

\[
u'(c) \geq \beta R \sum_{s'} \Pr(s'|s)\ u'(c(x'(y'), y'))
\]

with equality if \( c < x + b \)

4 Convergence of Markov Processes

• Starting from \( \pi_0 \) (which could be degenerate), a Markov chain will evolve according to

\[
\pi_t = \pi_0 P^t
\]

• \( \pi \) belongs to the set

\[
\Delta^N = \left\{ \pi \in \mathbb{R}_{+}^n : \sum_{i=1}^{N} \pi_i = 1 \right\}
\]
• Define the following norm in $\Delta^N$:

$$\|\pi\| = \sum_{i=1}^{N} |\pi_i|$$

• We say a Markov chain converges to $\pi^*$ if

$$\lim_{T \to \infty} \|\pi_t - \pi^*\| = 0$$

• In this case we call $\pi^*$ an invariant distribution and the support of $\pi^*$ an ergodic set

• Define the operator $T : \Delta^N \to \Delta^N$ as

$$T(\pi) = \pi P$$

**Proposition 1.** (SLP Lemma 11.3) Let $\epsilon_j = \min_i P_{ij}$ and $\epsilon = \sum_{j=1}^{N} \epsilon_j$. If $\epsilon > 0$ then $T$ is a contraction modulus $1 - \epsilon$.

• Example:

$$P = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0 & 0.3 & 0.7 \\ 0.7 & 0.1 & 0.2 \end{pmatrix}$$

Here

$$\epsilon_j = \begin{pmatrix} 0 & 0.1 & 0.2 \end{pmatrix}$$

$$\epsilon = 0.3$$

• Interpretation: there is one value of $x$ that you reach with positive probability from any other state

**Proof.** The distance between $T(\pi)$ and $T(\rho)$ is:

$$\rho(T(\pi), T(\mu)) = \|T(\pi) - T(\mu)\|$$

$$= \|(\pi - \mu)P\|$$

$$= \sum_j \left| \sum_i (\pi_i - \mu_i)P_{ij} \right|$$
• Example:

\[
\begin{align*}
\pi &= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \\
\mu &= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \\
\pi - \mu &= \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} \\
\sum_i (\pi_i - \mu_i) P_{ij} &= \begin{pmatrix} 0.2 & 0.5 & 0.3 & 0.7 \end{pmatrix} \\
\sum_j \left| \sum_i (\pi_i - \mu_i) P_{ij} \right| &= 0.2 + 0.2 + 0.4 = 0.8 \leq (1 - 0.3) \times 2
\end{align*}
\]

(continue proof)

\[
= \sum_j \left| \sum_i (\pi_i - \mu_i) (P_{ij} - \epsilon_j) + \sum_i (\pi_i - \mu_i) \epsilon_j \right| \\
\leq \sum_j \left| \sum_i (\pi_i - \mu_i) (P_{ij} - \epsilon_j) \right| + \sum_j \left| \sum_i (\pi_i - \mu_i) \epsilon_j \right| \\
= \sum_j \left| \sum_i (\pi_i - \mu_i) (P_{ij} - \epsilon_j) \right| \\
\leq \sum_i |\pi_i - \mu_i| \sum_j (P_{ij} - \epsilon_j) \\
= \rho (\pi, \mu) (1 - \epsilon)
\]

• Proposition 1 immediately implies:

- There is a unique fixed point, i.e. unique solution to

\[
\pi = T (\pi)
\]

(denote it \( \pi^* \))

- Starting from any \( \pi_0 \),

\[
\lim_{T \to \infty} T (\pi) = \pi^*
\]
• You can find $\pi^*$ by solving

$$\pi = \pi P$$

$$\pi (I - P) = 0$$

i.e. $\pi^*$ is the eigenvector corresponding to the eigenvalue 1 of matrix $P$, scaled so that $\sum_{i=1}^{N} \pi_i = 1$.

• Generalizations:

- If the condition $\epsilon > 0$ holds for the matrix $P^n$, then $T^n$ is a contraction mapping, so the limiting condition holds too
- (Interpretation: there is a state that you reach with positive probability in $n$ steps starting from anywhere)
- For states spaces that are not finite, “condition $M$” is

**Definition 2.** A Markov process satisfies Condition $M$ if there exist $\epsilon > 0$ and $n \geq 1$ such that for any set $A \subseteq X$, either:

1. $\Pr(x_{t+n} \in A|x_t) \geq \epsilon$ for all $x_t \in X$
   or
2. $\Pr(x_{t+n} \in A^C|x_t) \geq \epsilon$ for all $x_t \in X$

• Example: suppose there is a recurrent element $x_R \in X$ such that

$$\Pr(x_{t+n} = x_R|x_t) \geq \epsilon \ \forall x_t$$

Then condition $M$ is satisfied because for any $A$, either

- $x_R \in A$ or
- $x_R \in A^C$

**Proposition 2.** *(SLP Lemma 11.11).* If a Markov process satisfies condition $M$, then $T^n$ is a contraction of modulus $\epsilon$