Macro II: Stochastic Processes I

Professor Griffy

UAlbany

Spring 2026

Introduction

- Today: start talking about time series/stochastic processes.
- Homework due in one week.
- New system for cluster access this year, still working on it.
- Probably have access today.

Stochastic Processes

- Random variables
- Conditional distributions
- Markov processes

Preliminaries

- X is a random variable, x is its realization

- Support: smallest set S such that $\Pr(x \in S) = 1$

- Cumulative distribution function: $F(x) = \Pr(X \leq x)$

- Density function: $f(x) = \frac{d}{dx} F(x)$ implying that $f(x)\, dx = dF(x)$

The Expected Value

- Mean is the expectation

$$\bar{X} = E(X) = \int_{-\infty}^{\infty} x dF(x) = \int_{-\infty}^{\infty} x f(x)\, dx$$

- The expectation of a function of a random variable, $g(X)$, is

$$E(g(X)) = \int_{-\infty}^{\infty} g(X)\, dF(x)$$

- Note that $E(g(X)) \neq g(\bar{X})$ unless $g(X)$ is linear, i.e.

$$g(X) = b \cdot X$$

The Variance

- Variance

$$V(X) = E\left[\left(X - \bar{X}\right)^2\right]$$

- Standard deviation

$$[V(X)]^{\frac{1}{2}}$$

Jointly Distributed Random Variables

- Random vector (X, Y)
- Joint distribution function: $F(x, y) = \Pr(X \leq x, Y \leq y)$
- Covariance: $C(X, Y) = E\left[\left(X - \bar{X}\right) \cdot \left(Y - \bar{Y}\right)\right]$
- Cross-correlation $= \frac{C(X,Y)}{[V(X) \cdot V(Y)]^{\frac{1}{2}}}$
- Expectation of a linear combination

$$E(aX + bY) = aE(X) + bE(Y)$$

What is a Stochastic Process?

- Stochastic process is an infinite sequence of random variables $\{X_t\}_{t=-\infty}^{\infty}$
- j'th autocovariance $= \gamma_j = C\left(X_t, X_{t-j}\right)$
- Strict stationarity: distribution of $\left(X_t, X_{t+j_1}, X_{t+j_2}, ...X_{t+j_n},\right)$ does not depend on t
- Covariance stationarity: $\bar{X}_t$ and $C\left(X_t, X_{t-j}\right)$ do not depend on t

Defining a Conditional Density

- Work with random vector $\underline{x} = (X, Y) \sim F(x, y)$.
  - X and Y are random variables
  - x and y are realizations of the random variables
  - $F(x, y)$ is joint cumulative distribution
  - $f(x, y)$ is joint density function

Conditional Variables and Independence

- Conditional probability
  - when $\Pr(\underline{x} \in B) > 0$,

$$\Pr(\underline{x} \in A | \underline{x} \in B) = \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}.$$

- Conditional distribution $F(y|x)$ (handles $\Pr(B) = 0$)
  - Marginal distribution: $F_X(x) = \Pr(X \leq x)$
  - $F(y|x)$ is $\Pr(Y \leq y)$ conditional on $X \leq x$

Defining a Conditional Density

- Independence: The random variables X and Y are independent if
$$F(x, y) = F_X(x) F_Y(y)$$
  - If X and Y are independent, then
$$F(y|x) = F_Y(y)$$
    and
$$F(x|y) = F_X(x)$$
  - i.i.d means independent and identically distributed
- Conditional (mathematical, rational) expectation
$$E(Y|x) = \int_{-\infty}^{\infty} y dF(y|x) = \int_{-\infty}^{\infty} y f(y|x) dy.$$

Markov Property

- A particular conditional process is called a Markov chain.
- Markov Property: A stochastic process $\{x_t\}$ is said to have the Markov property if for all $k \geq 1$ and all t,

$$Prob(x_{t+1}|x_t, x_{t-1}, ..., x_{t-k}) = Prob(x_{t+1}|x_t) \tag{1}$$

- That is, the dependence between random events can be summarized exclusively with the previous event.
- This allows us to characterize this process with a Markov chain.
- Markov chains are a key way of characterizing stochastic events in our models.

Markov Chains

- For a stochastic process with the Markov property, we can characterize the process with a Markov chain.
- A time-invariant Markov chain is defined by the tuple:
  1. an n-dimensional state space of vectors $e_i, i = 1, ..., n$,
     - where e_i is an n x 1 vector where
     - the ith entry equals 1 and the vector contains 0s otherwise.
  2. a transiton matrix P (n x n), which records the conditional probability of transitioning between states
  3. a vector π_0 (n x 1), that records the unconditional probability of being in state i at time 0.
- The key object here is P. Elements of this matrix are given by

$$P_{ij} = Prob(x_{t+1} = e_j | x_t = e_i) \tag{2}$$

- In other words, if you're in state i, this is the probability you enter state j.

Markov Chains

- Some assumptions on P and π_0:
  - For $i = 1, ..., n$, P satisfies

$$\sum_{j=1}^{n} P_{ij} = 1 \tag{3}$$

  - π_0 satisfies

$$\sum_{i=1}^{n} \pi_{0i} = 1 \tag{4}$$

- Where does this first property become useful?
- How would you calculate $Prob(x_{t+2} = e_j | x_t = e_i)$?

$$= \sum_{h=1}^{n} Prob(x_{t+2} = e_j | x_{t+1} = e_h) Prob(x_{t+1} = e_h | x_t = e_i) \tag{5}$$

$$= \sum_{h=1}^{n} P_{ih} P_{hj} = P_{ij}^{(2)} \tag{6}$$

Markov Chains

- This is also true in general:

$$Prob(x_{t+k} = e_j | x_t = e_i) = P_{ij}^{(k)} \tag{7}$$

- Why is this useful? We can use π_0 with this transition matrix to characterize the probability distribution over time:

$$\pi_1' = \pi_0' P \tag{8}$$
$$\pi_2' = \pi_0' P^2 \tag{9}$$
$$\vdots \tag{10}$$

- Thus, by knowing the initial distribution and the transition matrix, P, we know the distribution at time t

Stationary Distributions

- Where does this trend to over time?
- We know that the transition of the distribution takes the form $\pi'_{t+1} = \pi'_t P$.
- This distribution is stationary if

$$\pi_{t+1} = \pi_t \tag{11}$$

- (we will relax this to t large enough momentarily)
- This means that for a stationary distribution, π, P satisfy

$$\pi' = \pi' P \text{ or} \tag{12}$$

$$(I - P')\pi = 0 \tag{13}$$

- Anyone recognize this?

Stationary Distributions

$$\pi' = \pi' P \text{ or} \tag{14}$$

$$(I - P')\pi = 0 \tag{15}$$

- A lot of linearizing dynamic systems is about
  - finding eigenvectors with corresponding eigenvalues of less than 1 (non-explosive).
  - solving for initial conditions that are orthogonal to the explosive eigenvectors (i.e., the system does not explode).
- Intuitive refresher:
  - eigenvector: tells me the direction a system moves (i.e., distance traveled)
  - eigenvalue: tells me how many times it traveled since I last saw it.

Stationary Distributions

$$\pi' = \pi' P \text{ or} \tag{14}$$

$$(I - P')\pi = 0 \tag{15}$$

- It is useful to note (and will be useful when we think of linearized solution techniques), that
  - π is the (normalized) eigenvector of the stochastic matrix P.
  - In this case, the eigenvalue (root) is 1.

Asymptotically Stationary Distributions

- What about when $\pi_0 \neq \pi_t$? Can it still have a notion of stationarity?
- Yes. Asymptotic stationarity.
- Asymptotic stationarity:

$$\lim_{t \to \infty} \pi_t = \pi_\infty \tag{16}$$

- where $\pi'_\infty = \pi'_\infty P$
- Next, is this ending point unique?
- Does π_∞ depend on π_0?
- If not, π_∞ is an invariant or stationary distribution of P.
- This will be very useful when we talk about heterogeneous agents.

Some Examples

- Let's pick a simple initial condition: $\pi_0' = [1\ 0\ 0]$.
- And a matrix

$$P = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.2 & 0.7 \end{bmatrix} \tag{17}$$

- Now use Matlab to iterate.

Preliminaries

```
>> piMat = MMat'*piMat

piMat =

    0.9000
    0.1000
         0
```

Figure: First iteration

```
>> piMat = MMat'*piMat

piMat =

    0.8300
    0.1500
    0.0200
```

Figure: 2nd iteration

```
>> piMat = MMat^(100)'*piMat

piMat =

    0.6154
    0.2308
    0.1538
```

Figure: First iteration

```
>> piMat = MMat^(1000)'*piMat

piMat =

    0.6154
    0.2308
    0.1538
```

Figure: Grid of k values

- This distribution (P) is asymptotically stationary!
- Unique? Try $\pi_0' = [0\ 0\ 1]$

Preliminaries

```
>> piMat = MMat'*piMat

piMat =

    0.9000
    0.1000
         0
```

Figure: First iteration

```
>> piMat = MMat'*piMat

piMat =

    0.8300
    0.1500
    0.0200
```

Figure: 2nd iteration

```
>> piMat = MMat^(100)'*piMat

piMat =

    0.6154
    0.2308
    0.1538
```

Figure: First iteration

```
>> piMat = MMat^(1000)'*piMat

piMat =

    0.6154
    0.2308
    0.1538
```

Figure: Grid of k values

- This distribution (P) is (probably) a unique invariant distribution.
- How would we prove this?

Ergodicity

- We would like to be able to replace conditional expectations with unconditional expectations.
- i.e., not indexed by time or initial conditions.
- Some preliminaries:
  - Invariant function: "a random variable $y_t = \bar{y}' x_t$ is said to be invariant if $y_t = y_0,\ t \geq 0$, for all realizations of $x_t, t \geq 0$ that occur with positive probability under (P, π)."
- i.e., the state x can move around, but the outcome y_t stays constant at y_0.

Ergodicity

- Ergodicity:
  - "Let (P, π) be a stationary Markov chain. The chain is said to be ergodic if the only invariant functions $\bar{y}$ are constant with probability 1 under the stationary unconditional probability distribution π."
- In other words, for any initial distribution, the only functions that satisfy the definition of an invariant function are the same.

Next Time

- More stochastic processes.

- Homework due in one week.