Topics for Today

1.) Comparing Models to Data

2.) Time Series & Stochastic Processes
   - Stationarity & Ergodicity

3.) Markov Chains & Stochastic Difference Equations

4.) Invariant Distributions

5.) Impulse Response Functions & Spectral Densities

Comparing Models to Data

- The models we develop in this course generally take one of two forms:

1.) Discrete-State Markov Chain

2.) Continuous-State Stochastic Difference Equation

How do we Know whether a model is any good?

2 comments

1.) Models are designed to address specific questions. They should be evaluated based on their answers to the question at hand. A model may be good for some purposes, but bad for others. (Of course, the more questions a model can address, the better is the model!).

2.) It is important to keep in mind that the equilibria of our models are stochastic processes, not fixed numbers. Hence, they cannot be judged based on a single outcome.

- Instead, we must compare the statistical properties of the data our model hypothetically would generate, with the same statistical properties of the observed data.

- The time-series properties of our models are determined by a (small) set of parameters characterizing preferences, technology, and market structure.

- There are 2 basic strategies for comparing models to data:

1.) Econometrics: Pick parameters to minimize the "distance" between model data + actual data, and test using χ^2 or LR statistics.

Problem: All models are false. Rejections can be uninformative.

2.) Calibration: Pick the parameters to match some pre-selected moments (usually long-run means) and/or take them from other studies. Then, compare the remaining moments of interest to those in the data.

Problem: Must rely on "eyeball metric". Sensitivity to selected moments.

- Either way, to carry out this program, we obviously need to know something about time series; in particular, about Markov Chains + Stochastic Difference Equations.

Key Concepts

① **STATE** : A complete description of all relevant aspects of the current position of a dynamical system.

Examples : 1.) Capital Stock
2.) Technology
3.) Financial Wealth
4.) Employment Status
5.) Beliefs

Different Kinds of State Variables

- Endogenous vs. Exogenous > capital vs. Technology
- Discrete vs. Continuous > Employment Status vs. Wealth
- Finite Dimensional vs. Infinite Dimensional > Wealth vs. Beliefs.

The State plays a role in 2 ways

- $s_{t+1} = f(s_t, x_t, \varepsilon_{t+1})$ > state transition eq. (e.g., budget constraint or Bayes Rule)
- $x_t = g(s_t)$ > policy function

2. TIME SERIES / STOCHASTIC PROCESS: A time-indexed sequence of random variables.

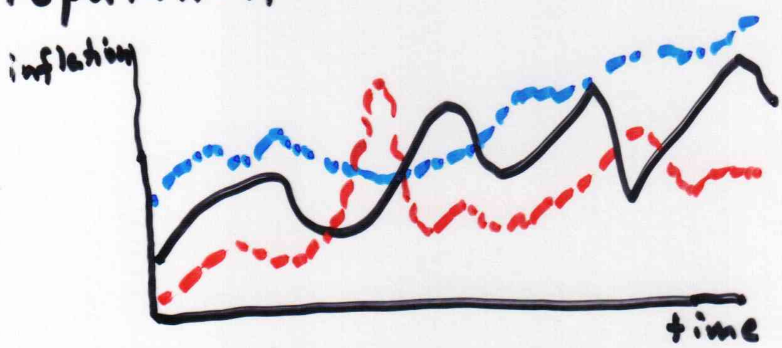
Stochastic Processes vs. Random Variables
(paths) (real numbers)

- We need to regard a time series as a single unit

Comments

1.) A realization is an entire path, not a number

2.) Population = Set of all possible realizations



3.) We observe only one realization (unless you're a believer in the "many-worlds" interpretation of quantum mechanics!).

4.) "Law of Large Numbers" and "Central Limit Theorem" arguments still applicable if observations over time provide enough new information about ensemble averages.

3.) STATIONARY STOCHASTIC PROCESS / COVARIANCE STATIONARITY

A stochastic process where all joint distributions are independent of time. "Initial conditions have worn off".

$$f(x_t, x_{t+1}, \cdots x_{t+j}) \text{ indpt. of } t \ \forall j$$

4.) MARKOV PROCESS : Restricts conditional distribution function

$$f(\text{future} \mid \text{present, past}) = f(\text{future} \mid \text{present})$$

$$\text{Prob}(x_{t+1} \mid x_t, x_{t-1}, \cdots x_{t-k}) = \text{Prob}(x_{t+1} \mid x_t) \quad \forall k \geq 1$$

- Whether a process is Markov depends on the dimensionality of the state

Example : $S_t = \alpha_1 S_{t-1} + \alpha_2 S_{t-2} + \alpha_3 S_{t-3} + \varepsilon_t$ } 3rd order Markov AR(3)

Redefine state as $Z_t = \begin{pmatrix} S_t \\ S_{t-1} \\ S_{t-2} \end{pmatrix}$

Then $Z_t = A Z_{t-1} + B \varepsilon_t$ } 1st order Markov

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad B = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

5. MARKOV CHAIN : A discrete-state Markov Process

- A Markov Chain is completely characterized by its Transition Matrix, P, where

$$p_{ij} = \text{Prob}(x_{t+1} = x_j \mid x_t = x_i)$$

Note, $\sum_{j=1}^{N} p_{ij} = 1$

6. STATIONARY (INVARIANT) DISTRIBUTION:

A distribution that remains invariant over time (stochastic analog of a limit point).

- Consider the case of a Markov Chain

Let π_0 = initial dist. over states $\begin{pmatrix} \pi_0(1) \\ \pi_0(2) \\ \pi_0(3) \end{pmatrix}$

Note, $\pi_1' = \pi_0' P$

$$\begin{pmatrix} \pi_0(1) & \pi_0(2) & \pi_0(3) \end{pmatrix} \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Hence,

$$\begin{pmatrix} \pi_1(1) \\ \pi_1(2) \\ \pi_1(3) \end{pmatrix} = \begin{pmatrix} \pi_0(1) p_{11} + \pi_0(2) p_{21} + \pi_0(3) p_{31} \\ \pi_0(1) p_{12} + \pi_0(2) p_{22} + \pi_0(3) p_{32} \\ \pi_0(1) p_{13} + \pi_0(2) p_{23} + \pi_0(3) p_{33} \end{pmatrix}$$

Note, if the Markov Chain is "homogeneous", then P does not depend on time, so

$$\pi_2' = \pi_0' P^2 \quad \text{or more generally} \quad \pi_K' = \pi_0' P^K$$

Clearly, the stationary (or invariant) dist. is characterized by:

$$\pi_\infty' = \pi_\infty' P \implies \pi_\infty'(I - P) = 0$$

$$\implies (I - P')\pi_\infty = 0$$

$\boxed{\pi_\infty = \text{eigenvector of } P' \text{ associated with a unit eigenvalue (normalized so that } \sum \pi_\infty(i) = 1\text{).}}$

Since P is a "stochastic matrix" [non-negative elements with row sum = 1] it alway has at least one unit eigenvalue. [It's unique if $p_{ij} > 0 \;\; \forall i,j$].

Example: Suppose $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$

Then, $\pi_\infty = \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta} \right)$

7. **ERGODIC MARKOV CHAIN**: A Markov Chain with a unique invariant distribution

- Having multiple stationary distributions is the stochastic analog of having multiple equilibria.
- Ergodicity is important because it guarantees sample averages over time converge to population moments (Law of Large Numbers).
- A Markov Chain is ergodic when it is possible to go from every state to every state (not necessarily in one jump). This property is sometimes called irreducibility.
- If $P_{ij} > 0 \ \forall i,j$ then obviously the chain is ergodic. However, this is not necessary.
- Example of non-ergodic Markov Chain:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}$$

8. LINEAR STOCHASTIC DIFFERENCE EQUATIONS:

$$Z_{t+1} = A Z_t + B \varepsilon_{t+1}$$

where ε_{t+1} is mean zero, iid, with continuous dist.

- A common specification for ε_{t+1} is $N(0, \Sigma)$. When ε_{t+1} is Gaussian, the first 2 moments of Z_t characterize the entire distribution. In particular, "covariance stationarity" implies stationarity.

First Moment

Let $\mu_t = E(Z_t)$. Then,

$$\mu_{t+1} = A \mu_t \implies (I - A)\mu_\infty = 0$$

$\implies$ Unconditional/Stationary mean is the eigenvector associated with (unique) unit eigenvalue.

Sufficient Condition: A has all eigenvalues inside unit circle.

Second Moment

Define

$$C(j) = E(X_{t+j} - \mu)(X_t - \mu)'$$ > covariogram (indpt. of time)

$$C(0) = \text{Variance}$$
$$= E[(AZ_t + B\varepsilon_{t+1})(Z_t'A' + \varepsilon_{t+1}'B')]$$

$$\Rightarrow C(0) = AC(0)A' + B\Sigma B'$$ > Lyapunov Eq.

$$C(j) = A^j C(0)$$

- There are 2 common ways to summarize the dynamics of a stochastic difference equation:

  a.) Impulse Response Function.

  b.) Spectral Density.

9. IMPULSE RESPONSE FUNCTION:

A plot of the system's moving average coeffici[ents]

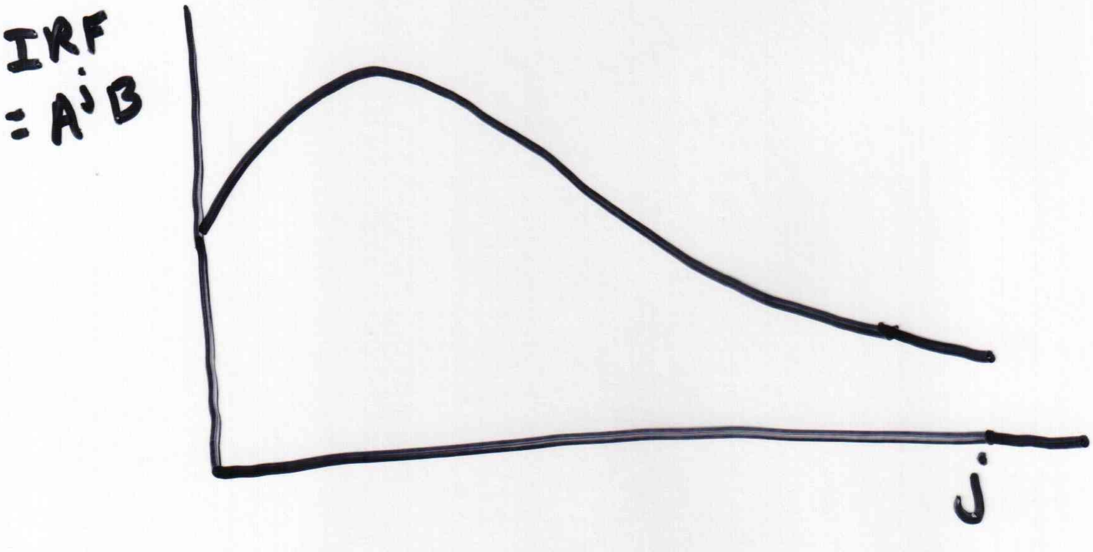
Write the stoch. diff. eq. using lag operators,

$$(I - AL) X_t = B \varepsilon_t \quad \text{where } L X_t = X_{t-1}$$

$$\Rightarrow X_t = (I - AL)^{-1} B \varepsilon_t$$

$$= \sum_{j=0}^{\infty} A^j B \varepsilon_{t-j}$$

> converges if A has all eigenvalues inside unit circle



10. SPECTRAL DENSITY : Fourier Transform of the covariogram

Basic idea : Decompose autocovariances into orthogonal components

Why do this ?

1.) Computational - Orthogonality $\Rightarrow$ Non-interference
No cross-terms

Converts functional equations into algebraic eqs.
Solve the algebraic eq. then transform back !

2.) Conceptual - Separate components may be interesting (e.g., business cycles, or seasonal cycles).

Fourier Transforms

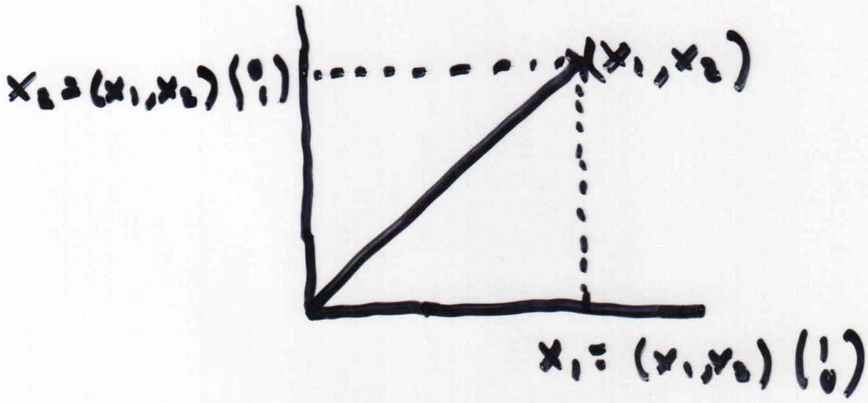
- Let's start slow. You know that any vector in $\mathbb{R}^2$ can be written :

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

sum of orthog. components

where the coefficients are given by an inner product:

$$x_1 = (x_1, x_2)\begin{pmatrix}1\\0\end{pmatrix} \qquad x_2 = (x_1, x_2)\begin{pmatrix}0\\1\end{pmatrix}$$



The idea behind Fourier Transforms is exactly the same, but rather than expressing a vector in R^n (a finite dimensional object) as an orthogonal sum, we want to express an "arbitrary" function as an orthogonal sum

what does this mean?

$$f(x) \hat{=} \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos nx + b_n \sin nx\right) \quad -\pi < x < \pi$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos nx \, dx$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \sin nx \, dx$$

Same as above with inner product defined as $(f, g) = \int_{-\pi}^{\pi} f(x) g(x) \, dx$

Note: (a_n, b_n) are like regression coefficients from projecting f onto sine & cosines.

Orthogonality

$$\int_{-\pi}^{\pi} \cos(mx)\sin(nx)\,dx = 0 \quad \forall m,n$$

$$\int_{-\pi}^{\pi} \cos(mx)\cos(nx)\,dx = 0 \quad m \neq n$$

$$\int_{-\pi}^{\pi} \sin(mx)\sin(nx)\,dx = 0 \quad m \neq n$$

- For functions defined on the whole real line (which are not periodic) we need a fourier integral

$$f(x) = \int_{-\infty}^{\infty} a(w)\, e^{-iwx}\, dw$$

Examples

$$S(w) = \sum_{j=-\infty}^{\infty} c(j)\, e^{-iwj} \qquad \frac{w}{2\pi}: \text{ frequency (cycles per period)}$$

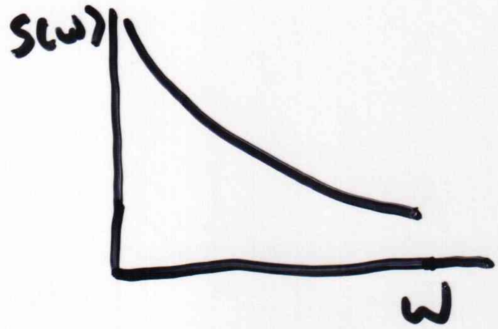
$$= (I - Ae^{-iw})^{-1} BB' (I - A'e^{iw})^{-1}$$

$$c(j) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(w)\, e^{iwj}\, dw$$

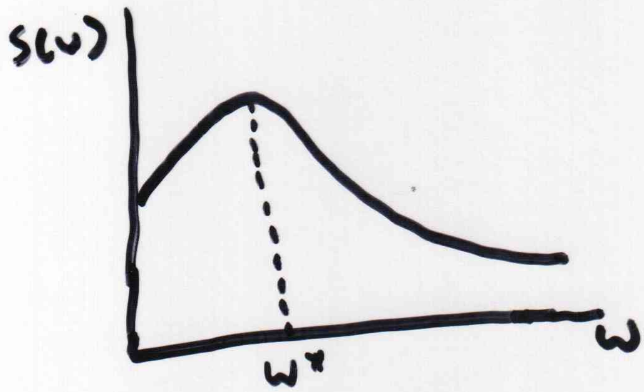
scalar case, $x_t = a x_{t-1} + \varepsilon_t \Rightarrow c_j = a^{|j|} c_0$

$$\Rightarrow S(w) = \frac{c_0(1-a^2)}{(1-ae^{-iw})(1-ae^{iw})} = \frac{c_0(1-a^2)}{1+a^2-a\cos w}$$

- Most economic time series have spectral densities that look like:



- A series with a cycle has a spectral density with a local peak



The implied cycle has period $\frac{2\pi}{w^*}$

The theoretical foundation of fourier analysis is the following theorem:

Riesz-Fischer Theorem: Let $\{c_n\}$ be a square-summable sequence of complex numbers, (i.e., $\sum |c_n|^2 < \infty$). Then there exists a complex-valued function, $g(w)$, defined for $w \in [-\pi, \pi]$, s.t.

$$\text{(A)} \qquad \boxed{g(w) = \sum_{j=-\infty}^{\infty} c_j e^{-iwj}}$$

where convergence is in the mean-square sense,

$$\lim_{n\to\infty} \int_{-\pi}^{\pi} \left| \sum_{j=-n}^{n} c_j e^{-i\omega j} - g(\omega) \right|^2 d\omega = 0$$

and $g(\omega)$ is square (Lebesgue) integrable

$$\int_{-\pi}^{\pi} |g(\omega)|^2 d\omega < \infty .$$

Conversely, given a square integrable $g(\omega)$, $\exists$ a square summable sequence s.t.

$$\boxed{(B) \qquad c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\omega) e^{i\omega k} d\omega}$$

- (A) + (B) form a "Fourier Transform pair" linked via the *Parseval Identity*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |g(\omega)|^2 d\omega = \sum_{j=-\infty}^{\infty} |c_j|^2$$

In words, $\exists$ an isometric isomorphism between ℓ_2 (time domain) and L^2 (freq. domain) (i.e., a mapping that is one-to-one and preserves distance + linear structure).