ECOS3904: Applied Macroeconometrics

Charles Christopher Hyland

Semester 2 2018

Abstract

Thank you for stopping by to read this. These are notes collated from lectures and tutorials as I took this course.

Contents

1. Introduction to Time Series

1.0.1 Basic Defintions

Definition 1.1 *(Random Variable) A random variable is a function $X = X(\omega)$: $\Omega \to \mathbb{R}$ where Ω is the sample space and $\mathbb{R}$ is the real line.*

Definition 1.2 *(Stochastic Process) For a given probability space $(\Omega, \mathbb{F}, \mathbb{A})$ and a measurable space (S, Σ), a stochatic process is a collection of S-valued random variables, which can be written as:*

$$\{X(t) : t \in T\}$$

Remark 1.3 *To elaborate, a stochastic process is a collection of random variables $X_1, X_2, X_3, ...$ defined on a common **probability space** $(\Omega, \mathbb{F}, \mathbb{A})$ where Ω is a sample space, $\mathbb{F}$ is a σ-algebra, and $\mathbb{F}$ is a probability measure. The random variables, indexed by an **index set T** all takes values in the same mathematical space S, which is **measurable** with respect to some σ-algebra Σ.*

Remark 1.4 *The random variable can take on values from either a **discrete** or a **continuous state** space.*

Remark 1.5 *A sequence is an enumerated collection, which does not require unique items in the collection.*

Definition 1.6 *(Time Series) A stochastic process indexed by time is known as a time series denoted by $\{X_t\}$ where t=1,2,3,... For each t, we have the random variable $X_t = X_t(\omega)$ where $\omega \in \Omega$.*

Remark 1.7 *If T consists of integers, then the process is a **Discrete Time Stochastic Process**. If T consists of the real numbers, then the process is a **Continuous Time Stochastic Process**.*

Definition 1.8 *(Realization/Sample Path) When a certain outcome $\omega_0 \in \Omega$ is realised, the random variable X_t will take on a real value $x_t = X_t(\omega_0)$ for each t. A **realization/sample path** of a time series X_t is a sequence of real numbers x_t.*

The issue with time series problems is that we only observe the realisation of the underlying process once. Hence, this makes it difficult to do inference, hypothesis testing, prediction, and more.

Definition 1.9 *(Strictly Stationarity) Let $\{X_t\}$ be a **stochastic process** and let $F_x(x_{t_1+\tau} + ... + x_{t_k+\tau})$ be the CDF of the **unconditional joint distribution** of $\{X_t\}$ at times $t_1+\tau, ..., t_k+\tau$. $\{X_t\}$ is **strictly/strongly** stationary if for all k, τ, and $t_1, ..., t_k$:*

$$F_x(x_{t_1+\tau} + ... + x_{t_k+\tau}) = F_x(x_{t_1} + ... + x_{t_k})$$

Remark 1.10 *Since τ does not affect the CDF $F_X(.)$, this means that the CDF F_X is not a function of time. In other words, the random variables all have the same probability distribution.*

A strictly stationary processes parameter such as mean and variance remain the same across t.

Furthermore, applying a **measurable function** f on the stationary process, still gives us a strictly stationary process. A measurable function is a function between two **measurable spaces** such that the preimage of any measurable set is also measurable.

Definition 1.11 *(Measure-Preserving Transformation) A transformation T where $\mu(T^{-1}(A)) = \mu(A)$.*

Definition 1.12 *(Ergodicity) Let $(X, \mathbb{A}, \mu)$ be a probability space and $T: X \to X$ be a **measure-preserving transformation**. We say that T is **ergodic** with respect to μ if the following condition hold:*

$$E \in \mathbb{A} \text{ with } T^{-1}(E) = E \text{ with either } \mu(E) = 0 \text{ or } \mu(E) = 1.$$

To elaborate, ergodicity refers to the fact that any set that is unchanged by a transformation must be either everything in the set or nothing in the set. We require this so that eventually, every point in the space is realised. Ergodicity implies that our dynamical system can't be further decomposed.

Definition 1.13 *(Serial Dependence/Autocorrelation) Random variables in a time series have **serial dependence** if the value at some time t in the series is statistically dependent on the value at another time.*

Definition 1.14 *(Serial Correlation) A time series has **serial correlation** if the condition holds that some pair of values are correlated.*

Definition 1.15 *(I.I.D Sequence) A sequence of independent and identically distributed (i.i.d.) random variables is a **strictly stationary process** which exhibits **no serial dependence**.*

Definition 1.16 *(Constant Series) A time series is a constant series if $X_t = X_1$ for all t=1,2,... A constant series is strictly stationary process which exhibits **perfect** serial dependence.*

Definition 1.17 *(Weakly/Covariance Stationary) A time series $\{X_t\}$ is **weakly/covariance stationary** if:*

1. *$E(X_t), E(X_t^2) < \infty$ for all t.*

2. *$E(X_t)$ does not depend on t.*

3. *$Cov(X_t, X_s) = Cov(X_{t+k}, X_{s+k})$ for any t,s, and k.*

Definition 1.18 *(Weakly Dependent) A **weakly dependent** time series is one where:*

$$Corr(x_t, x_{t+h}) \to 0 \quad h \to \infty$$

A weakly dependent time series in one where the process is asymptotically uncorrelated. An important example of a weakly stationary process is a white noise process.

Definition 1.19 *(White Noise) A time series $\{\epsilon_t\}$ is a white noise process if it satisfies:*

1. *$E(\epsilon_t) = 0$.*

2. $Var(\epsilon_t) = \sigma^2$.

3. $Cov(\epsilon_t, \epsilon_s) = 0$.

When modellign a serially correlated stationary time series. we can break it up into 2 additive components:

$$X_t = g(X_{t-1}, X_{t-2}, ...) + \epsilon_t$$

where $g(X_{t-1}, X_{t-2}, ...)$ is **predictable** whilst ϵ_t is a white noise process. The white noise process is unpredictable based on past history of the series.

When modelling, we tend to make the assumption that the function g(.) is **linear**. An example of this is:

$$X_t = \alpha_0 + \alpha_1 X_{t-1} + \epsilon_t$$

Definition 1.20 *(Integrated Process) An **integrated process** can be made stationary by taking differences over time.*

Definition 1.21 *(Stochastic Trend) A **stochastic trend** is an integrated process that is **highly persistent**.*

Definition 1.22 *(Random Walk) A random walk is a process where:*

$$Y_t = Y_{t-1} + \epsilon_t$$

where ϵ_t is a white noise process.

Note that the first difference of a random walk is a white noise process.

2. Difference Equations

2.0.2 Introduction to modelling time series

Time seires econometrics is concerned with estimation of difference equations containing stochastic components. With time series, we want to see if the data contains any important economic information.

Definition 2.23 *(Finite-difference Equations). An equation of the form*

$$\pi_{t+1} = f(\pi_t)$$

*which relate values at **discrete times** is called a **finite-difference Equations**.*

Remark 2.24 *π_t is known as the **state** of the system at time t. The function f is known as the **dynamics** of the system and how it captures the state changes in time. We seek to figure out this function f.*

In particular, a solution to the finite difference equation is a sequence of states that satisfies the finite difference equation.

2.0.3 Solution by Forward Iteration

Given a known value at a time point, we iterate from that period and obtain subsequent time path of the entire y sequence. Let us denote this value as y_0. Hence, we can have

$$y_1 = a_0 + a_1 y_0 + \epsilon_1.$$

Likewise for y_2, we have that

$$y_2 = a_0 + a_1 y_1 + \epsilon_t$$

$$= a_0 + a_1(a_0 + a_1 y_0 + \epsilon_1) + \epsilon_2$$

$$= a_0(1 + a_1) + a_1^2 y_0 + a_1 \epsilon_1 + \epsilon_2.$$

Resultantly, we have that for all $t > 0$

$$y_t = a_0 \sum_{i=0}^{t-1} a_i^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_i^i \epsilon_{t-i}.$$

The **forcing process** here is $x_t = \sum_{i=0}^{t-1} a_i^i \epsilon_{t-i}$ term. Note that y_0 is not the very beginning of the series but the first point that is known.

To verify solutions, we can use substitution whereby we sub in the solution at t for the t+1 equation and see if it holds. Furthermore, if you notice, if $a_i < 1$, we will observe **exponential decay**, if $a_i > 1$, we will observe **exponential growth**, and if $a_i = 1$, we then just have a steady state. Furthermore, if $-1 < a_i < 0$, we have alternating decay, for $a_i < -1$ we have alternating growth and for $a_i = -1$, we have a periodic cycle.

2.0.4 Solution by Iteration without initial condition

Suppose we no longer have the initial condition y_0, we wouldn't be able to employ the technique from the solution by forward iteration method. Now we no longer have a point to terminate the process of iteration once we reached it, instead, we keep on iterating backwards past that point. Hence, from

$$y_t = a_0 \sum_{i=0}^{t-1} a_i^i + a_1^t y_0 + \sum_{i=0}^{t-1} a_i^i \epsilon_{t-i}.$$

substitute in $y_0 = \alpha_0 + \alpha_1 y_{t-1} + \epsilon_0$ so that

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t(\alpha_0 + \alpha_1 y_{t-1} + \epsilon_0) + \sum_{i=0}^{t-1} a_i^i \epsilon_{t-i}$$

$$= a_0 \sum_{i=0}^{t-1} a_1^i + \sum_{i=0}^{t} a_i^i \epsilon_{t-i} + a_1^{t+1} y_{-1}.$$

We then iterate backwards another m periods, so we arrive at

$$y_t = a_0 \sum_{i=0}^{t+m} a_1^i + \sum_{i=0}^{t+m} a_i^i \epsilon_{t-i} + a_1^{t+m+1} y_{-m-1}.$$

2.0.5 Cobweb Method

Here, we have issues with solving difference equations for nonlinear difference equations. Hence, we use the iteration methods of either the Cobweb method or the numerical iteration method. To run the Cobweb method, we draw a 45 degree line on a graph and the function we are interested in looking at it. We then pick an initial point y_0 and draw a line vertically until we hit the function of interest, then we draw a horizontal line until it hits the 45 degree line. Then, we draw another vertical line until we hit the function of interest and we keep iterating this process. This allows us to see the behaviour of this nonlinear system.

2.0.6 Numerical Iteration

Here, we simply look at each point and take the difference between this and the last point of interest. We stop iterating until we reach a point of convergence for our algorithm to terminate.

2.0.7 Simple Difference Equation

Recall that a **first order homogenous difference equation** is simply

$$y_t = \alpha y_{t-1}.$$

Recall the homogenous here refers to the fact that there is no drift in the system. In particular, the the solution is

$$y_t = \alpha^k y_0.$$

Now we look at **non-homogenous difference equations**. In particular, we now have a drift as seen in

$$y_t = \alpha y_{t-1} + b$$

for a non-zero b. Through iteration, we arrive at

$$y_t = \alpha^t y_0 + b \sum_{i=1}^{t} \alpha^{i-1}.$$

We have different cases for this in terms of behaviour.

Case One:

$\alpha = 1$ leads to a steady behaviour with $y_t = y_0 + bt$ where y_t grows linearly with growth rate b;

Case Two:

We then have that

$$y_t = \alpha^t y_0 + \frac{b(1 - \alpha^t)}{1 - \alpha}$$

$$= (y_0 - \frac{b}{1 - \alpha})\alpha^t + \frac{b}{1 - \alpha}.$$

Hence, if $\alpha > 1$, we have an explosion for y_t whilst if $\alpha < 1$, we have a convergence to $\frac{b}{1-\alpha}$.

For $\alpha = -1$, we have a periodic cycle as a result so that

$$y_t = (y_0 - \frac{b}{1 - \alpha})(-1)^t + \frac{b}{1 - \alpha}.$$

2.0.8 Difference Equation with Forcing Process

If we now had a deterministic/random process w_t rather than b, this changes things up.

$$y_t = \alpha y_{t-1} + w_t.$$

Here, w are shocks that causes changes in the value y. We can look at this through multiple ways.

Definition 2.25 *(Dynamic Multipler).*

Given the solution g, the effect of w_{t-i} on y_t can be written as

$$\frac{\partial y_t}{\partial w_{t-i}} = \frac{\partial g(t, w_1, ..., w_t, y_0)}{\partial w_{t-i}}$$

and thisi s known as the i-th period dynamic multiplier.

Definition 2.26 *(Impulse Response Function).*

This is defined to be

$$\frac{\partial y_t}{\partial w_{t-i}}$$

which keeps track of time path of all dyanamic multipliers

$$\frac{\partial y_t}{\partial w_t}, \frac{\partial y_t}{\partial w_{t-1}}, ..., \frac{\partial y_t}{\partial w_{t-i}}, ...$$

This looks at the effect of the input process w_t on the output process y_t.

2.0.9   Lag Operator

Definition 2.27 *(Lag Operator).*

The lag operator is a transformation that turns one time series into another by shifting everything back by one period.

$$Lx_t = x_{t-1}.$$

*For multiple periods we define the **lag polynomial** as*

$$(a_0L^0 + a_1L^1 + ... + a_pL^p)x_t = a_0x_t + a_1x_{t-1} + ... + a_px_{t-p}.$$

We can express this as
$$\alpha(L) = (a_0L^0 + a_1L^1 + a_2L^2 + ... + a_pL^p).$$

We note some properties of the lag operator.

  (i) Commutative a(L)b(L) = b(L)a(L);

  (ii) We can raise it to positive powers $a(L)^2 = a(L)a(L)$;

 (iii) We can factorise them and invert each term;

 (iv) The lag operator applied to constant is just the constant;

  (v) We can multiply lag opeator terms through.

Properties of lag operators

  1. Linearity;

  2. Associative.

Definition 2.28 *(Basic Models).*

We have models of the following form:

  1. *AR(P): $a(L)x_t = \epsilon_t$*

  2. *MA(Q): $x_t = b(L)\epsilon_t$*

  3. *ARMA(P,Q): $a(L)x_t = b(L)\epsilon_t$*

We can define the L^{-1} to be the inverse of the lag operator. Furthermore, L^0 is known as the **identity operator**.

Example: With Non-Homoegenous Equation with 1st difference: Suppose we had

$$(1 - \alpha L)y_t = b$$

where $\alpha < 1$. We can take the inverse and recall lag on constant is just the constant

$$= \left(\sum_{i=0}^{\infty} \alpha^i L^i \right) b$$

$$= \left(\sum_{i=0}^{\infty} \alpha^i \right) b$$

$$= \frac{1}{1-\alpha} b.$$

Example: With Difference Equation with Forcing Process: Replace b with w_t. We now have that

$$(1 - \alpha L)y_t = w_t$$

where w_t needs to be stochastically bounded. Again, we have that

$$= \sum_{i=0}^{\infty} \alpha^i w_{t-i}.$$

2.0.10 Multiple Representations of Time Series

We can express our models in different forms depending on how we want to work with them. Expressing them in the form of the summation of independent random variables can be quite useful in cases of finding variance and autocovariance of terms. In particular, we we can turn AR(1) series into a MA(∞).

AR(1) to MA(∞) by recursive substitution

Start with AR(1) $= x_t = ax_{t-1} + \epsilon$. By recursive substitution, you should we can arrive at

$$x_t = a^k x_{t-k} + a^{k-1} \epsilon_{t-k+1} + ... + a^2 \epsilon_{t-2} + a\epsilon_{t-1} + \epsilon_t.$$

Hence, we can express an AR(1) model as an ARMA(k,k-1) model. Furthermore, if we had $|a| < 1$, which then implies that $\lim_{k \to \infty} a^k x_{t-k} = 0$, we then have that

$$x_t = \sum_{j=0}^{\infty} a^j \epsilon_{t-j}.$$

AR(1) to MA(∞) with lag operators Another idea is to use Lag opeators. If we had

$$(1 - aL)x_t = \epsilon_t$$

we can then invert this by going

$$x_t = (1 - aL)^{-1} \epsilon_t$$

provided that $|a| < 1$. To interpret this inverse, recall the geometric series

$$(1 - z)^{-1} = 1 + z + z^2 + ...$$

for $|z| < 1$. Hence, using the same idea, we have that

$$x_t = (1 + aL + a^2 L^2 + ...)\epsilon_t = \sum_{j=0}^{\infty} a^j \epsilon_{t-j}.$$

2.0.11 AR(P) with lag operators

Suppose we had the AR(P) model of

$$y_t = a_0 y_{t-1} + a_1 y_{t-2} + \dots + \epsilon_t,$$

where each $\epsilon \sim$ white noise process. We can express this as

$$a(L) y_t = \epsilon_t,$$

where $a(L)$ is the L-th order lag polynomial

$$a(L) = 1 - a_0 L - a_1 L^2 - \dots - a_p L^p.$$

If we wanted to factorise a(L), we know that (from the Fundamental theorem of algebra),

$$a(L) = (1 - \lambda_1 L)(1 - \lambda_2 L)\dots(1 - \lambda_p L)$$

where λ_j are called the **characteristic roots** and they are the inverse of roots of $a(z) = 0$.

For a stationary process of y_t, we require that these characteristic roots λ all lie within the unit circle $|\lambda_j| < 1$ for j=1,...,p. Meanwhile, all polynomial roots must lie **within** the unit circle.

To then get a $MA(\infty)$ representation, we take the inverse where each inverse is

$$(1 - \lambda_i L)^{-1} = \sum_{k=1}^{\infty} \lambda_i^k L^k.$$

2.0.12 AR(2) with Lag Operators

If $\lambda_i < 1$ for all i $= 1,...,$p, then the process is stable. In particular, we will have

$$y_t = \alpha(L)^{-1} \epsilon_t$$

$$= \left(\sum_{j=0}^{\infty} \lambda_1^j L^j \right)\left(\sum_{j=0}^{\infty} \lambda_2^j L^j \right)\epsilon_t$$

$$= \sum_{i=0}^{\infty} c_i \epsilon_t$$

where we have that

$$c_i = \sum_{k=0}^{i} \lambda_1^k \lambda_2^{i-k}.$$

2.0.13 Characteristic Polynomials

The characteristic equation is defined as

$$\lambda^2 - \alpha_1 \lambda - \alpha_2 = 0$$

whereby the λ_1, λ_2 roots to this equation is known as the **characteristic roots**.

Some important things to recall regarding roots

$b^2 + 4a > 0 \rightarrow$ roots are real numbers; $b^2 + 4a < 0 \rightarrow$ roots are complex numbers and conjugate pairs; $b^2 + 4a = 0 \rightarrow$ roots all same values.

ECOS3904: Applied Macroeconometrics

3. Forward Looking Models

Forward Looking Models

3.0.14 Forward Looking

What if we had an issue where we had that $\alpha > 1$? We can now instead express things in a different manner to assist us. Now we look at the **lead operator**, which we define as L_*. In particular, note that

$$\alpha^* = \alpha^{-1}$$

and we have that

$$L^* = L^{-1}.$$

Hence, $L_* y_t = y_{t+1}$. Hence, we now have that

$$(1 - \alpha L)^{-1} = (1 - (\alpha^* L^*)^{-1})^{-1}$$

$$= -\alpha^* L^* (1 - \alpha^* L^*)^{-1}$$

$$= -\alpha^* L^* \sum_{k=0}^{\infty} (\alpha^* L^*)^k$$

$$= -\sum_{k=1}^{\infty} (\alpha^* L^*)^k.$$

Hence, applying it to

$$y_t = (1 - \alpha L)^{-1}$$

we have that

$$y_t = -\sum_{k=1}^{\infty} \alpha^{-k} w_{t+k}.$$

This is a forward looking representation of the model. This is a different scenario to the backward looking representation whereby if $\alpha > 1$ the system is stable but if $\alpha < 1$, then the system explores. The issue with this is that you don't actually have future values when modelling this.

3.0.15 Prediction

We want to see $E_t(x_{t+j})) = E(x_{t+j}|x_t, x_{t-1}, ..., \epsilon_t, \epsilon_{t-1}, ...)$.

AR(1) For AR(1), $x_{t+1} = \phi x_t + \epsilon_{t+1}$, we have $E(x_{t+1}) = \phi x_t$ for one period ahead. For 2 period ahead, we have $\phi^2 x_t$. We can keep iterating for this.

We note that for the unconditional mean and variance, we have that

$$\lim_{k \to \infty} E_t(x_{t+k}) = 0 = E(x_t).$$

$$\lim_{k \to \infty} Var(x_{t+k}) = \frac{1}{1 - \phi^2} \sigma_\epsilon^2 = var(x_t).$$

In particular, we note that the limits of conditional momnets is the unconditional moment.

3.0.16 Impulse-Response Functions

The impulse response function is the path that x follows as a result of a shock ϵ_t. In particular, $\epsilon_t = 1$ whilst all other points is 0.

We note that the $MA(\infty)$ representation is the same thing as the impulse-response function. Furthermore, calculating an $MA(\infty)$ representation is equivalent to simulating the impulse-response function.

ECOS3904: Applied Macroeconometrics

4. Power Spectrum

Power Spectrum

4.0.17   Power Spectrum

Definition 4.29 *(Strongly/Strictly Stationary).*

The joint PDF of x_t is independent of t for all s.

Definition 4.30 *(Weakly/Covariance Stationary).*

A process x_t is weakly/covariance Stationary if its first 2 moments are finite and the covariance $E(x_t, x_{t-j})$ depends only on j, not t.

Note that strongly stationary does **not** imply weakly stationary, due to the 2nd moment being finite is not always satsfied.

Theorem 4.31 *(Wold Decomposition Theorem).*

Any mean zero covariance stationary process $\{x_t\}$ can be represented in the form

$$x_t = \sum_{j=0}^{\infty} \theta_j \epsilon_{t-j} + \eta_t.$$

4.0.18   Spectral Representation

We can also represent time series in its *spectral density* form. Succinctly put, this is the fourier transform of the autocorrelation/autocovariance function.

We can represent complex numbers in numerous ways

   (i) A + Bi;

  (ii) $C(cos\theta + isin\theta)$;

 (iii) $Ce^{i\theta}$.

Here, $C = (A^2 + B^2)^{\frac{1}{2}}$ is the amplitude/magnitude and $\theta = tan^{-1}\frac{B}{A}$ is known as the angle/phase.

For waves, if we had
$$y_t = cos(\omega t)$$

where ω is the frequency and tells you how large the oscillation is. Increasing ω increases the oscillation dramatically. The period is defined as $\frac{2\pi}{\omega}$.

We can now construct a stochastic series of the form

$$y_t = \mathbb{A}cos(\omega t) + \mathbb{B}sin(\omega t).$$

The random variables $\mathbb{A}$ and $\mathbb{B}$ have zero mean, are mutually uncorrelated, and have a common finite variance.

From this, we can also break y up into multiple components to have

$$y_t = \sum_{j=1}^{n} \left[\mathbb{A}_j cos(\omega_j t) + \mathbb{B}_j sin(\omega_j t)\right]$$

The set of random variables $\{\mathbb{A}_j, \mathbb{B}_j\}_{j=1}^{n}$. We can express this in complex exponential form as

$$y_t = ce^{i\omega t} + \bar{c}e^{-i\omega t}.$$

In particular, we have that

$$E(y_t) = 0;$$

$$Var(y_t) = \sum_{j=1}^{n} \sigma_j^2;$$

$$Cov(y_t, y_{t-k}) = \sum_{j=1}^{n} \sigma_j^2 cos(\omega_j k).$$

Here, we see that the variant contribution from different components to the variance and autocovariance. Hence, different ω_j represents the different components of a time series. We use the variance to see how large the random variable is, and in this case, it is called the **power/energy** of that component.

Now we take a final step of now have a continuous sort of distribution.

$$y_t = \int_0^\pi cos(\omega t)d\mathbb{A}(\omega) + \int_0^\pi sin(\omega t)d\mathbb{B}(\omega)$$

where $d(\mathbb{A}(\omega)$ and $d(\mathbb{B}(\omega)$ are random variables indexed by frequency ω. We can express this in complex exponential form.

$$y_t = \int_{-\pi}^\pi e^{i\omega t}dZ(\omega)$$

where we define $dZ(\omega) = \frac{1}{2}[dA(\omega) - idB(\omega)]$ for $\omega \geq 0$ and $dZ(\omega) = dZ(\bar{-\omega})$ for $\omega < 0$. This is known as the **spectral representation** of a weakly stationary process whilst $dZ(\omega)$ is the **random spectral measure** on $[-\pi, \pi]$.

Definition 4.32 *(Spectral Measure).*

We can take

$$Var[dZ(\omega)] = dF(\omega)$$

for $w \in [-\pi, \pi]$.

This can be interpreted as the **variance/power/energy** of the cyclical component of y corresponding to frequency ω.

Definition 4.33 *(Density).*

Given a $dF(\omega)$, we can define a nonnegative function $S(\omega)$ such that

$$dF(\omega) = S(\omega)\frac{d\omega}{2\pi}$$

*for $\omega \in [-\pi, \pi]$. Here $S(\omega)$ is known as the **density**.*

4.0.19   Autocorrelation and Autocovariance Functions

Definition 4.34 *(Autocovariance).*

The autocovariance of a series x_t is defined as

$$\gamma_j = cov(x_t, x_{t-j}) = \mathbb{E}\big(x - E(x_t)\big)\big(x_{t-j} - E(x_{t-j})\big)$$

For models where we do not have a mean, then we have $E(x_t) = 0$. In that case, we have that $\gamma_j = E(x_t x_{t-j})$. Furthermore, note that $\gamma_0 = Var(x_t)$.

Definition 4.35 *(Autocorrelation).*

We can normalise the autocvariance by dividing it by $\gamma(0)$. Hence, we have that

$$\rho_j = \frac{\gamma_j}{\gamma_0} = \frac{\gamma_j}{Var(x_t)}.$$

If given a spectral density $S(\omega)$, we can derive the autocovariance function as

$$\gamma(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega k} S(\omega) d\omega.$$

If we set k=0, we get

$$\gamma(0) = Var(y_t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S(\omega) d\omega.$$

The uses of autocovariance and autocovariance of a time series helps us to model the joint distribution of a time series $\{x_t\}$. Hence, from this we can see the persistence of a time series. Furthermore, the MA representation makes thing easy to compute variances as all covariance terms $(E(\epsilon_t, \epsilon_{t-k})) = 0$.

Definition 4.36 *(Fourier Transform).*

For any series of numbers $\{x_t\}$, the fourier transform is defined to be

$$x(\omega) = \sum_{t=-\infty}^{\infty} e^{-i\omega t} x_t.$$

This takes a series as a function of t and turns it into a complex valued function of ω.

Definition 4.37 *(Inverse Fourier Transform).*

For any complex valued function that is a function of ω, we can turn it into a function of t by

$$x_t = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{+i\omega t} x(\omega) d\omega.$$

As $e^{i\omega t} = cos\theta + isin\theta$, that means the inverse fourier transform expresses any x_t as a linear combination of sines and cosines functions.

Now, we let $x_t = \gamma_j$ or in other words, we sub in the autocovariance of the series into the Fourier transform.

Definition 4.38 *(Spectral Density).*

The spectral density $S(\omega)$ is the Fourier transform of the autocovariance function

$$S(\omega) = \sum_{j=-\infty}^{\infty} e^{-i\omega j} \gamma_j.$$

As γ_j is symmetric, we have that

$$S(\omega) = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j cos(j\omega).$$

Additionally, from $S(\omega)$, we can recover the autocovariance γ_j by the inverse Fourier transform of

$$\gamma_j = \frac{1}{2\pi} \int_{\pi}^{\pi} e^{+i\omega j} S(\omega) d\omega.$$

From this, we can divide everything of the Fourier transform to get the **autocorrelation** function and recall that $\rho_j = \gamma_j/\gamma_0$. Hence, we have that

$$f(\omega) = \frac{S(\omega)}{\gamma_0} = \sum_{j=-\infty}^{\infty} e^{-i\omega j} \rho_j.$$

Spectral Representation of White Noise

The power spectral density of ϵ_t is
$$S_\epsilon(\omega) = \sigma^2$$
for all $\omega \in [-\pi, \pi]$.

4.0.20 Filtering

Linear filters are **infinite-order lag polynomials**.

Definition 4.39 *(Filtering). Applying a moving average to a series x_t.*

$$y_t = \sum_{j=-\infty}^{\infty} b_j x_{t-j} = b(L) x_t.$$

Hence, y_t is a filtered version of x_t.

ECOS3904: Applied Macroeconometrics

13. Conditional Variance/Volatility Models

13.1 Conditional Variance/Volatility Models

13.1.1 Motivation

When we looked at ARMA models, these rely only on the conditional mean. However, in volatile data, it makes more sense to also consider the conditional variance of the data. This will give us a "better" information set and help us do inference and forecasting more accurately.

In particular, when conducting forecasts, using only the conditional mean gives us point forecasts whilst conditional variances gives us interval forecasts. In particular,

$$\text{Conditional Mean} : E(X_{T+1}|I_t)$$

$$\text{Conditional Variance} : Var(X_{T+1}|I_t)$$

13.1.2 AR(1) Model

Looking at the stationary AR(1) model

$$y_t = c + \alpha y_{t-1} + \epsilon_t$$

where $\epsilon_t \sim iid\ WN(0, \sigma^2)$.

Unconditional mean: $\frac{c}{1-\alpha}$ for all t.

Conditional mean: $E[y_t|I_{t-1}] = c + \alpha y_{t-1}$.

Notice now that conditional mean changes over time as it incorporate updated information.

We want to move from a point forecast (1st moment of conditional distribution) into an interval forecast, whereby we need to know the conditional variance of y_{t+1} given I_t.

$$Var(y_{t+1}) := E_t[y_{t+1} - E_t(y_{t+1})]^2.$$

Hence, for simple AR(1) model, the variance is

$$Var(y_{t+1}) := E_t[y_{t+1} - (c + \alpha y_t)]^2 = E_t(\epsilon_{t+1}^2) = \sigma^2$$

where σ^2 is the variance of the error term. So, there is an issue here with the AR (and ARMA) model that the conditional variance of y_t is **time-invariant** just like the unconditional variance. Hence, we call them **unconditionally and conditionally homoskedastic**.