ECOM90024 – FORECASTING IN ECONOMICS & BUSINESS

LECTURE 5: AUTOREGRESSIVE & MOVING AVERAGE MODELS

TODAY’S LECTURE

- Moving average processes
- The Lag Operator
- The AR(2) Process
- The AR(P) Process
- The Yule – Walker equations

MOVING AVERAGE PROCESSES

- In the previous lecture we derived the properties of an AR(1) process. Now let's consider,

$$Y_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

- Where ε_t is a sequence of iid errors with mean zero and variance σ^2.

- This series is called a ***first order moving average*** process or MA(1) as it species Y_t to be a weighted sum of the two most recent values of ε.

- The expectation of an MA(1) is easily computed as,

$$E[Y_t] = E[\varepsilon_t + \theta \varepsilon_{t-1}] = 0$$

MOVING AVERAGE PROCESSES

- The variance of an MA(1) is computed as,

$$E[(Y_t - \mu)^2] = E[(\varepsilon_t + \theta\varepsilon_{t-1})^2] = E[\varepsilon_t^2 - 2\theta\varepsilon_t\varepsilon_{t-1} + \theta^2\varepsilon_{t-1}^2]$$

- Simplification yields,

$$\gamma(0) = E[(Y_t - \mu)^2] = \sigma^2 + \theta^2\sigma^2 = (1 + \theta)\sigma^2$$

- Hence the variance of an MA(1) process is finite so long as the variance of the error process is finite.

MOVING AVERAGE PROCESSES

- The first order autocovariance of an MA(1) process is,

$$\gamma(1) = E[(Y_t - \mu)(Y_{t-1} - \mu)] = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-1} + \theta\varepsilon_{t-2})] = \theta\sigma^2$$

- All higher autocovariances of order $\tau > 1$ will be zero,

$$\gamma(\tau) = E[(Y_t - \mu)(Y_{t-\tau} - \mu)] = E[(\varepsilon_t + \theta\varepsilon_{t-1})(\varepsilon_{t-\tau} + \theta\varepsilon_{t-\tau-1})] = 0$$

- Therefore, it is clear than an MA(1) process is covariance stationary.

MOVING AVERAGE PROCESSES

- Similarly, the first order autocorrelation coefficient is given by,

$$\rho(1) = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta\sigma^2}{(1+\theta^2)\sigma^2} = \frac{\theta}{(1+\theta^2)}$$

- While all higher order autocorrelations are zero.

- Note that unlike the autoregressive model, we do not require that the parameter θ be less than 1 in absolute value in order for covariance stationarity to hold! θ can be any real number!

MOVING AVERAGE PROCESSES

- A q-th order moving average process, denoted MA(q), is characterized by,

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

- Where ε_t is a sequence of iid errors with mean zero and variance σ^2.
- As with the MA(1), the mean of the process is given by

$$E[Y_t] = E\left[\varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}\right] = 0$$

- While the variance is

$$\gamma(0) = \sigma^2 + \theta_1^2 \sigma^2 + \theta_2^2 \sigma^2 + \cdots + \theta_q^2 \sigma^2$$

MOVING AVERAGE PROCESSES

- The autocovariances for an MA(q) process are computed as,

$$\gamma(\tau) = E\begin{bmatrix} \left(\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \ \theta_q\varepsilon_{t-q}\right) \\ \times\left(\varepsilon_{t-\tau} + \theta_1\varepsilon_{t-\tau-1} + \theta_2\varepsilon_{t-\tau-2} + \cdots + \ \theta_q\varepsilon_{t-\tau-q}\right) \end{bmatrix}$$

- Dropping all the cross products, we have

$$\gamma(\tau) = E\left[\theta_\tau\varepsilon_{t-\tau}^2 + \theta_{\tau+1}\theta_1\varepsilon_{t-\tau-1}^2 + \theta_{\tau+2}\theta_2\varepsilon_{t-\tau-2}^2 + \cdots + \theta_q\theta_{q-\tau}\varepsilon_{t-\tau-2}^2\right]$$

- Hence,

$$\gamma(\tau) = \sigma^2\left(\theta_\tau + \theta_{\tau+1}\theta_1 + \theta_{\tau+2}\theta_2 + \cdots + \theta_q\theta_{q-\tau}\right) \text{ for } \tau = 1,2, \dots, q \text{ and } \theta_0 = 1$$

- The autocovariances of order higher than q will be zero.

MOVING AVERAGE PROCESSES

- For example, for an MA(2) process,

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

- We have that

$$\gamma(0) = (1 + \theta_1^2 + \theta_2^2)\sigma^2$$

$$\gamma(1) = (\theta_1 + \theta_2\theta_1)\sigma^2$$

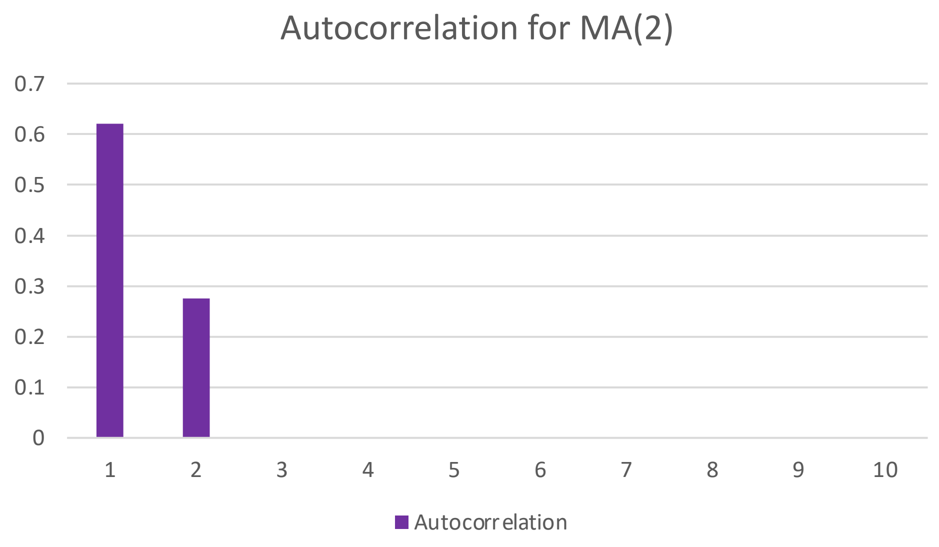
$$\gamma(2) = \theta_2\sigma^2$$

$$\gamma(3) = \gamma(4) = \cdots = 0$$

- Therefore, it is clear that a MA(q) process is covariance stationary.

MOVING AVERAGE PROCESSES

- The autocovariance and autocorrelation functions of moving average processes have clear cut-offs. For an MA(2), we have that



$$Y_t = \varepsilon_t + 1.5\varepsilon_{t-1} + 2\varepsilon_{t-2}$$

We will talk about the partial autocorrelations a little later on!

COVARIANCE STATIONARITY & STABILITY

- From our analysis of the AR(1) and MA models, we can see from their correlograms that their autocorrelations decay as we increase the displacement parameter τ.

- This decay tells us that the effect of a shock in period t will eventually die out as we move forward in time.

- We can contrast this to the random walk model in which shocks accumulate over time which causes the variance of the series to blow up.

- Therefore, the class of covariance stationary time series models can be thought of as models of stable dependence.

COVARIANCE STATIONARITY & STABILITY

- One way to think about a non-stationary time series is that it is an ***explosive*** series. Consider the following AR(1) process,

$$Y_t = 1.2Y_{t-1} + \varepsilon_t$$

- A shock in period t becomes accumulated and magnified as we move forward in time!

- When the variance of a time series is increasing over time, forecasts become increasingly poor as we increase the forecast horizon!

THE LAG OPERATOR

- As we move into a deeper analysis and application of autoregressive and moving average models, some of the manipulations can start to become rather cumbersome.

- A highly useful device which will make our lives a little easier is the ***lag operator*** which we will represent by the symbol L

$$Lx_t \equiv x_{t-1}$$

- Applying the lag operator twice yields,

$$L(Lx_t) = L^2 x_t = x_{t-2}$$

- In general, for any integer k,

$$L^k x_t = x_{t-k}$$

THE LAG OPERATOR

- The lag operator is distributive over the addition operator,

$$L(x_t + w_t) = Lx_t + Lw_t = x_{t-1} + w_{t-1}$$

- In fact, we are free to use the standard commutative, associative and distributive algebraic laws for multiplication and addition,

$$y_t = (a + bL)Lx_t = (aL + bL^2)x_t = ax_{t-1} + bx_{t-2}$$

- Note that the lag operator applied to a constant (i.e., a quantity that does not vary over time) simply returns that same constant,

$$Lc = L^2c = c$$

THE LAG OPERATOR

- Let’s now apply the lag operator to an AR(1) model,

$$Y_t = \phi Y_{t-1} + \varepsilon_t = \phi L Y_t + \varepsilon_t$$

- Then we can see that we can rewrite it as,

$$(1 - \phi L) Y_t = \varepsilon_t$$

- We know that when $|\phi| < 1$, we can use the property of a geometric sum to write,

$$\frac{1}{1 - \phi L} = \lim_{j \to \infty} \left(1 + \phi L + \phi^2 L^2 + \cdots + \phi^j L^j\right)$$

- So that we can easily see the result we obtained via recursive substitution that an AR(1) can be written as an infinite moving average process,

$$Y_t = \frac{1}{1 - \phi L} \varepsilon_t = \varepsilon_t + \phi \varepsilon_{t-1} + \phi^2 \varepsilon_{t-2} + \phi^3 \varepsilon_{t-3} + \cdots$$

COVARIANCE STATIONARITY & LAG OPERATOR

- Recall that the AR(1) model may be written as:

$$(1 - \phi L)Y_t = \varepsilon_t$$

- We have shown that if $|\phi| \geq 1$, the AR(1) process becomes nonstationary.

- This is equivalent to the root of the first order polynomial $(1 - \phi L)$ being less than 1 in absolute value!

- That is, for $(1 - \phi L) = 0$, $L = \frac{1}{\phi}$ so that if $|\phi| \geq 1$ it must be the case that $|L| \leq 1$

- This characterization of non-stationarity is useful when we move onto higher order AR models!

THE AR(2) MODEL

- A second order autoregression, denoted AR(2) satisfies,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

- Using the lag operator, we can write this as,

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = \varepsilon_t$$

- We know from our analysis of the AR(1) model that covariance stationarity is achieved when $\phi < 1$. But what about the AR(2) model?

- What are the conditions on ϕ_1 and ϕ_2 that we must impose in order to guarantee the covariance stationarity of an AR(2) process?

THE AR(2) MODEL

- When written in terms of lag operators, an AR(2) model is represented by a second order polynomial in the lag operator L.

- Suppose we factor this polynomial, that is, find numbers λ_1 and λ_2 such that,

$$(1 - \phi_1 L - \phi_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L) = 1 - (\lambda_1 + \lambda_2)L + \lambda_1 \lambda_2 L^2$$

- In other words, given values for ϕ_1 and ϕ_2, we seek numbers λ_1 and λ_2 such that,

$$\lambda_1 + \lambda_2 = \phi_1$$

$$\lambda_1 \lambda_2 = -\phi_2$$

THE AR(2) MODEL

- Therefore, as was the case with the AR(1) model, we need to find the roots of the lag polynomial!

- Since L is an operator, to make things super clear and simple, let's replace it with a scalar z so that we are trying to find λ_1 and λ_2 such that

$$(1 - \phi_1 z - \phi_2 z^2) = (1 - \lambda_1 z)(1 - \lambda_2 z)$$

- Dividing both sides by z^2, we obtain

$$(z^{-2} - \phi_1 z^{-1} - \phi_2) = (z^{-1} - \lambda_1)(z^{-1} - \lambda_2)$$

THE AR(2) MODEL

- Now define

$$\lambda \equiv z^{-1}$$

- Then we can rewrite our polynomial as

$$(\lambda^2 - \phi_1 \lambda - \phi_2) = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

- Now it must be true that the values of λ that set the right side to zero are $\lambda = \lambda_1$ and $\lambda = \lambda_2$. These same values must set the left side to zero as well

$$(\lambda^2 - \phi_1 \lambda - \phi_2) = 0$$

THE AR(2) MODEL

- Now recall from high school mathematics that given the quadratic equation,

$$ax^2 + bx + c$$

- The roots are given by

$$x = \frac{-b \pm \sqrt{b^2 + 4ac}}{2a}$$

- Applying this to our polynomial, $(\lambda^2 - \phi_1 \lambda - \phi_2)$, we obtain,

$$\lambda_1 = \frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

$$\lambda_2 = \frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

THE AR(2) MODEL

- An AR(2) process will be covariance stationary if the roots of

$$(\lambda^2 - \phi_1 \lambda - \phi_2)$$

lie ***inside*** the unit circle. Or equivalently, if the roots of

$$(1 - \phi_1 z - \phi_2 z^2)$$

lie ***outside*** the unit circle.

THE AR(2) MODEL

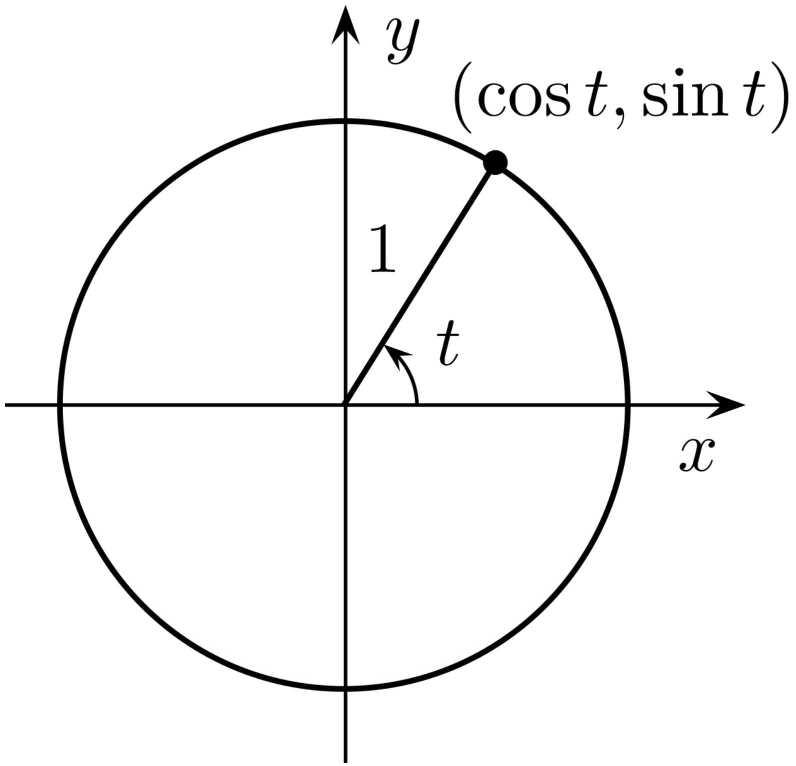
- A unit circle is a circle with a radius equal to 1.

- The reason why we use the term ***unit circle*** in our condition for covariance stationarity is due to the fact that there will exist values of ϕ_1 and ϕ_2 such that

$$\phi_1^2 + 4\phi_2 < 0$$

Which will produce roots that are complex numbers.

- For the case where $\phi_1^2 + 4\phi_2 > 0$, the AR(2) process will be covariance stationary when the roots of $(\lambda^2 - \phi_1\lambda - \phi_2)$ are less than $|1|$.



THE AR(2) MODEL

- Now let's translate this condition into a set of values for ϕ_1 and ϕ_2. Suppose that the arithmetically larger solution λ_1 is greater than 1

$$\frac{\phi_1 + \sqrt{\phi_1^2 + 4\phi_2}}{2} > 1$$

- This occurs when

$$\sqrt{\phi_1^2 + 4\phi_2} > 2 - \phi_1$$

- Assuming that the solution is real, i.e. $\sqrt{\phi_1^2 + 4\phi_2} > 0$, then the above inequality would be satisfied for any value of $\phi_1 > 2$

THE AR(2) MODEL

- If on the other hand, $\phi_1 < 2$, we can square both sides to conclude that λ_1 will be greater than 1 when

$$\phi_1^2 + 4\phi_2 > 4 - 4\phi_1 + \phi_1^2$$

- Which simplifies to

$$\phi_2 > 1 - \phi_1$$

- Thus for real solutions, λ_1 will be greater than 1 either if $\phi_1 > 2$ or if (ϕ_1, ϕ_2) lie northeast of the line $\phi_2 = 1 - \phi_1$

THE AR(2) MODEL

- Now let's consider the arithmetically smaller solution λ_2. This will be less than -1 when,

$$\frac{\phi_1 - \sqrt{\phi_1^2 + 4\phi_2}}{2} < -1$$

- Which we can simplify to

$$\sqrt{\phi_1^2 + 4\phi_2} > 2 + \phi_1$$

- Again, if we assume that the solution is real, this inequality will be satisfied if $\phi_1 < -2$

THE AR(2) MODEL

- In the case when $\phi_1 > -2$, we again square both sides of the inequality to obtain

$$\phi_1^2 + 4\phi_2 > 4 + 4\phi_1 + \phi_1^2$$

- Which we can simplify to

$$\phi_2 > 1 + \phi_1$$

- Thus in the real region, the solution λ_2 will be less than -1 either if $\phi_1 < -2$ or (ϕ_1, ϕ_2) lie northwest of the line $\phi_2 = 1 + \phi_1$

THE AR(2) MODEL

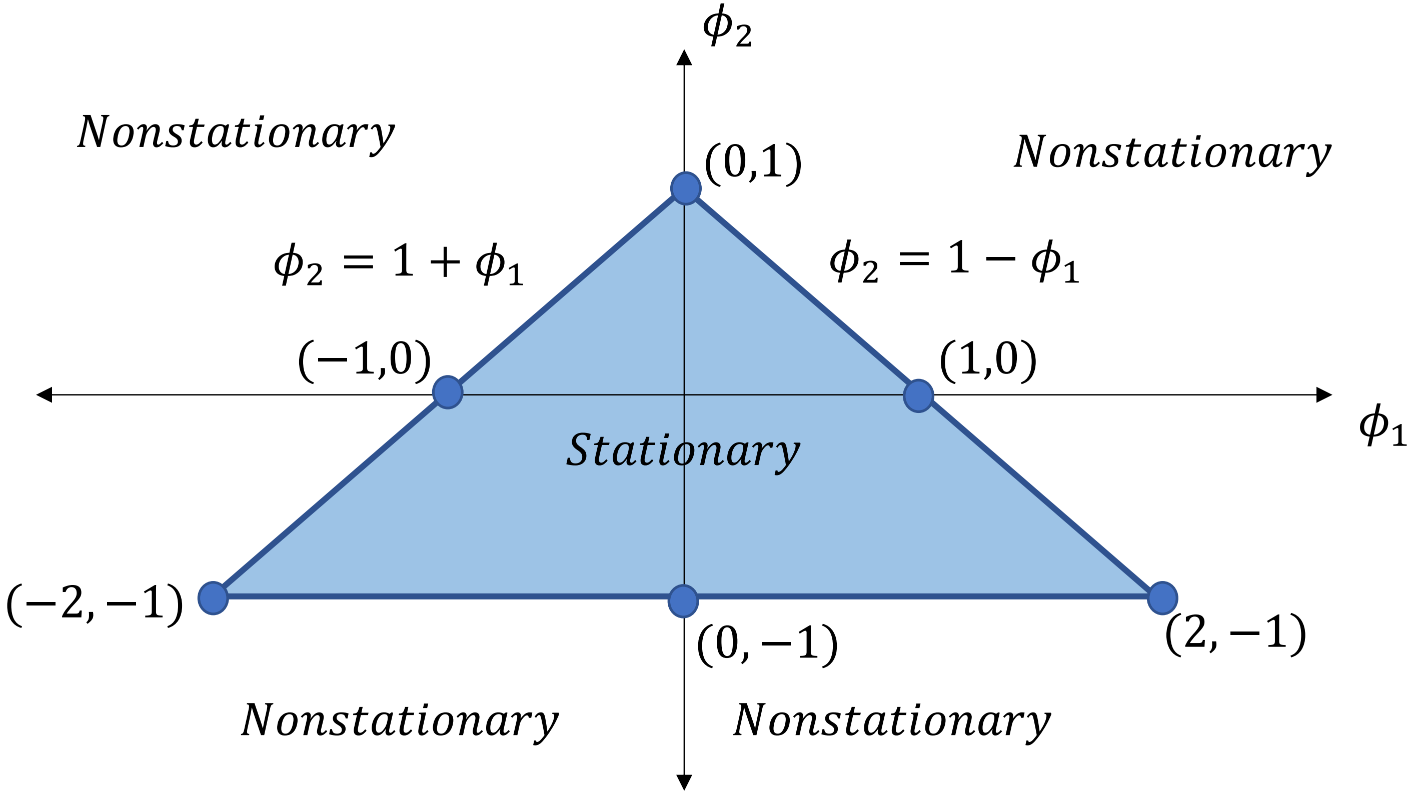
- Therefore we have the following conditions on our AR(2) parameters that must be satisfied in order for the process to be covariance-stationary:

1. $\phi_1 < 2$
2. $\phi_1 > -2$
3. $\phi_2 < 1 - \phi_1$
4. $\phi_2 < 1 + \phi_1$
5. $\phi_2 > -1$ (*This condition is required for the set of complex solutions*)

THE AR(2) MODEL

- Visually,

All the values of ϕ_1 and ϕ_2 that lie within this triangle sum to less than $|1|$.



THE AR(2) MODEL

- If we have an AR(2) process that is covariance stationary, then we can compute it’s mean as

$$E[Y_t] = \phi_1 E[Y_{t-1}] + \phi_2 E[Y_{t-2}] + E[\varepsilon_t]$$

- Writing $E[Y_t] = \mu$ we have that

$$\mu = \phi_1 \mu + \phi_2 \mu + 0$$

- Which we can rewrite as

$$\mu(1 - \phi_1 - \phi_2) = 0$$

- Therefore $E[Y_t] = \mu = 0$

THE AR(2) MODEL

- To compute the autocovariances, we first rewrite

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \varepsilon_t$$

- Then, multiplying both sides by $(Y_{t-j} - \mu)$ and taking expectations we obtain,

$$E\left[(Y_t - \mu)\left(Y_{t-j} - \mu\right)\right] = \gamma_j = \phi_1\,\gamma_{j-1} + \phi_2\,\gamma_{j-2}$$

- Thus, the autocovariances follow the same second-order difference equation as does the process for Y_t, with the difference equation for γ_j indexed by the lag j

THE AR(2) MODEL

- The autocorrelations are easily found by dividing both sides of the autocovariance function by γ_0

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2}$$

- In particular, setting $j = 1$ produces

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

- This is obtained from the symmetry of the autocovariances $\gamma_j = \gamma_{-j}$. Further simplification yields,

$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$

- Then for $j = 2$ we have that

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

This system of equations also tells me that I can recover the autoregressive coefficients if I know the autocorrelations!

THE AR(2) MODEL

- The variance of the AR(2) process can be found by multiplying both sides by $(Y_t - \mu)$ and taking expectations,

$$E[(Y_t - \mu)^2] = \phi_1 E[(Y_{t-1} - \mu)(Y_t - \mu)] + \phi_2 E[(Y_{t-1} - \mu)(Y_t - \mu)] + E[\varepsilon_t(Y_t - \mu)]$$

- This can be rewritten more compactly as

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + E[\varepsilon_t(Y_t - \mu)]$$

- Looking at the last term, we notice that

$$E[\varepsilon_t(Y_t - \mu)] = E[\varepsilon_t(\phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \varepsilon_t)] = \phi_1.0 + \phi_2.0 + \sigma^2$$

- Therefore,

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\gamma_2 + \sigma^2 = \phi_1\rho_1\gamma_0 + \phi_2\rho_2\gamma_0 + \sigma^2$$

THE AR(2) MODEL

- Since we have shown that,

$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$

$$\rho_2 = \phi_1\rho_1 + \phi_2$$

- We have that

$$\gamma_0 = \left[\frac{\phi_1^2}{1-\phi_2} + \frac{\phi_2\phi_1^2}{1-\phi_2} + \phi_2^2\right]\gamma_0 + \sigma^2$$

- Solving for γ_0 yields,

$$\gamma_0 = \frac{(1-\phi_2)\sigma^2}{(1+\phi_2)[(1-\phi_2)^2 - \phi_1^2]}$$

THE AR(p) MODEL

- A p-th order autoregression, denoted AR(p), satisfies

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$$

- Written in terms of lag operators, it is given by,

$$\left(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p\right) Y_t = \varepsilon_t$$

- Therefore, generalizing the result we derived from the AR(2) case, we can say that the AR(p) process will be covariance-stationary so long as the roots of

$$\left(\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \cdots - \phi_{p-1}\lambda - \phi_p\right) = 0$$

- ***Lie inside the unit circle*** or equivalently if the roots of

$$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$$

- ***Lie outside the unit circle***.

A necessary condition for covariance stationarity is that the sum of the coefficients $\phi_1, \phi_2, \dots, \phi_p$ be less than $|1|$.

THE AR(p) MODEL

- Assuming that the stationarity condition is satisfied, the mean of the AR(p) will be given by,

$$E[Y_t] = \phi_1\mu + \phi_2\mu + \cdots + \phi_p\mu$$

- While the autocovariances are given by

$$\gamma_j = \phi_1\gamma_{j-1} + \phi_2\,\gamma_{j-2} + \cdots + \phi_p\,\gamma_{j-p} \quad \text{for } j = 1,2,\ldots$$

- And the variance is given by

$$\gamma_0 = \phi_1\gamma_1 + \phi_2\,\gamma_2 + \cdots + \phi_p\,\gamma_p + \sigma^2$$

THE AR(p) MODEL

- Dividing the autocovariance function by the variance produces the autocorrelation function,

$$\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p} \qquad \text{for } j = 1,2,3 \ldots$$

- Where from the property of covariance stationarity we have that $\rho_{-k} = \rho_k$

- The set of equations that generates the autocorrelations is known as the ***Yule-Walker equations***

THE YULE-WALKER EQUATIONS

- To see why the Yule-Walker equations are useful and interesting, let's write out the set of autocorrelations

$$
\begin{gathered}
\rho_1 = \phi_1\rho_0 + \phi_2\rho_1 + \cdots + \phi_p\rho_{p-1} \\
\rho_2 = \phi_1\rho_1 + \phi_2\rho_0 + \cdots + \phi_p\rho_{p-2} \\
\vdots \\
\vdots \\
\rho_p = \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_p\rho_0
\end{gathered}
$$

THE YULE-WALKER EQUATIONS

- We can stack all these elements into matrices,

$$\begin{bmatrix}\rho_1\\ \rho_2\\ \vdots\\ \rho_p\end{bmatrix} = \begin{bmatrix}\rho_0 & \rho_1 & \cdots\\ \rho_1 & \rho_0 & \ddots\\ \vdots & \vdots & \ddots\\ \rho_{p-1} & \rho_{p-2} & \cdots\end{bmatrix}\begin{bmatrix}\phi_1\\ \phi_2\\ \vdots\\ \phi_p\end{bmatrix}$$

- Which can be represented as

$$\rho = \mathrm{P}\Phi$$

- As it turns out, the matrix P is symmetric, square and full rank which means that it’s invertibility is guaranteed. Therefore

$$\Phi = \mathrm{P}^{-1}\rho$$

- This tells us that we can recover the entire set of autoregressive coefficients from the autocorrelation coefficients. This means that we don’t even need OLS to estimate the set of autoregressive coefficients, all we need to compute are the sample autocorrelations! **THIS IS SO COOL**!