ECOM90024 – FORECASTING IN ECONOMICS & BUSINESS

LECTURE 6: THE GENERAL LINEAR MODEL (ARMA)

TODAY'S LECTURE

- Invertibility of an MA process
- Unconditional vs. Conditional Moments
- The General Linear Model
- A General Approach to Estimating a Time Series Model

AN INVERTIBLE MA PROCESS

- Recall that an MA(1) process is given by,

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

$$\varepsilon_t \sim_{iid} (0, \sigma^2)$$

- In the previous lecture we computed the ***unconditional*** mean and variance of the process as,

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

$$var(Y_t) = var(\varepsilon_t) + \theta^2 var(\varepsilon_{t-1}) = \sigma^2(1 + \theta^2)$$

AN INVERTIBLE MA PROCESS

- We also showed that the first autocovariance and autocorrelations are given by,

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

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

- While the higher order autocovariance and autocorrelations are zero.

AN INVERTIBLE MA PROCESS

- The requirements of covariance stationarity are met for any MA(1) process regardless of the value of it's parameter θ.

- However, if we impose the condition that $|\theta| < 1$, we can ***invert*** the MA(1) process express the current value of the series in terms of a current shock and lagged values of the series. That is, we can obtain a stable (i.e. convergent) autoregressive representation of the moving average!

- To see this we rewrite our MA(1) as,

$$\varepsilon_t = y_t - \theta\varepsilon_{t-1}$$

AN INVERTIBLE MA PROCESS

- This implies that,

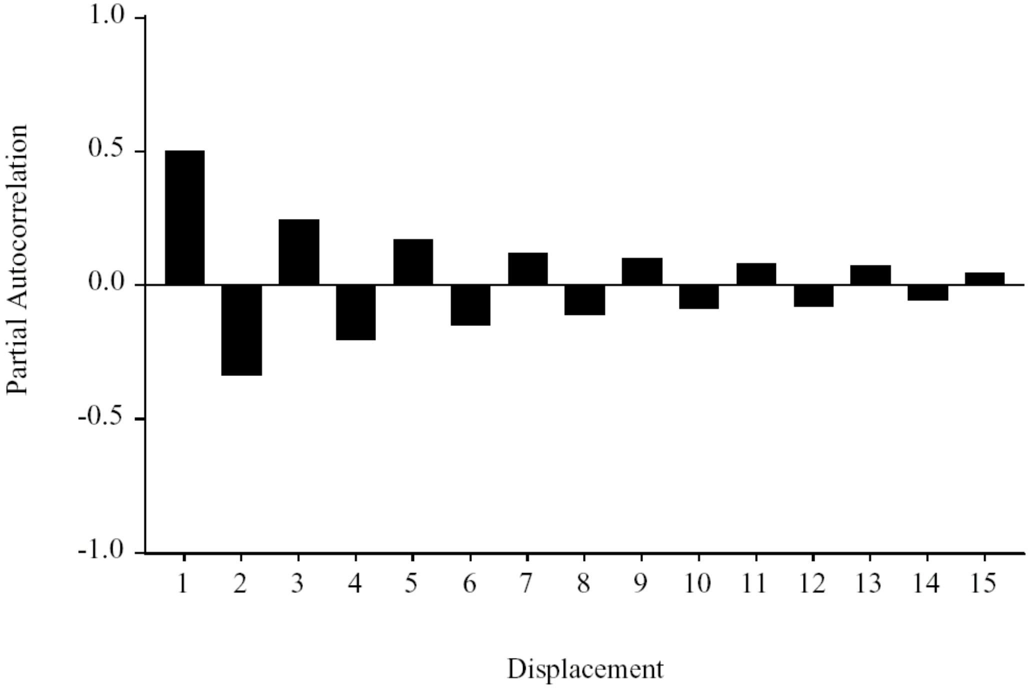
$$\varepsilon_{t-1} = y_{t-1} - \theta \varepsilon_{t-2}$$
$$\varepsilon_{t-2} = y_{t-2} - \theta \varepsilon_{t-3}$$
$$\vdots$$

- Making use of these expressions for lagged innovations, we can substitute backward in the MA(1) process, yielding an infinite order autoregressive process in which the autoregressive coefficients alternate in sign and converge to zero as we move back infinitely in time.

$$Y_t = \varepsilon_t + \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3 \, Y_{t-3} - \cdots$$

AN INVERTIBLE MA PROCESS

- From our understanding of autoregressive processes, we know that for $\theta > 0$, such a model will produce a partial autocorrelation function that oscillates and decays.



What will the partial autocorrelation function look like for $\theta < 0$?

AN INVERTIBLE MA PROCESS

- Now let's consider a general finite-order moving average process of order q,

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

- We can represent this process using a q-th order lag polynomial

$$\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

- So that

$$Y_t = \Theta(L)\ \varepsilon_t$$

AN INVERTIBLE MA PROCESS

- The MA(q) process will be invertible so long as the roots of

$$1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = 0$$

- Lie ***outside*** the unit circle. If this is the case, then the MA(q) can be expressed as infinite autoregressive model.

$$\frac{1}{\Theta(L)} Y_t = \varepsilon_t$$

AN INVERTIBLE MA PROCESS

- If the invertibility conditions are satisfied, both the MA(1) and MA(q) processes can be rewritten as AR(∞).

- From the previous lecture we saw that an MA(1) has an autocorrelation function that cuts off at $\tau = 1$, while an MA(q) process has an autocorrelation function that cuts off at $\tau = q$.

- If invertible, both the MA(1) and MA(q) processes will have partial autocorrelation functions that decay gradually.

UNCONDITIONAL VS. CONDITIONAL MOMENTS

- So far, when computing the means and variances of our AR and MA processes, we have restricted to our attention to the ***unconditional moments***.
- The unconditional moments describe the stochastic behaviour of our time series.
- We have seen that for covariance-stationary time series, the unconditional moments do not change over time!
- As forecasters however, it is also of interest to us to explore the behaviour of these time series with regards to an evolving information set!

CONDITIONAL MOMENTS FOR MOVING AVERAGE PROCESS

- Let the information set at time t be given by

$$\Omega_{t-1} = \varepsilon_{t-1}, \varepsilon_{t-2}, \dots,$$

- The conditional mean of an MA(q) is then given by

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

- Recall that when we compute conditional expectations, we treat all the objects in the in the conditioning set as known (i.e. not random). Thus, from the linearity of the expectations operator, we have that

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

- Since $\varepsilon_t \sim_{iid} (0, \sigma^2)$, it must be the case that $E[\varepsilon_t|\Omega_{t-1}] = 0$ so that,

$$E[Y_t|\Omega_{t-1}] = \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q}$$

We can think of the information set as comprising of all that is known (or observed) at the beginning of time t!

The conditional mean moves over time in response to the evolving information set!

CONDITIONAL MOMENTS FOR MOVING AVERAGE PROCESS

- The conditional variance of a moving average process of order q is given by,

$$var(Y_t|\Omega_{t-1}) = E[(Y_t - E[Y_t|\Omega_{t-1}])^2|\Omega_{t-1}]$$

- We've already computed the conditional mean, so we have that

$$Y_t - E[Y_t|\Omega_{t-1}] = \varepsilon_t$$

- Therefore,

$$var(Y_t|\Omega_{t-1}) = E[\varepsilon_t^2|\Omega_{t-1}] = \sigma^2$$

- From this we can see that while the conditional mean evolves according to the conditioning set, the conditional variance does not!

CONDITIONAL MOMENTS FOR AUTOREGRESSIVE PROCESS

- Now, let the information set at time t be given by

$$\Omega_{t-1} = Y_{t-1}, Y_{t-2}, \dots,$$

- The conditional mean of an AR(p) is then given by

$$E[Y_t|\Omega_{t-1}] = E\left[\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t \middle| \Omega_{t-1}\right]$$

- Again we treat all the conditioning variables as known, so that

$$E[Y_t|\Omega_{t-1}] = E[\varepsilon_t|\Omega_{t-1}] + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$$

- Since $\varepsilon_t \sim_{iid} (0, \sigma^2)$, it must be the case that $E[\varepsilon_t|\Omega_{t-1}] = 0$ so that,

$$E[Y_t|\Omega_{t-1}] = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$$

CONDITIONAL MOMENTS FOR AUTOREGRESSIVE PROCESS

- The process of deriving the conditional variance of an AR(p) process is the same as the MA(q) case,

$$var(Y_t|\Omega_{t-1}) = E[(Y_t - E[Y_t|\Omega_{t-1}])^2|\Omega_{t-1}]$$

- We've already computed the conditional mean, so we have that

$$Y_t - E[Y_t|\Omega_{t-1}] = \varepsilon_t$$

- Therefore,

$$var(Y_t|\Omega_{t-1}) = E[\varepsilon_t^2|\Omega_{t-1}] = \sigma^2$$

Later on in the course, we will discuss models that allow the conditional variance to evolve over time!

SUMMARY OF AR & MA PROCESSES

AR PROCESS	MA PROCESS
$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$	$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$
Covariance Stationary if the roots of: $$1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p = 0$$ Lie ***outside the unit circle*** or equivalently if the roots of: $$\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \cdots - \phi_{p-1}\lambda - \phi_p = 0$$ Lie ***inside the unit circle.***	Invertible if the roots of: $$1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q = 0$$ Lie ***outside the unit circle*** or equivalently if the roots of: $$\lambda^q + \theta_1 \lambda^{q-1} + \theta_2 \lambda^{q-2} + \cdots + \theta_{q-1}\lambda + \theta_q = 0$$ Lie ***inside the unit circle.***
If covariance stationary, an AR process can be written as an infinite order MA process.	If invertible, an MA process can be written as an infinite order AR process.

SUMMARY OF AR & MA PROCESSES

AR PROCESS	MA PROCESS
$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$	$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$
The autocovariance function decays gradually as $j \to \infty$ and follows a recursive structure where for $j = 1,2, \ldots$ $\gamma_j = \phi_1 \gamma_{j-1} + \phi_2\, \gamma_{j-2} + \cdots + \phi_p\, \gamma_{j-p}$ And the variance is given by: $\gamma_0 = \phi_1 \gamma_1 + \phi_2\, \gamma_2 + \cdots + \phi_p\, \gamma_p + \sigma^2$ So that the correlations also decay gradually have a recursive structure: $\rho_j = \phi_1 \rho_{j-1} + \phi_2 \rho_{j-2} + \cdots + \phi_p \rho_{j-p}$	The autocovariance function for $j = 1,2, \ldots, q$ is given by $\gamma_j = \sigma^2 \left(\theta_j + \theta_{j+1}\theta_1 + \theta_{j+2}\theta_2 + \cdots + \theta_q \theta_{q-j}\right)$ And the variance is given by: $\gamma_0 = \sigma^2 + \theta_1^2 \sigma^2 + \theta_2^2 \sigma^2 + \cdots + \theta_q^2 \sigma^2$ All autocovariances for $j > q$ will be zero. Therefore the autocorrelation structure cuts-off at lag q.

SUMMARY OF AR & MA PROCESSES

AR PROCESS	MA PROCESS
$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \varepsilon_t$	$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$
The partial autocorrelation function cuts off at lag p	The partial autocorrelation function decays gradually as $j \to \infty$
$E[Y_t] = 0$ $E[Y_t \vert \Omega_{t-1}] = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$ $var(Y_t \vert \Omega_{t-1}) = E[\varepsilon_t^2 \vert \Omega_{t-1}] = \sigma^2$	$E[Y_t] = 0$ $E[Y_t \vert \Omega_{t-1}] = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$ $var(Y_t \vert \Omega_{t-1}) = E[\varepsilon_t^2 \vert \Omega_{t-1}] = \sigma^2$

WOLD'S REPRESENTATION THEOREM

- Let Y_t be ***ANY*** zero mean, covariance-stationary process. Wold's representation theorem states that:

$$Y_t = \Psi(L)\varepsilon_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

$$\varepsilon_t \sim_{iid} (0, \sigma^2)$$

This model of infinite distributed lags is also called the ***general linear model.***

- Where

$$\psi_0 = 1$$

$$\sum_{i=0}^{\infty} \psi_i^2 < \infty$$

- That is to say, any zero mean, covariance stationary process can be expressed as an infinite order moving average. In other words, the correct model (or data generating process) for any covariance stationary series is ***some infinite distributed lag of white noise.***

THE GENERAL LINEAR MODEL

- The unconditional mean and variance of the general linear model are easily obtained as:

$$E[Y_t] = E\left[\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}\right] = \sum_{i=0}^{\infty} \psi_i E[\varepsilon_{t-i}] = 0$$

$$var(Y_t) = var\left(\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}\right) = \sum_{i=0}^{\infty} \psi_i^2 var(\varepsilon_{t-i}) = \sigma^2 \sum_{i=0}^{\infty} \psi_i^2$$

- While Wold's Representation Theorem says something very powerful about the nature of all covariance-stationary series, we will never be able to work with the General Linear Model directly since it comprises of an infinite number of parameters, $\psi_i, i = 0, \ldots, \infty$.

THE GENERAL LINEAR MODEL

- So we can’t work directly with an infinite distributed lag model. So what is the next best thing?
- Well, if we think about the properties of AR and MA processes that we just discussed, it should be clear that we can use a combination of AR and MA processes to approximate the general linear model!
- For instance, consider an MA(q),

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

- If we specify q to be large enough, we could obtain an arbitrarily close approximation.

THE GENERAL LINEAR MODEL

- Also recall that an autoregressive model can be written as an infinite order moving average. For instance, recall that for an AR(1), we have that:

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

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

- The AR(1) is equivalent to an MA of infinite order but only one parameter underlies it.
- Therefore, specifying higher order autoregressive models will allow for a closer approximation to the infinite sequence of parameters parameters, $\psi_i, i = 0, \dots, \infty$.

RATIONAL DISTRIBUTED LAGS AND THE ARMA MODEL

- If we were restricted to using only the MA or the AR model in isolation to approximate the general linear model, we would typically need to include a lot of lags in order to obtain a decent approximation! This would necessitate the estimation of a lot of parameters!
- We can cut down on the number of parameters needed by specifying an ARMA model which includes both autoregressive and moving average terms:

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

- The ARMA(1,1) model is expressed as:

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

$$\varepsilon_t \sim_{iid} (0, \sigma^2)$$

- Written in terms of lag operators,

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

- Thus the ARMA(1,1) is ***invertible*** if $|\theta| < 1$ and ***stationary*** if $|\phi| < 1$

AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

- The ***MA representation*** of the ARMA(1,1) process is:

$$Y_t = \frac{1 + \theta L}{1 - \phi L} \varepsilon_t$$

- The ***AR representation*** of the ARMA(1,1) process is:

$$\frac{1 - \phi L}{1 + \theta L} Y_t = \varepsilon_t$$

AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

- An ARMA process with p lags in the AR component and q lags in the MA component is referred to as an ARMA(p, q):

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

- Which again can be written in lag operator notation:

$$\Phi(L) Y_t = \Theta(L) \varepsilon_t$$

$$\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$

$$\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$$

- The ARMA(p,q) is invertible if all the roots of $\Theta(L)$ lie outside the unit circle.
- The ARMA(p,q) is stationary if all the roots of $\Phi(L)$ lie outside the unit circle.

AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

- If the invertibility and stationarity conditions are satisfied, we can write the following:

$$Y_t = \frac{\Theta(L)}{\Phi(L)}\varepsilon_t$$

This has $p + q$ parameters

- Recall that the general linear model is written as:

$$Y_t = \Psi(L)\varepsilon_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

This has ∞ parameters

- Therefore the ARMA model approximates an infinite distributed lag or current and past innovations using the ratio of two finite-order lag polynomials:

$$\Psi(L) \approx \frac{\Theta(L)}{\Phi(L)}$$

A GENERAL APPROACH TO ESTIMATING A TIME SERIES MODEL

1. Plot the time series data that you have collected.
2. Specify and estimate using OLS any deterministic components (i.e. trend and seasonality) that you believe to exist, along with the mean of the series. An example model of this sort could be:

$$Y_t = \alpha + \beta Trend_t + \sum_{i=1}^{k} \gamma_i D_{i,t} + \eta_t$$

3. Generate the residuals as

$$n_t = y_t - \hat{\alpha} - \hat{\beta}\, Trend_t - \sum_{i=1}^{k} \hat{\gamma}_i\, D_{i,t}$$

n_t can be thought of as the demeaned, detrended and seasonally adjusted series.

A GENERAL APPROACH OF ESTIMATING A TIME SERIES MODEL

4. Generate the sample ACF and PACF of the demeaned, detrended and seasonally adjusted series. These will give you a sense of the dependence structure of the time series.

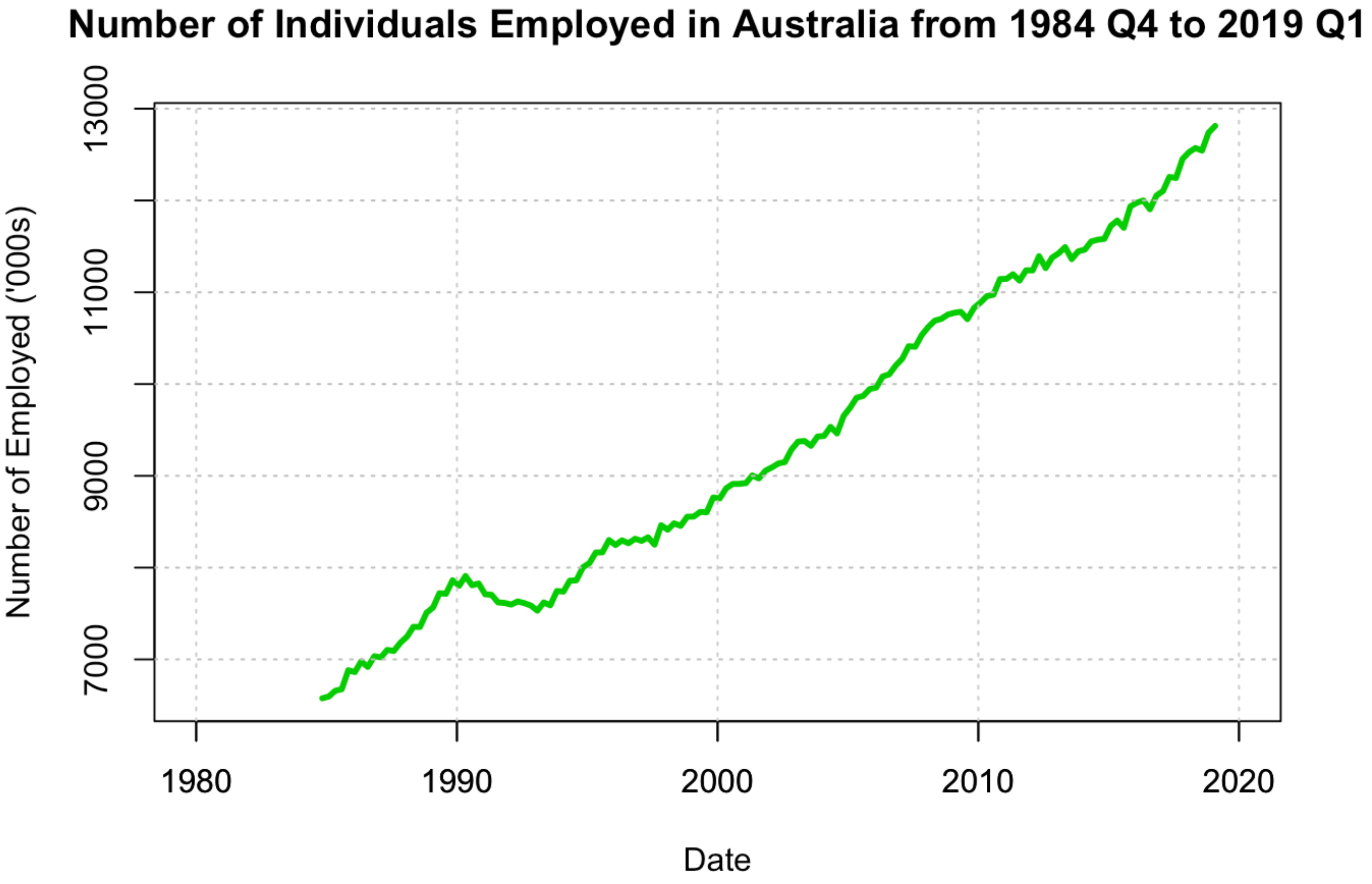
5. Estimate a range of ARMA(p, q) and choose your preferred model using the AIC and BIC.

6. Once you have chosen your preferred model, you can use it to compute point and interval forecasts!

EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

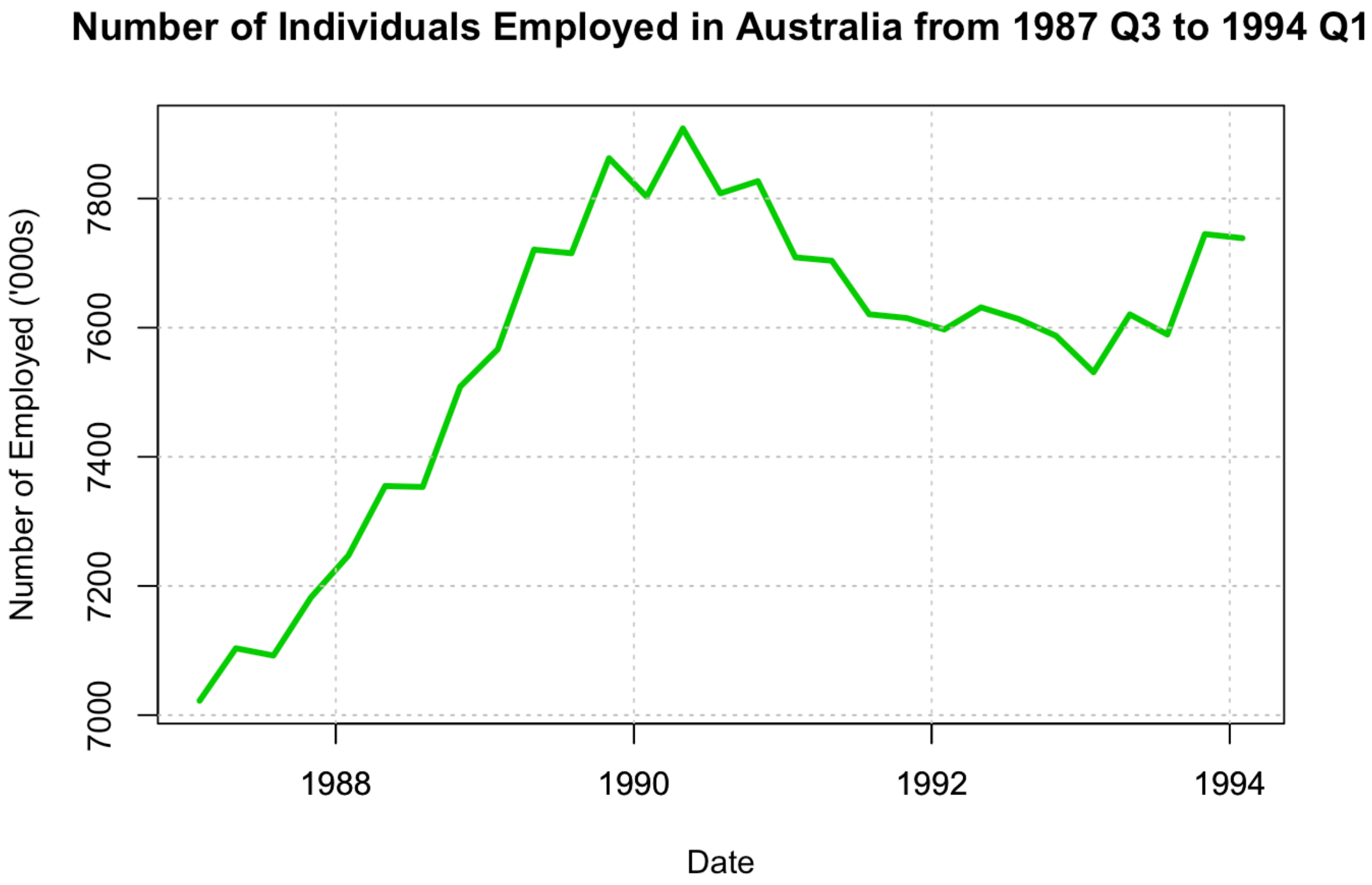
- To illustrate, let's consider a set of quarterly time series observations on the number of individuals who are employed in Australia from 1984:4 to 2019:1

- Looking at this plot, there is a clear upward trend. There appears to be a bit of curvature to the upward trend, so a good starting point would be a quadratic or exponential trend specification.



EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

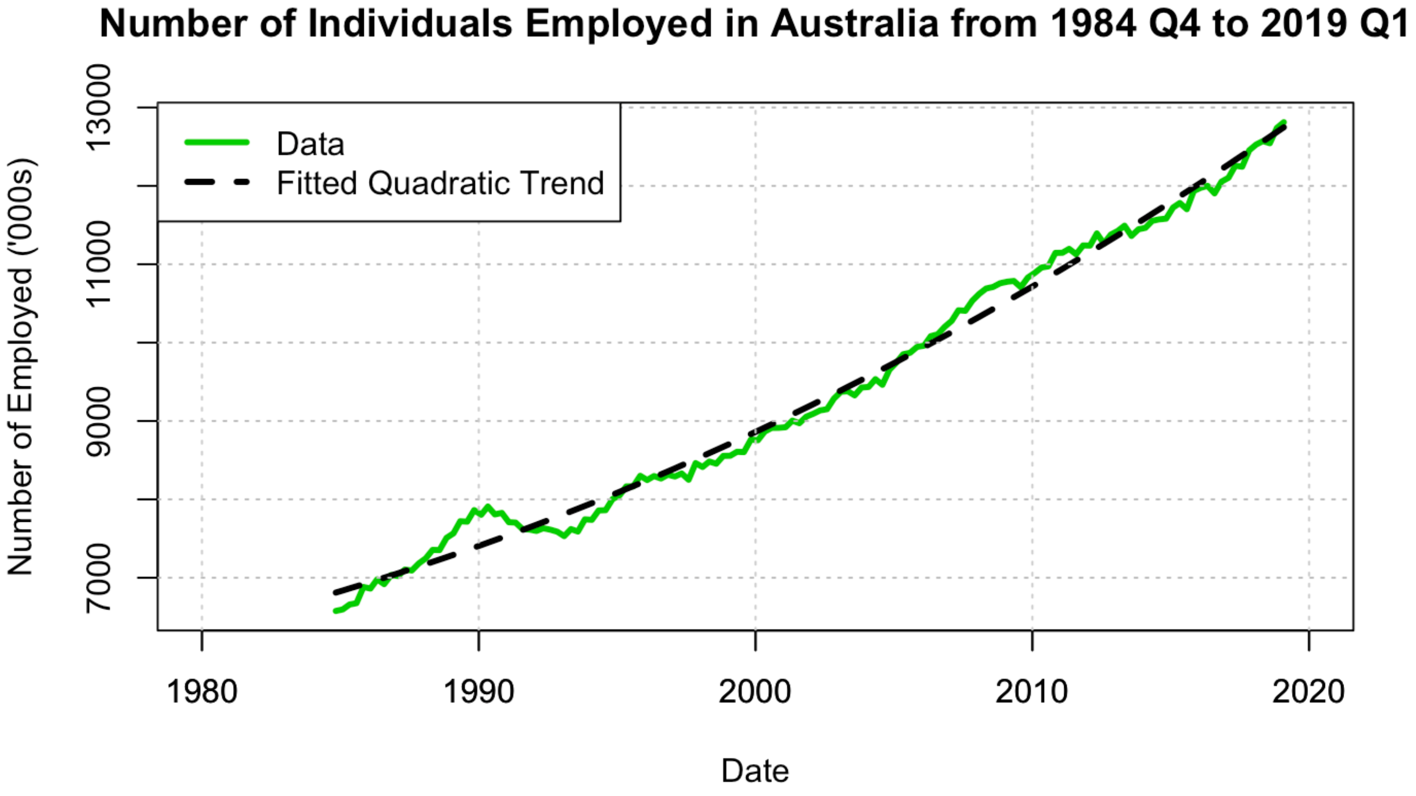
- If we zoom in a bit, we can see that there are some shorter-term fluctuations which could *potentially* be modeled as seasonal fluctuations:



EXAMPLE – NUMER OF EMPLOYED IN AUSTRALIA

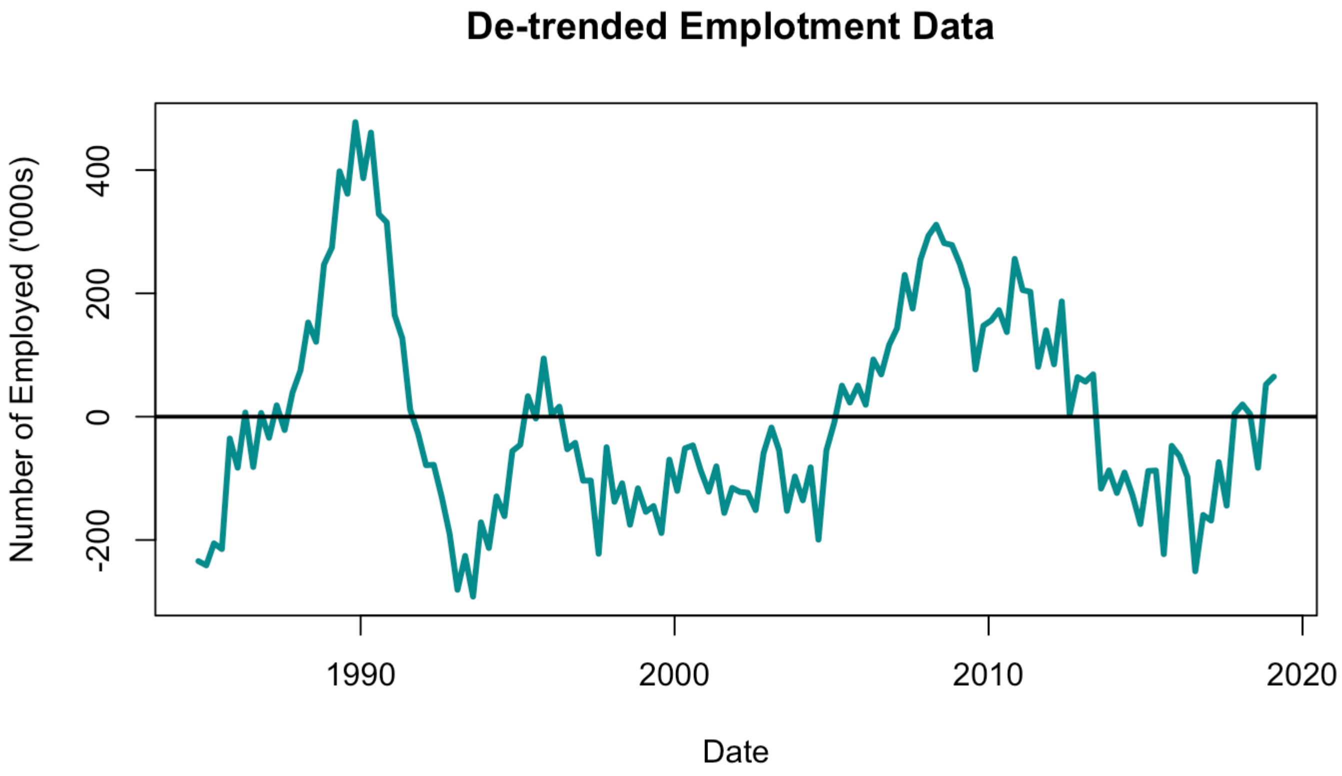
- Fitting a quadratic trend:

```
##
## Call:
## lm(formula = data$ausemp ~ time + timesq)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -291.78 -121.34  -44.00   90.81  477.59
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.784e+03  4.309e+01  157.45   <2e-16 ***
## time        2.597e+01  1.431e+00   18.15   <2e-16 ***
## timesq      1.249e-01  9.974e-03   12.53   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 166.3 on 135 degrees of freedom
## Multiple R-squared:  0.9911, Adjusted R-squared:  0.991
## F-statistic:  7515 on 2 and 135 DF,  p-value: < 2.2e-16
```



EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- Once we have estimated the quadratic trend model, we can obtain the de-trended data as the residuals of the trend model.
- These residuals definitely don't look like white noise! In fact, it looks awfully close to a random walk without drift! We will discuss how to deal with this in future lectures. For now, let's assume that this is a stationary process:
- To test this explicitly, we can compute Ljung-Box and Box-Pierce statistics. If the p-values are small, then we reject the null hypothesis that the data is generated from a white noise process



EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- We can then proceed to test whether there exists any significant deterministic seasonality using the same technique that we applied in Assignment 1.

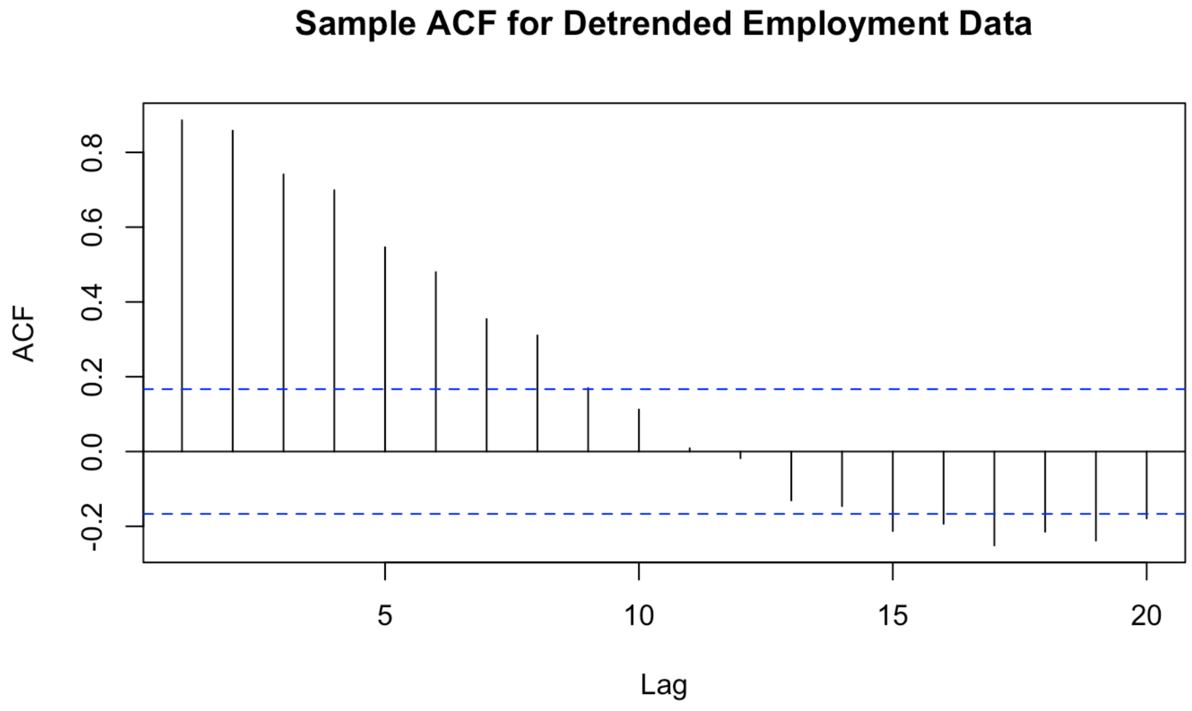
- By excluding one of seasonal dummy variables and including a constant in our regression, the seasonal average corresponding to the excluded seasonal dummy becomes the baseline average so that the regression coefficients multiplying the remaining dummy variables now have the interpretation of deviations from the baseline.

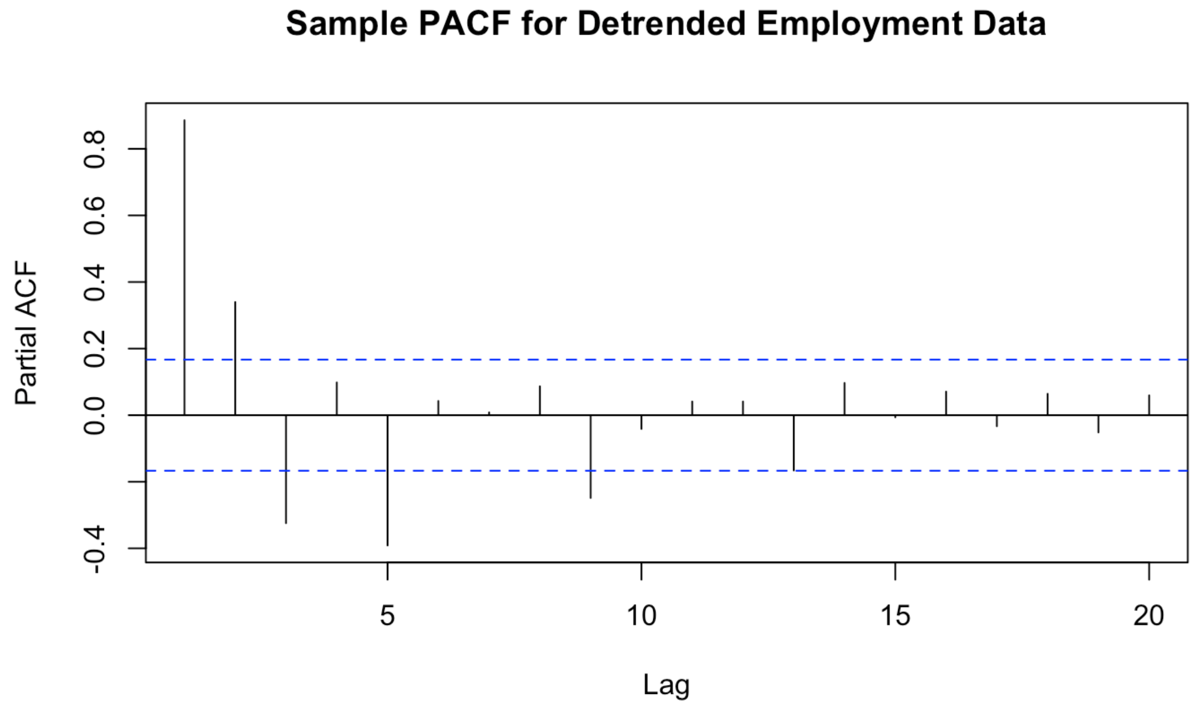
- Since none of the coefficients are statistically different from zero, we can exclude the seasonal dummy variables from the model.

```
Call:
lm(formula = data$detrended ~ Q2 + Q3 + Q4)

Residuals:
    Min      1Q  Median      3Q     Max 
-276.53 -116.75  -37.43   97.51  458.47 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   -4.223     27.789  -0.152    0.879
Q2            33.708     39.587   0.851    0.396
Q3           -40.592     39.587  -1.025    0.307
Q4            23.338     39.299   0.594    0.554

Residual standard error: 164.4 on 134 degrees of freedom
Multiple R-squared:  0.0298,	Adjusted R-squared:  0.008084 
F-statistic: 1.372 on 3 and 134 DF,  p-value: 0.254
```

EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- Looking at the sample ACF and PACF of the de-trended data, we can see that the partial autocorrelations cut off at the 5th (possibly 9th lag):





EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- We estimate the ARMA model for a range of AR and MA orders and compute the associated information criterion. Specifically, we will search from AR(1) – AR(9) and from MA(0) to MA(4). Here are the results for the AIC:

AIC					
	ma0	**ma1**	**ma2**	**ma3**	**ma4**
ar1	1585.133	1576.281	1563.519	1563.591	1541.354
ar2	1566.788	1516.895	1517.056	1521.448	1515.832
ar3	1549.459	1517.054	1509.753	1511.539	1511.407
ar4	1549.894	1518.931	1511.558	NA	1510.738
ar5	1515.490	1517.170	1514.735	NA	1493.999
ar6	1517.157	1504.838	NA	NA	NA
ar7	1519.062	1505.967	1512.764	1494.275	NA
ar8	1517.649	1504.010	1512.409	1539.534	NA
ar9	1507.797	1505.992	1509.860	NA	NA

EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- Here are the results for the BIC:

- Both the AIC and BIC are minimized when $p = 5, q = 4$

- When the AIC and BIC disagree, the BIC will typically choose the smaller model due to the heavier penalty it imposes on the number parameters in the model.

BIC					
	ma0	**ma1**	**ma2**	**ma3**	**ma4**
ar1	1590.987	1585.062	1575.228	1578.227	1558.918
ar2	1575.570	1528.604	1531.692	1539.011	1536.323
ar3	1561.168	1531.691	1527.317	1532.030	1534.825
ar4	1564.530	1536.495	1532.048	*NA*	1537.083
ar5	1533.054	1537.661	1538.153	*NA*	1523.272
ar6	1537.648	1528.256	*NA*	*NA*	*NA*
ar7	1542.481	1532.312	1542.036	1526.475	*NA*
ar8	1543.994	1533.283	1544.609	1574.661	*NA*
ar9	1537.069	1538.192	1544.987	*NA*	*NA*

EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- Here are the estimates of our chosen model:

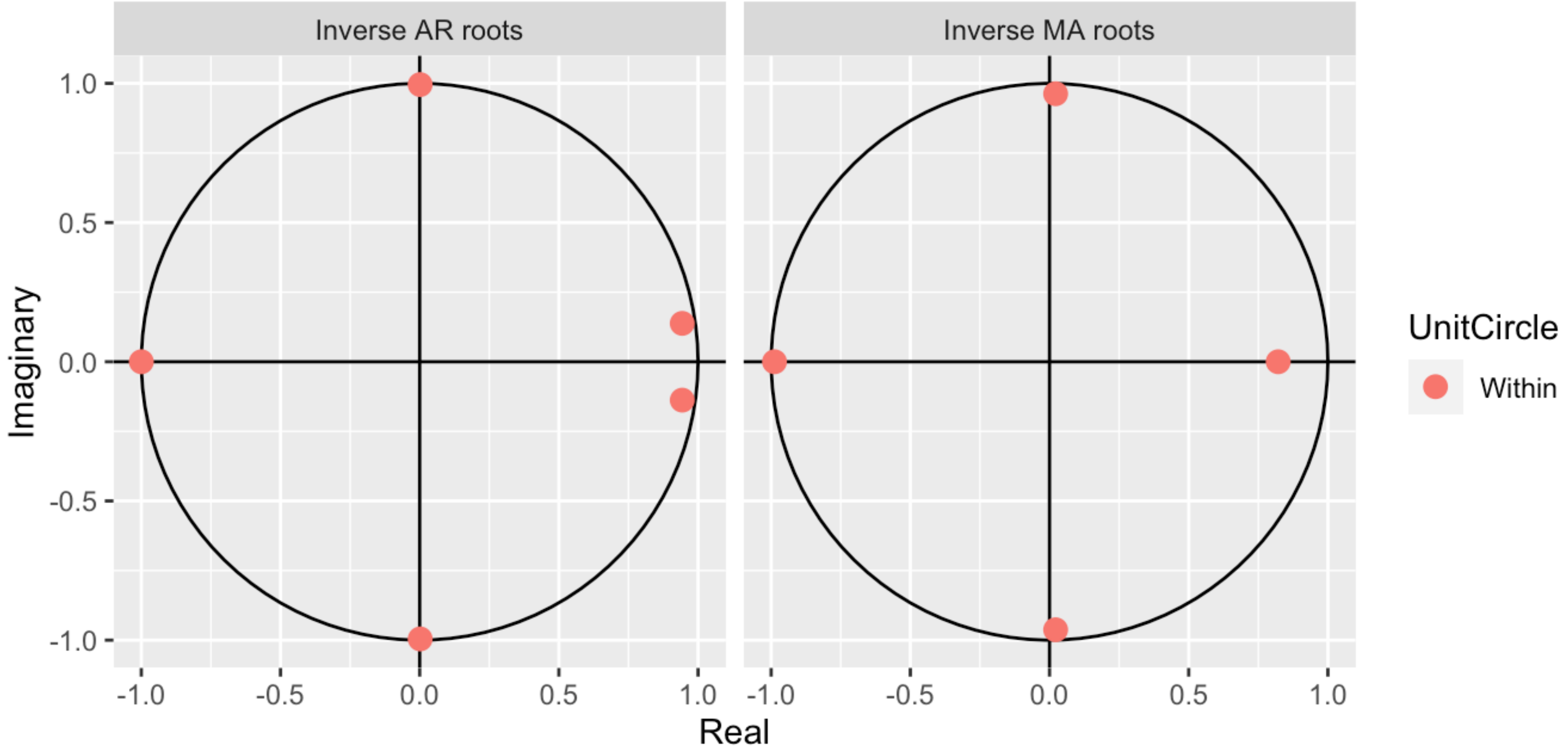
```
##
## Call:
## arima(x = data$detrended, order = c(5, 0, 4), include.mean = FALSE, method = "ML")
##
## Coefficients:
##          ar1      ar2      ar3     ar4      ar5     ma1     ma2     ma3
##       0.8898  -0.0166  -0.0334  0.9721  -0.9004  0.1225  0.1083  0.1900
## s.e.  0.0758   0.0219   0.0252  0.0246   0.0690  0.1467  0.1203  0.1528
##           ma4
##       -0.7525
## s.e.   0.1276
##
## sigma^2 estimated as 2388:  log likelihood = -737,  aic = 1494
##
## Training set error measures:
##                     ME     RMSE      MAE       MPE     MAPE      MASE
## Training set 0.6042093 48.87115 39.27138 -39.74189 104.4956 0.6378159
##                    ACF1
## Training set -0.0755377
```

The magnitude of the estimated AR coefficients reflect the high degree of dependence we observe in the data.

EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- We can also have a look at the estimated roots the lag polynomial:

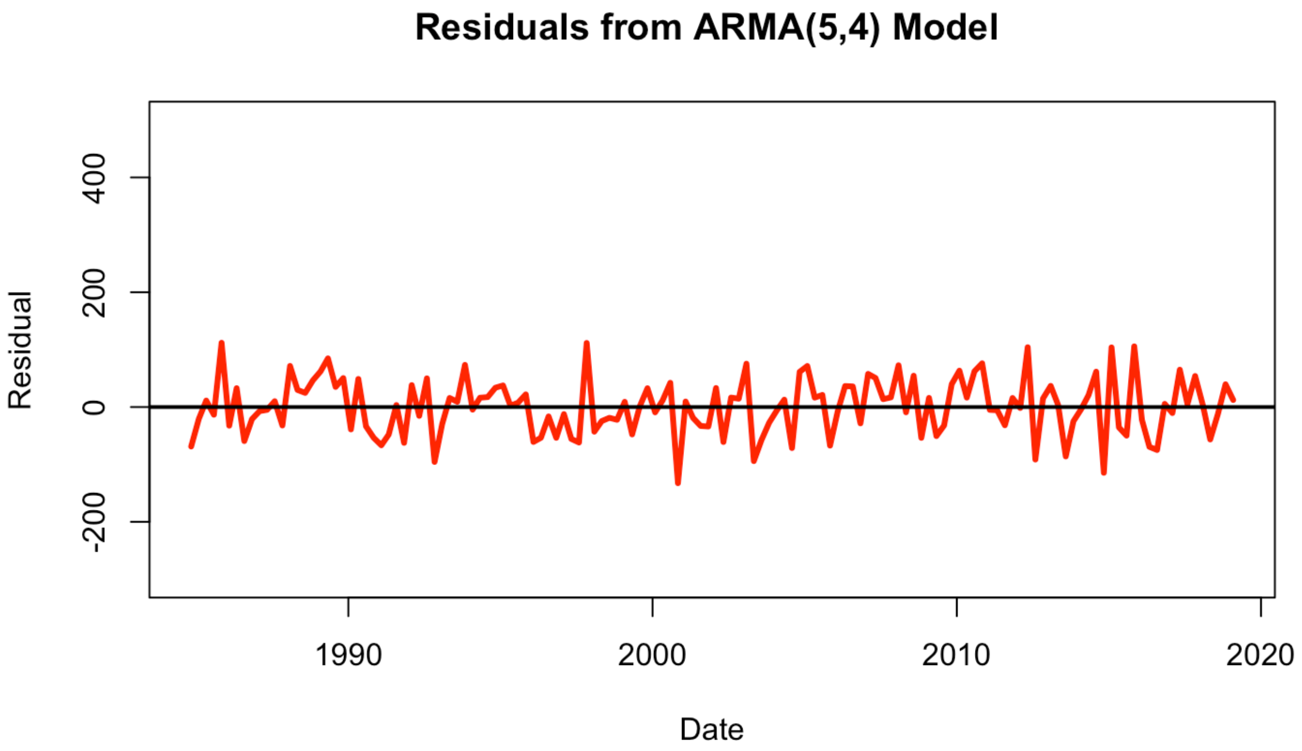
Note that some of these roots lie right on the boundary. Again, this could be due to the detrended series being non-stationary!



EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

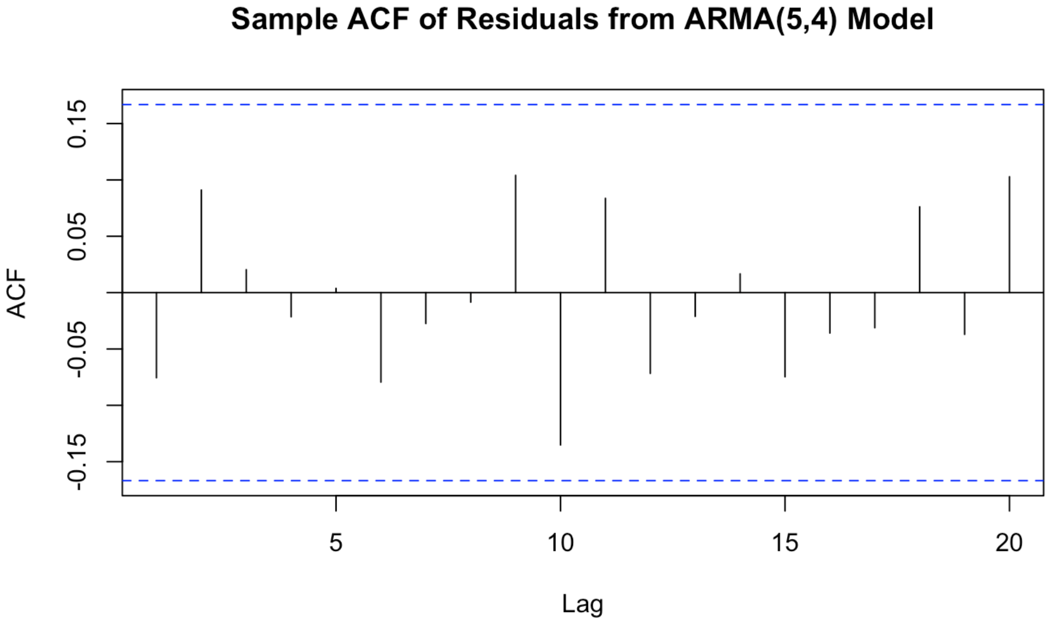
- Looking at the residual plot for the ARMA(5,4) we can see that this looks much closer to something produced by a white noise process:

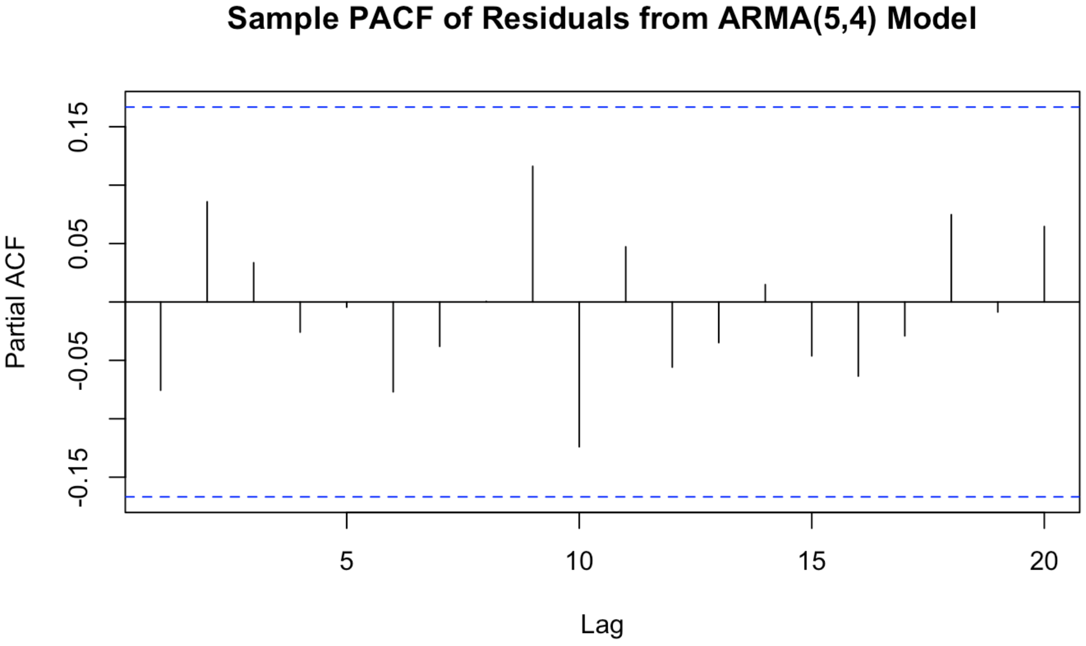
```
        Box-Pierce test

data:  bestmod$residuals
X-squared = 8.7255, df = 12, p-value = 0.7262


        Box-Ljung test

data:  bestmod$residuals
X-squared = 9.3834, df = 12, p-value = 0.6699
```



EXAMPLE – NUMBER OF EMPLOYED IN AUSTRALIA

- Looking at the sample ACF and PACF of the residuals tells the same story!





NEXT WEEK

- How do we actually estimate the parameters of an ARMA model?
- How do we compute appropriate point and interval forecasts?