Pulling Things Together

Ka-fu WONG

University of Hong Kong

March 22, 2019

Contents

1. A recap of the unobserved components model

According to the unobserved components model of a time series, the series y_t, is made up of the sum of three independent components

- a time trend component
- a seasonal component
- an irregular or cyclical component.

$$y_t = \text{time trend} + \text{seasonal} + \text{cyclical} = T_t + S_t + C_t$$

The forecast of the y_{T+h} is simply the sum of the forecast of the individual components.

$$\hat{y}_{T+h,T} = \hat{T}_{T+h,T} + \hat{S}_{T+h,T} + \hat{C}_{T+h,T}$$

2. Obtaining forecast in two steps

To reduce the unnecessary complication in our discussion, we assume that the time series consists of only two components: Trend and cyclical.

$$y_t = T_t + c_t$$

where T_t and c_t are respectively the trend component and cylical component of y_t.

We make additional two assumptions of these two components.

(1) The trend component to be well approximated by some polynomial trend. That is, for some positive integer s

$$T_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots \beta_s t^s$$

(2) The deviations from trend, c_t, (which we also refer to as they cyclical component of y_t) are assumed to be a zero-mean covariance stationary time series with an $ARMA(p, q)$ representation, i.e.,

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \ldots + \phi_p c_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

where $\epsilon_t \sim WN(0, \sigma^2)$ and the ϕ's satisfy the usual stationarity condition and θ's satisfy the usual invertibility condition.

The h-step ahead forecast of y_{T+h} given information available at time T and when s and p, as well as the parameters, are known is

$$y_{T+h,T} = T_{T+h} + c_{T+h,T} = \beta_0 + \beta_1(T+h) + \ldots + \beta_s(T+h)^s + c_{T+h,T}$$

where $c_{T+h,T}$ is the h-step ahead forecast of c implied by the $ARMA(p, q)$ model.

Obtaining forecast in two steps

In order to make these forecasts operational, we need to select s and p and then estimate the parameters, β_0, β_1, ..., β_s, ϕ_1, ϕ_2, ..., ϕ_p,θ_1, θ_2, ...,θ_q in two steps.

(1) Select s using some model selection criteria such as AIC and SIC. Given the selected s, estimate the β's, and produce the in-sample predicted value of $c_1, \ldots, c_T$:

$$\hat{c}_t = y_t - (\hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 + \ldots \hat{\beta}_s t^s)$$

(2) Select p and q, using the $\hat{c}_t$'s in place of c_t's, again using some model selection criteria such as AIC and SIC. Given the selected p and q, estimate the ϕ's and θ's by fitting the $\hat{c}_t$ to an $ARMA(p, q)$ model.

Then the point forecast $\hat{y}_{T+h,T}$ is simply the sum of the forecasts of the two components.

$$\hat{y}_{T+h,T} = \hat{T}_{T+h,T} + \hat{c}_{T+h,T}$$

The 95% forecast interval for y_{T+h} will be

$$\hat{y}_{T+h,T} + 1.96\sigma_h$$

where σ_h is the standard error of the h-step ahead forecast. σ_h can be a complicated object.

3. Obtaining forecast in one step

3.1. Linear trend plus MA(1). To reduce our burden of notations, we assume that $s = 1$, $p = 0$ and $q = 1$. That is, we assume a linear trend plus MA(1) in cyclical component

$$y_t = \beta_0 + \beta_1 t + c_t \tag{3.1}$$

$$c_t = \epsilon_t + \theta_1 \epsilon_{t-1} \tag{3.2}$$

It is easy to see that, we can estimate the coefficients in one step (either with nonlinear least squares or maximum likelihood).

$$y_t = \beta_0 + \beta_1 t + \epsilon_t + \theta_1 \epsilon_{t-1}$$

With NLS, we would start with a set of initial values of β_0, β_1 and θ_1, and the assumption that ϵ_0 equals to the unconditional expectation of zero, and recursively compute the implied residuals.

t	e_t
0	$e_0 = 0$
1	$e_1 = y_1 - b_0 - b_1 \times 1 - d_1 e_0$
2	$e_2 = y_2 - b_0 - b_1 \times 2 - d_1 e_1$
$\vdots$	$\vdots$
T	$e_T = y_T - b_0 - b_1 \times T - d_1 e_{T-1}$

The parameters are chosen to minimized the sum of squared residuals numerically.

$$(\hat{\beta}_0, \hat{\beta}_1, \hat{\theta}_1) = \arg \min_{(b_0, b_1, d_1)} \sum_{t=1}^{T} e_t^2$$

(1) Forecast y_{T+1}, ..., y_{T+H}

(a) $E(y_{T+1} \mid y_T, \ldots)$

$$\begin{aligned} E(y_{T+1} \mid y_T, \ldots) &= E\left[\beta_0 + \beta_1(T+1) + \epsilon_{T+1} + \theta_1 \epsilon_T \mid y_T, \ldots, \epsilon_T, \ldots\right] \\ &= \beta_0 + \beta_1(T+1) + \theta_1 \epsilon_T \end{aligned}$$

(b) $E(y_{T+2} \mid y_T, \ldots)$

$$\begin{aligned} E(y_{T+2} \mid y_T, \ldots) &= E\left[\beta_0 + \beta_1(T+2) + \epsilon_{T+2} + \theta_1 \epsilon_{T+1} \mid y_T, \ldots, \epsilon_T, \ldots\right] \\ &= \beta_0 + \beta_1(T+2) \end{aligned}$$

h	$\hat{y}_{T+h,T}$
1	$\beta_0 + \beta_1(T+1) + \theta_1\epsilon_T$
2	$\beta_0 + \beta_1(T+2)$
$\vdots$	$\vdots$
H	$\beta_0 + \beta_1(T+H)$

Obtaining forecast in one step

The extension to the more general trend and more general MA model should be straightforward.

$$
\begin{aligned}
y_t &= \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_s t^s + c_t \\
c_t &= \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}
\end{aligned}
$$

The model can be rewritten into

$$
y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_s t^s + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}
$$

EXERCISE:

(1) Describe how to estimate the coefficients in the model in one step?

(2) What are the h-step-ahead forecast given the coefficients β_0, β_1, ..., β_s, θ_1, θ_2, ..., θ_q?

3.2. **Linear trend plus AR(1).** To reduce our burden of notations, we assume that $s = 1$, $p = 1$ and $q = 0$. That is, we assume a linear trend plus AR(1) in cyclical component

$$y_t = \beta_0 + \beta_1 t + c_t \tag{3.3}$$

$$c_t = \phi c_{t-1} + \epsilon_t \tag{3.4}$$

Note that equation (3.3) can be written as

$$c_t = y_t - (\beta_0 + \beta_1 t) \tag{3.5}$$

Substitute (3.5) into equation (3.4), we have

$$\begin{aligned}
c_t &= \phi c_{t-1} + \epsilon_t \\
y_t - (\beta_0 + \beta_1 t) &= \phi[y_{t-1} - (\beta_0 + \beta_1 (t-1))] + \epsilon_t \\
y_t &= \phi y_{t-1} + (\beta_0 - \phi\beta_0) + [\beta_1 t - \phi\beta_1 (t-1)] + \epsilon_t \\
y_t &= \phi y_{t-1} + (1-\phi)\beta_0 + (1-\phi)\beta_1 t + \phi\beta_1 + \epsilon_t \\
y_t &= \phi y_{t-1} + [(1-\phi)\beta_0 + \phi\beta_1] + (1-\phi)\beta_1 t + \epsilon_t \\
y_t &= \phi y_{t-1} + \alpha_0 + \alpha_1 t + \epsilon_t \qquad (3.6)
\end{aligned}$$

where $\alpha_0 = [(1-\phi)\beta_0 + \phi\beta_1]$, and $\alpha_1 = (1-\phi)\beta_1$.

Alternatively, we can use lag operators to obtain a similar conclusion. We demonstrate this approach here because it is cleaner and easier to apply when we have a more general ARMA cyclical component. Equation (3.4) can be rewritten as

$$(1-\phi L)c_t = \epsilon_t$$

Apply $(1-\phi L)$ through equation (3.3), we have

$$\begin{aligned}
(1-\phi L)y_t &= (1-\phi L)\beta_0 + (1-\phi L)\beta_1 t + (1-\phi L)c_t \\
y_t - \phi y_{t-1} &= (1-\phi)\beta_0 + \beta_1 t - \phi\beta_1(t-1) + \epsilon_t \\
y_t &= \phi y_{t-1} + [(1-\phi)\beta_0 + \phi\beta_1] + (1-\phi)\beta_1 t + \epsilon_t
\end{aligned}$$

In short, both approaches will reach identical conclusion:
a model with linear trend plus AR(1) in cyclical component can be rewritten into a model with linear trend and AR(1) in y_t.

With this understanding, we can do the estimation in one step instead of two.

(1) Estimate α_0, α_1 and ϕ in equation (3.6).

(2) Forecast y_{T+1}, ..., y_{T+H}

(a) $E(y_{T+1} \mid y_T, ...)$

$$\begin{aligned} E(y_{T+1} \mid y_T, ...) &= E\left[\phi y_T + \alpha_0 + \alpha_1(T+1) + \epsilon_{T+1} \mid y_T, ...\right] \\ &= \phi y_T + \alpha_0 + \alpha_1(T+1) \end{aligned}$$

(b) $E(y_{T+2} \mid y_T, ...)$

$$\begin{aligned} E(y_{T+2} \mid y_T, ...) &= E\left[\phi y_{T+1} + \alpha_0 + \alpha_1(T+2) + \epsilon_{T+2} \mid y_T, ...\right] \\ &= \phi E\left[y_{T+1} \mid y_T, ...\right] + \alpha_0 + \alpha_1(T+2) \end{aligned}$$

h	$\hat{y}_{T+h,T}$
1	$\hat{\phi} y_T + \hat{\alpha}_0 + \hat{\alpha}_1(T+1)$
2	$\hat{\phi} \hat{y}_{T+1,T} + \hat{\alpha}_0 + \hat{\alpha}_1(T+2)$
$\vdots$	$\vdots$
H	$\hat{\phi} \hat{y}_{T+H-1,T} + \hat{\alpha}_0 + \hat{\alpha}_1(T+H)$

Now, let's consider the case with more general trend and more general AR model.

$$\begin{aligned} y_t &= \beta_0 + \beta_1 t + \beta_2 t^2 + \ldots + \beta_s t^s + c_t \\ c_t &= \phi_1 c_{t-1} + \phi_2 c_{t-1} + \ldots + \phi_p c_{t-p} + \epsilon_t \end{aligned}$$

The model can be rewritten into

$$y_t = \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \alpha_0 + \alpha_1 t + \ldots + \alpha_s t^s + \epsilon_t$$

where the α's are functions of the β's and the ϕ's.

Thus, given p and s, we would

(1) Fit this model with OLS to estimate the α's and ϕ's.

(2) Generate $\hat{y}_{T+h,T}$ recursively according to

$$\hat{y}_{T+h,T} = \phi_1 \hat{y}_{T+h-1,T} + \dots + \phi_p \hat{y}_{T+h-p,T} + \alpha_0 + \alpha_1(T+h) + \dots + \alpha_s(T+h)^s$$

where $\hat{y}_{T+h-s,T} = y_{T+h-s}$ if $T+h-s \leq T$.

In practice, s and p have to be chosen based on data. To choose the optimal model, we would experiment with different s and p and choose the model with the smallest AIC and SIC.

3.3. **Seasonality plus MA(1).** Can we obtain a single equation for estimation when we have seasonality and cyclical components? To simplify the burden of notations, suppose we have no trend component (i.e., $s = 0$), quarterly seasonality and MA(1) (i.e., $p = 0$).

(3.7) $$y_t = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} + c_t$$

(3.8) $$c_t = \epsilon_t + \theta_1 \epsilon_{t-1}$$

It is easy to see that, we can estimate the coefficients in one step (either with nonlinear least squares or maximum likelihood).

$$y_t = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} + \epsilon_t + \theta_1 \epsilon_{t-1}$$

EXERCISE:

(1) Describe how to estimate the coefficients in the model in one step?
(2) What are the h-step-ahead forecast given the coefficients α_1, ..., α_4, θ_1?

3.4. **Seasonality plus AR(1).**

$$y_t = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} + c_t \tag{3.9}$$

$$(1-\phi L)c_t = \epsilon_t \tag{3.10}$$

We can write

$$\begin{aligned}(1-\phi L)y_t &= (1-\phi L)\left[\alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t}\right] + \epsilon_t \\ &= \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t} + \epsilon_t\end{aligned}$$

and hence,

$$y_t = \phi y_{t-1} + \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t} + \epsilon_t$$

To understand the equivalence of $(1-\phi L)\left[\alpha_1 D_{1t}+\alpha_2 D_{2t}+\alpha_3 D_{3t}+\alpha_4 D_{4t}\right]$ and $\gamma_1 D_{1t}+\gamma_2 D_{2t}+\gamma_3 D_{3t}+\gamma_4 D_{4t}$, let's look at the data matrix of the seasonal dummies.

t	Q	D_{1t}	D_{2t}	D_{3t}	D_{4t}
1	1	1	0	0	0
2	2	0	1	0	0
3	3	0	0	1	0
4	4	0	0	0	1
5	1	1	0	0	0
6	2	0	1	0	0
7	3	0	0	1	0
8	4	0	0	0	1
9	1	1	0	0	0
...	...	...	...	...	...

t	Q	D_{1t}	D_{2t}	D_{3t}	D_{4t}	D_{1t-1}	D_{2t-1}	D_{3t-1}	D_{4t-1}
1	1	1	0	0	0	0	0	0	1
2	2	0	1	0	0	1	0	0	0
3	3	0	0	1	0	0	1	0	0
4	4	0	0	0	1	0	0	1	0
5	1	1	0	0	0	0	0	0	1
6	2	0	1	0	0	1	0	0	0
7	3	0	0	1	0	0	1	0	0
8	4	0	0	0	1	0	0	1	0
9	1	1	0	0	0	0	0	0	1
…	…	…	…	…	…	…	…	…	…

We conclude $D_{1t-1} = D_{2t}$, $D_{2t-1} = D_{3t}$, $D_{3t-1} = D_{4t}$, and $D_{4t-1} = D_{1t}$.

With $D_{1t-1} = D_{2t}$, $D_{2t-1} = D_{3t}$, $D_{3t-1} = D_{4t}$, $D_{4t-1} = D_{1t}$, we can write

$$
\begin{aligned}
& (1-\phi L)\left[\alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t}\right] \\
= \; & \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} \\
& -\phi\alpha_1 D_{1t-1} - \phi\alpha_2 D_{2t-1} - \phi\alpha_3 D_{3t-1} - \phi\alpha_4 D_{4t-1} \\
= \; & \alpha_1 D_{1t} + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \alpha_4 D_{4t} \\
& -\phi\alpha_1 D_{2t} - \phi\alpha_2 D_{3t} - \phi\alpha_3 D_{4t} - \phi\alpha_4 D_{1t} \\
= \; & (\alpha_1 - \phi\alpha_4) D_{1t} + (\alpha_2 - \phi\alpha_1) D_{2t} + (\alpha_3 - \phi\alpha_2) D_{3t} + (\alpha_4 - \phi\alpha_3) D_{4t} \\
= \; & \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t}
\end{aligned}
$$

4. Full model

In a more general model, we will have seasonal component, and the cyclical component may be better modelled as ARMA process.

$$
\begin{aligned}
y_t &= T_t(\beta) + \sum_{i=1}^{s} \gamma_i D_{it} + c_t \\
\Phi(L) c_t &= \Theta(L) \epsilon_t
\end{aligned}
$$

where $T_t(\beta)$ is a trend component, D_{it} are seasonal dummies, $\Phi(L)$ and $\Theta(L)$ are polynomials of lag operators.

The notations of the above model may look complicated. It is easy to check, however, that it reduces to the special case of linear trend with AR(1) cyclical component we have discussed earlier when

$$\begin{aligned} T_t(\beta) &= \beta_0 + \beta_1 t \\ \gamma_i &= 0 \\ \Phi(L) &= 1 - \phi L \\ \Theta(L) &= 1 \end{aligned}$$

Using lag operators, it is easy to see that the general trend and general ARMA model can be written in one equation.

$$\begin{aligned} \Phi(L)y_t &= \Phi(L)T_t(\beta) + \sum_{i=1}^{s} \gamma_i \Phi(L) D_{it} + \Phi(L) c_t \\ \Phi(L)y_t &= \Phi(L)T_t(\beta) + \sum_{i=1}^{s} \gamma_i \Phi(L) D_{it} + \Theta(L) \epsilon_t \end{aligned}$$

EXERCISE:

(1) Consider a model with linear trend, quarterly seasonality and ARMA(1,1). Describe how to estimate the coefficients in the model in one step?

(2) What are the h-step-ahead forecast given the coefficients?

5. Example: Forecasting Electricity

5.1. Data.

- Data: *monthly* electricity consumption data *from* 1970:01 *to* 2013:11 (*a total of* 527 observations) from Hong Kong Census and Statistics Department.
(https://www.censtatd.gov.hk/hkstat/sub/sp90.jsp?tableID=127&ID=0&productType=8)
- Estimation and model selection: data from 1970:01 to 2011:12 (a total of 504 observations).
- Out-of-sample comparison: 2012:01 to 2013:11 (a total of 23 observations)

5.2. **Estimation and forecast strategy.** We will use the conveniently available ARMA function of R instead of writing our own script for the purpose. However, ARMA function *can only deal with covariance stationary series*.

- Allowed: Series with constant mean or series with seasonality.
- Not allowed: Series with trend or series that is non-stationary.

So *our strategy is* to model trend and seasonality first.

$$y_t = T_t(\beta) + \sum_{i=1}^{s} \gamma_i D_{it} + c_t$$

From the best (selected) trend-and-seasonality model, we obtain the residuals

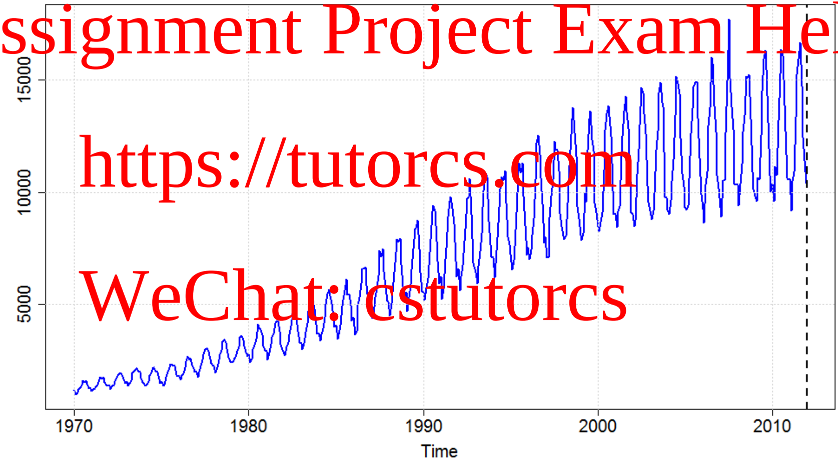
$$\hat{c}_t = y_t - T_t(\hat{\beta}) - \sum_{i=1}^{s} \hat{\gamma}_i D_{it}$$

and estimate various ARMA models of $\hat{c}_t$.

$$\begin{aligned} \Phi(L)\hat{c}_t &= \Theta(L)\epsilon_t \\ & \epsilon_t \sim NWN(0, \sigma^2) \end{aligned}$$

Throughout, we use Box test (for white noise) and the AIC and SIC as guidance for our decision of best ARMA models.

FIGURE 5.1. Time series plot of our estimation sample



From Figure 5.1, we can easily see an increasing volatility/variance with levels. Thus, we *decide to try* a log-transformation of the data.

FIGURE 5.2. Time series plot of our estimation sample

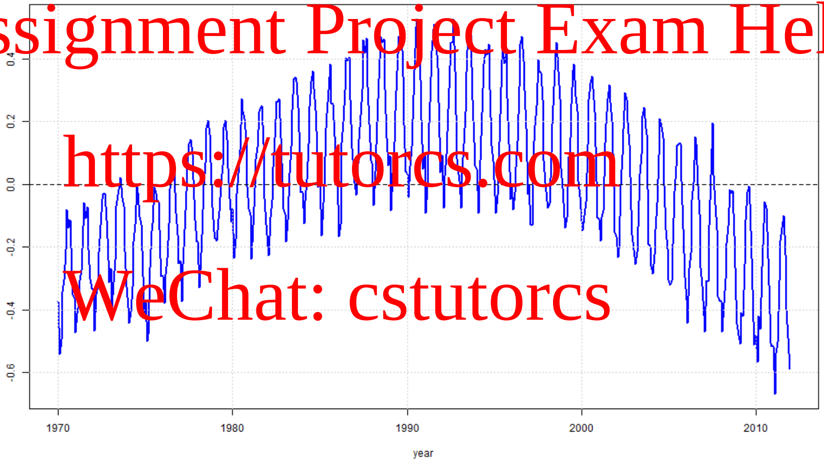
From Figure 5.2, we can see that the log-transformed data have a much more stable variance across levels. Therefore, we *decide to use* log-transformed data for our most of our analysis.

The *main reason* for our use of log-transformed data is efficiency in our estimation and hence forecast. Experience tells us that working with data of similar variance will likely yield more efficient estimators of the coefficients (via some form of Gauss-Markov Theorem). Such efficiency will often translate into forecast with less uncertainty.

Nevertheless, we can easily see that the data appears to have strong seasonality and a upward trend. The trend is *likely quadratic*.

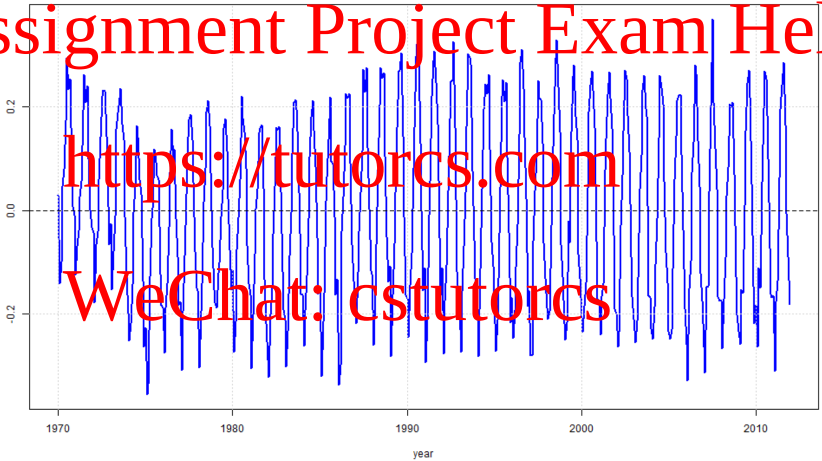
We will use residual plots, as well as AIC and SIC, to *guide our choice of best trend models*.

FIGURE 5.3. Residuals from model with pure Linear Trend



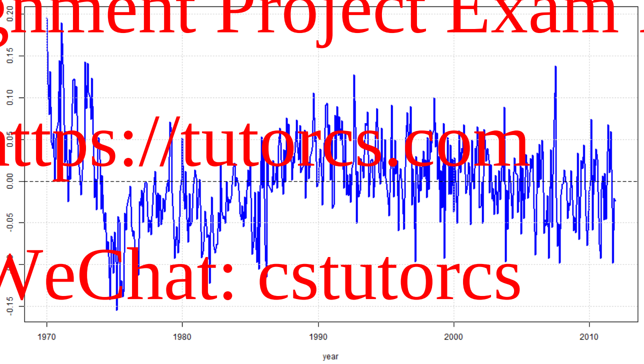
The residual plot (Figure 5.3) *shows a hump shape trend / pattern*. Residuals are generally below zero in early years, above zero mid way and below zero again near the end of the sample. This clearly *suggests* that the linear trend is inadequate and a quadratic trend is more appropriate.

FIGURE 5.4. Residuals from model with pure Quadratic Trend



The residual plot (Figure 5.4) shows no specific trend *but the seasonality pattern stands out*. This clearly *suggests* that we should consider a quadratic trend model with seasonlity. *Since we have relatively large sample*, we can afford to model monthly seasonality, the most one can do with monthly data.

FIGURE 5.5. Residuals from model with Quadratic Trend and Monthly Seasonality



The residual plot (Figure 5.5) shows neither specific trend nor seasonality. However, we do see *some mild persistence in the residuals*, more substantial in the early part of the sample than the later part of the sample.

Since our purpose is forecast out of sample, the almost absence of pattern in the latter part of the sample suggest that the additional modelling of ARMA of the residuals *will unlikely yield better forecast* than without such additional work.

Table 5.1 shows the AIC and SIC for the models we have considered so far. Both criteria points to quadratic trend with seasonality, *the same conclusion* we have made based on the residual plots.

TABLE 5.1. Comparison of Selection Critieria among Models with Trend and Seasonality

	AIC	SIC
Linear Trend Only	54.258	66.926
Quadratic Trend Only	-307.608	-290.717
Quadratic Trend with Seasonality	-1493.361	-1430.023

We proceed to model the cyclical components (i.e., ARMA models) *to explore the possibility of improving our forecasts*. That is, we will model $\hat{c}_t$

$$\hat{c}_t = y_t - \hat{\beta}_1 TIME_t - \hat{\beta}_2 TIME_t^2 - \sum_{i=1}^{12} \hat{\gamma}_i D_{it}$$

where D_{it} takes the value of one if t falls into month i and zero otherwise.

We use Box Test of white noise as a *major criteria*. Then, among those models which we cannot reject the null of white noise at conventional significance levels, we look for models that have smallest AIC and SIC.

TABLE 5.2. p-values of Box Test for Different Combination of $ARMA(p,q)$

	q = 0	q = 1	q = 2	q = 3	q = 4	q = 5	q = 6	q = 7	q = 8	q = 9	q = 10
p = 0	0	0	0	0	0	0	0	0	0	0	0
p = 1	0	0	0	0	0	0	0	0	0	0	0
p = 2	0	0	0	0	0	0	0	0	0	0	0
p = 3	0	0	0	0	0	0	0	0	0	0	0
p = 4	0	0	0	0	0	0	0	0	0.85	0	0.01
p = 5	0	0	0	0	0	0	0	0.45	0.97	0.99	0.99
p = 6	0	0	0	0	0.29	0.41	0.62	0.62	0.67	0.96	1
p = 7	0	0	0	0	0	0.14	0	0.91	0.99	0.99	1
p = 8	0	0.03	*0.68*	0.7	0.88	0.88	0.95	0.93	0.95	1	1
p = 9	0	0.08	0.7	0.89	0.79	0.99	0.99	0.99	0.99	0.77	1
p = 10	0.01	0.1	0.25	0.89	0.61	0.97	1	1	1	1	—

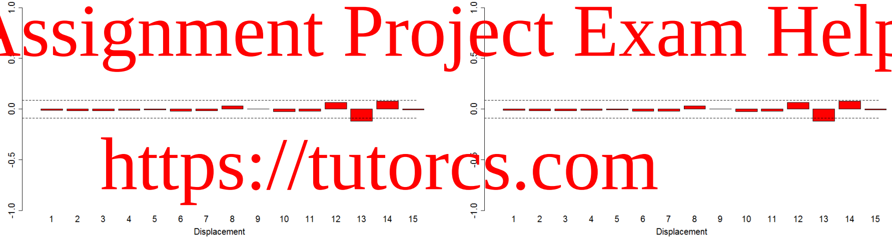
TABLE 5.3. AIC for Different Combination of $ARMA(p, q)$

	q = 0	q = 1	q = 2	q = 3	q = 4	q = 5	q = 6	q = 7	q = 8	q = 9	q = 10
p = 0	-1519	-1636	-1694	-1722	-1723	-1733	-1746	-1744	-1743	-1741	-1740
p = 1	-1717	-1753	-1756	-1768	-1770	-1769	-1767	-1765	-1777	-1779	-1783
p = 2	-1749	-1763	-1765	-1768	-1769	-1774	-1773	-1770	-1783	-1782	-1784
p = 3	-1750	-1748	-1764	-1751	-1768	-1788	-1787	-1781	-1782	-1783	-1782
p = 4	-1748	-1746	-1764	-1775	-1773	-1787	-1788	-1790	-1842	-1792	-1789
p = 5	-1750	-1772	-1752	-1773	-1782	-1784	-1785	-1848	-1854	-1856	-1854
p = 6	-1750	-1753	-1773	-1787	-1811	-1849	-1851	-1848	-1847	-1839	-1857
p = 7	-1748	-1751	-1771	-1787	-1817	-1810	-1777	-1845	-1853	-1853	-1862
p = 8	-1758	-1791	-1842	-1840	-1848	-1841	-1840	-1851	-1849	-1853	-1862
p = 9	-1774	-1796	-1840	-1842	-1852	-1857	-1855	-1853	-1851	-1853	-1856
p = 10	-1786	-1798	-1799	-1840	-1846	-1844	-1857	-1860	-1859	-1858	—

TABLE 5.4. SIC for Different Combination of $ARMA(p,q)$

	q = 0	q = 1	q = 2	q = 3	q = 4	q = 5	q = 6	q = 7	q = 8	q = 9	q = 10
p = 0	-1511	-1623	-1677	-1700	-1698	-1704	-1712	-1706	-1701	-1695	-1689
p = 1	-1704	-1736	-1735	-1743	-1740	-1735	-1729	-1723	-1730	-1728	-1728
p = 2	-1732	-1742	-1740	-1739	-1735	-1736	-1730	-1724	-1732	-1727	-1725
p = 3	-1729	-1722	-1735	-1717	-1730	-1746	-1740	-1730	-1727	-1724	-1719
p = 4	-1722	-1716	-1730	-1737	-1731	-1740	-1737	-1735	-1783	-1729	-1721
p = 5	-1721	-1738	-1714	-1731	-1736	-1734	-1730	-1789	-1790	-1789	-1783
p = 6	-1716	-1715	-1731	-1740	-1760	-1794	-1791	-1785	-1779	-1767	-1781
p = 7	-1710	-1709	-1724	-1736	-1762	-1751	-1714	-1777	-1782	-1777	-1781
p = 8	-1716	-1745	-1791	-1785	-1788	-1777	-1772	-1779	-1773	-1772	-1778
p = 9	-1727	-1746	-1785	-1783	-1789	-1789	-1783	-1777	-1771	-1769	-1768
p = 10	-1735	-1743	-1740	-1777	-1778	-1772	-1781	-1780	-1775	-1769	—

FIGURE 5.6. ACF and PACF of the Residuals from ARMA(8,2)



(A) ACF (B) PACF

The ACF and PACF verify that the residuals of the ARMA models look like white noises. Thus, the ARMA(8,2) appears adequate.

Given the Box test (Table 5.2) , AIC (Table 5.3) and SIC (Table 5.4), as well as the plots of ACF and PACF (Figure 5.6), we *decide* that ARMA(8,2) is appropriate.

$$\hat{c}_t = \alpha_0 + \alpha_1 \hat{c}_{t-1} + \ldots + \alpha_8 \hat{c}_{t-8} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2}$$

Note that there are other ARMA combinations that yield a smaller AIC and SIC but they tend to have many more parameters, and *sometimes the estimation appears unstable*.

Our forecast is then the sum of the forecast of trend and seasonality from an earlier model, and the forecast from the ARMA(8,2) model.

$$\hat{y}_{T+h,T} = \hat{\beta}_1 TIME_{T+h} + \hat{\beta}_2 TIME^2_{T+h} + \sum_{i=1}^{12} \hat{\gamma}_i D_{iT+h} + \hat{\hat{\varepsilon}}_{T+h,T}$$

The standard error from the ARMA(8,2) model is used to construct the 95% forecast interval.

$$\hat{y}_{T+h,T} \pm 1.96 \times se(\hat{\varepsilon})$$

Note that the *parameter uncertanty is ignored* in the construction of confidence interval because there apepars no simple way to make such corrections.

FIGURE 5.7. Forecast from the Model with Trend and Seasonality only

FIGURE 5.8. Forecast from the Full Model (with Trend and Seasonality and ARMA(8,2))

FIGURE 5.9. A Comparison of the Two Forecast Intervals (with and without ARMA)

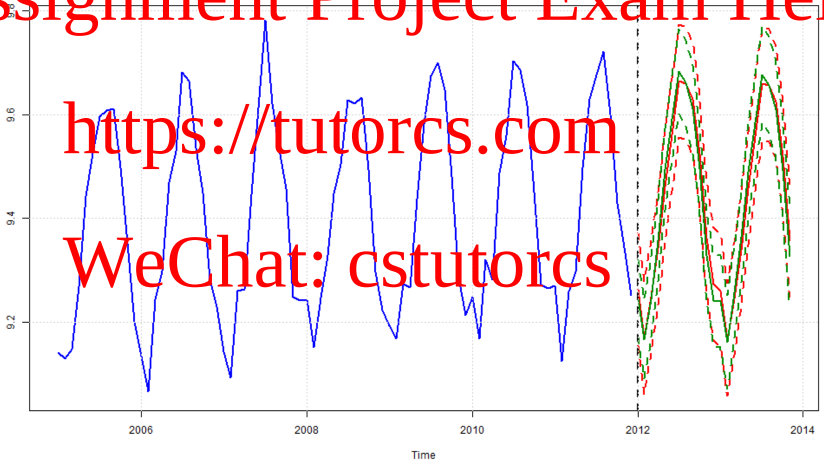


FIGURE 5.10. Forecasts from the Model with Trend and Seasonality only, with Realized Values

FIGURE 5.11. Forecast from the Full Model, with Realized Values

FIGURE 5.12. A Comparison of Point Forecasts from the Two Models (with and without ARMA), with realized values

FIGURE 5.13. A Comparison of the two Forecast Intervals (with and without ARMA), with realized values

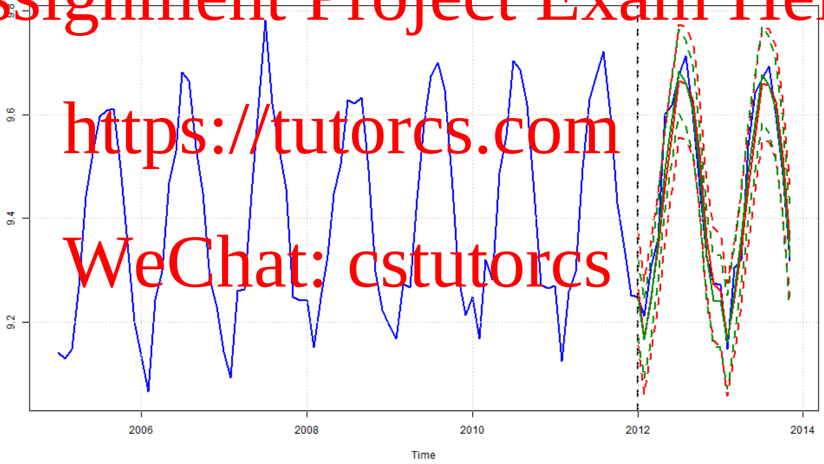


FIGURE 5.14. Forecast from the Full Model on the Original Scale, with realized values