Assignment 1 Solutions

$\boxed{1.1}$ (a) Note that $E[(Y-c)^2] = E(Y^2 - 2Yc + c^2] = E(Y^2) - 2c\mu + c^2$

Find the extreme point by differentiating with respect to c:

$$\frac{d}{dc}\left(E(Y^2) - 2c\mu^2 + c^2\right) = -2\mu + 2c$$

Setting equal to zero yields $c=\mu$ is an extremum

Since $\frac{d^2}{dc^2}\left(E(Y^2) - 2c\mu^2 + c^2\right) = 2 > 0$ $c=\mu$ in a minumum

(b) We have $E[(Y-f(X))^2 | X] = E(Y^2 - 2Y f(X) + f^2(X) | X]$

$$= E(Y^2|X) - 2f(X)E(Y|X) + f^2(X)$$

which is minimized by $f(X) = E(Y|X)$. To see this, take $c = f(X)$ and $\mu = E(Y|X)$ in (a).

(c) $E[(Y-f(X))^2] = E\left[E[(Y-f(X))^2 | X]\right]$ by Law of total Expectation

so the result follows from (b).

$\boxed{1.3}$ It is sufficient to show that the conditions of weak stationary hold true, when $\{X_t\}$ is assumed strongly stationary with $E(X_t^2) < \infty$.

- Strict Stationarity implies that $X_t \overset{D}{=} X_{t-h}$ (equally distributed) for all t and all h. Taking $h = t-1$, we have $X_t \overset{D}{=} X_1$ for all t.

  $\Rightarrow E(X_t) = E(X_1)$

  Let us call $E(X_1) = \mu$. Thus $E(X_t) = \mu$ for any t and the first condition of weak stationarity is satisfied

- Strict Stationarity implies that $(X_t, X_{t+k}) \overset{D}{=} (X_{t-h}, X_{t+k-h})$ for all t, h and k.

$\Rightarrow Cov(X_t, X_{t+k}) = Cov(X_{t-h}, X_{t+k-h}) \quad \forall t, \forall h, \forall k$

If $h=t$, $Cov(X_t, X_{t+k}) = Cov(X_0, X_k)$

If $h=t+k$, $Cov(X_t, X_{t+k}) = Cov(X_{-k}, X_0)$

Jointly these expressions tell us $Cov(X_t, X_{t+k}) = Cov(X_0, X_k)$
$= Cov(X_{-k}, X_0)$

which is independent of t for all k. This satisfies the second condition of weak stationarity

$\therefore \{X_t\}$ is weakly stationary

$\boxed{1.4}$ (a) $X_t = a + bZ_t + cZ_{t-2}$

- $E(X_t) = E(a + bZ_t + cZ_{t-2}) = a$ ← independent of t
- $Cov(X_t, X_{t+h}) = Cov(a + bZ_t + cZ_{t-2}, a + bZ_{t+h} + cZ_{t+h-2})$

$$= b^2 Cov(Z_t, Z_{t+h}) + bc\, Cov(Z_t, Z_{t+h-2}) + cb\, Cov(Z_{t-2}, Z_{t+h}) + c^2 Cov(Z_{t-2}, Z_{t+h-2})$$

$$= \begin{cases} b^2 Var(Z_t) + c^2 Var(Z_{t-2}) & \text{if } h=0 \\ bc\, Var(Z_t) & \text{if } h=2 \\ cb\, Var(Z_{t-2}) & \text{if } h=-2 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \sigma^2(b^2+c^2) & \text{if } h=0 \\ \sigma^2 bc & \text{if } |h|=2 \\ 0 & \text{otherwise} \end{cases}$$

which does not depend on t for any h. Since $\mu_X(t) = a$ and $\gamma_X(h)$ as above are independent of t, $\{X_t\}$ is stationary.

(b) $X_t = Z_1 \cos(ct) + Z_2 \sin(ct)$

- $E(X_t) = E[Z_1 \cos(ct) + Z_2 \sin(ct)] = 0$ ← independent of t
- $Cov(X_t, X_{t+h}) = Cov[Z_1 \cos(ct) + Z_2 \sin(ct), Z_1 \cos(c(t+h)) + Z_2 \sin(c(t+h))]$

$= \cos(ct)\cos(c(t+h)) Var(Z_1) + \sin(ct)\sin(c(t+h)) Var(Z_2)$ * $Cov(Z_1,Z_2) = Cov(Z_2,Z_1) = 0$

$= \sigma^2[\cos(ct)\cos(c(t+h)) + \sin(ct)\sin(c(t+h))]$

$= \sigma^2 \cos(ch)$ ← does not depend on t

*note: $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$

Since $\mu_X(t) = 0$ and $\gamma_X(h)$ as above do not depend on t, $\{X_t\}$ is stationary

(c) $X_t = Z_t \cos(ct) + Z_{t-1} \sin(ct)$

- $E(X_t) = E[Z_t \cos(ct) + Z_{t-1} \sin(ct)] = 0$
- $Cov(X_t, X_{t+h}) = Cov[Z_t \cos(ct) + Z_{t-1} \sin(ct), Z_{t+h} \cos(c(t+h)) + Z_{t+h-1} \sin(c(t+h))]$

$= \cos(ct)\cos(c(t+h)) Cov(Z_t, Z_{t+h}) + \cos(ct)\sin(c(t+h)) Cov(Z_t, Z_{t+h-1})$
$+ \sin(ct)\cos(c(t+h)) Cov(Z_{t-1}, Z_{t+h}) + \sin(ct)\sin(c(t+h)) Cov(Z_{t-1}, Z_{t+h-1})$

$$= \begin{cases} \cos(ct)^2 Var(Z_t) + \sin(ct)^2 Var(Z_{t-1}) & \text{if } h=0 \\ \cos(ct)\sin(c(t+h)) Var(Z_t) & \text{if } h=1 \\ \sin(ct)\cos(c(t+h)) Var(Z_{t-1}) & \text{if } h=-1 \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \sigma^2 & \text{if } h=0 \\ \sigma^2 \cos(ct)\sin(c(t+h)) & \text{if } h=1 \\ \sigma^2 \sin(ct)\cos(c(t+h)) & \text{if } h=-1 \\ 0 & \text{otherwise} \end{cases}$$

Notice that if $c = k\pi$, $k \in \mathbb{Z}$ then the $\sin()\cos()$ terms are zero for any t. Thus, if $c = k\pi$, then $\{X_t\}$ is stationary with $\mu_X(t) = 0$ and $\gamma_X(h) = \sigma^2$ if $h=0$ and $\gamma_X(h) = 0$ otherwise. If $c \neq k\pi$, then $\{X_t\}$ is not stationary.

(d) $X_t = a + bZ_0$

- $E(X_t) = E(a+bZ_0) = a$ ← independent of t
- $Cov(X_t, X_{t+h}) = Cov(a+bZ_0, a+bZ_0) = b^2 Var(Z_0) = b^2\sigma^2$

Since $\mu_x(t)$ and $\gamma_x(h)$ do not depend on t, $\{X_t\}$ is stationary

(e) $X_t = Z_0 \cos(ct)$

- $E(X_t) = E(Z_0 \cos(ct)) = 0$ ← independent of t
- $Cov(X_t, X_{t+h}) = Cov(Z_0 \cos(ct), Z_0 \cos(c(t+h)))$

$$= \cos(ct)\cos(c(t+h)) Var(Z_0)$$
$$= \sigma^2 \cos(ct)\cos(c(t+h))$$

Notice that $\gamma_x(t, t+h)$ does not depend on t if the coefficient of t is an odd multiple of $\frac{\pi}{2}$ i.e. $c = (2k+1)(\frac{\pi}{2})$ where $k \in \mathbb{Z}$. Thus, if $c = (2k+1)(\pi/2)$ for $k \in \mathbb{Z}$, then $\{X_t\}$ is stationary, but if not, $\{X_t\}$ is not stationary.

(f) $X_t = Z_t Z_{t-1}$

- $E(X_t) = E(Z_t Z_{t-1}) = 0$ ← independent of t
- $Cov(X_t, X_{t+h}) = Cov(Z_t Z_{t-1}, Z_{t+h} Z_{t+h-1})$

$$= E(Z_t Z_{t-1} Z_{t+h} Z_{t+h-1}) - E(Z_t \cancel{Z_{t-1}}^{0}) E(Z_{t+h} \cancel{Z_{t+h-1}}^{0})$$
$$= \begin{cases} E(Z_t^2 Z_{t-1}^2) & \text{if } h=0 \\ 0 & \text{otherwise} \end{cases}$$
$$= \begin{cases} E(Z_t^2)E(Z_{t-1}^2) = \sigma^4 & \text{if } h=0 \\ 0 & \text{otherwise} \end{cases} \quad \leftarrow \text{independent of } t$$

Since $\mu_x(t) = 0$ and $\gamma_x(h) = \gamma_x(t, t+h)$ do not depend on t, $\{X_t\}$ is a stationary time series.

$\boxed{1.7}$ Consider the time series $\{W_t\} = \{X_t + Y_t\}$ where $\{X_t\}$ and $\{Y_t\}$ are both uncorrelated stationary time series.

- $\mu_w(t) = E(W_t) = E(X_t) + E(Y_t)$

$$= \mu_x(t) + \mu_y(t)$$

Since $\{X_t\}$ and $\{Y_t\}$ are both stationary, $\mu_x(t)$ and $\mu_y(t)$ both do not depend on t. As such $\mu_w(t)$ does not dependent on t.

- $Cov(W_t, W_{t+h}) = Cov(X_t + Y_t, X_{t+h} + Y_{t+h})$

$$= Cov(X_t, X_{t+h}) + \cancel{Cov(X_t, Y_{t+h})}^{0}$$
$$\cancel{Cov(Y_t, X_{t+h})}^{0} + Cov(Y_t, Y_{t+h})$$
$$= \gamma_x(h) + \gamma_y(h)$$
$$= \gamma_w(h)$$

Since $\{X_t\}$ and $\{Y_t\}$ are both stationary, their ACVF's do not depend on t. As such $\gamma_w(h)$ does not depend on t either. Therefore $\{W_t\} = \{X_t + Y_t\}$ is stationary with $\gamma_w(h) = \gamma_x(h) + \gamma_y(h)$

$\boxed{2.3}$ (a) $X_t = Z_t + 0.3 Z_{t-1} - 0.4 Z_{t-2}$, $\{Z_t\} \sim WN(0,1)$

$Cov(X_t, X_{t+h}) = Cov(Z_t + 0.3 Z_{t-1} - 0.4 Z_{t-2}, Z_{t+h} + 0.3 Z_{t+h-1} - 0.4 Z_{t+h-2})$

$= Cov(Z_t, Z_{t+h}) + 0.3\, Cov(Z_t, Z_{t+h-1}) - 0.4\, Cov(Z_t, Z_{t+h-2})$
$+ 0.3\, Cov(Z_{t-1}, Z_{t+h}) + 0.09\, Cov(Z_{t-1}, Z_{t+h-1})$
$- 0.12\, Cov(Z_{t-1}, Z_{t+h-2}) - 0.4\, Cov(Z_{t-2}, Z_{t+h})$
$- 0.12\, Cov(Z_{t-2}, Z_{t+h-1}) + 0.16\, Cov(Z_{t-2}, Z_{t+h-2})$

$$= \begin{cases} 1 + 0.09 + 0.16 & \text{if } h = 0 \\ 0.3 - 0.12 & \text{if } |h| = 1 \\ -0.4 & \text{if } |h| = 2 \\ 0 & \text{o.w.} \end{cases}$$

$$= \begin{cases} 1.25 & \text{if } h = 0 \\ 0.18 & \text{if } |h| = 1 \\ -0.4 & \text{if } |h| = 2 \\ 0 & \text{if } h > 2 \text{ or } h < -2 \end{cases}$$

(b) $X_t = Z_t - 1.2 Z_{t-1} - 1.6 Z_{t-2}$, $\{Z_t\} \sim WN(0, 0.25)$

$Cov(X_t, X_{t+h}) = Cov(Z_t - 1.2 Z_{t-1} - 1.6 Z_{t-2}, Z_{t+h} - 1.2 Z_{t+h-1} - 1.6 Z_{t+h-2})$

$= Cov(Z_t, Z_{t+h}) - 1.2\, Cov(Z_t, Z_{t+h-1}) - 1.6\, Cov(Z_t, Z_{t+h-2}) -$
$1.2\, Cov(Z_{t-1}, Z_{t+h}) + 1.44\, Cov(Z_{t-1}, Z_{t+h-1}) + 1.92\, Cov(Z_{t-1}, Z_{t+h-2})$
$-1.6\, Cov(Z_{t-2}, Z_{t+h}) + 1.92\, Cov(Z_{t-2}, Z_{t+h-1}) + 2.56\, Cov(Z_{t-2}, Z_{t+h-2})$

$$= \begin{cases} (0.25)(1+1.44+2.56) & \text{if } h=0 \\ (0.25)(-1.2+1.92) & \text{if } |h|=1 \\ (0.25)(-1.6) & \text{if } |h|=2 \\ 0 & \text{o.w.} \end{cases}$$

$$= \begin{cases} 1.25 & \text{if } h=0 \\ 0.18 & \text{if } |h|=1 \\ -0.4 & \text{if } |h|=2 \\ 0 & \text{if } h>2 \text{ or } h<-2 \end{cases}$$

Additional Problem #1

(a) Find the ACVF of an MA(q) process.

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \quad \{\varepsilon_t\} \sim WN(0, \sigma^2)$$

$$= \sum_{i=0}^{q} \theta_i \varepsilon_{t-i} \quad \text{where } \theta_0 \equiv 1$$

* In what follows assume, without loss of generality that $h \geq 0$.

$$\gamma_X(h) = Cov(X_t, X_{t-h})$$

$$= E(X_t X_{t-h}) - E(X_t)E(X_{t-h})$$

$$= E(X_t X_{t-h}) \quad \text{since } E(X_t) = 0 \ \forall t.$$

$$= E\left[\left(\sum_{i=0}^{q}\theta_i \varepsilon_{t-i}\right)\left(\sum_{j=0}^{q}\theta_j \varepsilon_{t-h-j}\right)\right] \quad \text{where } \theta_0=1 \text{ by definition}$$

$$= \sum_{i=0}^{q}\sum_{j=0}^{q}\theta_i\theta_j E(\varepsilon_{t-i}\,\varepsilon_{t-h-j})$$

Note that if $t-i \neq t-h-j$ $E(\varepsilon_{t-i}\,\varepsilon_{t+h-j}) = 0$. This happens when $i=j+h$. In light of this, the expression above can be simplified:

$$\gamma_x(h) = \sum_{j=0}^{q}\theta_j\theta_{j+h}\,E(\varepsilon_t\,\varepsilon_{t-h-j})$$

Note that if $h>q$ $E(\varepsilon_t\,\varepsilon_{t-h-j})=0$. In light of this, $\gamma_x(h)$ can be simplified further:

$$\gamma_x(h) = \begin{cases} \sigma^2\sum_{j=0}^{q-h}\theta_j\theta_{j+h} & \text{if } 0\le h\le q \\ 0 & \text{if } h>q \end{cases}$$

(b) $\rho_x(h) = \dfrac{\gamma_x(h)}{\gamma_x(0)}$

$$\gamma_x(0) = \sigma^2\sum_{j=1}^{q}\theta_j^2 \;\rightarrow\; \rho_x(h) = \begin{cases} 1 & h=0 \\ \sum_{j=1}^{q-h}\theta_j\theta_{j+h}\Big/\sum_{j=1}^{q}\theta_j^2 & 0<h\le q \\ 0 & h>q \end{cases}$$

(c) Since $\rho_x(h)=0$ $\forall h>q$ and $\rho_x(h)\neq 0$ $\forall h\le q$ it is immediately obvious that the MA(q) process is q-correlated.

AS1 R solutions

Solutions to Question 2

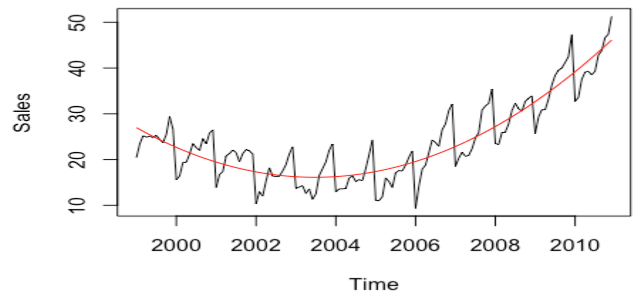
(a) Relevant Code and Output:

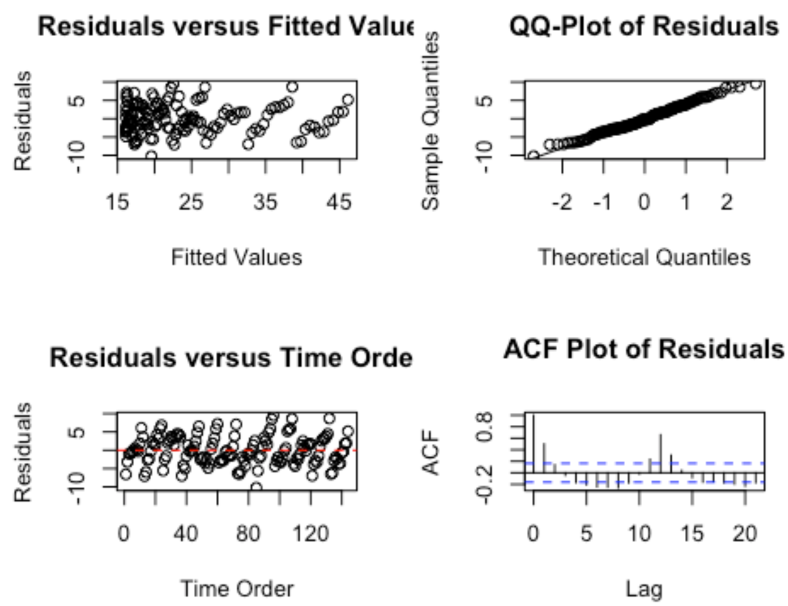
```
setwd("/Users/ntstevens/Documents/Teaching/MSAN 604/Assignments")
sales <- read.table('SALES.txt') #get the data
sales <- ts(sales,start=1999,frequency=12) #make it a time series objec
t
tim <- time(sales) #extract time covariate
tim2 <- tim^2 #create quadratic term
reg1 <- lm(sales~tim+tim2)  #fit model
summary(reg1) #model summary

##
## Call:
## lm(formula = sales ~ tim + tim2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.2493  -2.7326  -0.2823   2.6100   9.5576
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.175e+06  1.211e+05   17.95   <2e-16 ***
## tim         -2.171e+03  1.208e+02  -17.97   <2e-16 ***
## tim2         5.419e-01  3.014e-02   17.98   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.881 on 141 degrees of freedom
## Multiple R-squared:  0.8146, Adjusted R-squared:  0.812
## F-statistic: 309.8 on 2 and 141 DF,  p-value: < 2.2e-16

par(mfrow=c(1,1))
plot(sales,ylab='Sales') #plot the data
points(tim, predict.lm(reg1), type='l', col="red") #overlay fitted mode
l

#Residual Analysis:
par(mfrow=c(2,2))
plot(reg1$fitted, reg1$residuals, main = "Residuals versus Fitted Value
s", ylab = "Residuals", xlab = "Fitted Values")
qqnorm(reg1$residuals, main = "QQ-Plot of Residuals")
qqline(reg1$residuals)
plot(reg1$residuals, main = "Residuals versus Time Order", ylab = "Resi
duals", xlab = "Time Order")
abline(h=0,col='red',lty=2)
acf(reg1$residuals, main = "ACF Plot of Residuals")
```





The R^2 of the model is 0.8146, which means 81.46% of the total variability in sales is explained by the model. Looking at the plot of the data, with the fitted curve superimposed, we see that the general quadratic trend is captured, but the fluctuations around the curve have not been modeled.

Now let us look at the diagnostics plots based on residuals (see plots above).

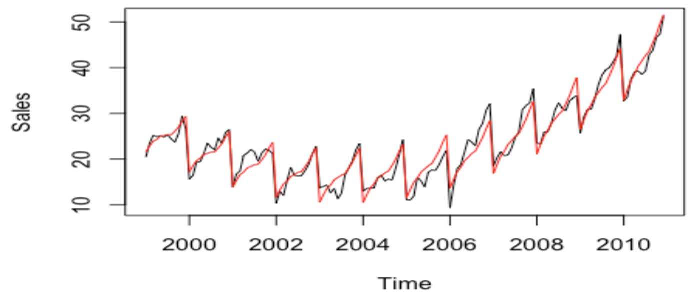
1. The residuals vs. fitted values plot (top left) shows no particular trend. There is no indication of non-constant variance either.
2. The QQ-plot (top right panel) checks the normality assumption. The points in this plot lie along the straight line, so there is no indication the normality assumption for the residuals is violated.
3. The plot of residuals vs. time (bottom left panel) checks for constant mean (=0), constant variance and can also reveal dependencies among residuals. The points are randomly scattered about zero, there does appear to be a seasonal pattern in the residuals suggesting that one is not independent of the next.
4. The ACF plot (bottom right panel) checks for autocorrelation among residuals. We see many ACF values outside the 95% confidence limits for different lags, which confirms the existence of autocorrelation among residuals, violating the independence assumption.

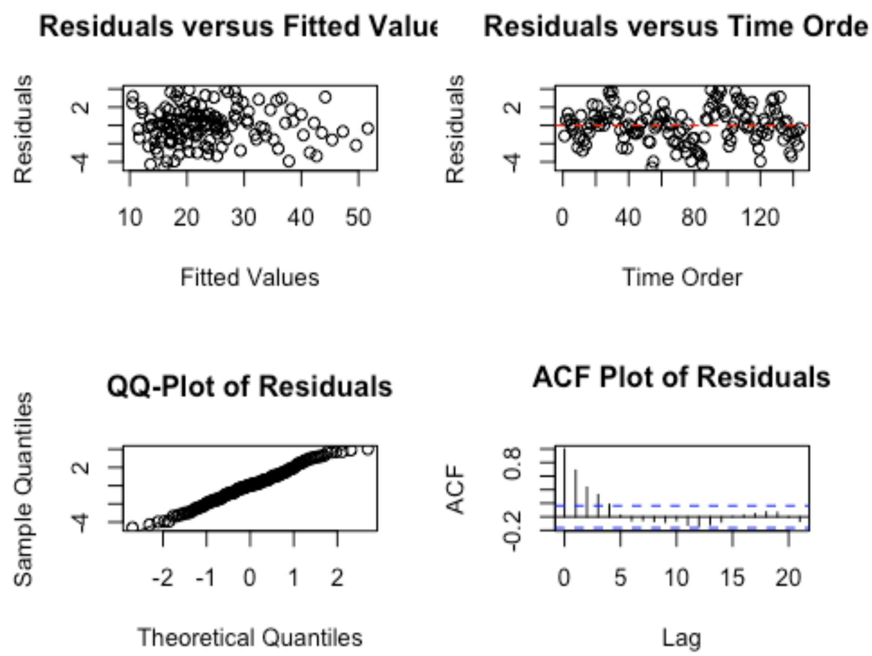
(b) Relevant Code and Output:

```
month <- as.factor(cycle(sales)) #extract month covariate (indicator va
riables)
reg2 <- lm(sales~tim+tim2+month) #fit model
summary(reg2) #model summary

##
## Call:
## lm(formula = sales ~ tim + tim2 + month)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.6296 -1.3720  0.0598  1.2164  4.0276
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.175e+06  6.356e+04  34.219  < 2e-16 ***
## tim         -2.171e+03  6.340e+01 -34.243  < 2e-16 ***
## tim2         5.418e-01  1.581e-02  34.267  < 2e-16 ***
## month2       1.674e+00  8.313e-01   2.014 0.046047 *
## month3       3.144e+00  8.313e-01   3.782 0.000236 ***
## month4       3.922e+00  8.314e-01   4.718 6.06e-06 ***
## month5       5.060e+00  8.315e-01   6.086 1.21e-08 ***
## month6       5.565e+00  8.316e-01   6.693 5.93e-10 ***
## month7       6.121e+00  8.317e-01   7.360 1.86e-11 ***
## month8       6.428e+00  8.318e-01   7.728 2.61e-12 ***
## month9       7.577e+00  8.319e-01   9.108 1.29e-15 ***
## month10      8.811e+00  8.321e-01  10.589  < 2e-16 ***
## month11      1.047e+01  8.323e-01  12.580  < 2e-16 ***
## month12      1.186e+01  8.325e-01  14.240  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.036 on 130 degrees of freedom
## Multiple R-squared:  0.9529, Adjusted R-squared:  0.9482
## F-statistic: 202.5 on 13 and 130 DF,  p-value: < 2.2e-16

par(mfrow=c(1,1))
plot(sales, ylab="Sales") #plot the data
points(tim, predict.lm(reg2), type='l', col="red") #overlay fitted mode
l
#Residual Analysis:
par(mfcol=c(2,2))
plot(reg2$fitted, reg2$residuals, main = "Residuals versus Fitted Value
s", ylab = "Residuals", xlab = "Fitted Values")
qqnorm(reg2$residuals, main = "QQ-Plot of Residuals")
qqline(reg2$residuals)
plot(reg2$residuals, main = "Residuals versus Time Order", ylab = "Resi
duals", xlab = "Time Order")
abline(h=0,col='red',lty=2)
acf(reg2$residuals, main = "ACF Plot of Residuals")
```





The R^2 of this model is 0.9529 which means that 95.29% of the total variability in sales is explained by this model. Plotting the data and the fitted model (above), we can see that both the quadratic trend as well as the periodic fluctuations around it, are captured by the model.

Now we look at the diagnostics plots based on the residuals.

1. The residuals vs. fitted values plot (top left) shows no particular trend. There is no indication of non-constant variance either.
2. The QQ-plot (bottom left panel) checks the normality assumption. The points in this plot lie along the straight line, so there is no indication the normality assumption for the residuals is violated.
3. The plot of residuals vs. time (top right panel) checks for constant mean (=0), constant variance and can also reveal dependencies among residuals. The points are randomly scattered about zero and no particular trend is observed.
4. The ACF plot (bottom right panel) checks for autocorrelation among residuals. We see ACF values outside the 95% confidence limits for lags 1, 2, 3 and 4, which suggests the residuals are autocorrelated, violating the independence assumption.

(c) Since the two models have different number of parameters, we should compare their fits using the adjusted R^2. While for the first model $R^2_{adj} = 0.812$, in the second model $R^2_{adj} = 0.9482$. This shows that the second model provides a better fit to the data.

(d) Neither of the models satisfy all OLS assumptions as in both cases the independence assumption for the residuals is violated.

(e) Relevant Code and Output:

```
t.new <- seq(2011,2012,length=13)[1:12] #new time for forecasting 2011
(notice that it is 1:12 because 2012 should not be included)
t2.new <- t.new^2 #new quadratic term
month.new <- factor(rep(1:12,1)) #new seasonal values for forecasting
new <- data.frame(tim=t.new, tim2=t2.new, month=month.new) #putting the
values for forecasting into a dataframe
pred <- predict.lm(reg2,new,interval='prediction') #computing the predi
ctions as well as prediction intervals
par(mfrow=c(1,1))
plot(sales,xlim=c(1999,2012),ylim=c(0,65),ylab="Sales") #plotting the d
ata
abline(v=2011,col='blue',lty=2) #adding a vertical line at the point wh
ere prediction starts
lines(pred[,1]~t.new,type='l',col='red')# plotting the predictions
lines(pred[,2]~t.new,col='green') # plotting lower limit of the predict
ion intervals
lines(pred[,3]~t.new,col='green') # plotting upper limit of the  predic
tion intervals
```