Lecture 6

Discrete Time Series

9/21/2018

Discrete Time Series

Stationary Processes

A stocastic process (i.e. a time series) is considered to be *strictly stationary* if the properties of the process are not changed by a shift in origin.

A stocastic process (i.e. a time series) is considered to be *strictly stationary* if the properties of the process are not changed by a shift in origin.

In the time series context this means that the joint distribution of $\{y_{t_1}, \ldots, y_{t_n}\}$ must be identical to the distribution of $\{y_{t_1+k}, \ldots, y_{t_n+k}\}$ for any value of n and k.

Weak Stationary

Strict stationary is unnecessarily strong / restrictive for many applications, so instead we often opt for *weak stationary* which requires the following,

1. The process has finite variance

$$E(y_t^2) < \infty \text{ for all } t$$

2. The mean of the process is constant

$$E(y_t) = \mu \text{ for all } t$$

3. The second moment only depends on the lag

$$Cov(y_t, y_s) = Cov(y_{t+k}, y_{s+k}) \text{ for all } t, s, k$$

Weak Stationary

Strict stationary is unnecessarily strong / restrictive for many applications, so instead we often opt for *weak stationary* which requires the following,

1. The process has finite variance

$$E(y_t^2) < \infty \text{ for all } t$$

2. The mean of the process is constant

$$E(y_t) = \mu \text{ for all } t$$

3. The second moment only depends on the lag

$$Cov(y_t, y_s) = Cov(y_{t+k}, y_{s+k}) \text{ for all } t, s, k$$

When we say stationary in class we will almost always mean *weakly stationary*.

For a stationary time series, where $E(y_t) = \mu$ and Var$(y_t) = \sigma^2$ for all t, we define the autocorrelation at lag k as

$$\begin{aligned}\rho_k &= Cor(y_t,\, y_{t+k}) \\ &= \frac{Cov(y_t, y_{t+k})}{\sqrt{Var(y_t)Var(y_{t+k})}} \\ &= \frac{E\left((y_t - \mu)(y_{t+k} - \mu)\right)}{\sigma^2}\end{aligned}$$

For a stationary time series, where $E(y_t) = \mu$ and Var$(y_t) = \sigma^2$ for all t, we define the autocorrelation at lag k as

$$\begin{aligned}
\rho_k &= Cor(y_t,\, y_{t+k}) \\
&= \frac{Cov(y_t, y_{t+k})}{\sqrt{Var(y_t)Var(y_{t+k})}} \\
&= \frac{E\left((y_t - \mu)(y_{t+k} - \mu)\right)}{\sigma^2}
\end{aligned}$$

this is also sometimes written in terms of the autocovariance function (γ_k) as

$$\begin{aligned}
\gamma_k &= \gamma(t, t+k) = Cov(y_t, y_{t+k}) \\
\rho_k &= \frac{\gamma(t, t+k)}{\sqrt{\gamma(t,t)\gamma(t+k, t+k)}} = \frac{\gamma(k)}{\gamma(0)}
\end{aligned}$$

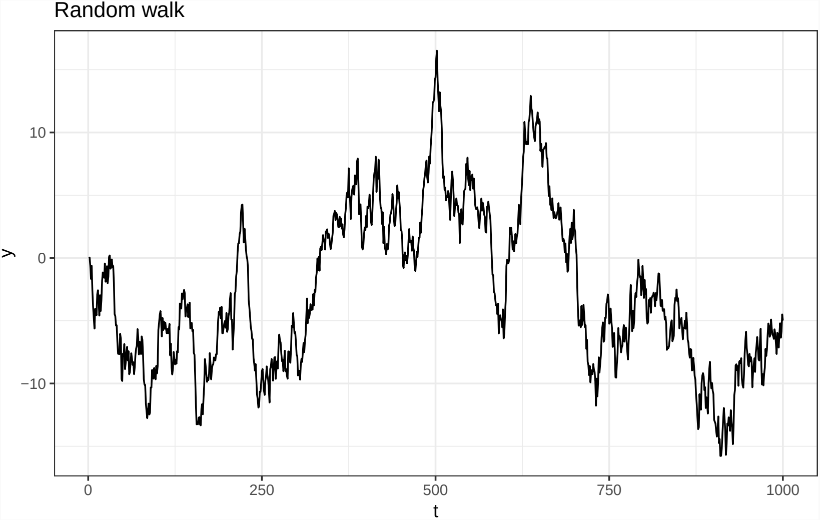
Based on our definition of a (weakly) stationary process, it implies a covariance of the following structure,

$$\mathbf{\Sigma} = \begin{pmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) & \cdots & \gamma(n-1) & \gamma(n) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(n-2) & \gamma(n-1) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(n-3) & \gamma(n-2) \\ \gamma(3) & \gamma(2) & \gamma(1) & \gamma(0) & \cdots & \gamma(n-4) & \gamma(n-3) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \gamma(n-3) & \gamma(n-4) & \cdots & \gamma(0) & \gamma(1) \\ \gamma(n) & \gamma(n-1) & \gamma(n-2) & \gamma(n-3) & \cdots & \gamma(1) & \gamma(0) \end{pmatrix}$$

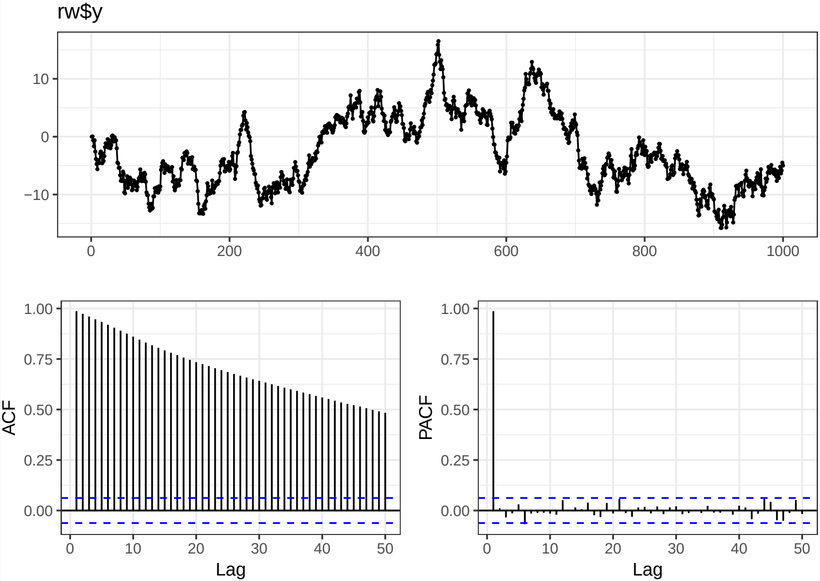
where $P_{t,k}(y)$ is the project of y onto the space spanned by $y_{t+1}, \ldots, y_{t+k-1}$.

Example - Random walk

Let $y_t = y_{t-1} + w_t$ with $y_0 = 0$ and $w_t \sim \mathcal{N}(0, 1)$.



ACF + PACF


rw$y
10
0
−10
0
200
400
600
800
1000
ACF
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
Lag
PACF
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
Lag

Stationary?

Is y_t stationary?

Partial Autocorrelation - pACF

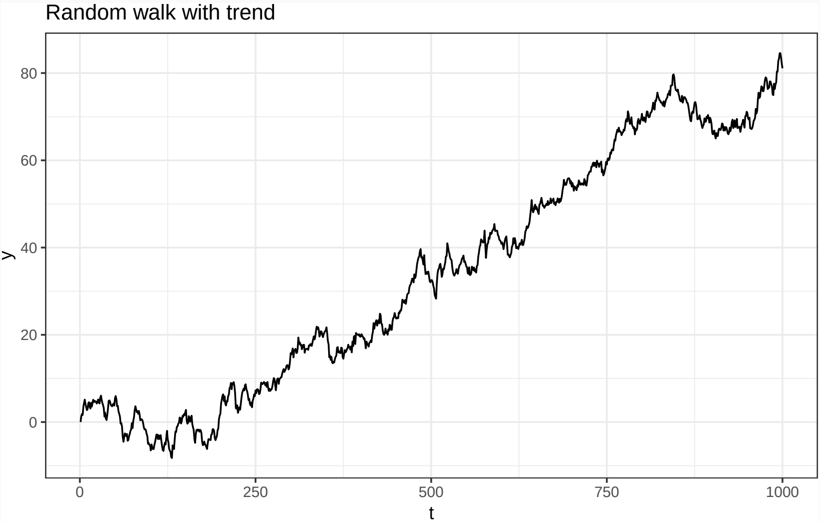
Given these type of patterns in the autocorrelation we often want to examine the relationship between y_t and y_{t+k} with the (linear) dependence of y_t on y_{t+1} through y_{t+k-1} removed.

This is done through the calculation of a partial autocorrelation ($\alpha(k)$), which is defined as follows:

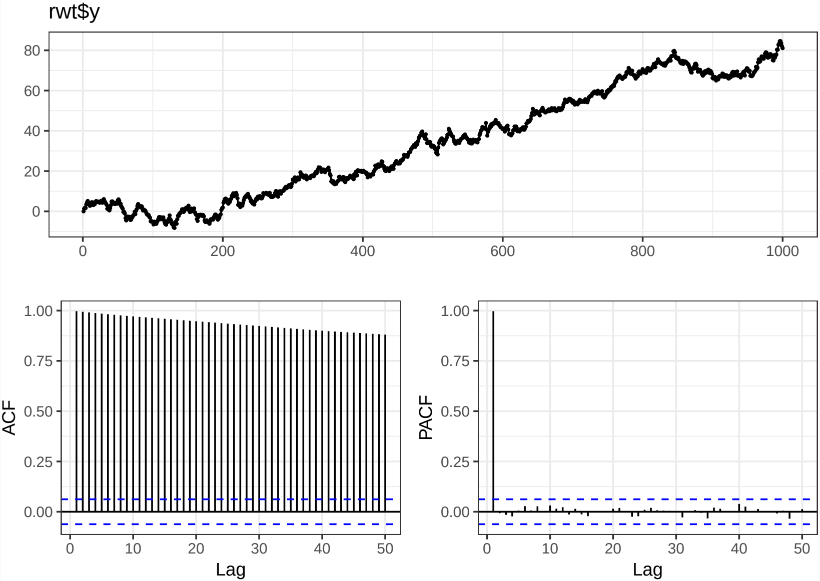
$$\begin{aligned}
\alpha(0) &= 1 \\
\alpha(1) &= \rho(1) = Cor(y_t, y_{t+1}) \\
&\vdots \\
\alpha(k) &= Cor(y_t - P_{t,k}(y_t),\ y_{t+k} - P_{t,k}(y_{t+k}))
\end{aligned}$$

Example - Random walk with drift

Let $y_t = \delta + y_{t-1} + w_t$ with $y_0 = 0$ and $w_t \sim \mathcal{N}(0,1)$.



ACF + PACF



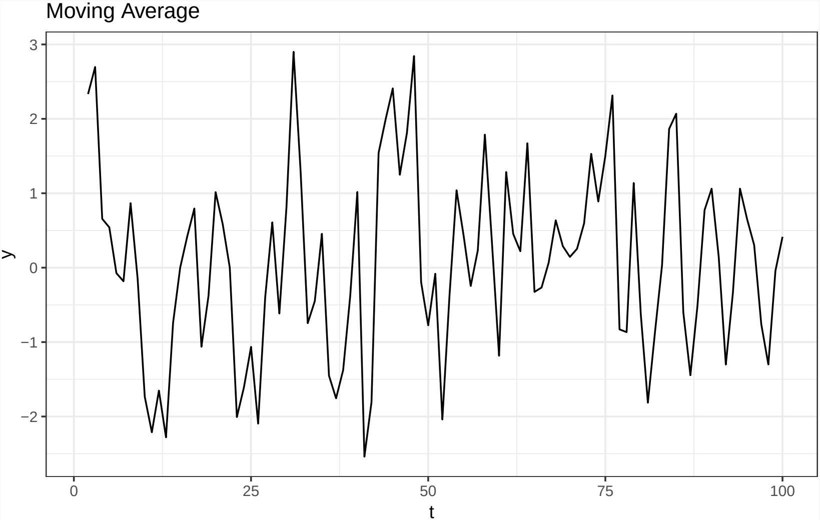
rwt$y
80
60
40
20
0
0
200
400
600
800
1000
ACF
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
Lag
PACF
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
Lag

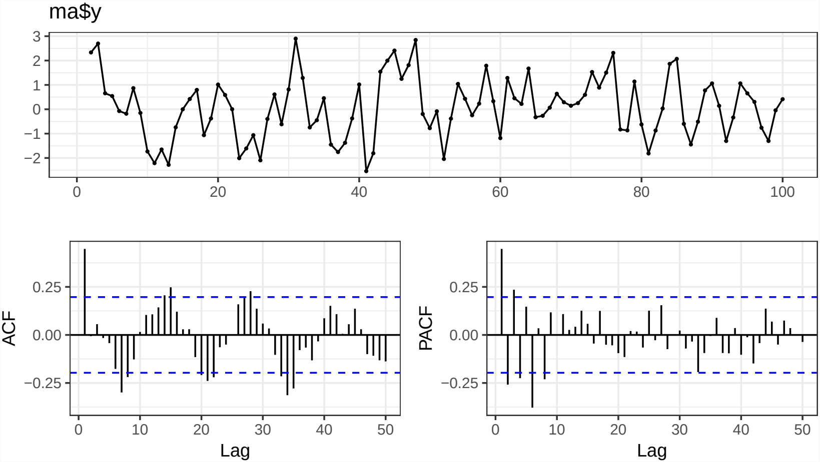
Stationary?

Is y_t stationary?

Example - Moving Average

Let $w_t \sim \mathcal{N}(0, 1)$ and $y_t = w_{t-1} + w_t$.

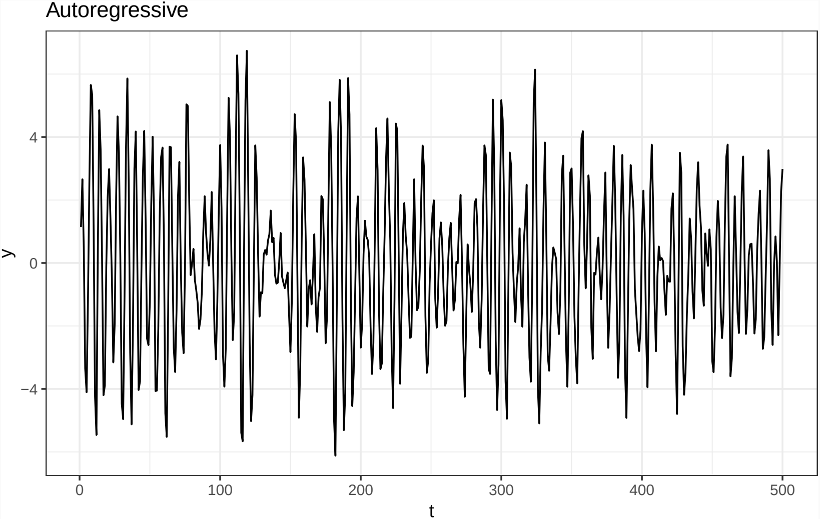



ma$y
3
2
1
0
−1
−2
0
20
40
60
80
100
ACF
0.25
0.00
−0.25
0
10
20
30
40
50
Lag
PACF
0.25
0.00
−0.25
0
10
20
30
40
50
Lag

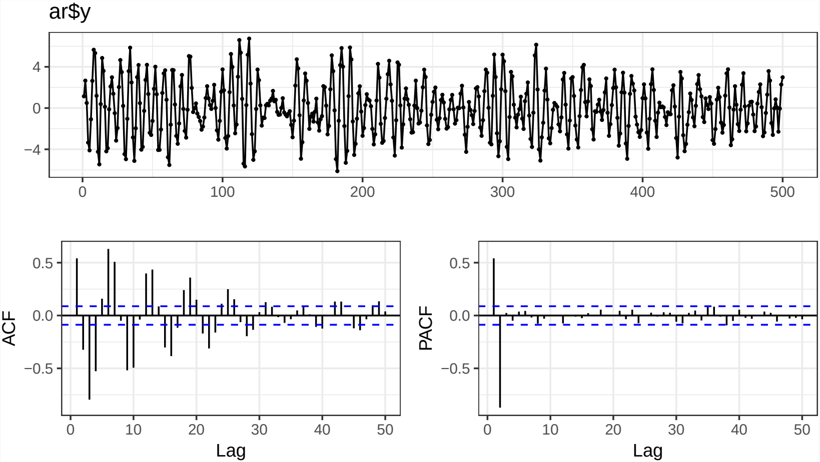
Stationary?

Is y_t stationary?

Let $w_t \sim \mathcal{N}(0,1)$ and $y_t = y_{t-1} - 0.9y_{t-2} + w_t$ with $y_t = 0$ for $t < 1$.



ACF + PACF



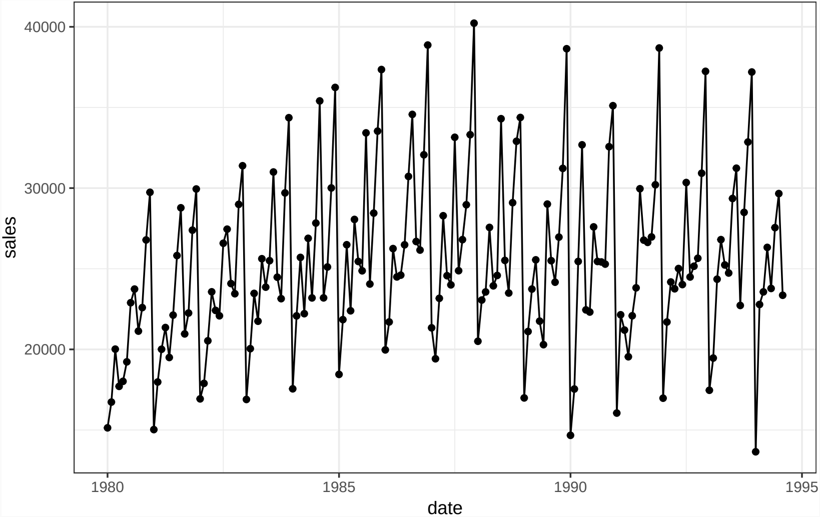
ar$y
4
0
−4
0
100
200
300
400
500
ACF
0.5
0.0
−0.5
0
10
20
30
40
50
Lag
PACF
0.5
0.0
−0.5
0
10
20
30
40
50
Lag

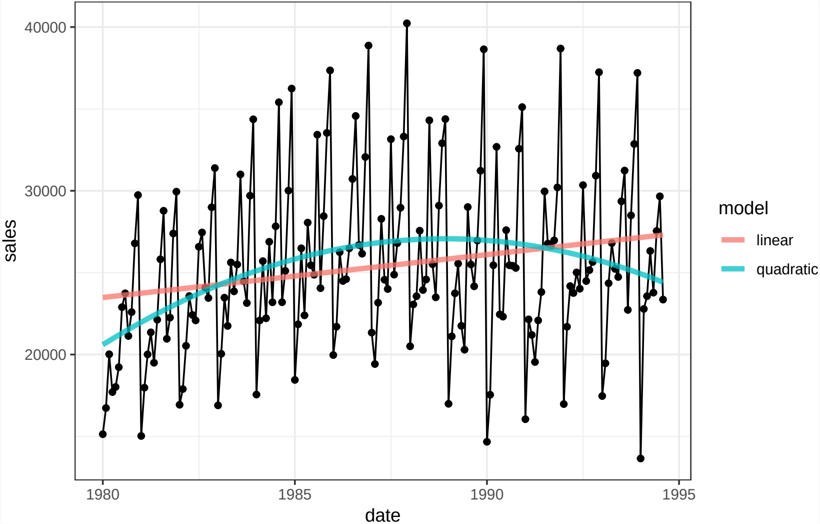
Example - Australian Wine Sales

Australian total wine sales by wine makers in bottles <= 1 litre. Jan 1980 – Aug 1994.

```
aus_wine = readRDS("../data/aus_wine.rds")
aus_wine
## # A tibble: 176 x 2
##     date sales
##    <dbl> <dbl>
##  1 1980  15136
##  2 1980. 16733
##  3 1980. 20016
##  4 1980. 17708
##  5 1980. 18019
##  6 1980. 19227
##  7 1980. 22893
##  8 1981. 23739
##  9 1981. 21133
## 10 1981. 22591
## # ... with 166 more rows
```

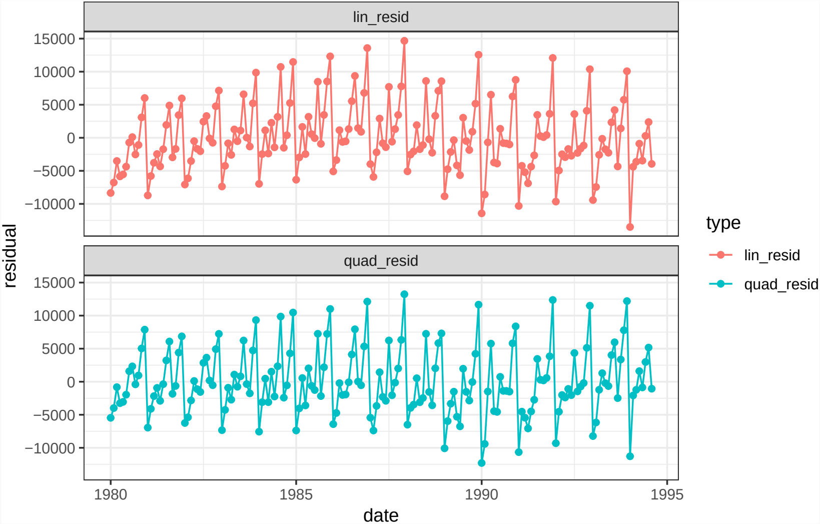
Time series





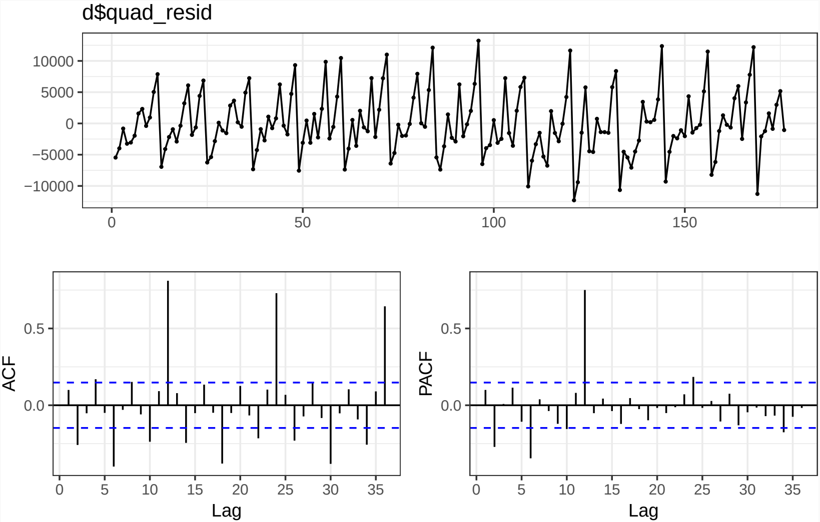
40000
30000
20000
sales
1980
1985
1990
1995
date
model
linear
quadratic

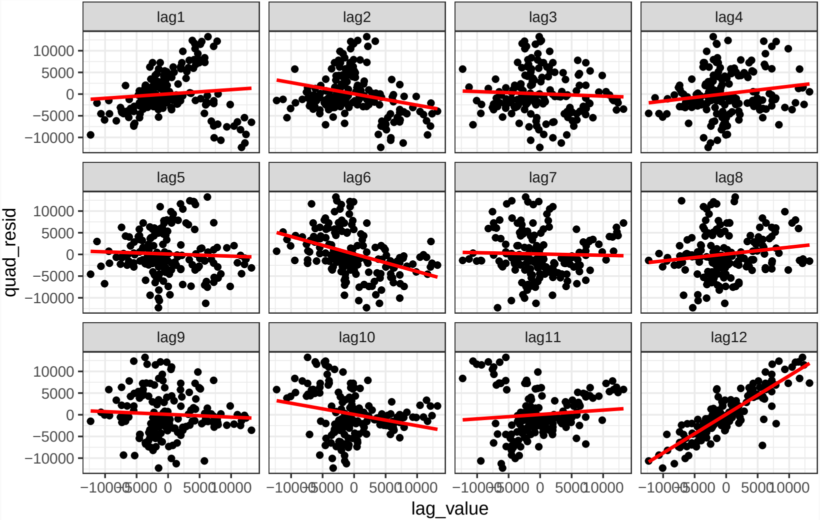
Residuals



lin_resid
quad_resid
15000
10000
5000
0
−5000
−10000
residual
1980
1985
1990
1995
date
type
lin_resid
quad_resid

Autocorrelation Plot

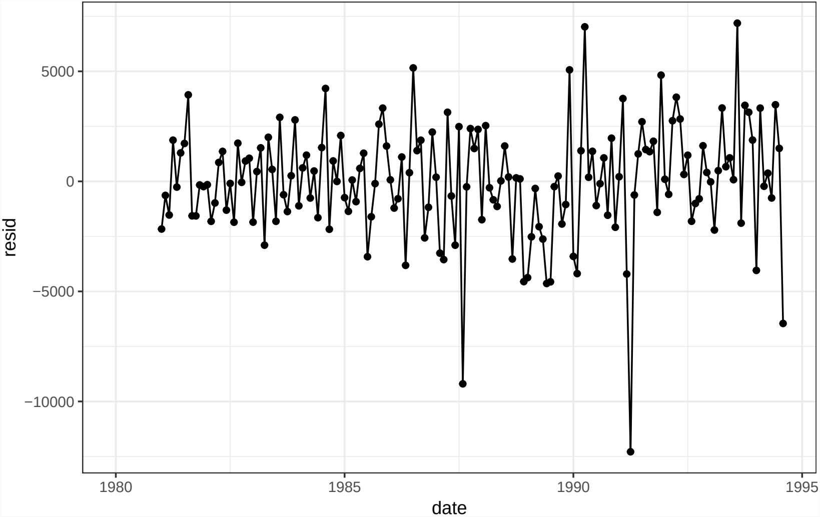



lag1
lag2
lag3
lag4
lag5
lag6
lag7
lag8
lag9
lag10
lag11
lag12
quad_resid
10000
5000
0
−5000
−10000
−10000
−5000
0
5000
10000
lag_value

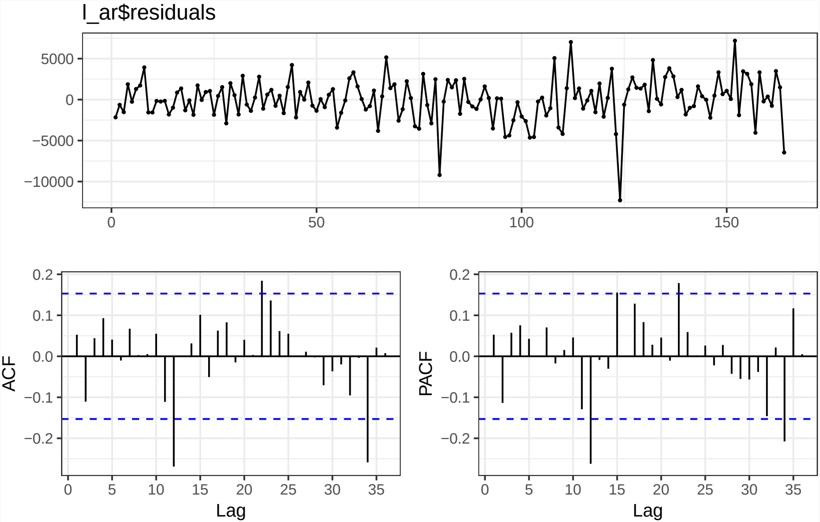
Auto regressive errors

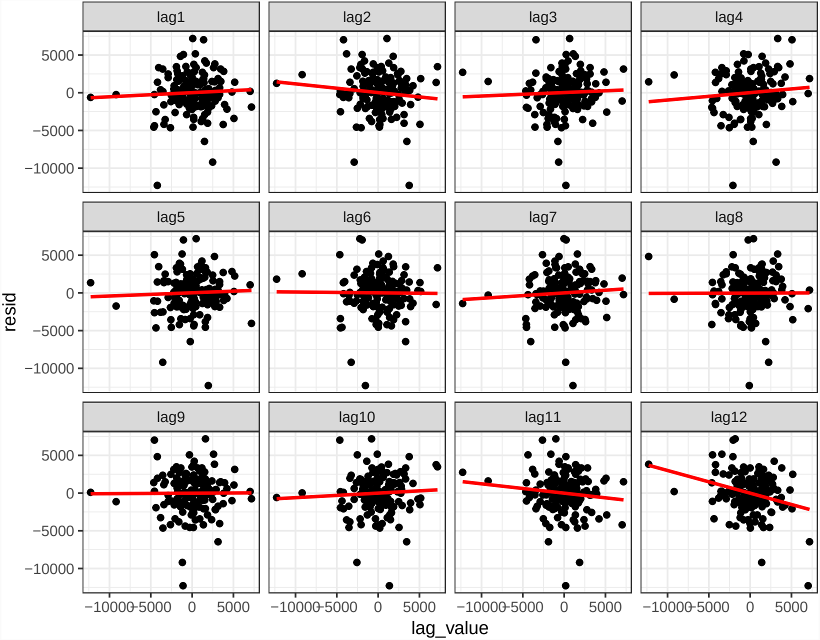
```
##
## Call:
## lm(formula = quad_resid ~ lag_12, data = d_ar)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12286.5  -1380.5     73.4   1505.2   7188.1
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  83.65080  201.58416   0.415    0.679
## lag_12        0.89024    0.04045  22.006   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2581 on 162 degrees of freedom
##   (12 observations deleted due to missingness)
## Multiple R-squared:  0.7493, Adjusted R-squared:  0.7478
## F-statistic: 484.3 on 1 and 162 DF,  p-value: < 2.2e-16
```

Residual residuals



Residual residuals - acf




lag1
lag2
lag3
lag4
lag5
lag6
lag7
lag8
lag9
lag10
lag11
lag12
resid
5000
0
−5000
−10000
−10000
−5000
0
5000
lag_value

Writing down the model?

So, is our EDA suggesting that we fit the following model?

$$\text{sales}(t) = \beta_0 + \beta_1\, t + \beta_2\, t^2 + \beta_3\, \text{sales}(t-12) + \epsilon_t$$

Writing down the model?

So, is our EDA suggesting that we fit the following model?

$$\text{sales}(t) = \beta_0 + \beta_1\, t + \beta_2\, t^2 + \beta_3\, \text{sales}(t - 12) + \epsilon_t$$

the model we actually fit is,

$$\text{sales}(t) = \beta_0 + \beta_1\, t + \beta_2\, t^2 + w_t$$

where

$$w_t = \delta\, w_{t-12} + \epsilon_t$$