Discrete Time Series

Lecture 07

Dr. Colin Rundel

Random variable review

Mean and variance of RVs

- Expected Value

$$E(X) = \begin{cases} \sum_x x \cdot P(X = x) & X \text{ is discrete} \\ \int_{-\infty}^{\infty} x \cdot f(x)\, dx & X \text{ is continuous} \end{cases}$$

- Variance

$$Var(X) = E\left(\left(X - E(X)\right)^2\right) = E(X^2) - E(X)^2$$

$$= \begin{cases} \sum_x \left(x - E(X)\right)^2 \cdot P(X = x) & X \text{ is discrete} \\ \int_{-\infty}^{\infty} \left(x - E(X)\right)^2 \cdot f(x)\, dx & X \text{ is continuous} \end{cases}$$

Covariance of RVs

$$\text{Cov}(X, Y) = E\Big(\big(X - E(X)\big)\big(Y - E(Y)\big)\Big) = E(XY) - E(X)E(Y)$$

$$= \begin{cases} \sum_x \big(x - E(X)\big)\big(y - E(Y)\big) \cdot P(X = x, Y = y) & X \text{ is discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \big(x - E(X)\big)\big(y - E(Y)\big) \cdot f(x, y)\, dx\, dy & X \text{ is continuous} \end{cases}$$

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\,\text{Var}(Y)}}$$

Properties of Expected Value

- *Constant*

  $E(c) = c$ if c is constant

- *Constant Multiplication*

  $E(cX) = cE(X)$

- *Constant Addition*

  $E(X + c) = E(X) + c$

- *Addition*

  $E(X + Y) = E(X) + E(Y)$

- *Subtraction*

  $E(X - Y) = E(X) - E(Y)$$

- *Multiplication*

  $E(XY) = E(X)\,E(Y)$

  if X and Y are independent

Properties of Variance

- *Constant*

  $Var(c) = 0$ if c is constant

- *Constant Multiplication*

  $Var(cX) = c^2\ Var(x)$

- *Constant Addition*

  $Var(X + c) = Var(X)$

- *Addition*

  $Var(X + Y) = Var(X) + Var(Y)$

  if X and Y are independent.

- *Subtraction*

  $Var(X - Y) = Var(X) + Var(Y)$

  if X and Y are independent.

Properties of Covariance

- *Constant*

  $\text{Cov}(X, c) = 0$ if c is constant

- *Constant Multiplication*

  $\text{Cov}(aX, bY) = ab\,\text{Cov}(X, Y)$

- *Identity*

  $\text{Cov}(X, X) = Var(X)$

- *Constant Addition*

  $\text{Cov}(X + a, Y + b) = \text{Cov}(X, Y)$

- *Symmetric*

  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

- *Distribution*

  $\text{Cov}(aX + bY, cV + dW) = ac\,\text{Cov}(X, V) + ad\,\text{Cov}(X, W) + bc\,\text{Cov}(Y, V) + bd$

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.

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.

Weakly 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 must have finite variance / second moment

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

2. The mean of the process must be constant

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

3. The cross moment (covariance) may only depends on the lag (i.e. $t - s$ for y_t and y_s)

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

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

Autocorrelation

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

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

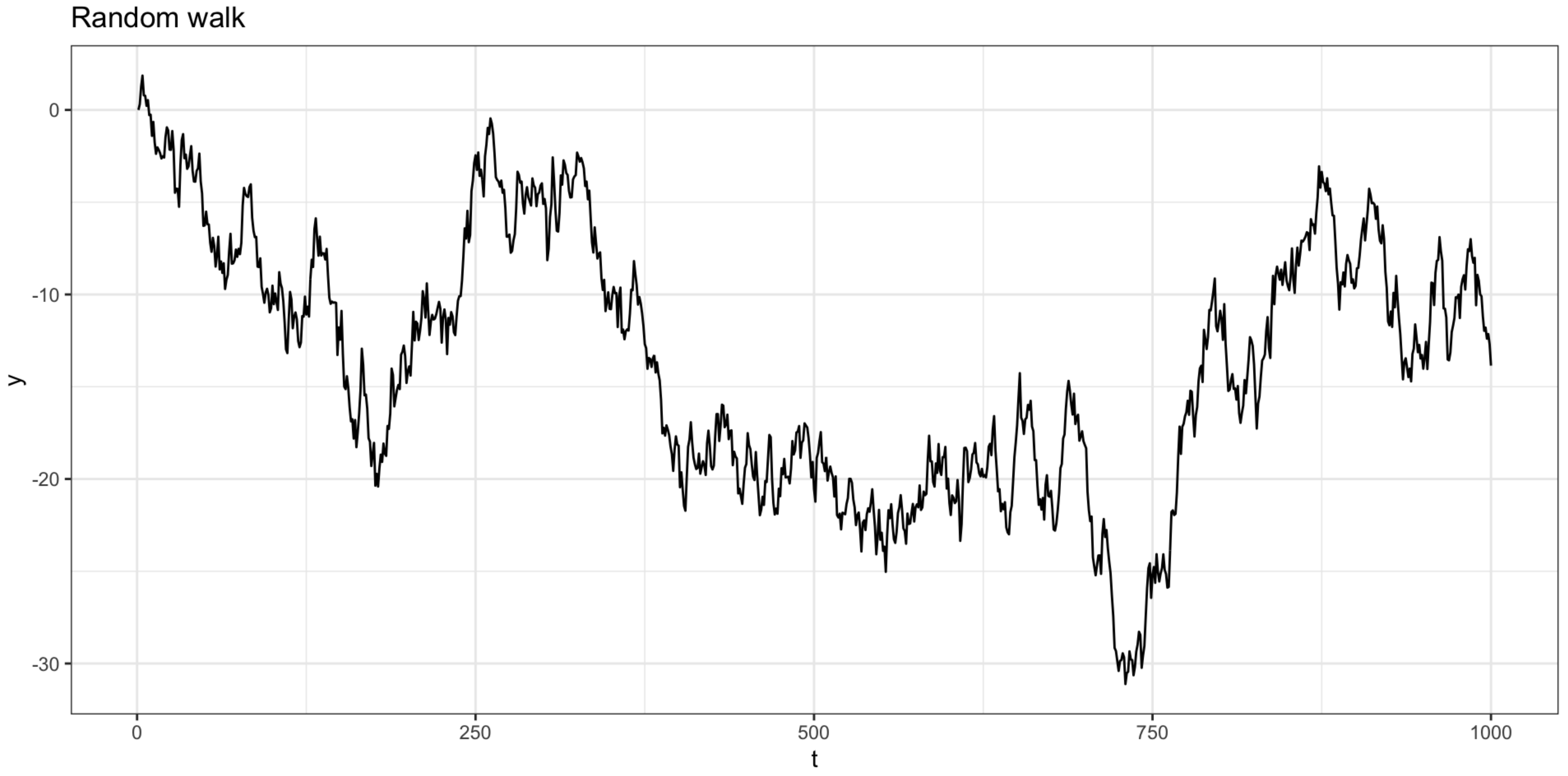
Covariance Structure

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

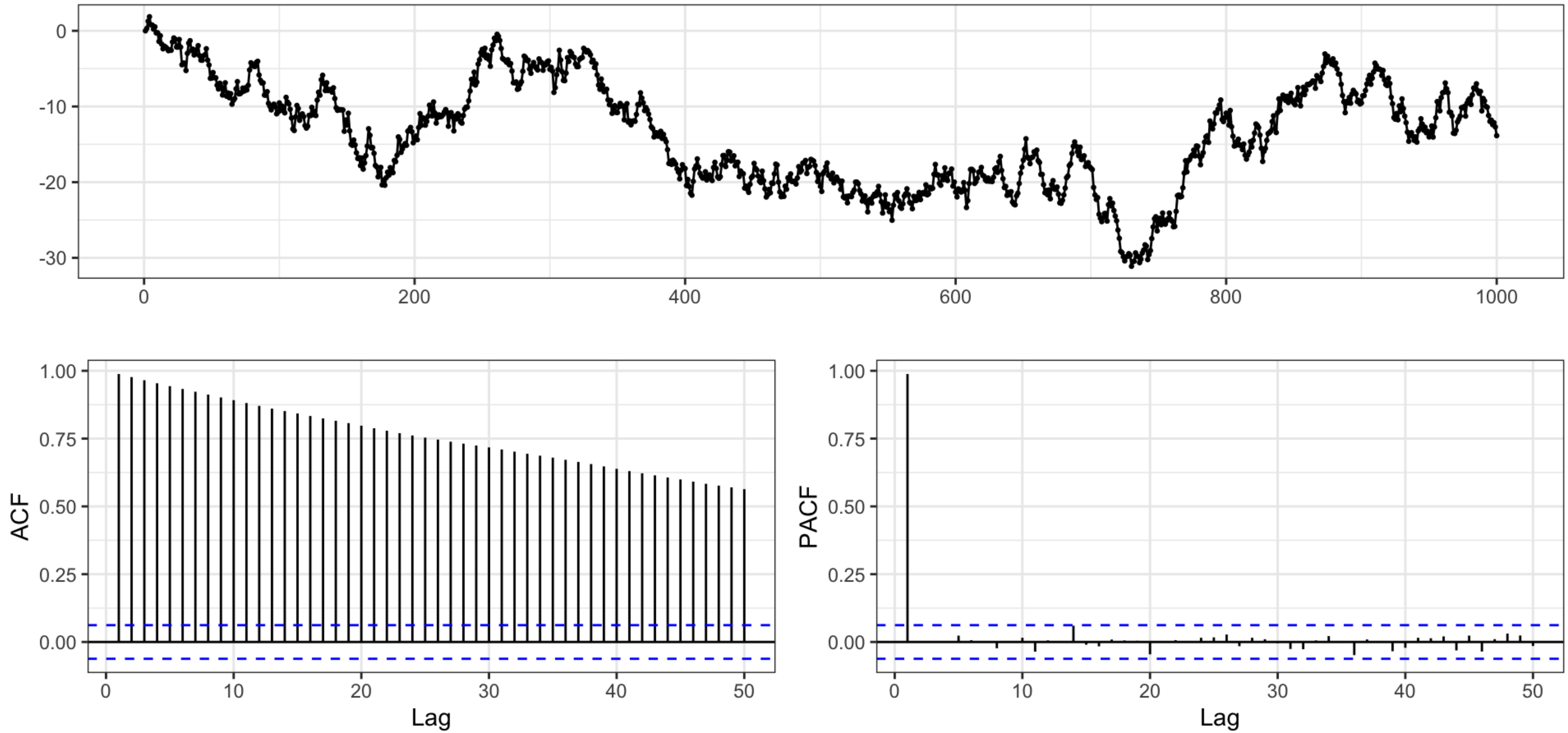
$$\boldsymbol{\Sigma} = \begin{pmatrix} \gamma(0) & \gamma(1) & \gamma(2) & \gamma(3) & \cdots & \gamma(n-1) & \gamma(n) \\ \gamma(1) & \gamma(0) & \gamma(1) & \gamma(2) & \cdots & \gamma(n-2) & \gamma(n-1) \\ \gamma(2) & \gamma(1) & \gamma(0) & \gamma(1) & \cdots & \gamma(n-3) & \gamma(n-2) \\ \gamma(3)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

Example - Random walk

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



ACF + PACF



0
-10
-20
-30
0
200
400
600
800
1000
1.00
0.75
0.50
0.25
0.00
ACF
0
10
20
30
40
50
Lag
1.00
0.75
0.50
0.25
0.00
PACF
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:

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

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

pACF - Calculation

Let $\rho(k)$ be the autocorrelation for the process at lag k then the partial autocorrelation at lag k will be $\phi(k, k)$ given by the Durbin-Levinson algorithm,

$$\phi(k, k) = \frac{\rho(k) - \sum_{t=1}^{k-1} \phi(k-1, t)\, \rho(k-t)}{1 - \sum_{t=1}^{k-1} \phi(k-1, t)\, \rho(t)}$$

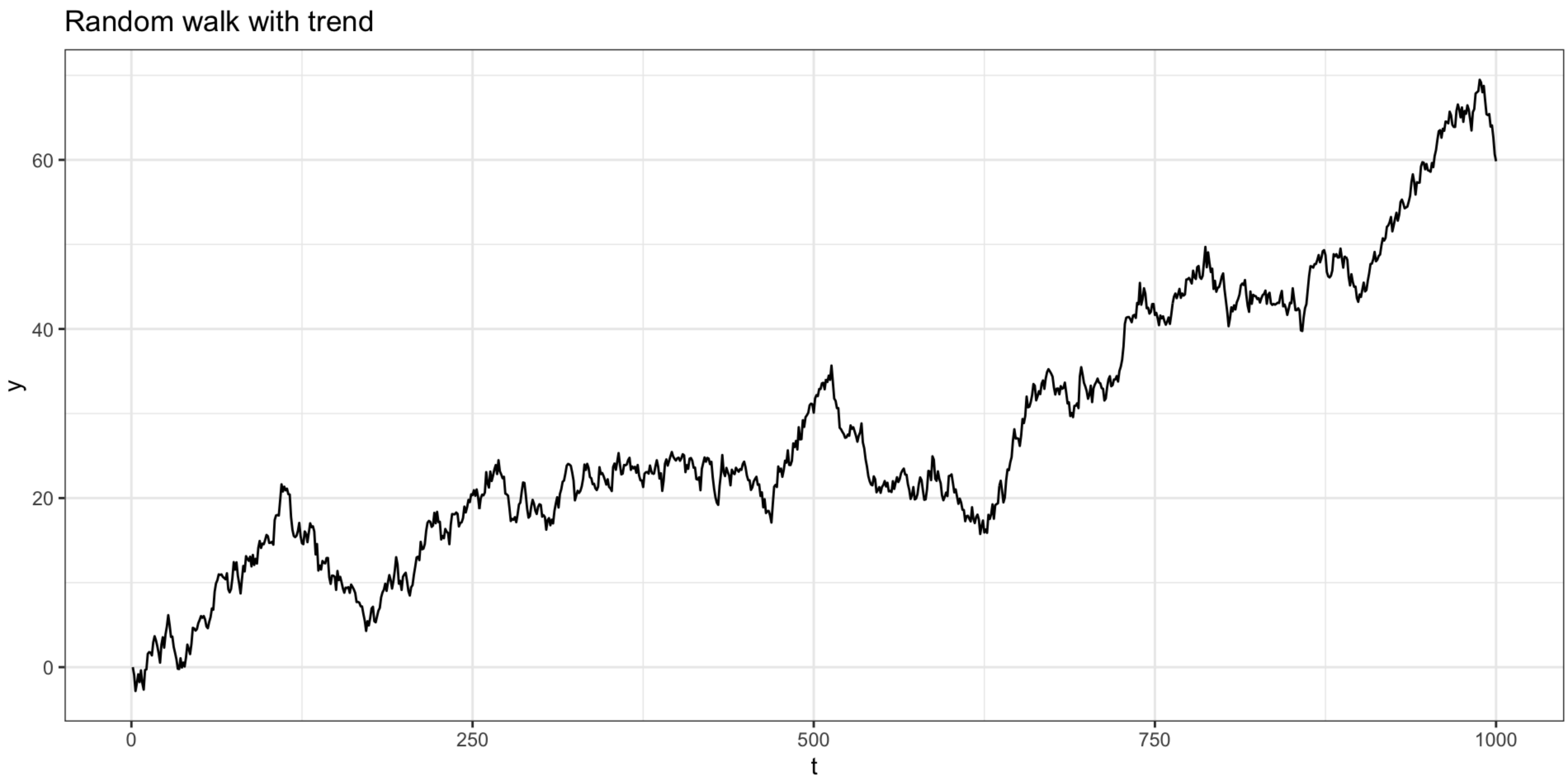
where

$$\phi(k, t) = \phi(k-1, t) - \phi(k, k)\, \phi(k-1, k-t)$$

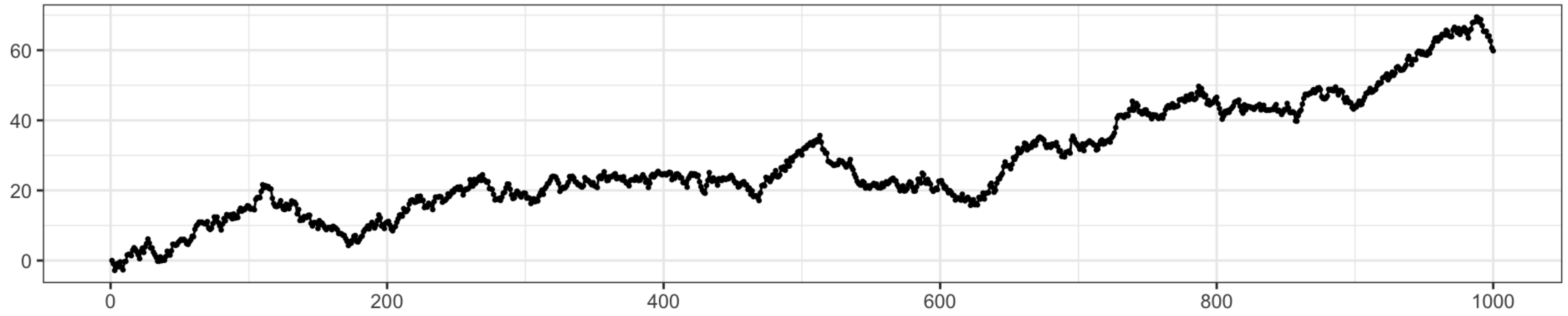
Starting with $\phi(1, 1) = \rho(1)$ we can solve iteratively for $\phi(2, 2), \ldots, \phi(k, k)$.

Example - Random walk with drift

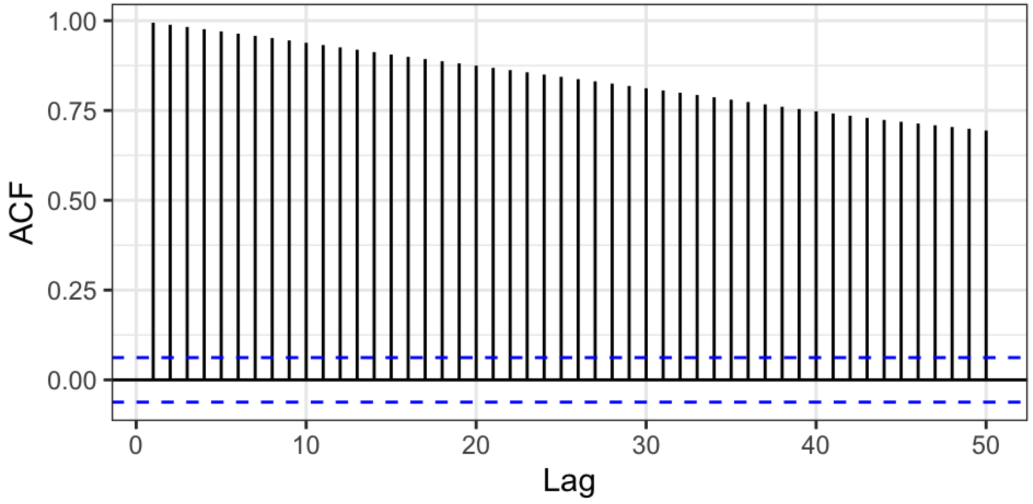
Let $y_t = \delta + y_{t-1} + w_t$ with $y_0 = 0$ and $w_t \sim N(0, 1)$.



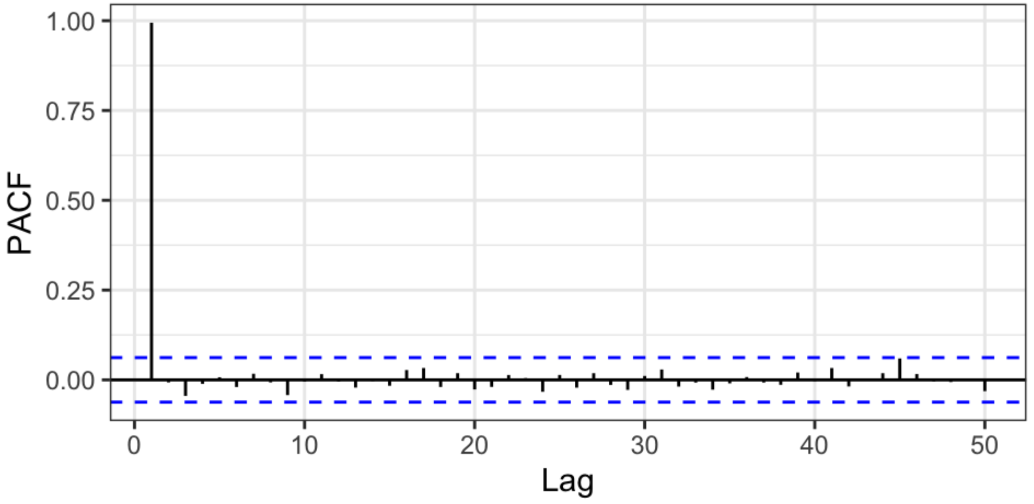
ACF + PACF



60
40
20
0
0
200
400
600
800
1000



1.00
0.75
0.50
0.25
0.00
ACF
0
10
20
30
40
50
Lag



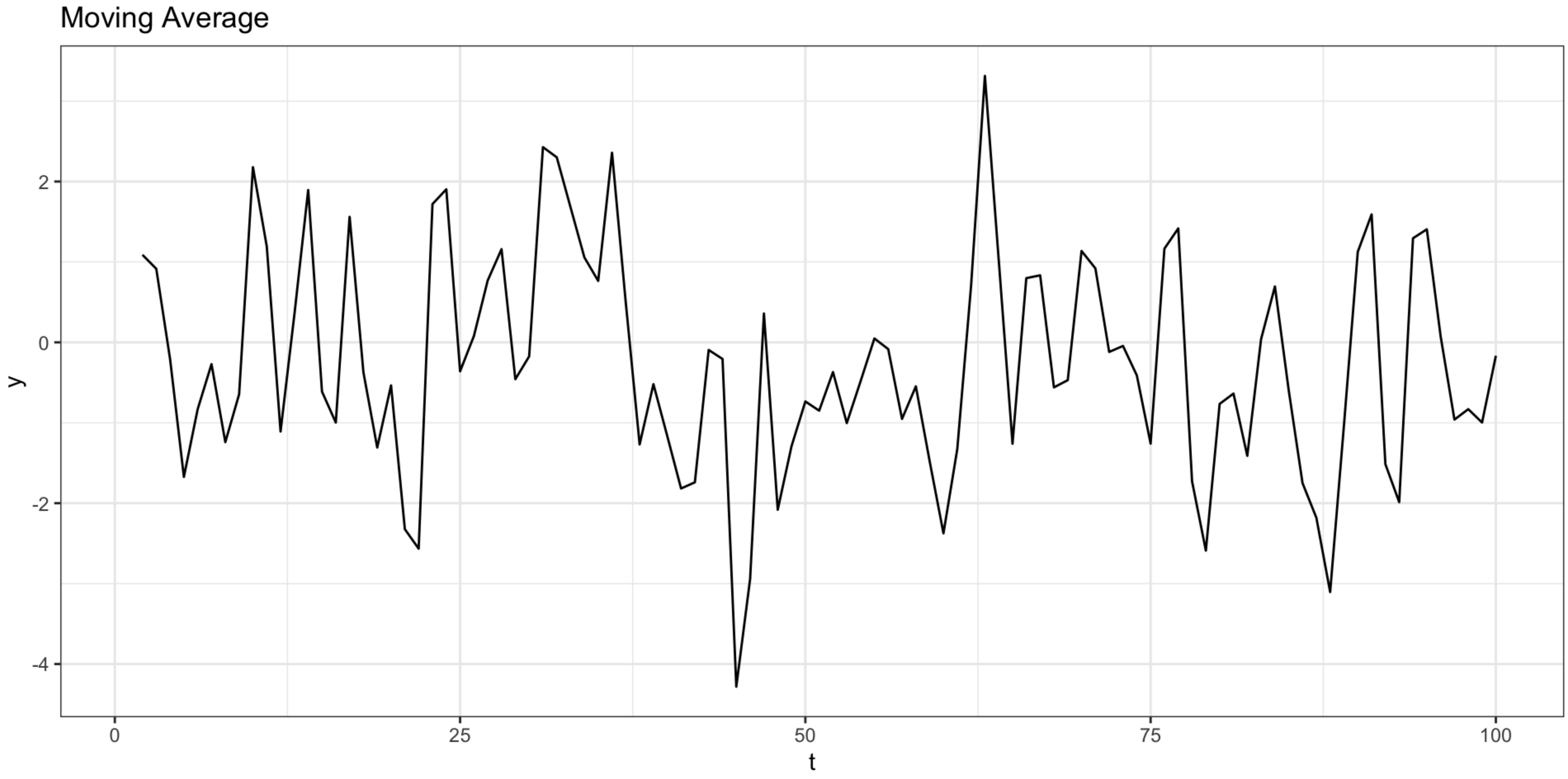
1.00
0.75
0.50
0.25
0.00
PACF
0
10
20
30
40
50
Lag

Stationary?

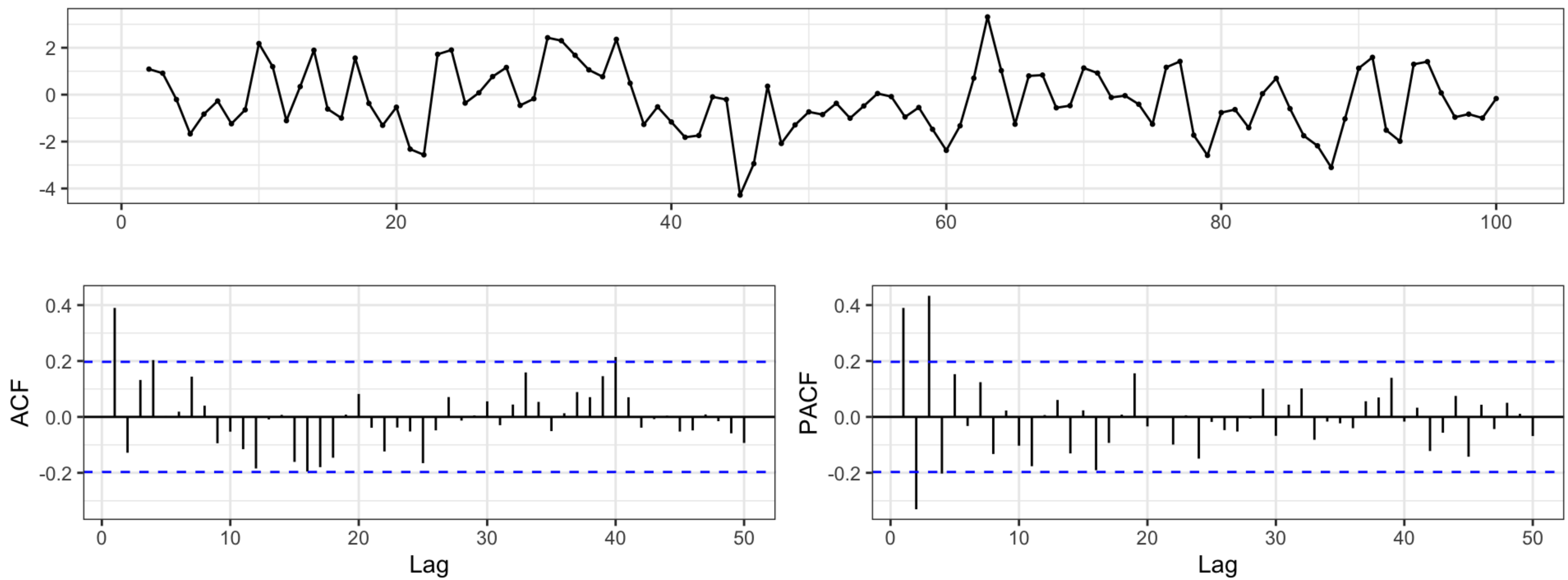
Is y_t stationary?

Example - Moving Average

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



ACF + PACF



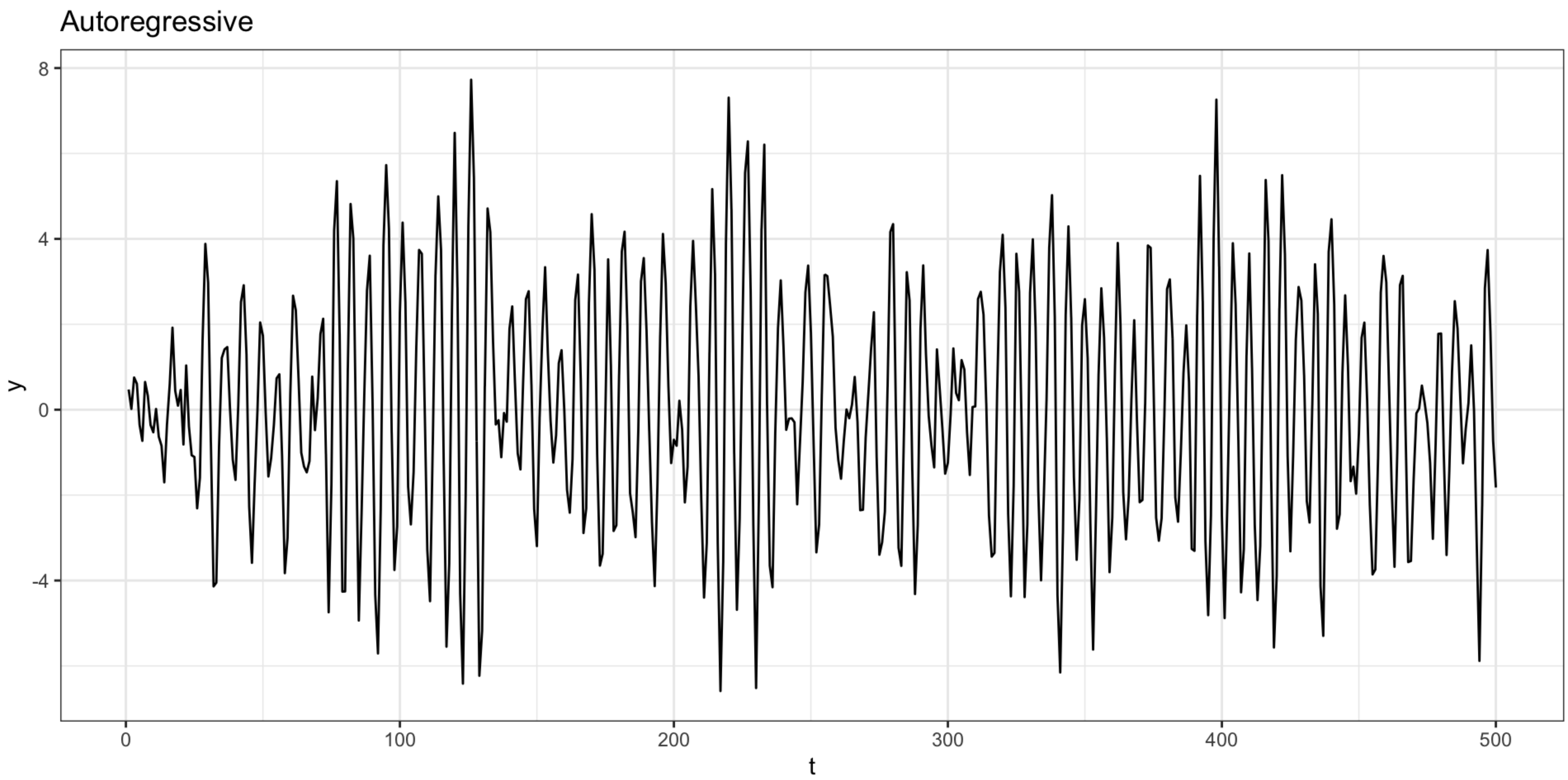
2
0
-2
-4
0
20
40
60
80
100
ACF
0.4
0.2
0.0
-0.2
0
10
20
30
40
50
Lag
PACF
0.4
0.2
0.0
-0.2
0
10
20
30
40
50
Lag

Stationary?

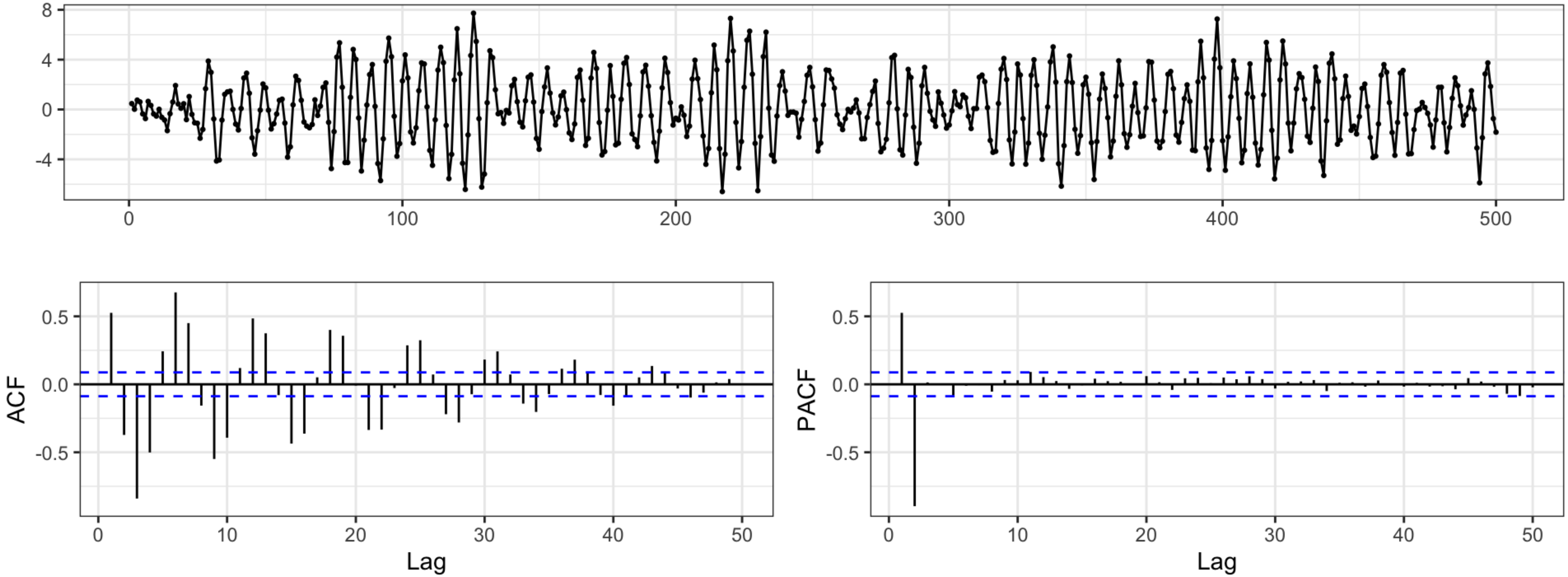
Is y_t stationary?

Autoregressive

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



ACF + PACF



8
4
0
-4
0
100
200
300
400
500
0.5
0.0
-0.5
ACF
0
10
20
30
40
50
Lag
0.5
0.0
-0.5
PACF
0
10
20
30
40
50
Lag

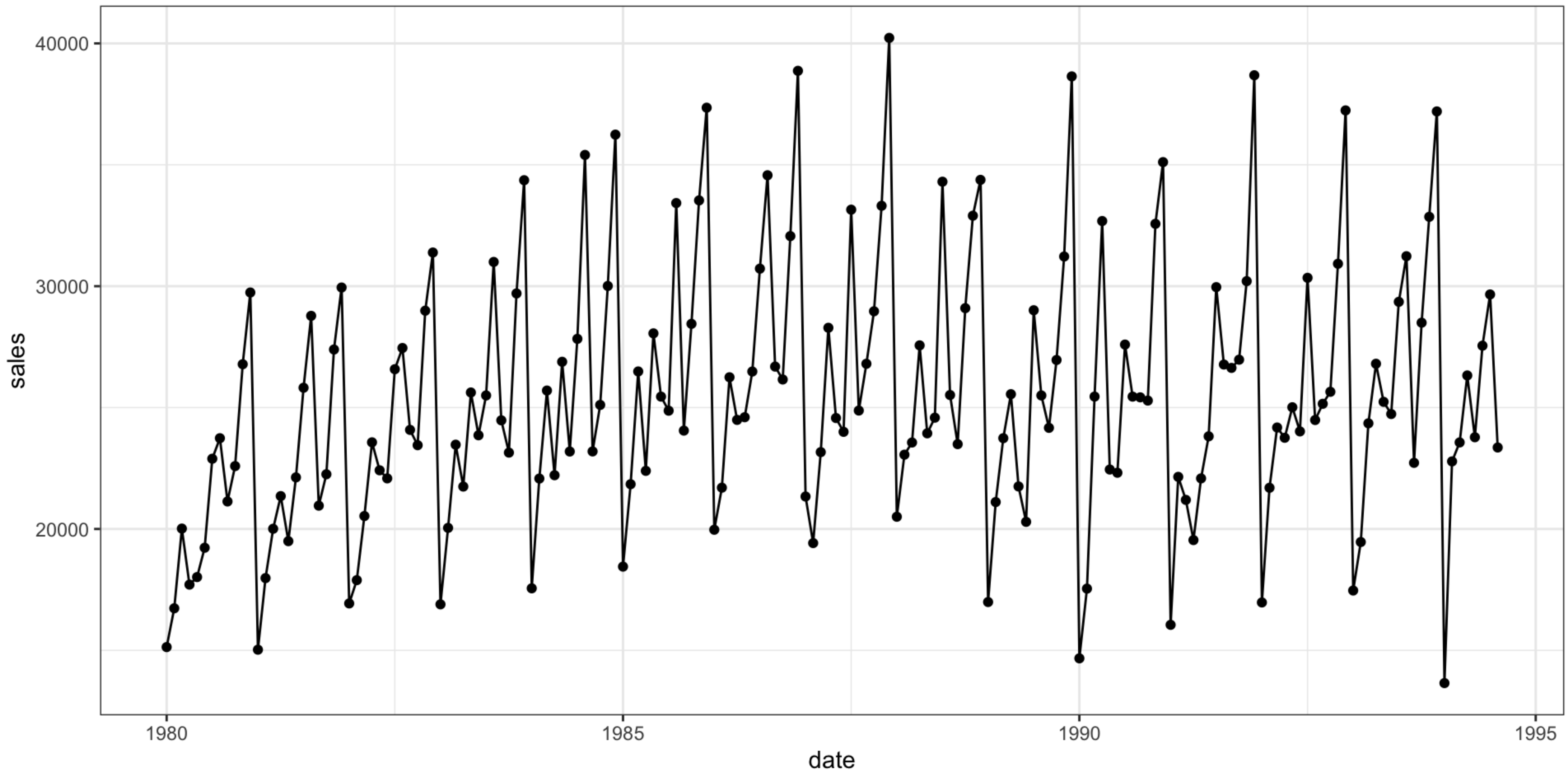
Example - Australian Wine Sales

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

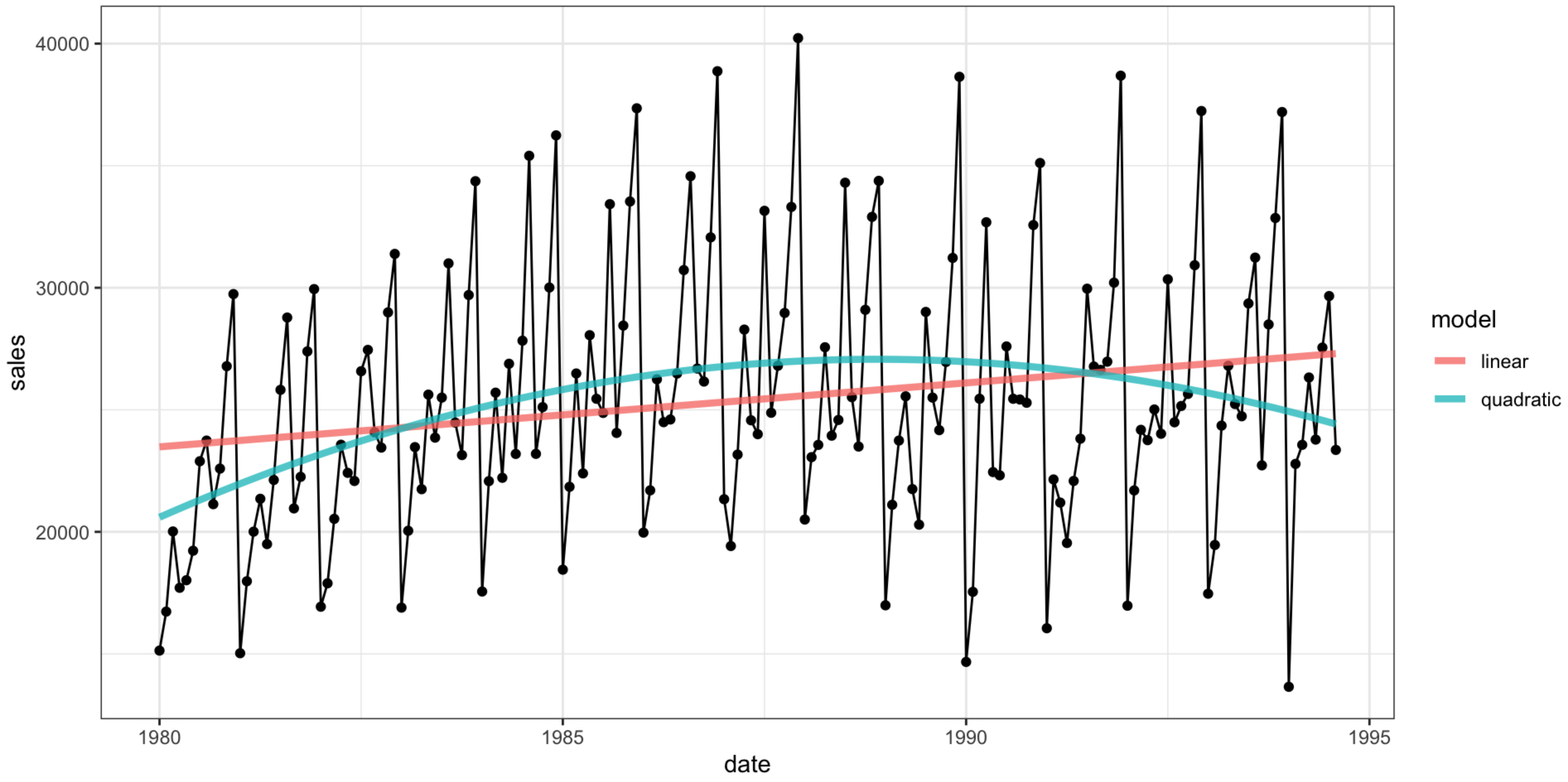
```r
aus_wine = readRDS("data/aus_wine.rds")
aus_wine
```

```
# A tibble: 176 × 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
```

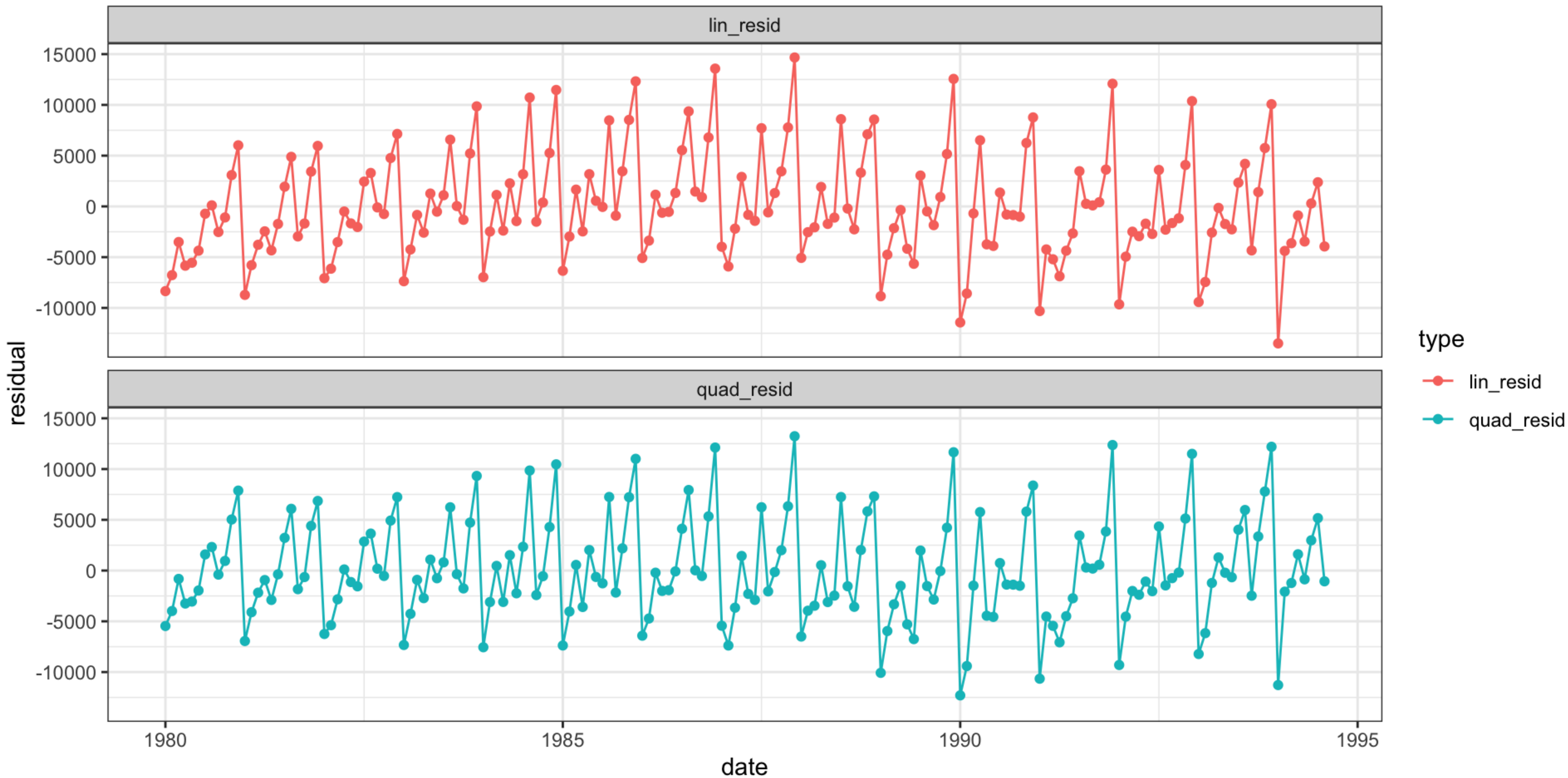
Time series



Basic Model Fit

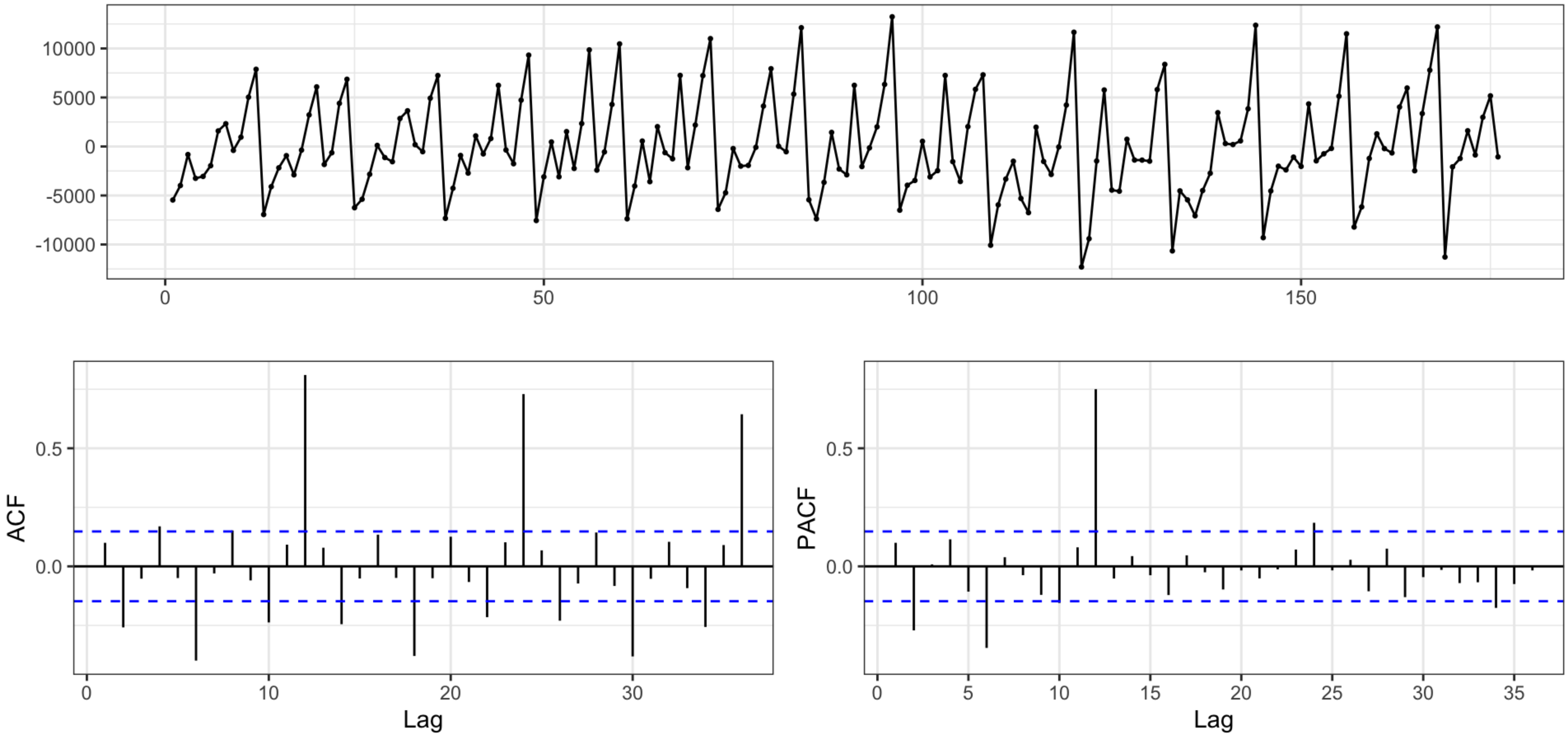


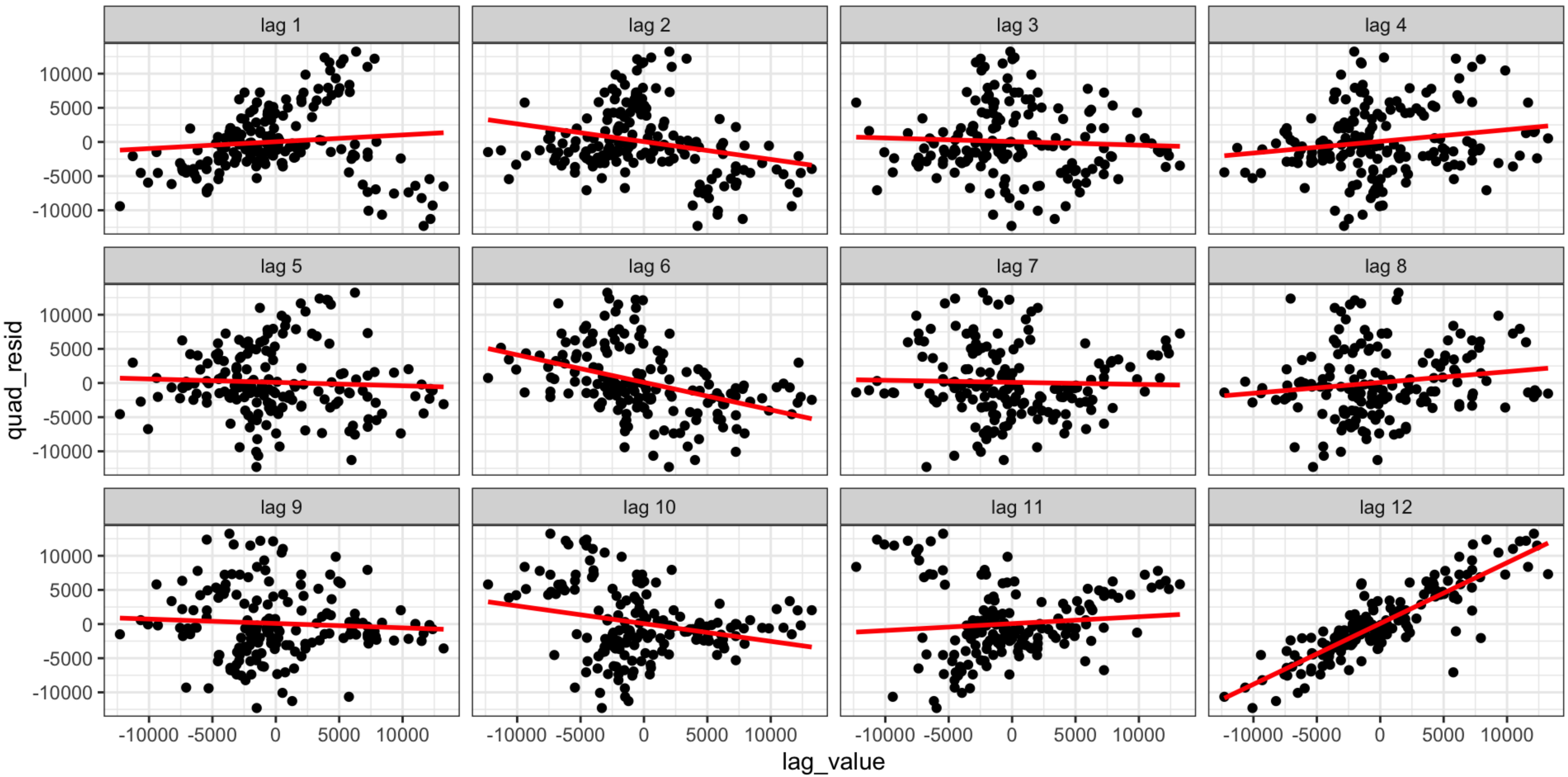
Residuals



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

Autocorrelation Plot



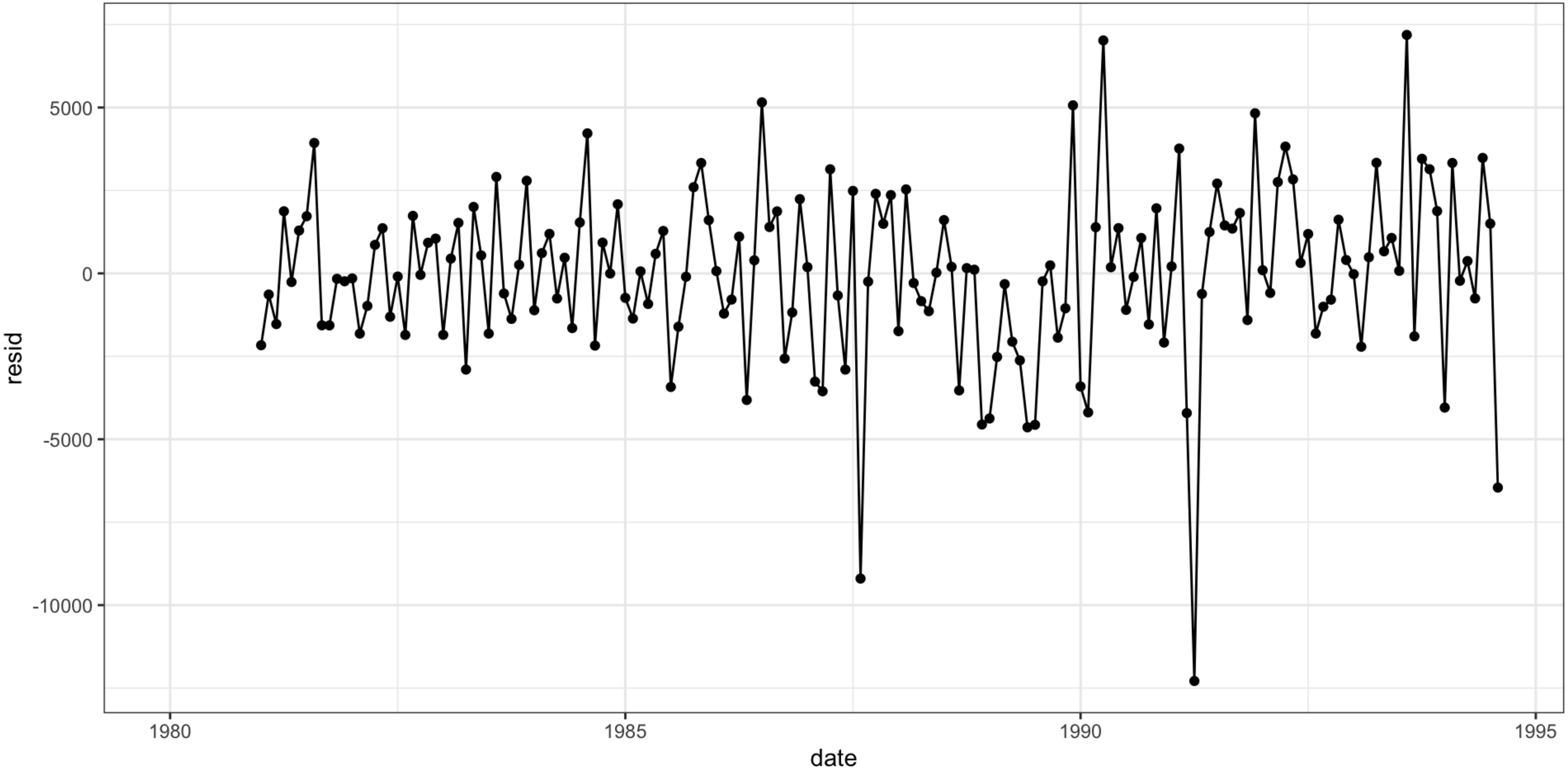


lag 1
lag 2
lag 3
lag 4
lag 5
lag 6
lag 7
lag 8
lag 9
lag 10
lag 11
lag 12
quad_resid
10000
5000
0
-5000
-10000
-10000 -5000 0 5000 10000
lag_value

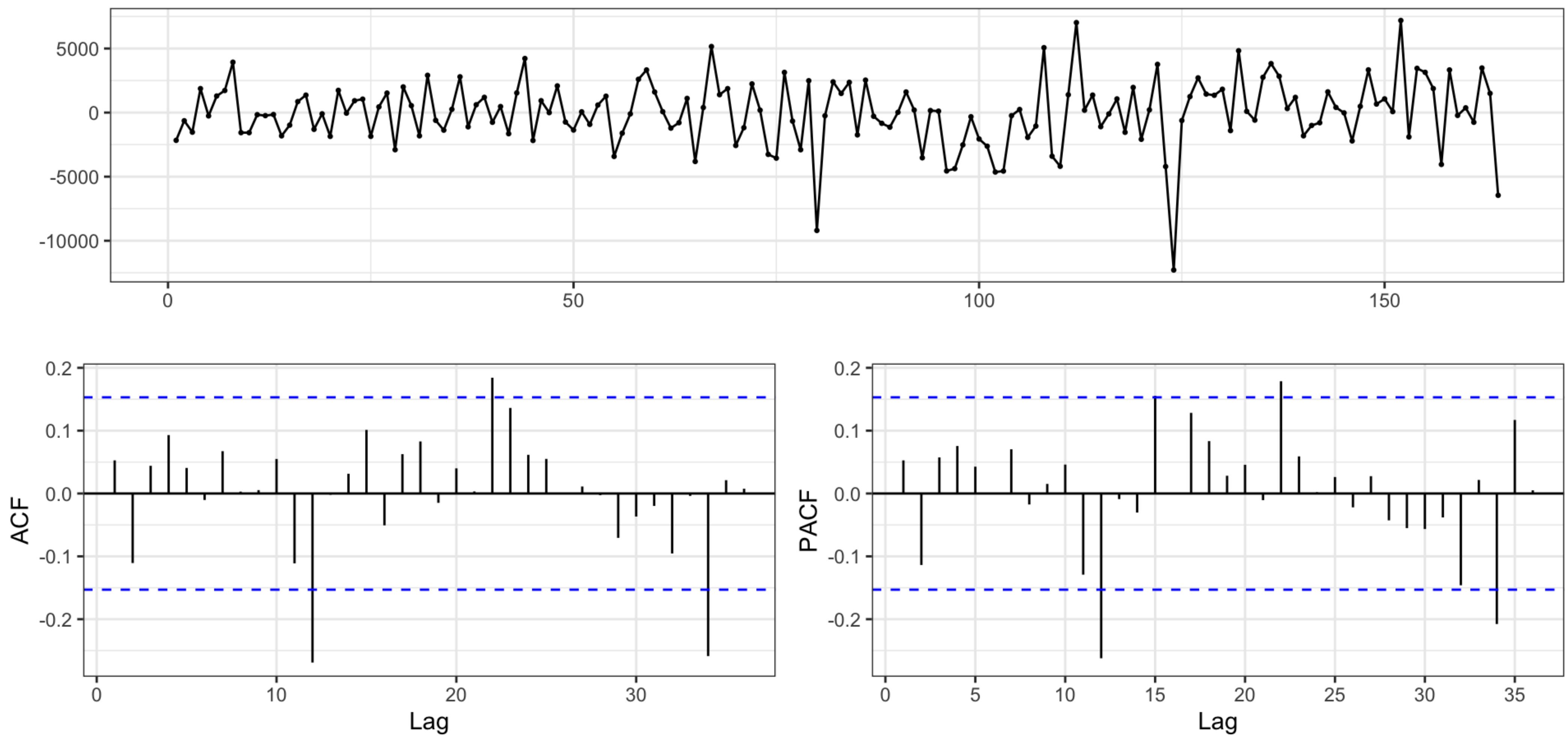
Auto regressive errors

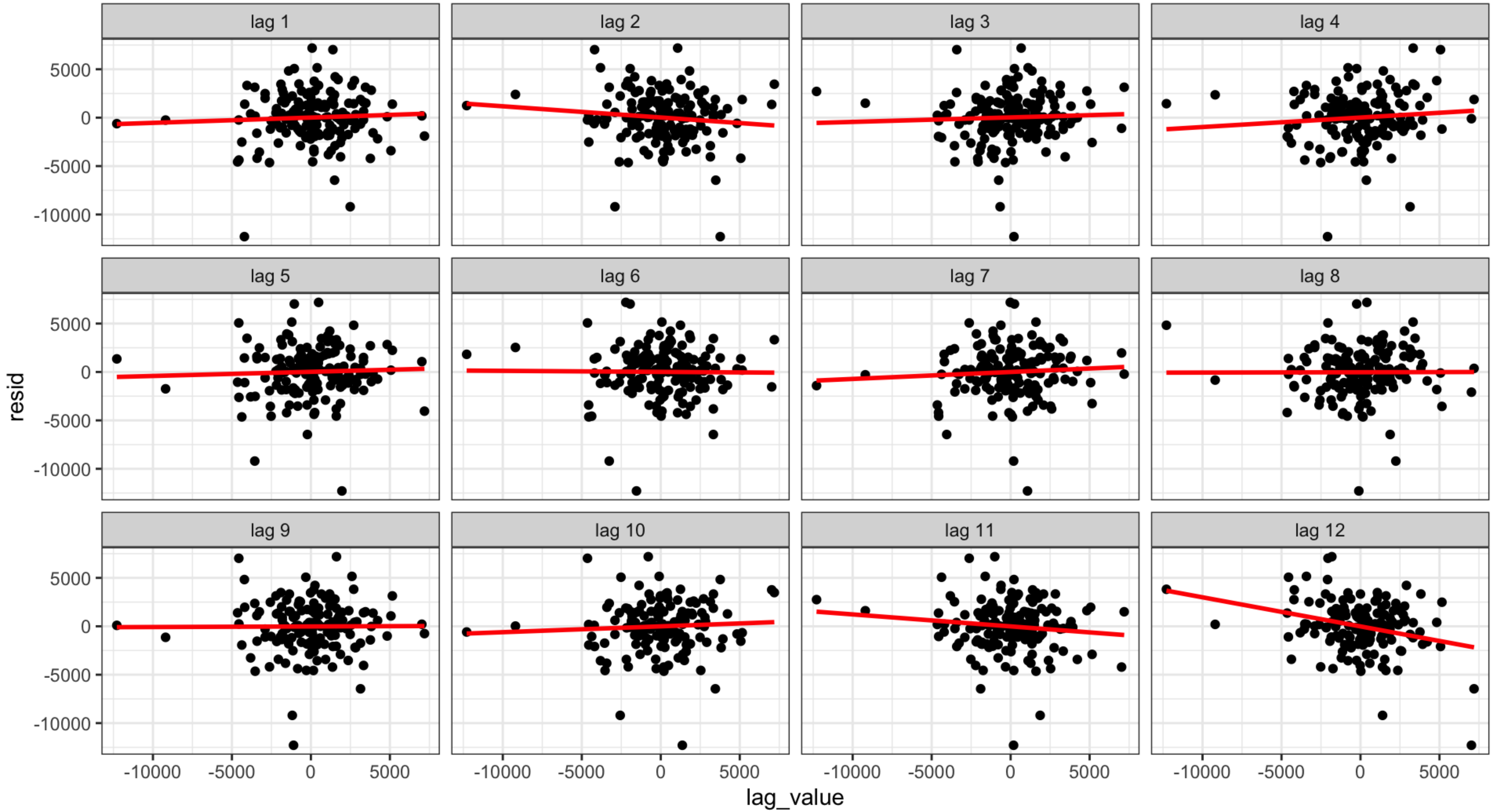
```
Call:
lm(formula = quad_resid ~ lag_12, data = d_ar)

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 ***
```

Residual residuals



Residual residuals - acf




lag 1
lag 2
lag 3
lag 4
lag 5
lag 6
lag 7
lag 8
lag 9
lag 10
lag 11
lag 12
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?

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

the model we actually fit is,

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

where

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