Model Specification:

Sample Partial Autocorrelation Function

Week VIII: Video 24

STAT 485/685, Fall 2020, SFU

Sonja Isberg

Video 24 Learning Objectives

By the end of this video, we should be able to:

- Define the partial autocorrelation function ϕ_{kk}
- Describe some important properties of ϕ_{kk} for the models we are familiar with
- Calculate the sample partial autocorrelation function $\hat{\phi}_{kk}$ for a given process, at any lag k
- Use a plot of $\hat{\phi}_{kk}$ vs. k to determine whether or not a dataset appears to come from an AR process, and give an estimate of the order p

Partial Autocorrelation Function: Motivation

MA(q) models have an autocorrelation function (ACF), ρ_k, that cuts off after lag $k = q$.

Therefore, a useful tool for identifying MA processes is the sample autocorrelation function (sample ACF), r_k.

- In Video 23, we learned how to use the sample ACF to identify an MA process, and how to identify a possible order q.

Partial Autocorrelation Function: Motivation

MA(q) models have an autocorrelation function (ACF), ρ_k, that cuts off after lag $k = q$.

Therefore, a useful tool for identifying MA processes is the sample autocorrelation function (sample ACF), r_k.

- In Video 23, we learned how to use the sample ACF to identify an MA process, and how to identify a possible order q.

However, AR processes are harder to identify using the ACF.

Instead, we need a different function: the *partial* autocorrelation function (PACF), and its counterpart – the sample partial autocorrelation function (sample PACF).

Partial Autocorrelation Function: Background

The **partial correlation** between any two random variables X and Y is a measure of their degree of association, with the effect of a set of other random variables removed.

This helps to control for the effect of any potential *confounding variables* in the dataset.

Partial Autocorrelation Function: Background

The **partial correlation** between any two random variables X and Y is a measure of their degree of association, with the effect of a set of other random variables removed.

This helps to control for the effect of any potential *confounding variables* in the dataset.

Example: How does education affect income?

There is a positive correlation between degree of education and an individual's income.

However, individuals from wealthy families might be more likely to acquire both more education and more income.

So, if we wish to get a more accurate representation of the relationship between education and income, we need to control for wealth.

Partial Autocorrelation Function: Background (cont'd)

The partial correlation lies on $[-1, 1]$:

- Close to 1: positive linear relationship, controlling for the other variables
- Close to -1: negative linear relationship, controlling for the other variables
- Close to 0: no linear relationship, once the other variables are accounted for

Partial Autocorrelation Function: Background (cont'd)

The partial correlation lies on $[-1, 1]$:

- Close to 1: positive linear relationship, controlling for the other variables
- Close to -1: negative linear relationship, controlling for the other variables
- Close to 0: no linear relationship, once the other variables are accounted for

Partial correlation between X and Y given a set of controlling variables $\boldsymbol{Z} = \{Z_1, \ldots, Z_m\}$:

$$\rho_{XY \cdot \boldsymbol{Z}} = Corr(e_X, e_Y)$$

where

$$e_X = \text{residuals resulting from fitting a linear regression of } X \text{ with } \boldsymbol{Z}$$

$$e_Y = \text{residuals resulting from fitting a linear regression of } Y \text{ with } \boldsymbol{Z}$$

Partial Autocorrelation Function: Definition

The **partial autocorrelation function (PACF)** for a process $\{Y_t\}$ is defined as the partial correlation between any Y_t and Y_{t-k}, controlling for all the observations between them:

$$\begin{aligned}\phi_{kk} = Corr(&Y_t - \beta_1 Y_{t-1} - \beta_2 Y_{t-2} - \cdots - \beta_{k-1} Y_{t-k+1},\\ &Y_{t-k} - \beta_1 Y_{t-k+1} - \beta_2 Y_{t-k+2} - \cdots - \beta_{k-1} Y_{t-1})\\ = Corr(&\text{residuals from a linear regression of } Y_t \text{ with } Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1},\\ &\text{residuals from a linear regression of } Y_{t-k} \text{ with } Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})\end{aligned}$$

where $\phi_{11} = \rho_1$ (typo in the textbook on pg. 113!).

In other words, it is just the autocorrelation between Y_t and Y_{t-k} that is not accounted for by the observations in between them.

Partial Autocorrelation Function: Definition

The **partial autocorrelation function (PACF)** for a process $\{Y_t\}$ is defined as the partial correlation between any Y_t and Y_{t-k}, controlling for all the observations between them:

$$\begin{aligned}\phi_{kk} = Corr(&Y_t - \beta_1 Y_{t-1} - \beta_2 Y_{t-2} - \cdots - \beta_{k-1} Y_{t-k+1},\\ &Y_{t-k} - \beta_1 Y_{t-k+1} - \beta_2 Y_{t-k+2} - \cdots - \beta_{k-1} Y_{t-1})\\ = Corr(&\text{residuals from a linear regression of } Y_t \text{ with } Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1},\\ &\text{residuals from a linear regression of } Y_{t-k} \text{ with } Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})\end{aligned}$$

where $\phi_{11} = \rho_1$ (typo in the textbook on pg. 113!).

In other words, it is just the autocorrelation between Y_t and Y_{t-k} that is not accounted for by the observations in between them.

If $\{Y_t\}$ is normally distributed, then this is equivalent to the conditional autocorrelation:

$$\phi_{kk} = Corr(Y_t, Y_{t-k} | Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})$$

Partial Autocorrelation Function: AR(1)

In general, for any stationary process $\{Y_t\}$, it can be shown that

$$\phi_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

For an AR(1) process, recall that $\rho_k = \phi^k$, and so

$$\phi_{22} = \frac{\phi^2 - (\phi)^2}{1 - (\phi)^2} = 0$$

Actually, for an AR(1) process: $\phi_{kk} = 0$ for all $k > 1$!

Partial Autocorrelation Function: AR(p)

We won't derive the expressions for ϕ_{kk} at lags $k \leq p$.

However, for $k > p$:

We'll learn in Ch. 9 that the best linear predictor of Y_t using the intermediary variables $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}$ is just

$$\hat{Y}_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$$

In other words, the regression cuts off after Y_{t-p} (it doesn't go all the way out to Y_{t-k+1}), and the β-coefficients are just the AR parameters!

As for the best linear predictor of Y_{t-k} using the intermediary variables $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}$, we'll just call it some unknown function $h(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})$.

Partial Autocorrelation Function: AR(p) (cont'd)

The best linear predictor of Y_t using the intermediary variables $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}$ is just

$$\hat{Y}_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p}$$

The best linear predictor of Y_{t-k} using the intermediary variables $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}$ is some unknown function $h(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})$.

Now:

$$\begin{aligned}
\phi_{kk} = Corr(&\text{residuals from a linear regression of } Y_t \text{ with } Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}, \\
&\text{residuals from a linear regression of } Y_{t-k} \text{ with } Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1}) \\
= Corr(&Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p}, \\
&Y_{t-k} - h(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})) \\
= Corr(&e_t, Y_{t-k} - h(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k+1})) \\
= 0 \qquad &(\text{since } e_t \perp\!\!\!\perp Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k})
\end{aligned}$$

Partial Autocorrelation Function: AR(p) (cont'd)

Therefore, for an AR(p) process,

$$\phi_{kk} = 0 \qquad \text{for all } k > p$$

This is analogous to the way in which the autocorrelation function cuts off after lag q for an MA(q) process!

This will be very useful to us for identifying AR processes and their orders.

Partial Autocorrelation Function: MA Processes

Expressions for the partial autocorrelation function ϕ_{kk} can also be derived for MA processes, but since they don't cut off after a certain lag they will not be as useful to us.

For an MA(1) process:

$$\phi_{11} = \rho_1 = -\frac{\theta}{1+\theta^2}$$

$$\phi_{kk} = -\frac{\theta^k(1-\theta^2)}{1-\theta^{2(k+1)}}$$

Notice that the PACF never cuts off for an MA(1) process, but instead decays exponentially.

For an MA(q) process: It can be shown that the PACF behaves very much like the ACF of an AR(q) process!

Partial Autocorrelation Function: General Formulation

General expressions can be written out for obtaining ϕ_{kk} at any lag k for any stationary process.

These are just Yule-Walker equations, and they are shown on pg. 114.

It can be shown that the Yule-Walker equations can be solved recursively using the expression

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

where

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

and $\phi_{11} = \rho_1$, as before.

Partial Autocorrelation Function: General Formulation (cont'd)

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

Example: $\phi_{11} = \rho_1$

Then:

$$\phi_{22} = \frac{\rho_2 - \sum_{j=1}^{1} \phi_{1,j}\rho_{2-j}}{1 - \sum_{j=1}^{1} \phi_{1,j}\rho_j} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

Also:

$$\phi_{21} = \phi_{11} - \phi_{22}\phi_{11}$$

Partial Autocorrelation Function: General Formulation (cont'd)

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}\phi_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

Example: $\phi_{11} = \rho_1$

Then:

$$\phi_{22} = \frac{\rho_2 - \sum_{j=1}^{1} \phi_{1,j}\rho_{2-j}}{1 - \sum_{j=1}^{1} \phi_{1,j}\rho_j} = \frac{\rho_2 - \phi_{11}\rho_1}{1 - \phi_{11}\rho_1} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

Also:

$$\phi_{21} = \phi_{11} - \phi_{22}\phi_{11}$$

Then:

$$\phi_{33} = \frac{\rho_3 - \sum_{j=1}^{2} \phi_{2,j}\rho_{3-j}}{1 - \sum_{j=1}^{2} \phi_{2,j}\rho_j} = \frac{\rho_3 - \phi_{21}\rho_2 - \phi_{22}\rho_1}{1 - \phi_{21}\rho_1 - \phi_{22}\rho_2} = \ldots$$

$\vdots$

Sample Partial Autocorrelation Function

For a time series $\{Y_1, Y_2, \ldots, Y_n\}$, we can estimate its partial autocorrelation function ϕ_{kk} using the **sample partial autocorrelation function (sample PACF)** $\hat{\phi}_{kk}$.

Values of $\hat{\phi}_{kk}$ at any k can be obtained using the same general formulation as before, except with ρ_k's replaced by r_k's:

$$\hat{\phi}_{11} = r_1 \quad \text{and}$$

$$\hat{\phi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j} \quad \text{where}$$

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

Sample Partial Autocorrelation Function: Example

$$\hat{\phi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j} \quad \text{where}$$

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

Example: $\hat{\phi}_{11} = r_1$

Then:

$$\hat{\phi}_{22} = \frac{r_2 - \sum_{j=1}^{1} \hat{\phi}_{1,j} r_{2-j}}{1 - \sum_{j=1}^{1} \hat{\phi}_{1,j} r_j} = \frac{r_2 - \hat{\phi}_{11} r_1}{1 - \hat{\phi}_{11} r_1} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

Also:

$$\hat{\phi}_{21} = \hat{\phi}_{11} - \hat{\phi}_{22} \hat{\phi}_{11}$$

Sample Partial Autocorrelation Function: Example

$$\hat{\phi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j} \quad \text{where}$$

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j} \quad \text{for } j = 1, 2, \ldots, k-1$$

Example: $\hat{\phi}_{11} = r_1$

Then:

$$\hat{\phi}_{22} = \frac{r_2 - \sum_{j=1}^{1} \hat{\phi}_{1,j} r_{2-j}}{1 - \sum_{j=1}^{1} \hat{\phi}_{1,j} r_j} = \frac{r_2 - \hat{\phi}_{11} r_1}{1 - \hat{\phi}_{11} r_1} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

Also:

$$\hat{\phi}_{21} = \hat{\phi}_{11} - \hat{\phi}_{22}\hat{\phi}_{11}$$

Then:

$$\hat{\phi}_{33} = \frac{r_3 - \sum_{j=1}^{2} \hat{\phi}_{2,j} r_{3-j}}{1 - \sum_{j=1}^{2} \hat{\phi}_{2,j} r_j} = \frac{r_3 - \hat{\phi}_{21} r_2 - \hat{\phi}_{22} r_1}{1 - \hat{\phi}_{21} r_1 - \hat{\phi}_{22} r_2} = \ldots$$

$\vdots$

Distribution of the Sample Partial ACF

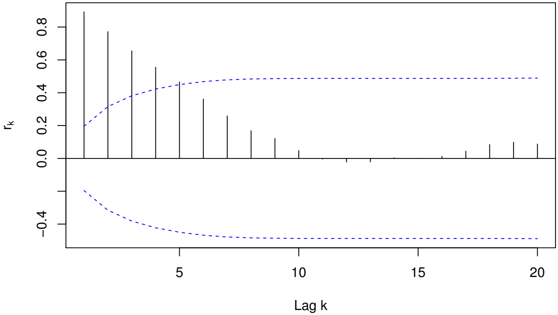
If $\{Y_t\}$ is an AR(p) process:

$$\hat{\phi}_{kk} \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{n}\right) \quad \text{for } k > p$$

So, for testing the null hypothesis that an AR(p) model is correct, we can construct our critical limits (i.e., "dashed lines") at any $k > p$ using

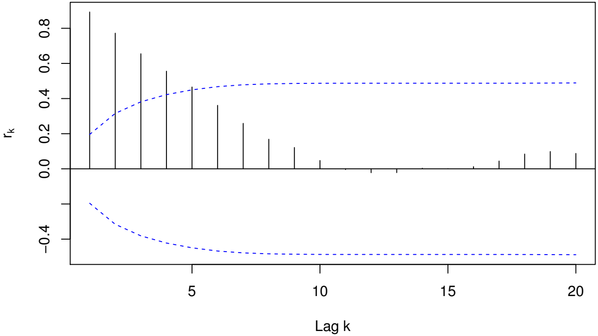
$$\pm 2\, SE(\hat{\phi}_{kk}) = \pm \frac{2}{\sqrt{n}}$$

Example 1: AR(1) Process




```
> e.vec <- rnorm(n=100, mean=0, sd=1)
> y.vec <- rep(NA, times=100)
> y.vec[1] <- 0
> for (t in 2:100)
  {
     y.vec[t] <- 0.9*y.vec[t-1] + e.vec[t]
  }
> y.vec.ts <- ts(data=y.vec)
> acf(y.vec.ts, ci.type='ma', ylab=expression(r[k]))
```

Example 1: AR(1) Process (cont'd)

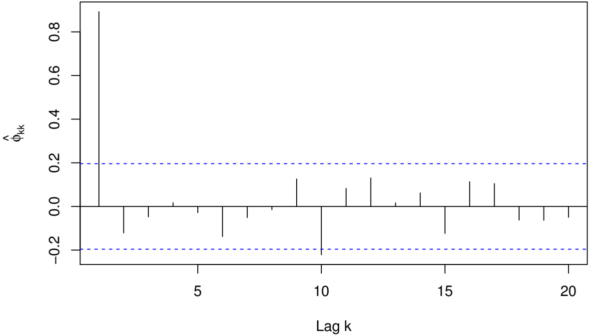


Sample ACF plot:

At each lag k, we are testing for whether or not the r_k cuts off after this lag (as it would if this were an MA($k-1$) process).

As expected, this does not match an MA process, except perhaps for $q > 5$.

Example 1: AR(1) Process (cont'd)

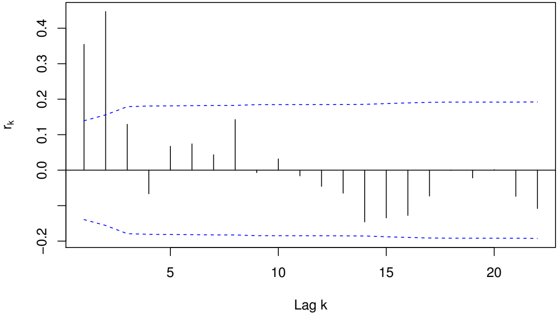



```
> pacf(y.vec.ts, ylab=expression(hat(phi)[kk]))
```

The sample PACF plot gives a much clearer picture about the underlying process.

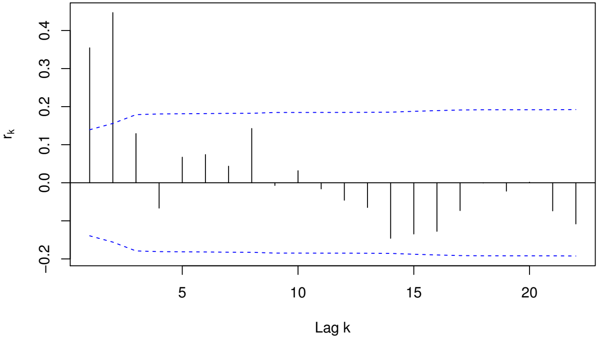
This plot certainly appears to support the possibility of an AR(1) model.

Example 2: MA(2) Process




```
> e.vec <- rnorm(n=200, mean=0, sd=1)
> y.vec <- rep(NA, times=200)
> for (t in 3:200)
  {
     y.vec[t] <- e.vec[t] + 0.2*e.vec[t-1] + 0.9*e.vec[t-2]
  }
> y.vec.ts <- ts(data=y.vec[-(1:2)])
> acf(y.vec.ts, ci.type='ma', ylab=expression(r[k]))
```

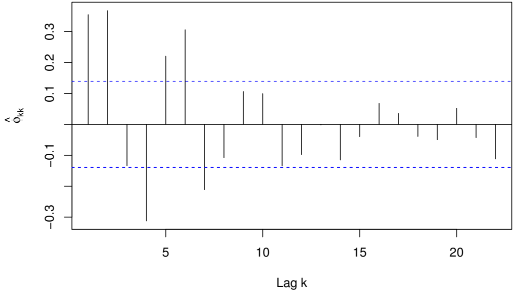
Example 2: MA(2) Process (cont'd)



The sample ACF plot appears to support the hypothesis that this data comes from an MA(2) process.

There does appear to be some sinusoidal behaviour in the plot, but we are not sure if this is due to some underlying AR behaviour, or if it is just due to sampling error.

Example 2: MA(2) Process (cont'd)



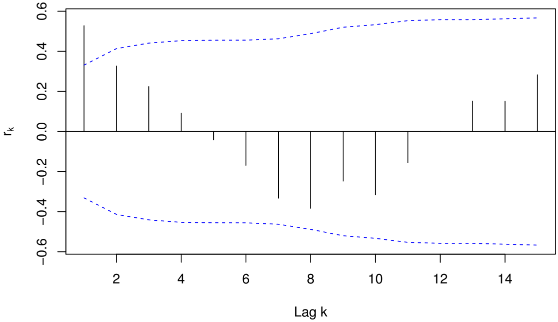

```
> pacf(y.vec.ts, ylab=expression(hat(phi)[kk]))
```

The sample PACF plot doesn't appear to give a clearer picture than the sample ACF plot, for this dataset.

One could potentially argue that this might look like an AR(7) process. However, due to the *principle of parsimony*, we would be more inclined to accept the much simpler MA(2) model.

Example 3: Color Dataset

The `color` dataset in the `TSA` package gives the color property from 35 consecutive batches in an industrial process. Its sample ACF is:

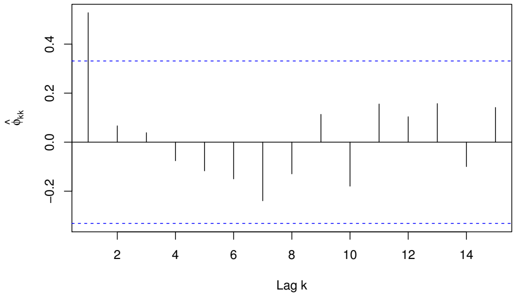



```
> data(color)
> acf(color, ci.type='ma', ylab=expression(r[k]))
```

The sample ACF plot suggests that an MA(1) model may potentially be appropriate here.

Example 3: Color Dataset (cont'd)

Let's create the sample PACF plot:




```
> pacf(color, ylab=expression(hat(phi)[kk]))
```

This plot suggests that perhaps an AR(1) model be appropriate instead!

As we can see, the "right" choice of model is not always clear. We choose one, and then review this choice in our *model diagnostics* step. If needed, we go back and try a different model.

Final Comments

That's all for now!

In this video, we've learned about the partial autocorrelation function ϕ_{kk}, and about some of its important properties for the models we are familiar with.

We also learned how to calculate the sample PACF, $\hat{\phi}_{kk}$, at any lag k.

Finally, we looked at a few examples of how plots of the sample PACF can help us determine whether a process appears to be an AR process.

Coming Up Next: The sample extended autocorrelation function (sample EACF).

Thank you!

References:

[1] Cryer, J. D., & Chan, K. S. (2008). *Time series analysis: with applications in R.* Springer Science and Business Media.

[2] Chan, K. S., & Ripley, B. (2020). TSA: Time Series Analysis. R package version 1.2.1.