

Lista II: Q1

Alberson Miranda

2022-11-22

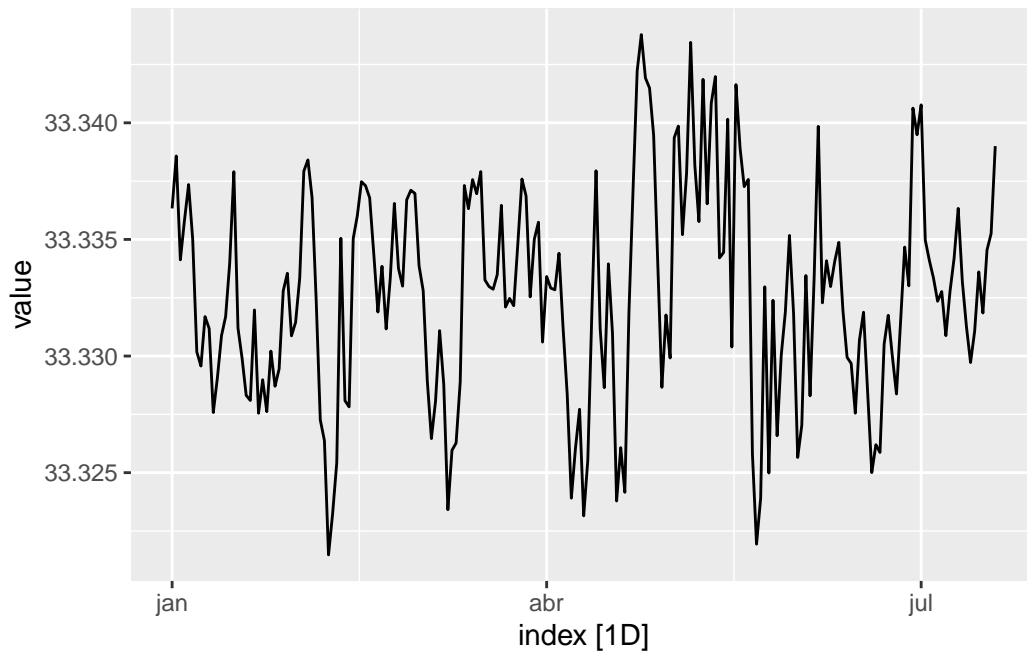
1 MODELAGEM BOX-JENKINS: SÉRIE I

O primeiro passo é a importação e visualização da série. Como não há informação sobre o período, usarei diário e tentarei identificar a partir de um padrão sazonal, se houver.

```
# importando dados
load("data/lista II.RData")
data = data.frame(
  value = conjunto1[, 1],
  index = seq(
    as.Date("2000-01-01"),
    by = 1,
    length.out = length(conjunto1[, 1])
  )
) |> tsibble(index = index)
```

A série é compacta, ou seja, de amplitude baixa, não requerendo transformação para estabilização de variância.

```
# plot série
autoplot(data, .vars = value)
```



O segundo passo é testar se a série é estacionária no primeiro momento. Não há evidências de raiz unitária tanto nos testes quanto nos gráficos de autocorrelação.

```
# KPSS test
data |>
  features(value, unitroot_kpss)
```

```
# A tibble: 1 x 2
  kpss_stat kpss_pvalue
    <dbl>      <dbl>
1    0.108        0.1
```

```
# Phillips-Perron test
data |>
  features(value, unitroot_pp)
```

```
# A tibble: 1 x 2
  pp_stat pp_pvalue
    <dbl>      <dbl>
1   -6.43        0.01
```

```
# Augmented-Dickey-Fuller test
data |>
  (\(x) ur.df(x$value, selectlags = "AIC", type = "drift", lags = 12))() |>
  summary()
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression drift

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.0101264	-0.0022878	-0.0000252	0.0020877	0.0103523

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	11.27905	2.04454	5.517	1.15e-07 ***
z.lag.1	-0.33838	0.06134	-5.517	1.15e-07 ***
z.diff.lag	-0.01060	0.07380	-0.144	0.886

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.003555 on 185 degrees of freedom

Multiple R-squared: 0.1688, Adjusted R-squared: 0.1598

F-statistic: 18.79 on 2 and 185 DF, p-value: 3.73e-08

Value of test-statistic is: -5.5166 15.2347

Critical values for test statistics:

	1pct	5pct	10pct
tau2	-3.46	-2.88	-2.57
phi1	6.52	4.63	3.81

```
data |>
  (\(x) ur.df(x$value, selectlags = "AIC", type = "trend", lags = 12))() |>
  summary()
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression trend

Call:

```
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.0102211	-0.0022799	0.0000216	0.0020829	0.0102631

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.141e+01	2.057e+00	5.549	9.87e-08 ***
z.lag.1	-3.424e-01	6.171e-02	-5.549	9.87e-08 ***
tt	3.313e-06	4.807e-06	0.689	0.492
z.diff.lag	-8.474e-03	7.397e-02	-0.115	0.909

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.00356 on 184 degrees of freedom

Multiple R-squared: 0.171, Adjusted R-squared: 0.1574

F-statistic: 12.65 on 3 and 184 DF, p-value: 1.487e-07

Value of test-statistic is: -5.5495 10.2859 15.411

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-3.99	-3.43	-3.13
phi2	6.22	4.75	4.07
phi3	8.43	6.49	5.47

```
data |>
  (\(x) ur.df(x$value, selectlags = "AIC", type = "none", lags = 12))() |>
  summary()
```

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
```

Test regression none

Call:

```
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-0.0126568	-0.0024009	0.0001073	0.0021428	0.0090807

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
z.lag.1	1.691e-06	8.054e-06	0.210	0.833965
z.diff.lag1	-2.703e-01	7.373e-02	-3.666	0.000325 ***
z.diff.lag2	-2.039e-01	7.549e-02	-2.701	0.007580 **
z.diff.lag3	-1.276e-01	7.622e-02	-1.674	0.095896 .
z.diff.lag4	-2.248e-01	7.510e-02	-2.993	0.003153 **
z.diff.lag5	-2.176e-01	7.532e-02	-2.889	0.004338 **
z.diff.lag6	-1.352e-01	7.648e-02	-1.767	0.078909 .
z.diff.lag7	-1.657e-01	7.501e-02	-2.209	0.028467 *
z.diff.lag8	-1.859e-01	7.320e-02	-2.540	0.011941 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.00368 on 179 degrees of freedom

Multiple R-squared: 0.1383, Adjusted R-squared: 0.095

F-statistic: 3.193 on 9 and 179 DF, p-value: 0.001316

Value of test-statistic is: 0.2099

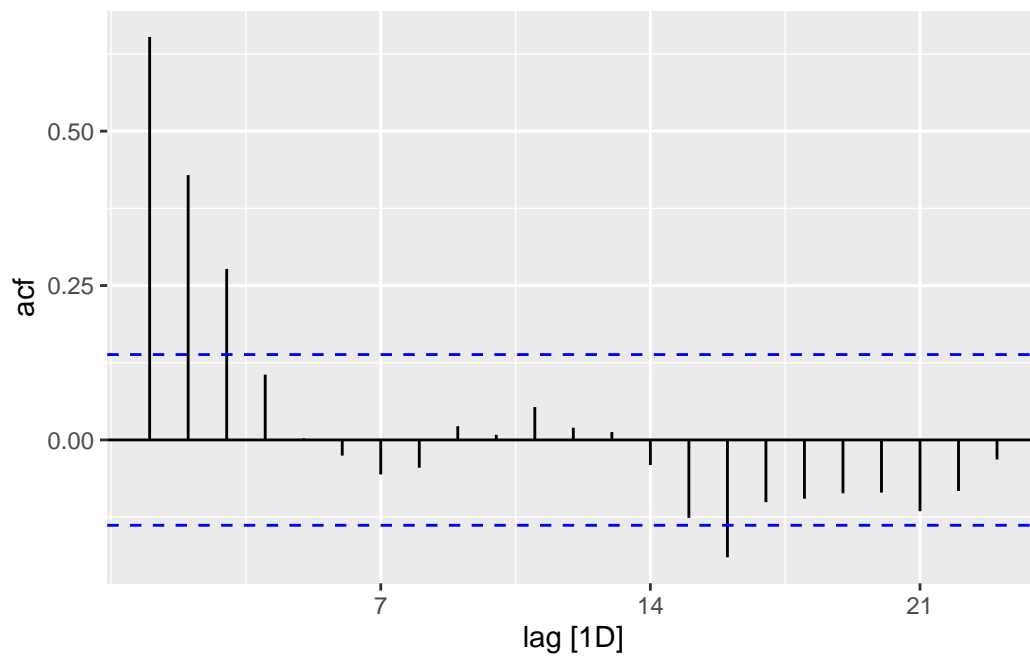
Critical values for test statistics:

1pct 5pct 10pct

```
tau1 -2.58 -1.95 -1.62
```

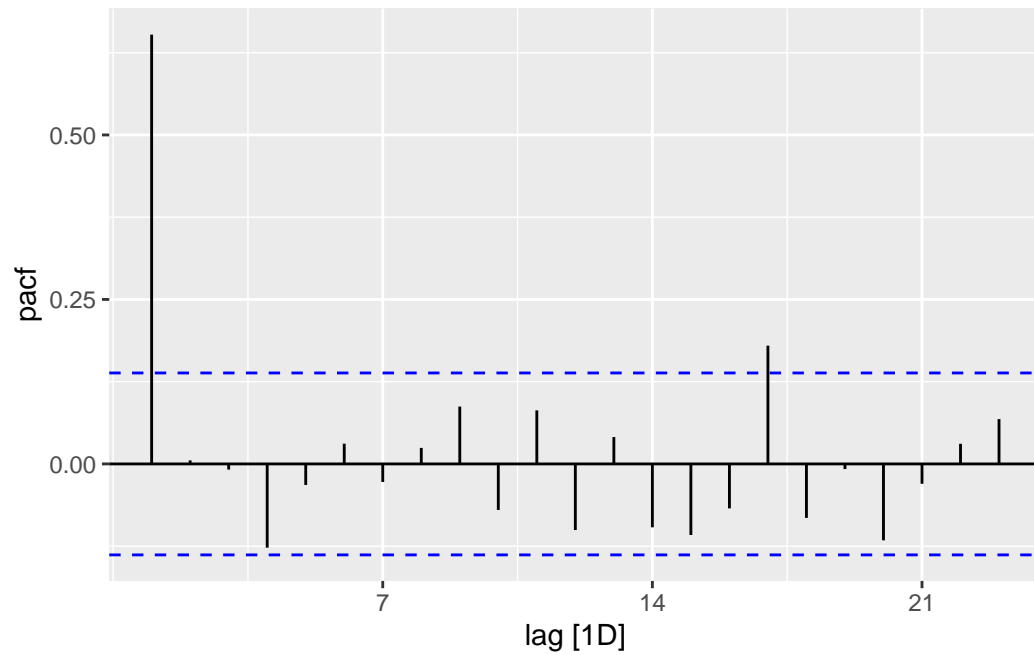
```
# ACF  
data |> ACF() |> autoplot()
```

Response variable not specified, automatically selected `var = value`



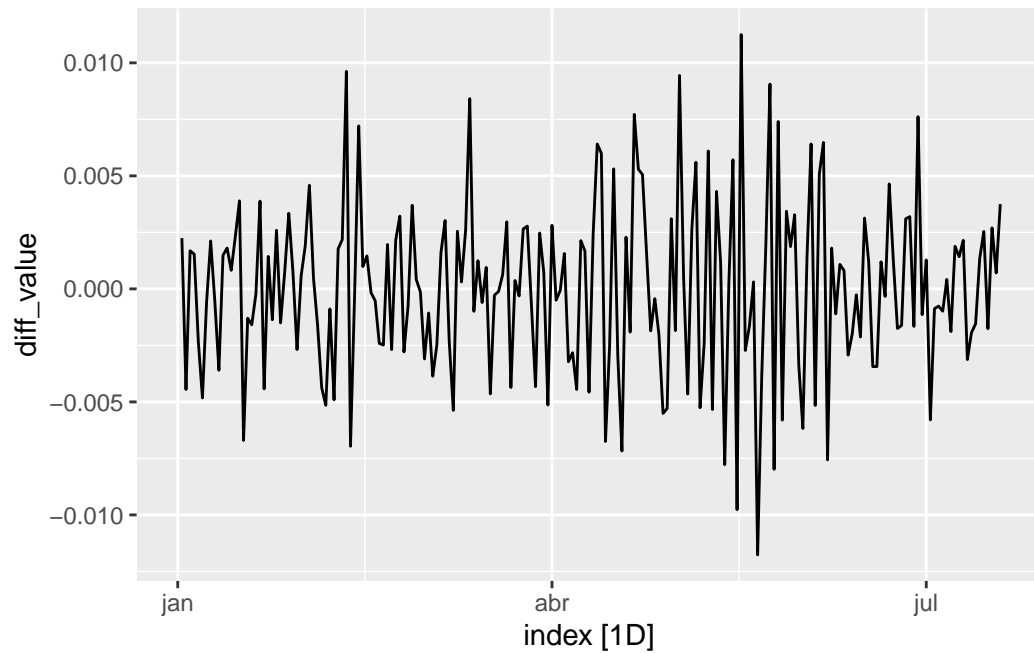
```
# ACF  
data |> PACF() |> autoplot()
```

Response variable not specified, automatically selected `var = value`



```
# diferenciando
data |>
  dplyr::mutate(diff_value = difference(value)) |>
  autoplot(.vars = diff_value)
```

Warning: Removed 1 row containing missing values (`geom_line()`).



A MODELAGEM BOX-JENKINS: SÉRIE II

A série não parece estável. Recomendável transformação para estabilização da variância.

```
lambda = data |>
  features(value, features = guerrero) |>
  (\(x) x[["lambda_guerrero"]])()

data |>
  autoplot(box_cox(value, lambda))
```