Análise de intervenção e previsão de atividade econômica

Lucas Emanuel Resck Domingues* Lucas Machado Moschen†

TODO

(DONE) 1. Fazer o processo de identificação manualmente para obter a primeira opção de modelo (mostrar que aprendemos isso).

(DONE) 2. Usar auto.arima para um segundo modelo. Se eles forem diferentes (o drift faz variar), comparamos os modelos usando checkresiduals. Escolher o "melhor modelo"

(DONE) 3. Fazer previsões do modelo e comparar com pósintervenção. Vamos comparar com a realidade e verificar que de fato houve efeito.

4. Fitar o modelo em todos os dados: propor um modelo para a crise (talvez dois e repetir a escolha do melhor?)

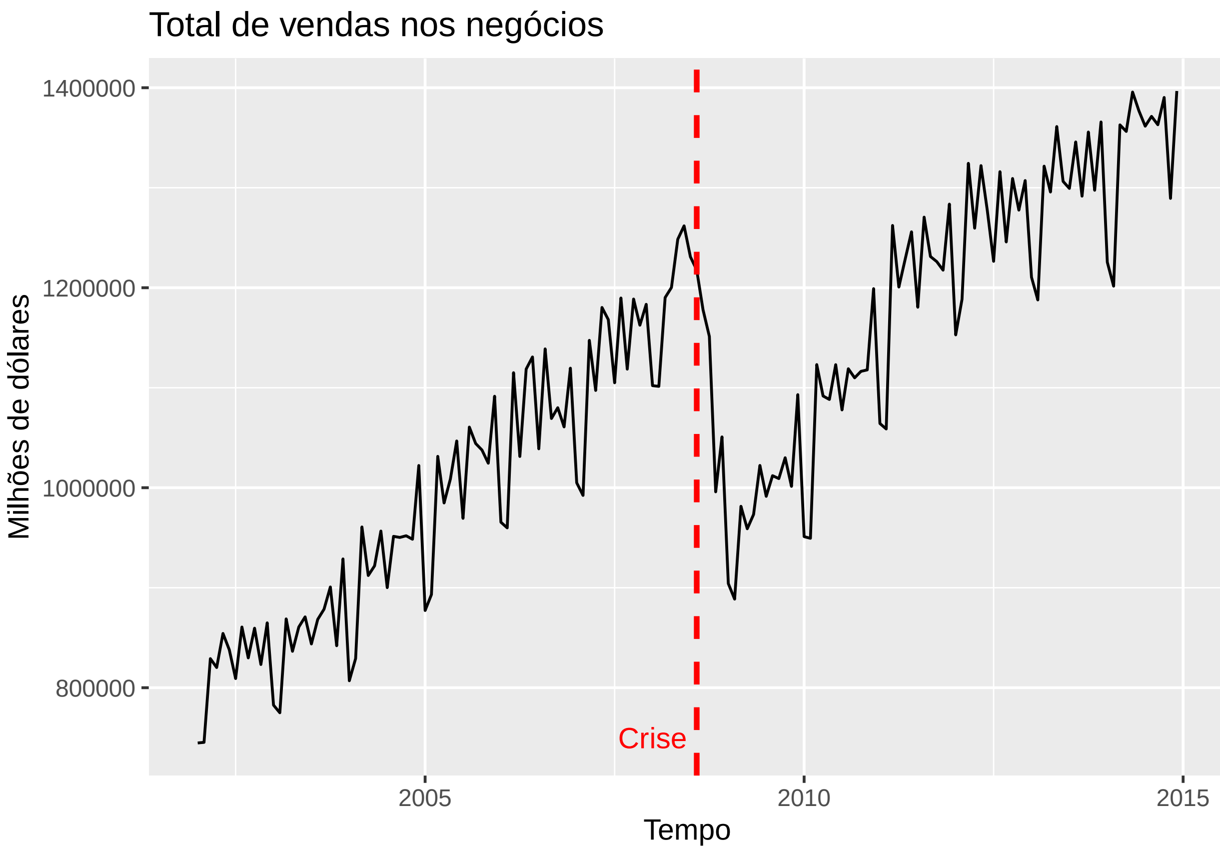
5. Fazemos as previsões e avaliamos o modelo no teste (podemos comparar com um naive).

Obs.: Esses 5 passos combrem a Metodologia, mas podemos adicionar algumas cerejas para deixar mais rico, dependendo do tempo.

Dados

A série temporal a seguir é o total de vendas mensais nos negócios em milhões de dólares, obtido em Fred, Federal Reserve Bank of St. Louis. A janela de observações está entre Janeiro de 2002 e Dezembro de 2014. Os últimos dois anos serão utilizados para a validação do modelo de previsão.

*Escola de Matemática Aplicada
†Escola de Matemática Aplicada



Vamos separar os últimos dois anos para validação do modelo e portanto, não usaremos no treinamento do modelo.

Vamos considerar que houve uma intervenção (a crise do subprime) em julho de 2008.

1. Modelagem pré-intervenção

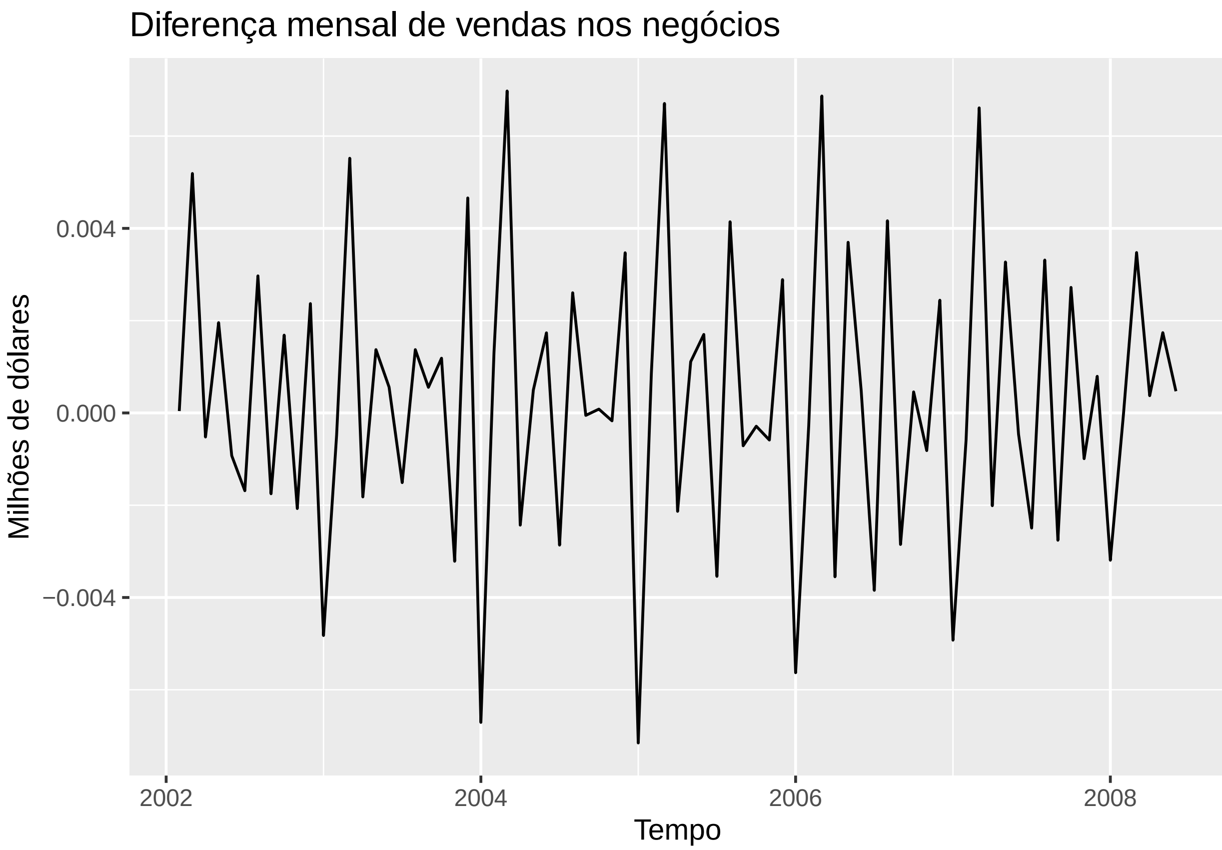
Vamos realizar uma modelagem pré-intervenção, isso é, vamos fazer a modelagem do processo antes da crise. Assim, poderemos verificar que de fato houve efeito pós intervenção.

1. Modelagem manual Box-Jenkins

Vamos seguir a metodologia Box-Jenkins. Primeiro, faremos uma transformação Box-Cox na série, se utilizando do λ que minimiza o coeficiente de variação para subséries da série.

```
## [1] "Obtemos lambda =  -0.222584674825719"
```

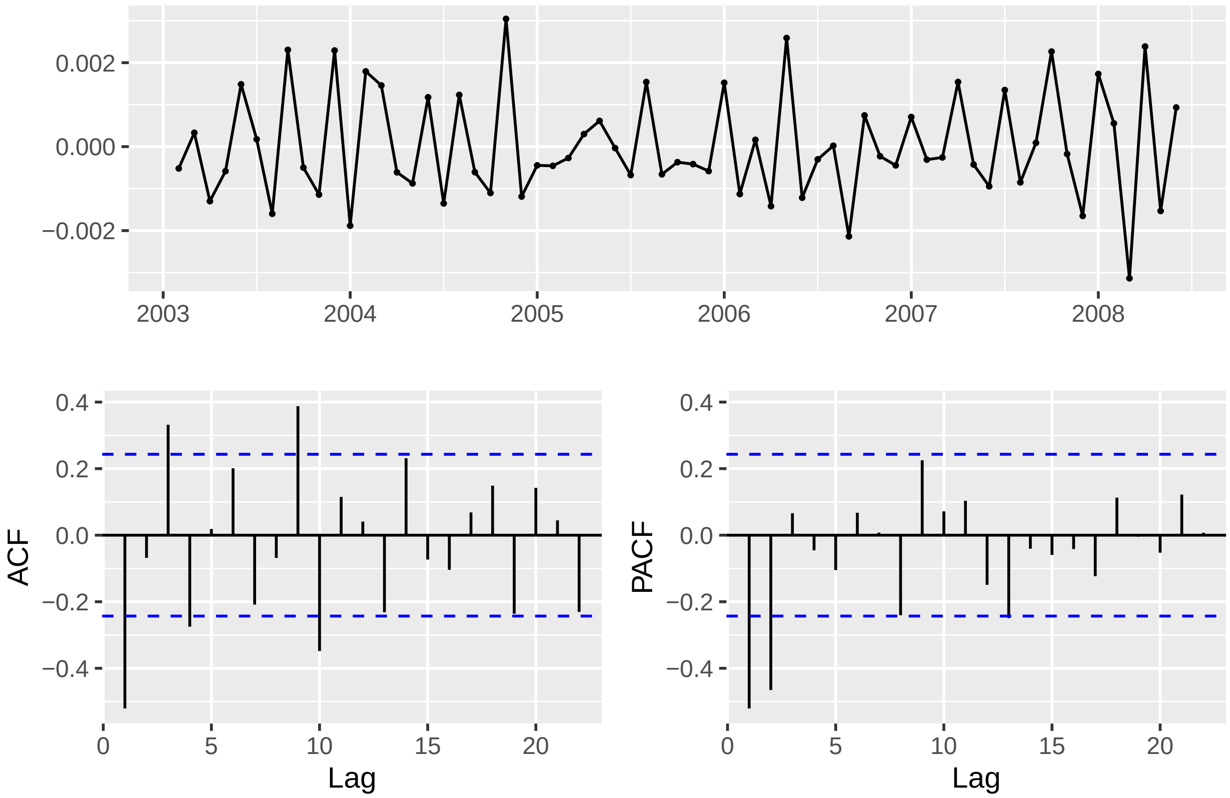
Agora, vamos remover a tendência diferenciando a série.



Com a série diferenciada, vamos checar a sazonalidade anual, como sugerido:

```
##
##  Kruskal-Wallis rank sum test
##
## data:  pre_intervention.d and cycle(pre_intervention.d)
## Kruskal-Wallis chi-squared = 66.041, df = 11, p-value = 6.862e-10
```

O p-valor é pequeno, então rejeitamos a hipótese nula do teste, o que nos dá suporte para diferenciar a série sazonalmente. Após a diferenciação, podemos ver que a ACF e a PACF não apresentam picos nos lags múltiplos de 12.



Assim, com o teste ADF rejeitando a não estacionaridade, podemos partir para a identificação do modelo.

```
## Warning in adf.test(pre_intervention.bc): p-value smaller than printed p-value
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  pre_intervention.bc
## Dickey-Fuller = -4.191, Lag order = 4, p-value = 0.01
## alternative hypothesis: stationary
```

Vemos uma ACF decaindo exponencialmente (em senóides amortecidas) e uma PACF morrendo após o lag 2. Assim imaginamos um modelo AR com grau não maior do que 3. Vamos testar essa ideia com os critérios de informação. Além disso, não percebemos nenhum LAG significativo sazonal, o que pode indicar um MA(1) ou não ter componentes autorregressivos ou de média móvel sazonais. Vamos testar esses modelos também com os critérios de informação. O modelo de teste ARIMA(3,1,3) teve problema de estacionariedade e foi retirado da análise.

A tabela mostra vários critérios de informação para os modelos ARMA(p, q)[12] e ARMA(p, q)[12] com uma média movel sazonal (S).

Podemos observar que, sem componente MA sazonalidade, o melhor modelo segundo os três critérios de informação foi ARMA(2,0), algo esperado dadas as nossas considerações anteriormente. Considerando MA sazonal, o melhor modelo fica, entre todos, ARIMA(2,1,0)(0,1,1)[12].

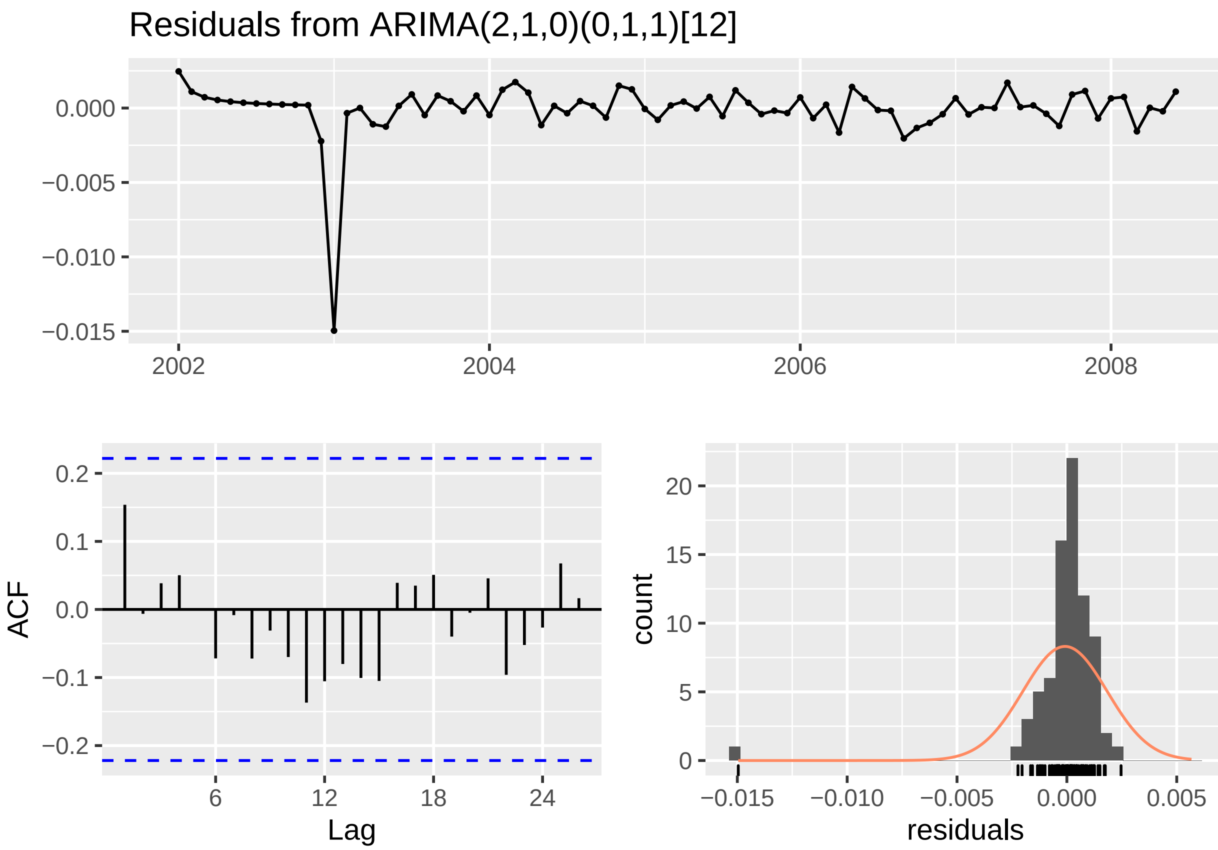
```
##    p q       AIC       BIC      AICc   AIC (S)   BIC (S)  AICc (S)
## 1  0 0 -679.1115 -676.9371 -679.0480 -677.4750 -673.1262 -677.2815
## 2  0 1 -706.1459 -701.7971 -705.9524 -705.4041 -698.8809 -705.0106
## 3  0 2 -706.7266 -700.2034 -706.3331 -708.4334 -699.7359 -707.7668
## 4  1 0 -697.5741 -693.2253 -697.3805 -695.6333 -689.1101 -695.2398
## 5  1 1 -705.8174 -699.2942 -705.4240 -706.0822 -697.3846 -705.4155
```

```
## 6  1 2 -704.7680 -696.0705 -704.1014 -706.4932 -695.6212 -705.4762
## 7  2 0 -710.9575 -704.4343 -710.5640 -717.9966 -709.2990 -717.3299
## 8  2 1 -709.3628 -700.6653 -708.6962 -716.5123 -705.6404 -715.4953
## 9  2 2 -707.3870 -696.5151 -706.3700 -714.5490 -701.5027 -713.1007
## 10 3 0 -709.2816 -700.5841 -708.6150 -716.4277 -705.5558 -715.4108
## 11 3 1 -707.3690 -696.4970 -706.3520 -714.5186 -701.4722 -713.0703
## 12 3 2 -707.3121 -694.2658 -705.8638 -712.8482 -697.6275 -710.8833
```

Agora que já identificamos o modelo, podemos estimá-lo.

```
## Series: pre_intervention
## ARIMA(2,1,0)(0,1,1)[12]
## Box Cox transformation: lambda= -0.2225847
##
## Coefficients:
##           ar1      ar2     sma1
##       -0.8633  -0.6094  -0.6911
## s.e.   0.0961   0.0994   0.2242
##
## sigma^2 estimated as 4.585e-06:  log likelihood=363
## AIC=-718   AICc=-717.33   BIC=-709.3
##
## Training set error measures:
##                     ME     RMSE      MAE        MPE     MAPE      MASE
## Training set -1902.353 38156.23 18285.63 -0.2765078 1.947999 0.2891937
##                   ACF1
## Training set 0.1325941
```

Agora com o modelo treinado, podemos fazer uma checagem sobre os resíduos:




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,1,0)(0,1,1)[12]
## Q* = 9.3584, df = 13, p-value = 0.7454
##
## Model df: 3.   Total lags used: 16
```

```
##
##  Jarque Bera Test
##
## data:  model1$residuals
## X-squared = 6926.5, df = 2, p-value < 2.2e-16
```

Observamos que existe um resíduo bem deslocado que no ano de 2003 quando houve uma queda que o modelo não conseguiu capturar. Além disso, o resíduo não aparenta normalidade; inclusive, a curtose e o *skewness* da normal foram rejeitados pelo teste Jarque-Bera. Por fim, as correlações estão bem interessantes, dentro das margens e, o teste de Ljung-Box não rejeitou a hipótese de descorrelação. Isso é um bom indício, porém a não normalidade dos dados indica algum problema.

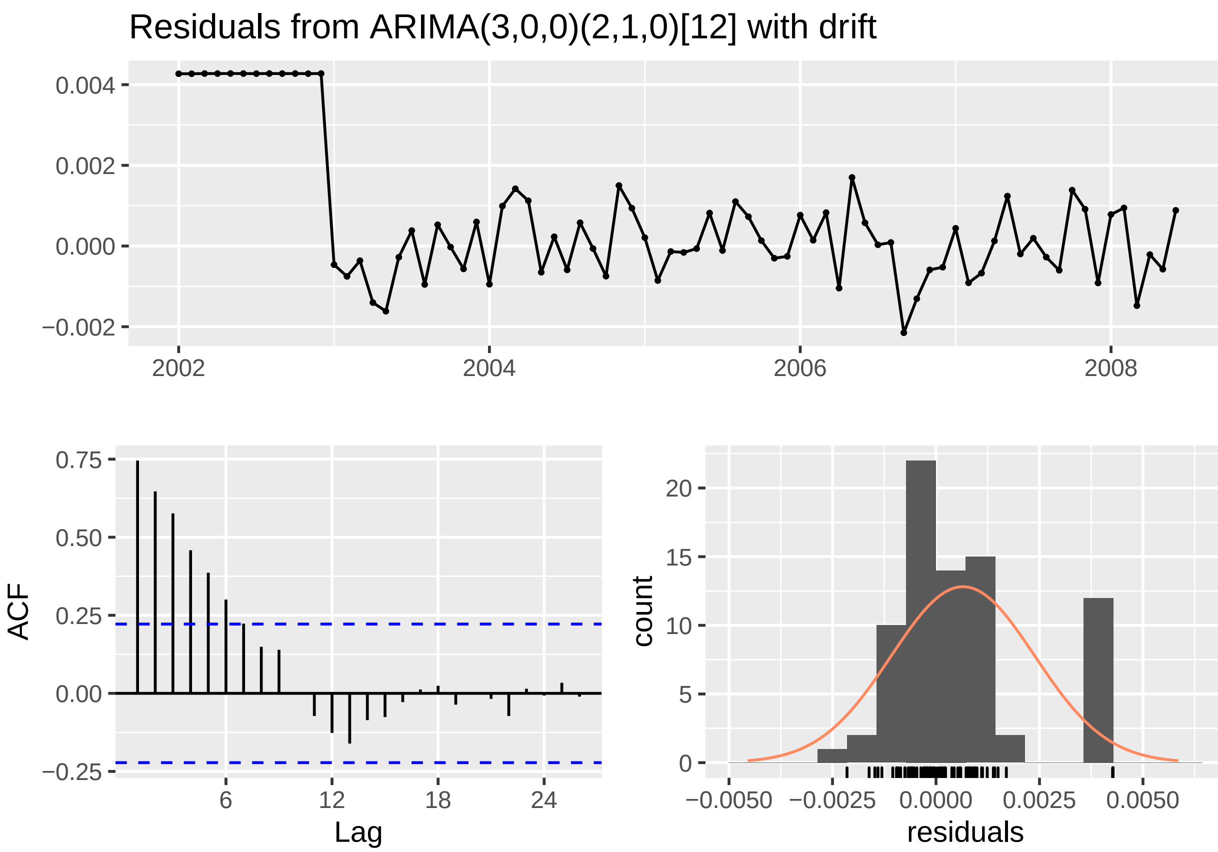
2. Modelo auto.arima

Vamos averiguar uma segunda opção de modelo que pode ser encontrada automaticamente, utilizando a função `auto.arima`. Assim, vemos que o modelo estimado tem drift e é do tipo

ARIMA(3,0,0)(2,1,0)[12].

```
## Series: pre_intervention
## ARIMA(3,0,0)(2,1,0)[12] with drift
```

```
## Box Cox transformation: lambda= -0.2225847
##
## Coefficients:
##          ar1     ar2     ar3     sar1     sar2  drift
##       0.1015  0.1235  0.5118  -0.3434  -0.4384  3e-04
## s.e.  0.1104  0.1156  0.1119   0.1419   0.1616  2e-04
##
## sigma^2 estimated as 4.413e-06:  log likelihood=371.16
## AIC=-728.32   AICc=-726.39   BIC=-712.99
##
## Training set error measures:
##                    ME     RMSE      MAE      MPE     MAPE     MASE      ACF1
## Training set 10721.76 32202.53 23457.97 1.270498 2.547099 0.370996 0.6357849
```




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(3,0,0)(2,1,0)[12] with drift
## Q* = 159.32, df = 10, p-value < 2.2e-16
##
## Model df: 6.   Total lags used: 16
```

```
##
##  Jarque Bera Test
##
## data:  model2$residuals
## X-squared = 17.463, df = 2, p-value = 0.0001614
```

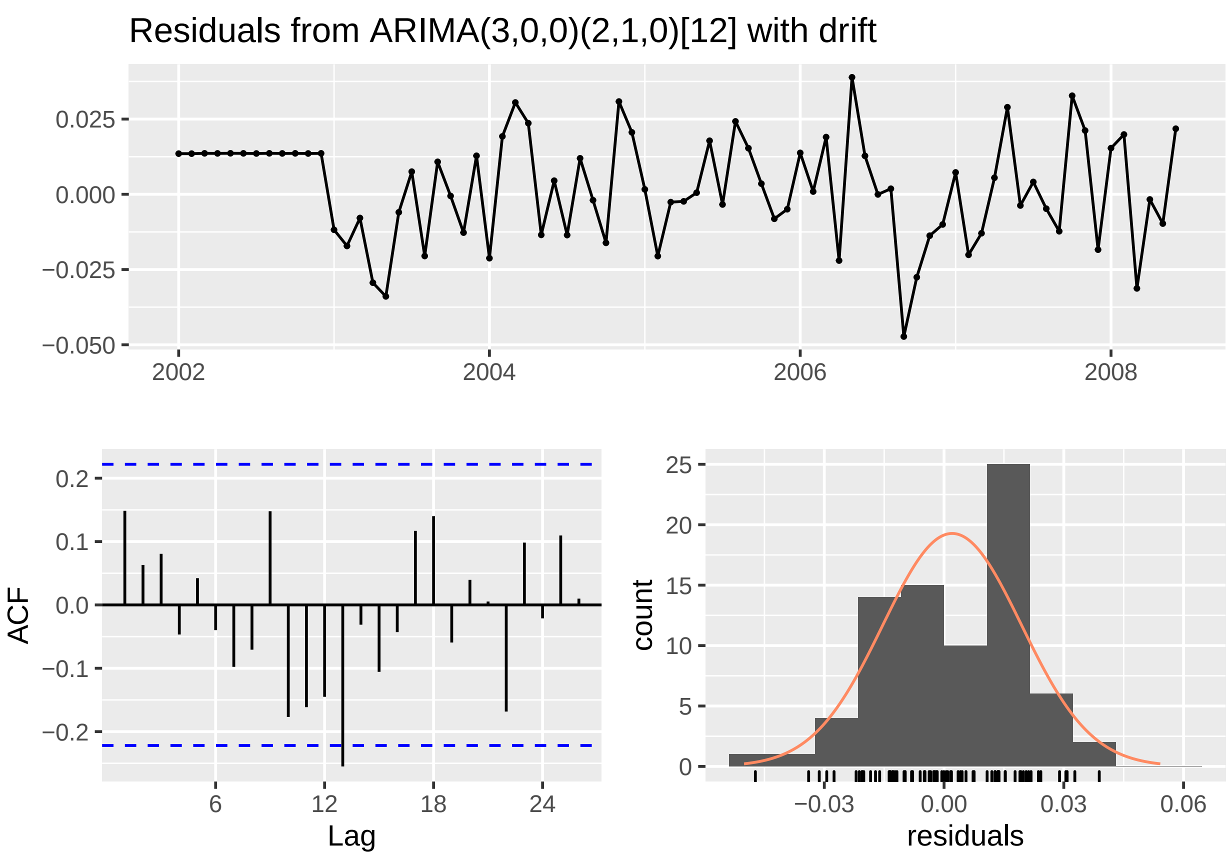
Pela ACF, os erros certamente tem correlação, a princípio uma correlação autorregressiva, e também existe uma assimetria no histograma. Esses fatos são confirmados com os testes, dado que rejeitamos ambas as hipóteses nulas, isto é, temos evidência para assegurar que não há normalidade e existe correlação nos resíduos. As medidas dos erros de treinamento também não são melhores do que o modelo 1.

3. Modelo auto.arima adaptado manualmente

Fazendo algumas experimentações com o `auto.arima` podemos chegar num fato interessante colocando $\lambda = 0$.

```
## Series: pre_intervention
## ARIMA(3,0,0)(2,1,0)[12] with drift
## Box Cox transformation: lambda= 0
##
## Coefficients:
##          ar1     ar2     ar3     sar1     sar2   drift
##       0.1055  0.1302  0.5103  -0.3279  -0.4297  0.0055
## s.e.  0.1105  0.1088  0.1116   0.1425   0.1418  0.0004
##
## sigma^2 estimated as 0.0003932:  log likelihood=168.3
## AIC=-322.59   AICc=-320.66   BIC=-307.27
##
## Training set error measures:
##                    ME     RMSE      MAE       MPE     MAPE     MASE     ACF1
## Training set 1907.311 17686.81 14286.31 0.1881717 1.442845 0.225943 0.106949
```

Conseguimos um modelo com todas as medidas de erro menores e, além disso:




```
##
##  Ljung-Box test
```

```
##
## data:  Residuals from ARIMA(3,0,0)(2,1,0)[12] with drift
## Q* = 21.302, df = 10, p-value = 0.01908
##
## Model df: 6.   Total lags used: 16
```

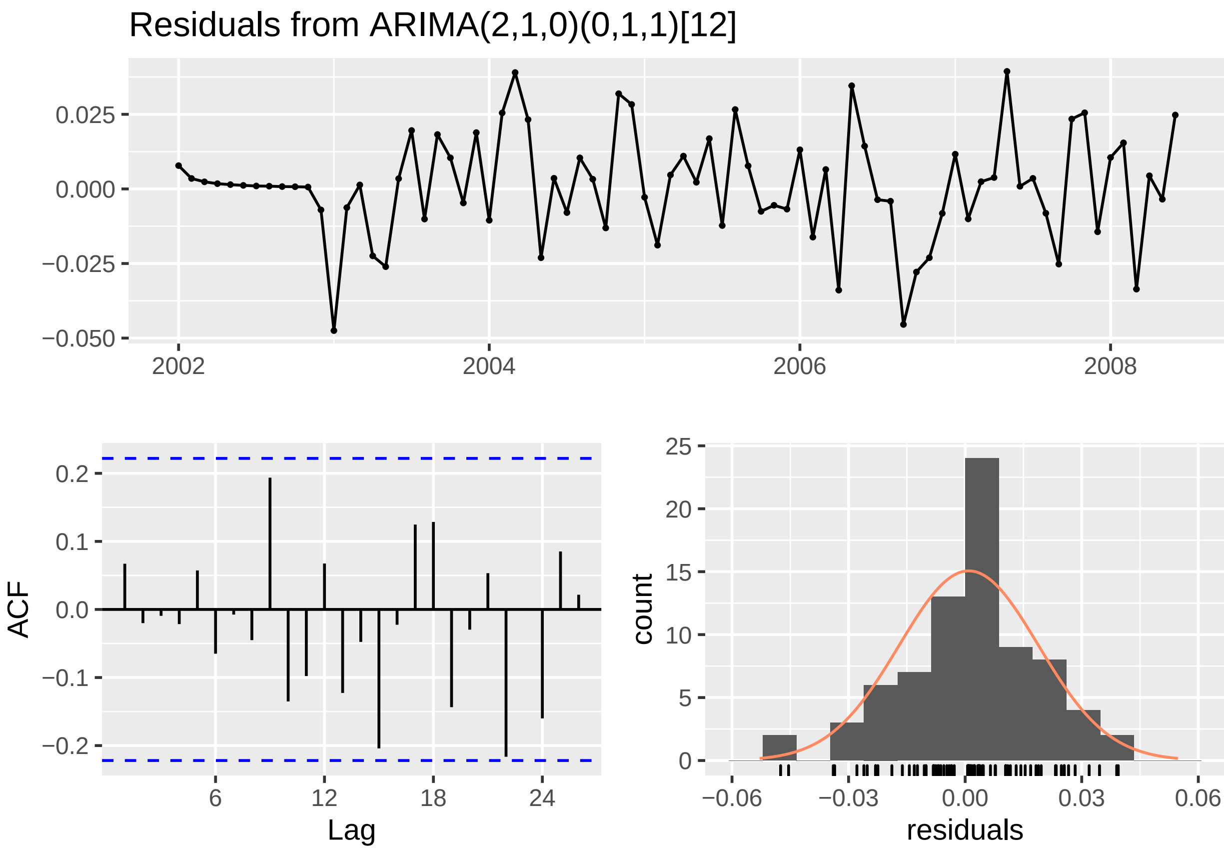
```
##
##  Jarque Bera Test
##
## data:  model3$residuals
## X-squared = 1.7383, df = 2, p-value = 0.4193
```

Note que os resíduos aparentam muito mais uma normalidade do que antes, o que também é indicado pelo teste Jarque Bera. Porém o teste Ljung-Box rejeita a hipótese nula, o que mostra que ainda existe uma correlação nos resíduos.

4. Modelo manual com variação

Utilizando o modelo inicial com a variação de λ, temos:

```
## Series: pre_intervention
## ARIMA(2,1,0)(0,1,1)[12]
## Box Cox transformation: lambda= 0
##
## Coefficients:
##           ar1      ar2     sma1
##       -0.8617  -0.6038  -0.6408
## s.e.   0.0971   0.1003   0.2063
##
## sigma^2 estimated as 0.0004012:  log likelihood=163.05
## AIC=-318.1   AICc=-317.44   BIC=-309.41
##
## Training set error measures:
##                    ME     RMSE      MAE        MPE   MAPE      MASE       ACF1
## Training set 1017.575 18230.79 13580.72 0.07746272 1.3489 0.2147838 0.05014901
```



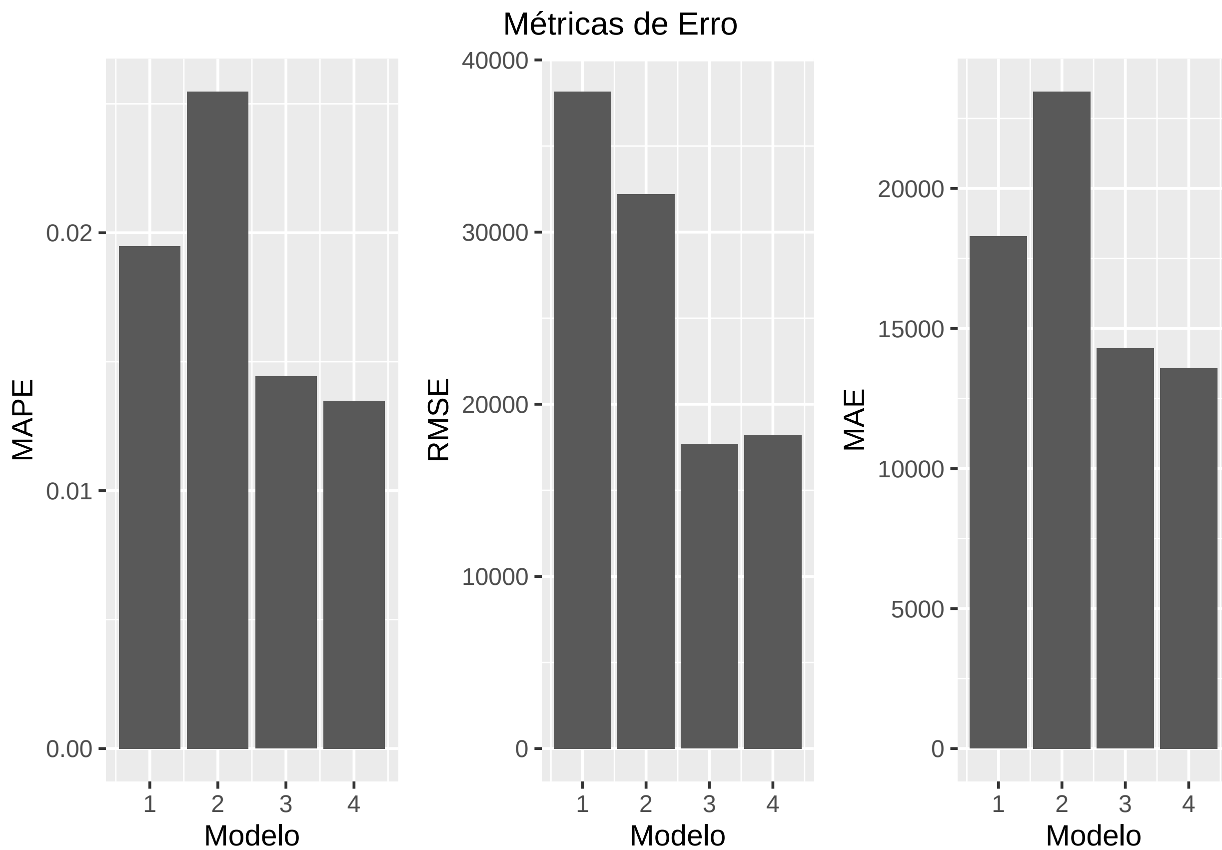

```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,1,0)(0,1,1)[12]
## Q* = 13.505, df = 13, p-value = 0.4096
##
## Model df: 3.   Total lags used: 16
```

```
##
##  Jarque Bera Test
##
## data:  model4$residuals
## X-squared = 1.0435, df = 2, p-value = 0.5935
```

Temos resíduos mais parecidos com uma normal, se comparados aos casos anteriores, e a ACF esá dentro das faixas de confiança, indicando um processo descorrelacionado. Tanto Jarque Bera quando Ljung Box não rejeitam suas hipóteses nulas, o que nos dá evidência para confirmar que os resíduos formam algo similar a um ruído branco. Os erros de treino também aparentam estar bem em relação ao modelo inicial.

Desta forma, fica claro que o modelo 4 é o mais indicado para representar a série antes da intervenção.

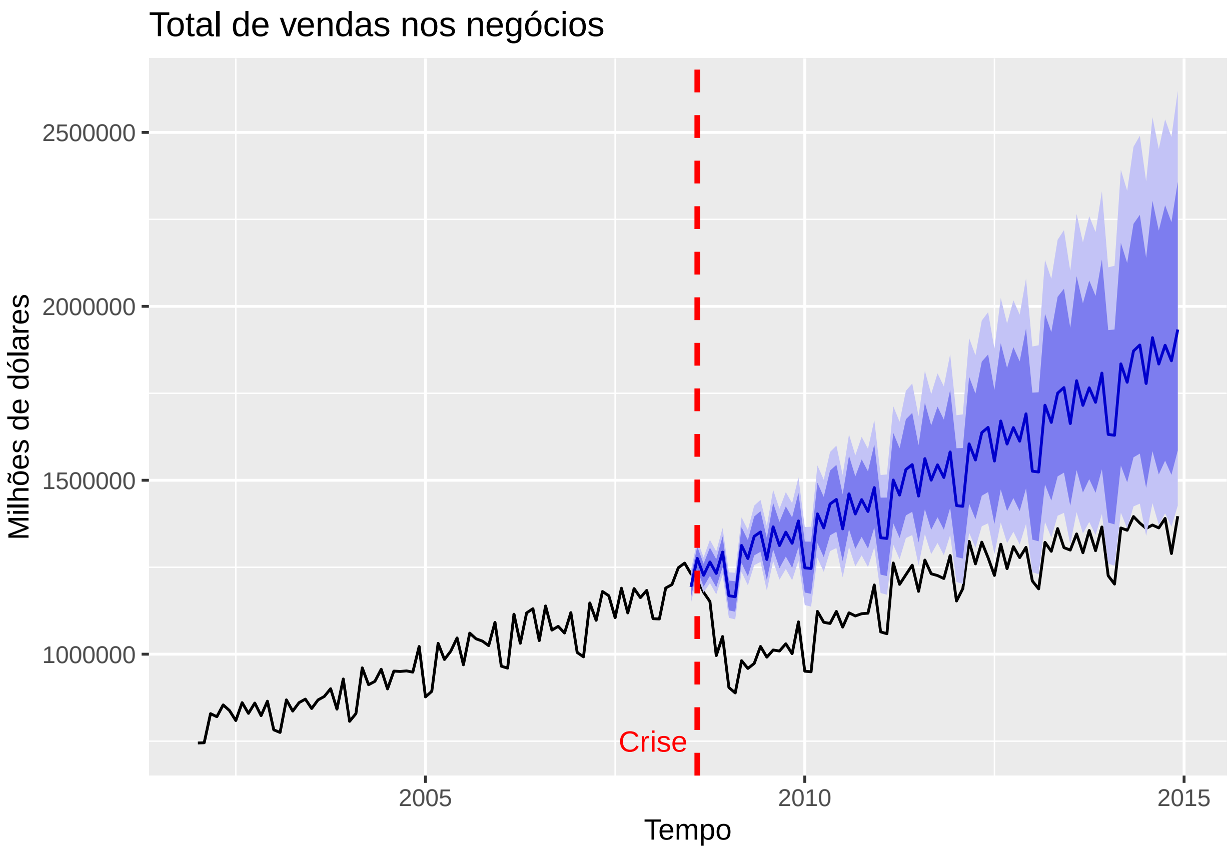
Podemos averiguar a comparação de três métricas para os erros:



Assim, além de o terceiro e o quarto modelos terem os menores erros, o quarto modelo vence em duas das métricas e, portanto, faz sentido ser escolhido (em adição aos seus resíduos terem um comportamento levemente melhor, como vimos).

Previsões do modelo e comparação pós-intervenção

Vejamos uma comparação entre o dado real e aquele previsto pelo modelo, ambos pós-crise.



Observamos que as previsões de nosso modelo ficam em um patamar mais alto do que os dados reais e por isso tomamos como hipótese nula de que a crise não foi significante. Consideramos os procedimentos sugeridos por Box e Tiao (1976) com a checagem do erro de previsão.

Considere a estatística

$$Q = \sum_{j=1}^{m} a_j^2/\hat{\sigma}_a^2$$

onde $a_j = Z_j - \hat{Z}_{j-1}(1), j = 1, ..., m$. Sabemos que Q seque uma distribuição $\chi^2(m)$ quando m é grande. Portanto rejeitamos a hipótese nula com nível de significância α_0 se $Q > F^{-1}(1-\alpha_0)$, onde F é a distribuição acumulada de uma chi-quadrado com m graus de liberdade. Ao observar $Q = q$, o p-valor será dado por $\alpha_0 = 1 - F(q)$.

```
## [1] "p-valor =  0 , Estatística Q =  262043736328174"
```

Como o p-valor é pequeno, em particular < 0.05, rejeitamos a hipótese nula, o que nos dá evidência para confirmar o efeito da crise subprime. Podemos, então, partir para a modelagem da intervenção.

Modelagem com intervenção

Consideraremos o modelo de intervenção dado por uma série que muda de nível abruptamente, dado a quebra do sistema bancário que tem efeito rápido e com um decaímento do efeito exponencial, dado que a economia se autoregula. Vamos supor que não haja outras intervenções, como governamental, e que elas devem ser incorporadas no modelo. Se nossa série atual é u_t, vamos modelar Z_t:

$$Z_t = v(B)X_t + u_t$$

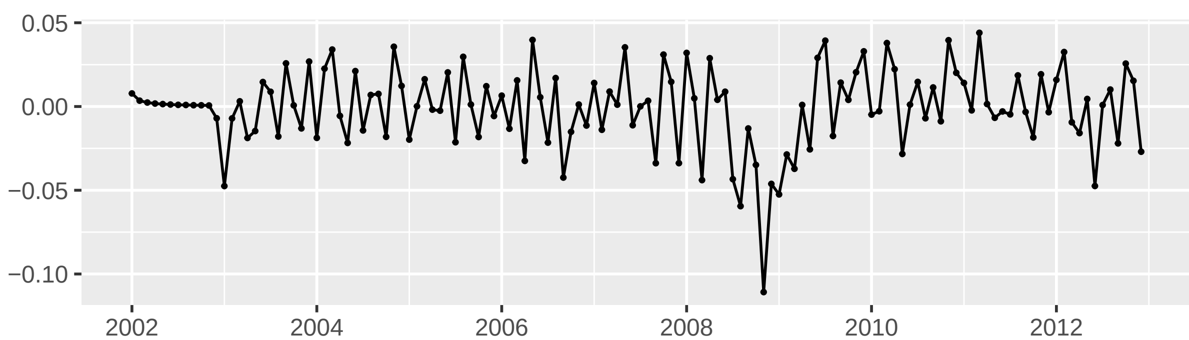
sendo X_t “temporário” e $v(B) = \dfrac{\omega_0}{1 - \delta B}$. Portanto,

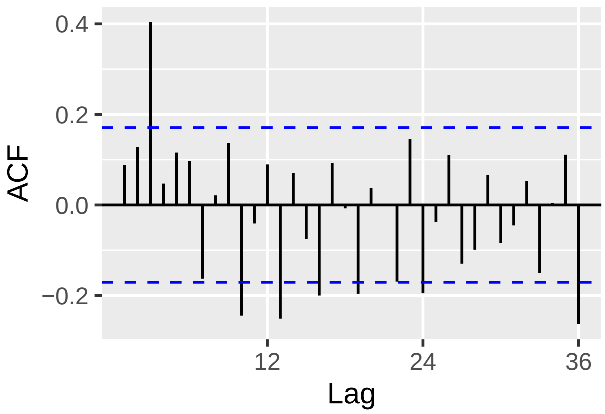
$$h_t = v(B)X_t = \begin{cases} 0, & t < \text{jul } 2018 \\ \delta^k \omega_0, & t = \text{jul } 2008 + k, \quad k = 0, 1, 2, \cdots \end{cases}$$

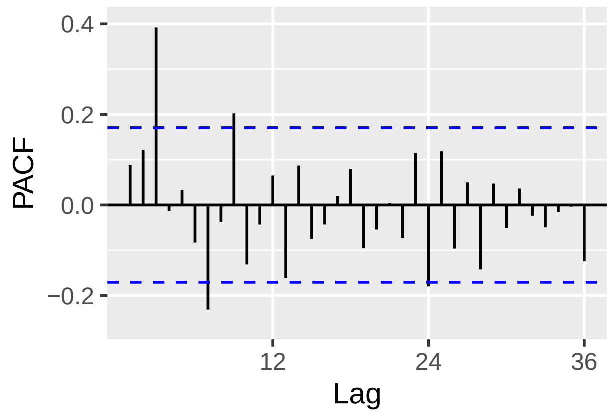
Estimaremos a série com intervenção, ou seja, também estimaremos os parâmetros δ e ω_0.

```
##
## Call:
## arimax(x = tbs_train.bc, order = c(2, 1, 0), seasonal = c(0, 1, 1), method = "ML",
##     xtransf = data.frame(Crise = 1 * (seq(tbs_train) == (length(pre_intervention) +
##         1))), transfer = list(c(1, 0)))
##
## Coefficients:
##           ar1      ar2     sma1  Crise-AR1  Crise-MA0
##       -0.3768  -0.2128  -1.0000     0.6068     0.0785
## s.e.   0.0905   0.0934   0.0993     0.1346     0.0245
##
## sigma^2 estimated as 0.0006043:  log likelihood = 257.69,  aic = -505.38
##
## Training set error measures:
##                        ME       RMSE        MAE          MPE      MAPE     MASE
## Training set -0.001111388 0.02372598 0.01754766 -0.008154253 0.1265103 0.347873
##                    ACF1
## Training set 0.08800144
```

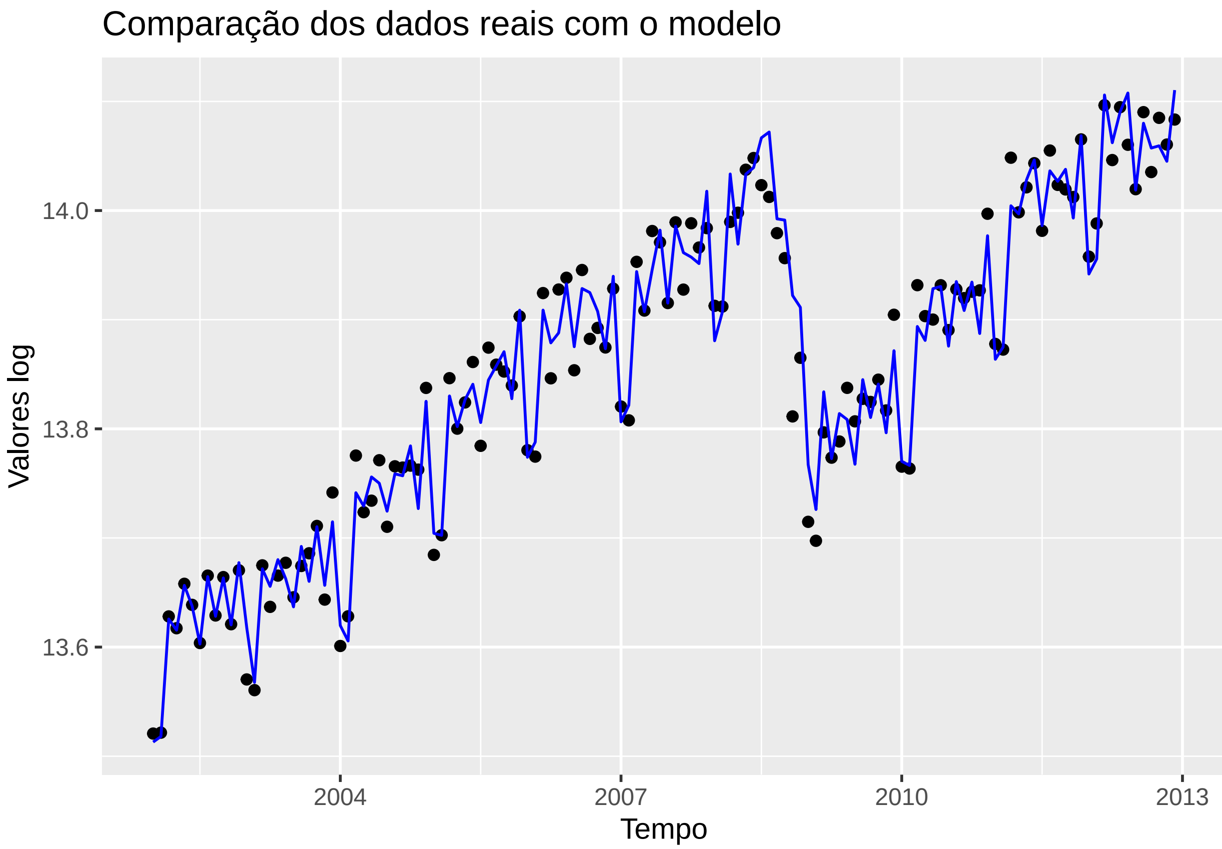
Vamos ver como se comportam os resíduos do modelo.







O resíduo não parece ser resultado de uma normal, pois justamente na intervenção há uma distorção bem grande. A ACF dos resíduos e a PACF estão com picos, o que indica que ainda existe uma correlação não capturada. Vamos ver como o modelo fitado se compara com os dados reais na escala log, dada a transformação BoxCox com $\lambda = 0$ que fizemos.

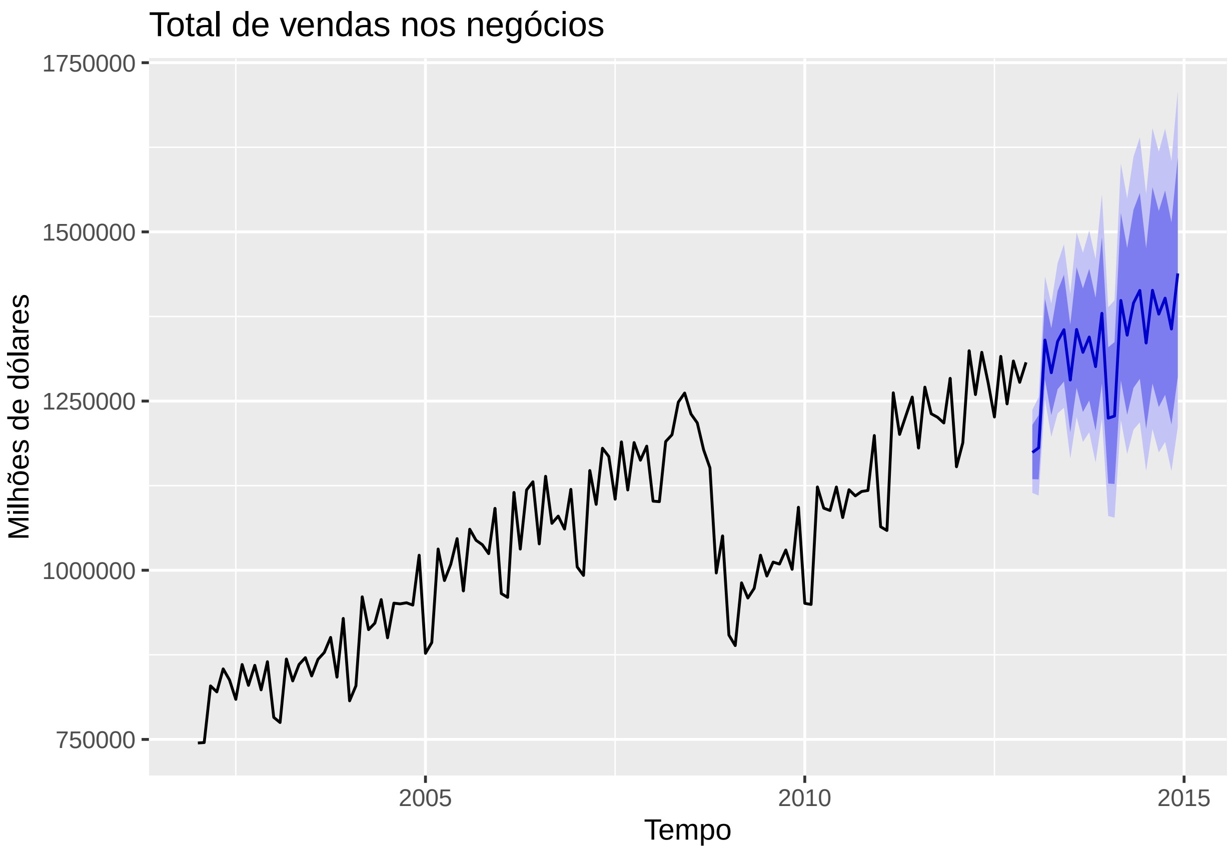


Em geral o modelo parece capturar bem a intervenção.

Previsão com o modelo de intervenção

Vamos agora realizar a previsão até 24 passos à frente. Para isso precisamos usar as estimativas de δ e ω_0 dados pelo arimax e calculamos h_t. Esses valores são usados como regressores da função `Arima`, a qual podemos fazemos fazer predições.

```
autoplot(forecast(mod.arima, h=24, xreg = tf[(n+m+1):(n+m+K)]),
         level = c(90, 95),
         main = "Total de vendas nos negócios",
         xlab = "Tempo",
         ylab = "Milhões de dólares")
```

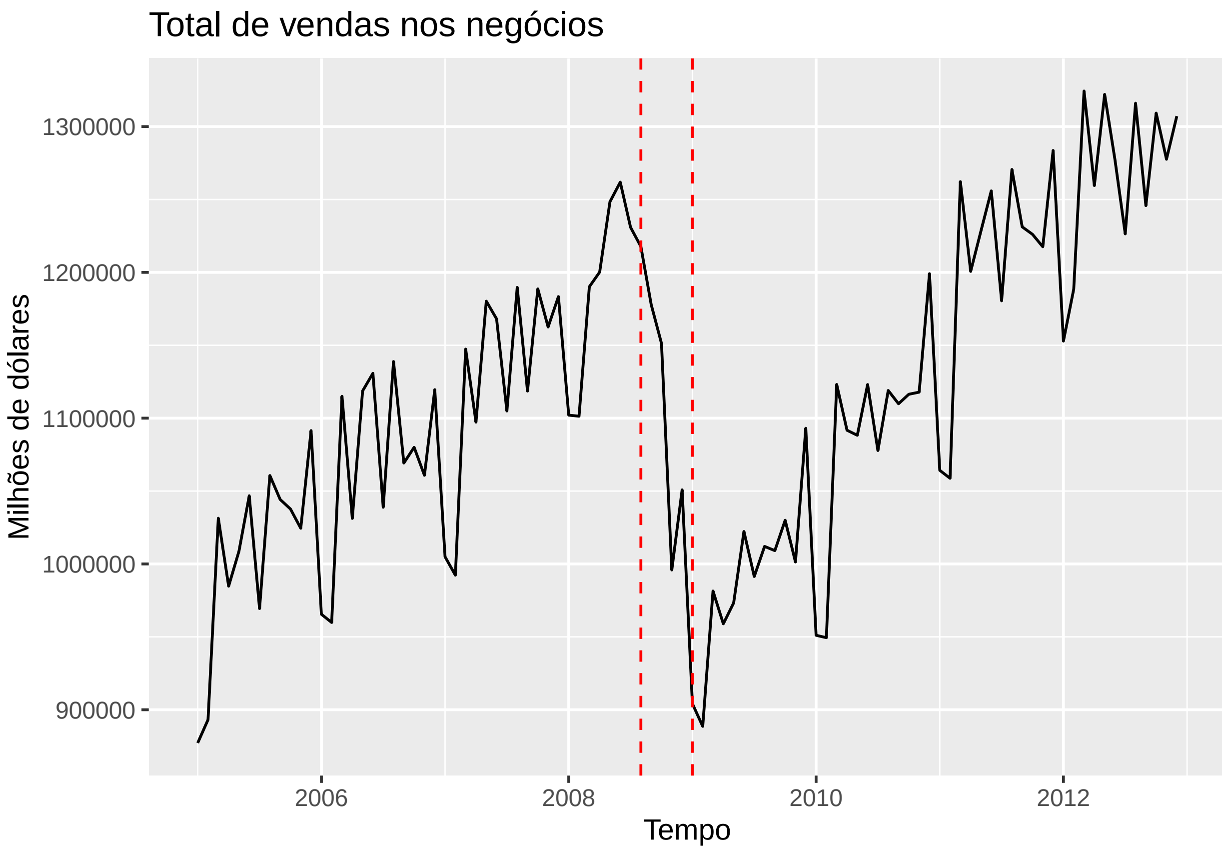


Vamos verificar o MAPE da nossa previsão, comparando com os dados reais de teste:

```
## [1] 0.01699805
```

Nos parece um ótimo resultado.

Porém, acreditamos que podemos melhorar. É óbvio que o efeito da intervenção não é permanente, por isso optamos pelo que chamamos de efeito temporário. Porém, observe que o efeito da intervenção também não é imediato:

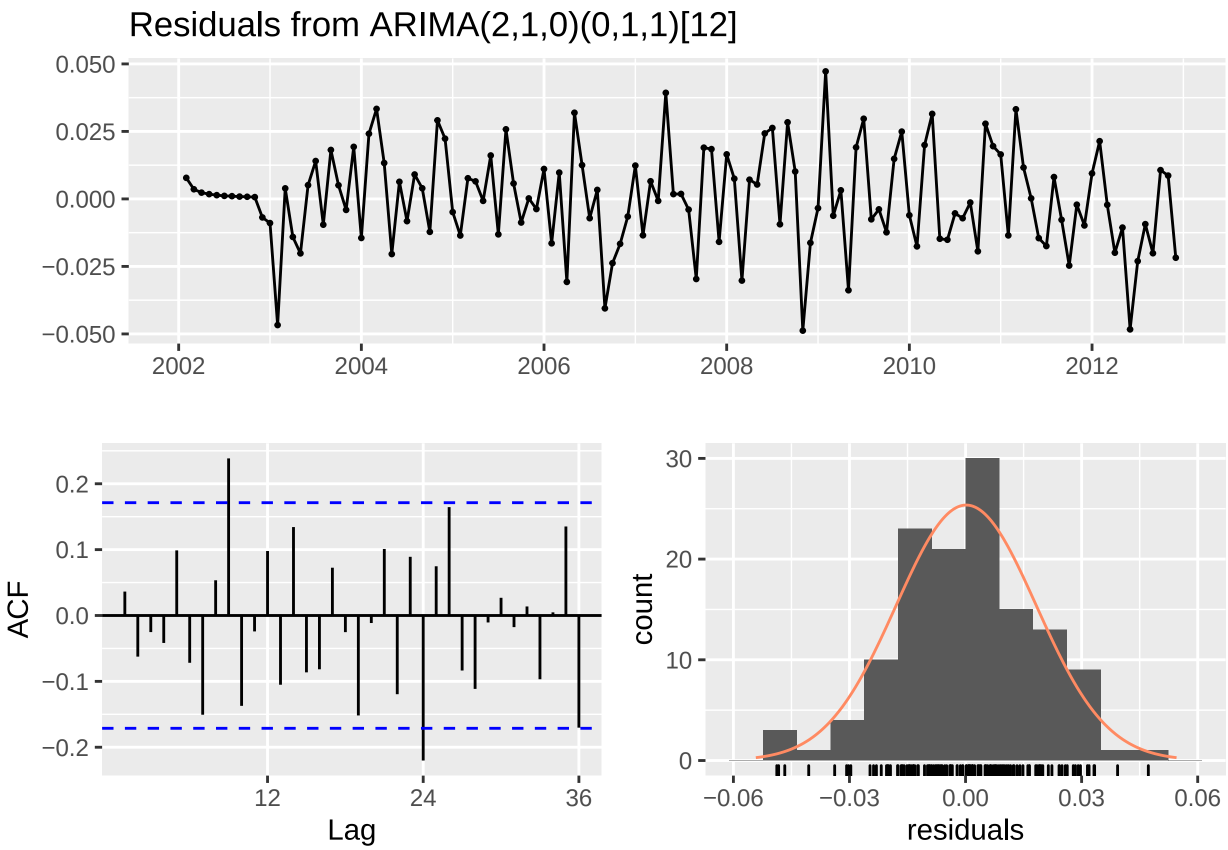


Isso nos indica que poderíamos considerar um efeito temporário, mas não imediato, de alguns meses.

Vamos modelar a mesma intervenção, porém permitindo que ela dure mais tempo, por 6 meses (desde jul 2008 até dez 2008).

```
##
## Call:
## arimax(x = tbs_train.bc, order = c(2, 1, 0), seasonal = c(0, 1, 1), method = "ML",
##     xtransf = data.frame(Crise = 1 * (seq(tbs_train) >= length(pre_intervention) +
##         1 & seq(tbs_train) <= length(pre_intervention) + 1 + 6)), transfer = list(c(1,
##         1)))
##
## Coefficients:
##           ar1      ar2     sma1  Crise-AR1  Crise-MA0  Crise-MA1
##       -0.7628  -0.5484  -0.9998     0.9937     0.0018    -0.0462
## s.e.   0.0776   0.0789   0.1271     0.0086     0.0106     0.0106
##
## sigma^2 estimated as 0.0003415:  log likelihood = 288.78,  aic = -565.56
##
## Training set error measures:
##                        ME      RMSE        MAE         MPE      MAPE      MASE
## Training set 0.0001661301 0.0180541 0.01420959 0.001178527 0.1024515 0.2816975
##                    ACF1
## Training set 0.03633095
```

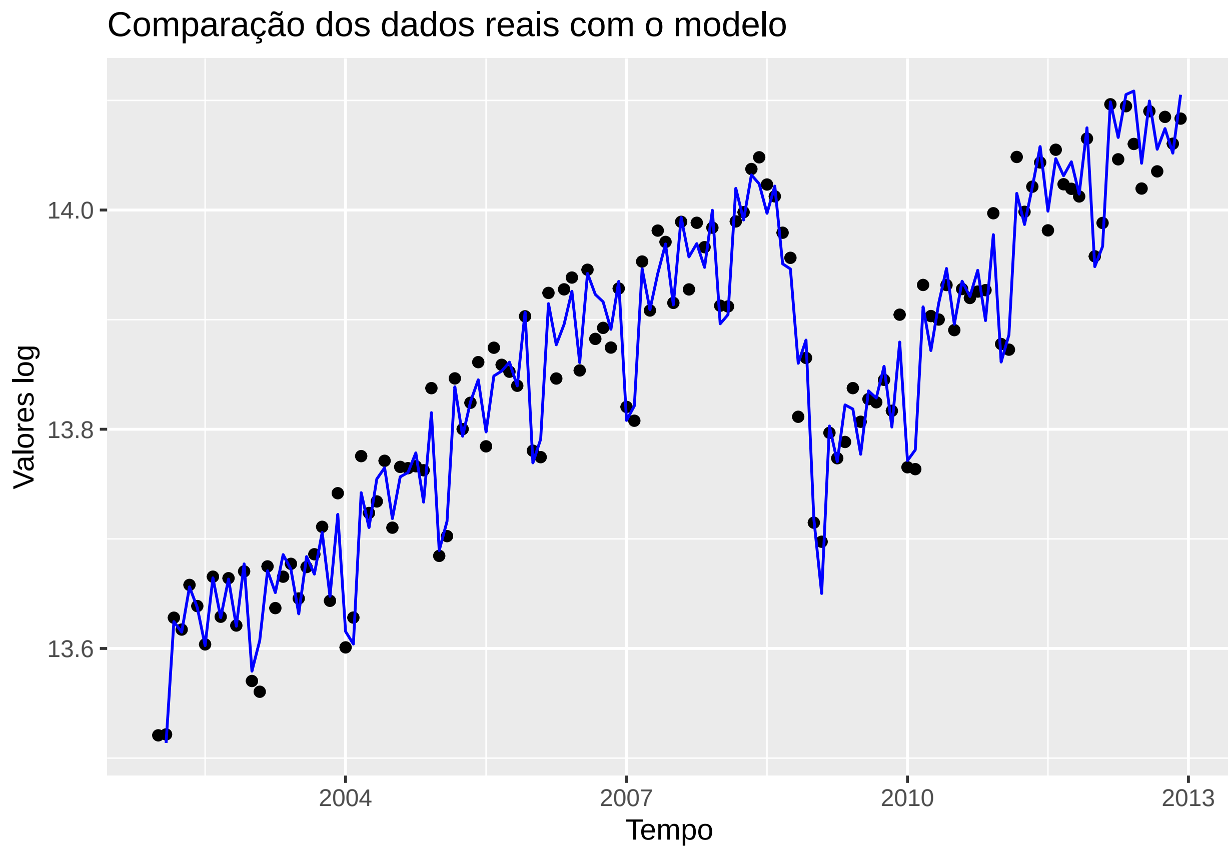
Vamos ver como se comportam os resíduos do modelo.




```
##
##  Ljung-Box test
##
## data:  Residuals from ARIMA(2,1,0)(0,1,1)[12]
## Q* = 43.034, df = 18, p-value = 0.0007915
##
## Model df: 6.   Total lags used: 24
```

Visualmente, os resíduos se comportam de forma melhor que anteriormente: a distorção na intervenção foi amenizada e a ACF se comporta melhor. Observamos um histograma razoável, porém ainda temos alguns picos na ACF. O teste de Ljung-Box rejeitou a hipótese de descorrelação, o que não é um bom sinal.

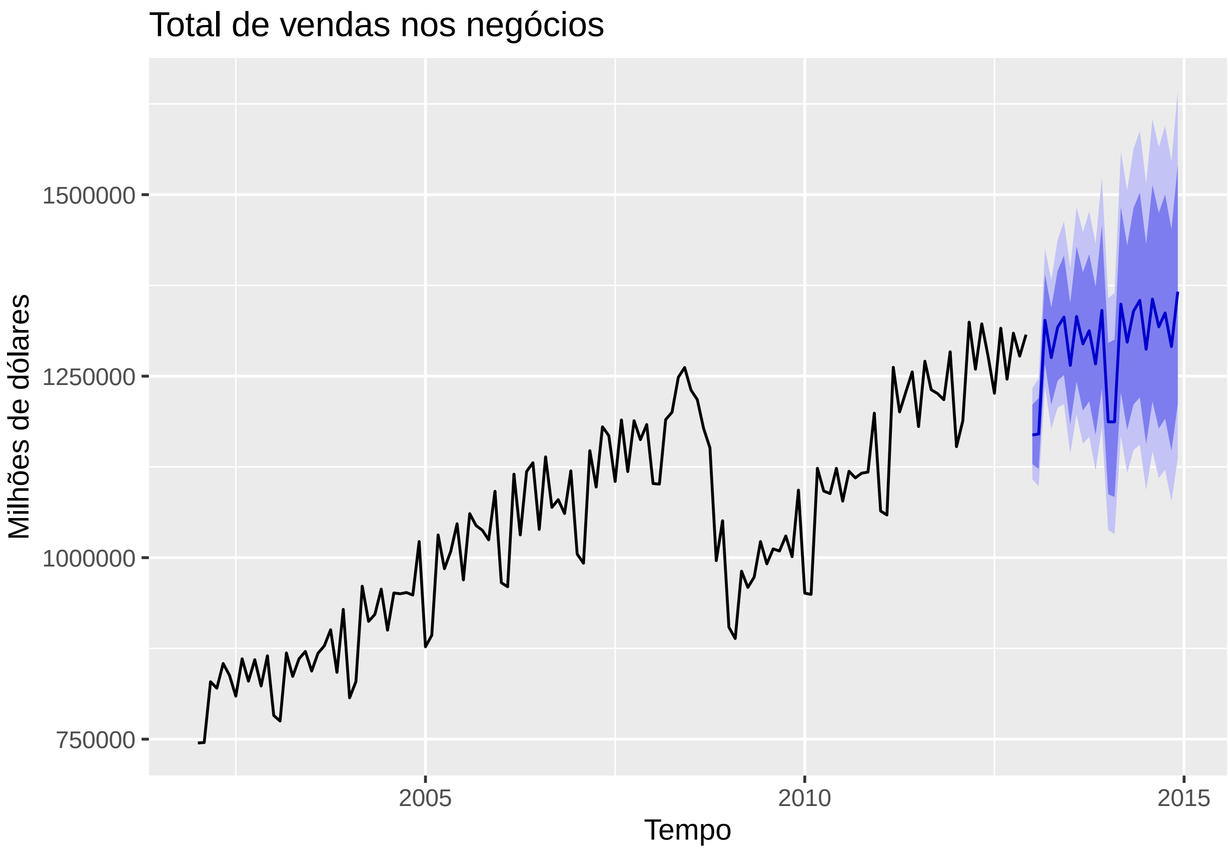
Ainda assim, percebemos que esse modelo reflete melhor a intervenção na crise do subprime. Podemos conferir como o modelo se comporta com os daos reais.



Previsão com o modelo de intervenção

O gráfico abaixo nos permite visualizar a predição do modelo até até 24 passos à frente.

```
autoplot(forecast(mod2.arima, h=24, xreg = tf[(n+m+1):(n+m+K)]),
         level = c(90, 95),
         main = "Total de vendas nos negócios",
         xlab = "Tempo",
         ylab = "Milhões de dólares")
```



E o MAPE para o erro de previsão é o seguinte:

```
## [1] 0.02276858
```

Percebemos que o MAPE de previsão é ligeiramente maior nesse modelo, e, portanto, o primeiro modelo estimado ainda parece ter um bom comportamento.