



Clube do Código 17

PEC 241: desagregando os gastos do governo com o R.

Vítor Wilher
Mestre em Economia
analisemacro.com.br

Pacotes e Scripts Externos

```
source('bcb.R')

library(forecast)
library(ggplot2)
library(ggthemes)
library(easyGgplot2)
library(grid)
library(TStools)
library(XLConnect)
library(ggthemes)
library(stargazer)
library(urca)
library(seasonal)
```

1 Coleta e tratamento de dados

Na edição anterior do [Clube do Código](#), resolvi abrir a caixa preta dos gastos e receitas primárias do governo central. Mostrei ao leitor como pegar os dados do Tesouro e levá-los diretamente para o R. Vimos que os gastos do governo crescem em média próximo a 6% ao ano, em termos reais, o que, se nada for feito, nos levará ao caos em pouco tempo. Ou seja, esse é o maior problema das contas públicas: o gasto não parou de crescer nas últimas décadas. Para continuar nossa análise, vamos aproveitar o script e abrir a despesa primária, olhando um pouco mais de perto como se comportaram os gastos do governo na amostra que temos disponível. Apenas para lembrar, os códigos abaixo importam os dados de gastos e receitas primárias do governo central diretamente do site do Tesouro para o R.¹

```
### Importar dados e tratá-los
www <- 'http://www.tesouro.fazenda.gov.br/documents/10180/246449/Anexos+RTN+Ago+2016.xlsx/4a211f60-4
temp <- tempfile()
download.file(www, temp, mode='wb')
data <- loadWorkbook(temp)
data <- readWorksheet(data, sheet = "1.1", header = TRUE,
                      colTypes = 'numeric')
```

Um tratamento inicial dos dados é posto abaixo.

```
data <- t(data[-c(1:3, 77:81),-1])
data <- ts((data[1:236,]), start=c(1997,01), freq=12)

names <- t(read.table('names.csv', header=F))
colnames(data) <- names
```

A importação do PIB mensal do Banco Central é feita abaixo.

```
pib <- ts(getSeries(4382, data.ini='01/12/1997',
                  data.fim = '01/08/2016')$valor,
          start=c(1997,12), freq=12)
```

O IPCA é importado abaixo.

¹Compreendem o governo central, o Tesouro, a Previdência Social e o Banco Central.

```
### Importar IPCA
ipca <- ts(read.csv2('ipca.csv', header=T, sep=';', dec=',')[,-1],
          start=c(1997,01), freq=12)
```

Deflacionamos e acumulamos em 12 meses as séries abaixo.

```
### Deflacionar Séries
real <- tail(ipca,1)*data/ipca
colnames(real) <- colnames(data)

pib.real <- tail(ipca,1)*pib/ipca

### Acumular em 12 meses
real12 <- real+lag(real,-1)+lag(real,-2)+lag(real,-3)+
  lag(real,-4)+lag(real,-5)+lag(real,-6)+lag(real,-7)+
  lag(real,-8)+lag(real,-9)+lag(real,-10)+lag(real,-11)
colnames(real12) <- colnames(data)

data12 <- data+lag(data,-1)+lag(data,-2)+lag(data,-3)+
  lag(data,-4)+lag(data,-5)+lag(data,-6)+lag(data,-7)+
  lag(data,-8)+lag(data,-9)+lag(data,-10)+lag(data,-11)
colnames(data12) <- colnames(data)
```

Abaixo, colocamos os dados como percentual do PIB.

```
### Percentual do PIB
ppib <- data12/pib*100
colnames(ppib) <- colnames(data)
```

Por fim, fazemos a variação mensal, interreal12 e acumulada em 12 meses como abaixo.

```
## Variação mensal
mensal <- (real12/lag(real12,-1)-1)*100
colnames(mensal) <- colnames(data)

## Variação interreal12
interanual <- (real12/lag(real12,-12)-1)*100
colnames(interanual) <- colnames(data)
```

```
## Variação acumulada em 12 meses
anual <- (((real12+lag(real12,-1)+lag(real12,-2)+lag(real12,-3)
          +lag(real12,-4)+lag(real12,-5)+lag(real12,-6)+lag(real12,-7)
          +lag(real12,-8)+lag(real12,-9)+lag(real12,-10)+
          lag(real12,-11))/12)/
          ((lag(real12,-12)+lag(real12,-13)+
            lag(real12,-14)+lag(real12,-15)
            +lag(real12,-16)+lag(real12,-17)+lag(real12,-18)
            +lag(real12,-19)+lag(real12,-20)+lag(real12,-21)
            +lag(real12,-22)+lag(real12,-23))/12)-1)*100

colnames(anual) <- colnames(data)
```

2 Análise gráfica

Lembre o leitor que temos um objeto criado no nosso environment do RStudio chamado `real12`, que traz 73 séries, deflacionadas (a preços de agosto de 2016) e acumuladas em 12 meses. Entre elas, 32 referem-se às despesas do governo central. Nessas despesas primárias estão quatro grandes grupos: gastos com previdência (INSS), gastos com pessoal, outras despesas obrigatórias e despesas discricionárias. Podemos querer saber quanto cada um desses grupos toma em média da despesa primária total e apresentar em termos gráficos. Para isso, rodamos o código abaixo.

```
despesa <- real12[,c(36:38, 61)]/real12[,35]*100

colnames(despesa) <- c('Previdência', 'Pessoal',
                      'Outras Despesas Obrigatórias',
                      'Despesas Discricionárias')

pie(colMeans(despesa), col=c('darkblue', 'red', '367', 'blue'),
    main='Distribuição dos gastos do governo central (%)')
```

Distribuição dos gastos do governo central (%)



Somados, os gastos do governo central com previdência e pessoal somam, em média, 65.47% da despesa primária total. As outras despesas obrigatórias somam 10.48% e as despesas discricionárias, 24.03%. Ou seja, leitor, a maior parte do gasto do governo central é engessado, i.e., ele não pode simplesmente cortar, caso, por exemplo, a receita de um determinado ano resolva cair, em função do baixo crescimento do PIB.

Nós podemos, inclusive, verificar como esses grandes grupos de despesa primária evoluíram ao longo da nossa amostra. Para tal, vamos ver primeiro a evolução do gasto mensal, já deflacionado pelo IPCA acima, dessas rubricas. Estou usando a função `autoplot` para quem se interessar e o pacote `ggthemes` para tornar o gráfico um pouco mais apresentável.

```
g1 <- autoplot(real[,36]/1000)+  
  geom_line(colour="darkblue", size=1)+  
  xlab('')+ylab('R$ bilhões')+  
  ggtitle('Gastos com Previdência')+  
  theme_economist()  
  
g2 <- autoplot(real[,37]/1000)+  
  geom_line(colour="red", size=1)+  
  xlab('')+ylab('R$ bilhões')+  
  ggtitle('Gastos com Pessoal')+  
  theme_economist()
```

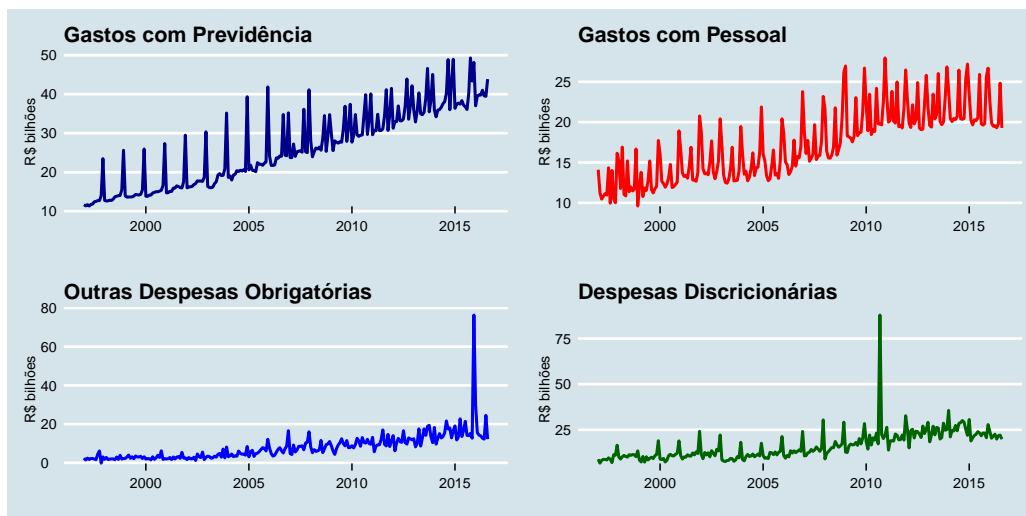
```

g3 <- autoplot(real[,38]/1000)+
  geom_line(colour="blue", size=1)+
  xlab('')+ylab('R$ bilhões')+
  ggtitle('Outras Despesas Obrigatórias')+
  theme_economist()

g4 <- autoplot(real[,61]/1000)+
  geom_line(colour="darkgreen", size=1)+
  xlab('')+ylab('R$ bilhões')+
  ggtitle('Despesas Discricionárias')+
  theme_economist()

ggplot2.multiplot(g1, g2, g3, g4, cols=2)

```



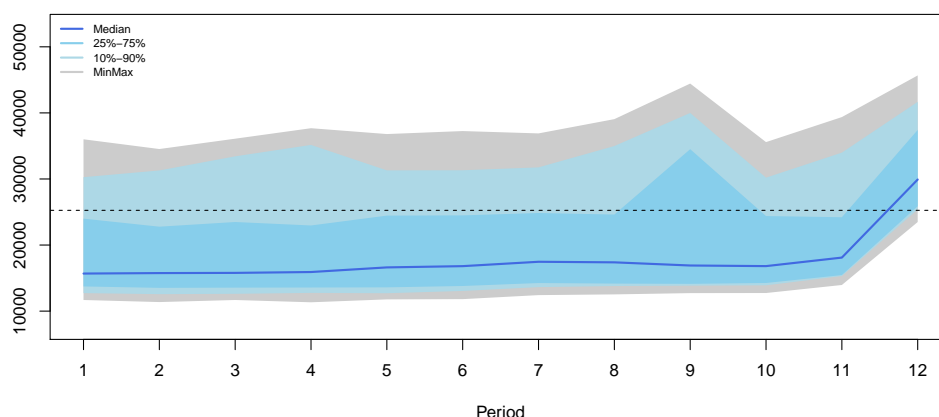
Como é possível observar nos gráficos acima, há uma sazonalidade nos dados. Podemos ver essa sazonalidade mais de perto com a função `seasplot` do pacote `TStools`. Como exemplo, pegamos os gastos com previdência, abaixo.

```

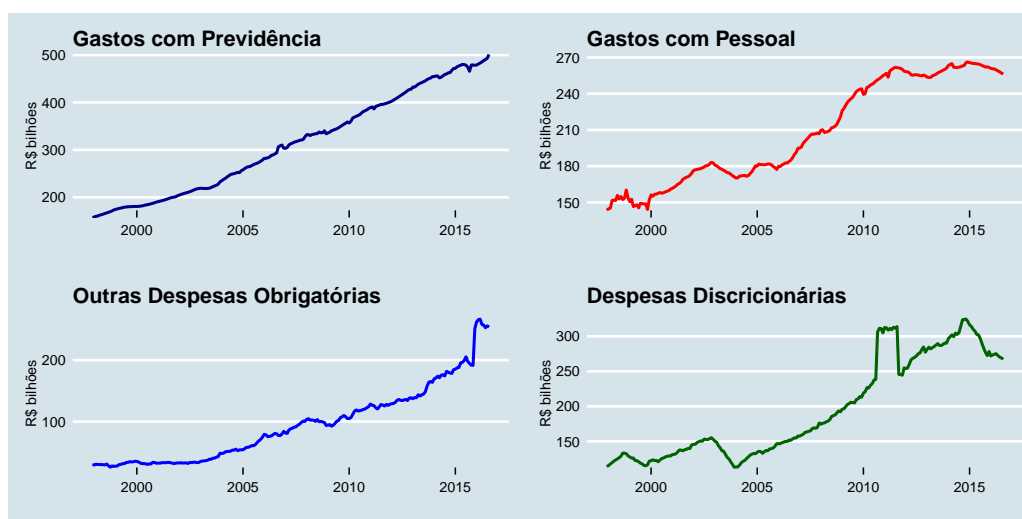
seasplot(real[,36], outplot=4, trend=F, plot.title = 'Sazonalidade dos Gastos com Previdência')

```

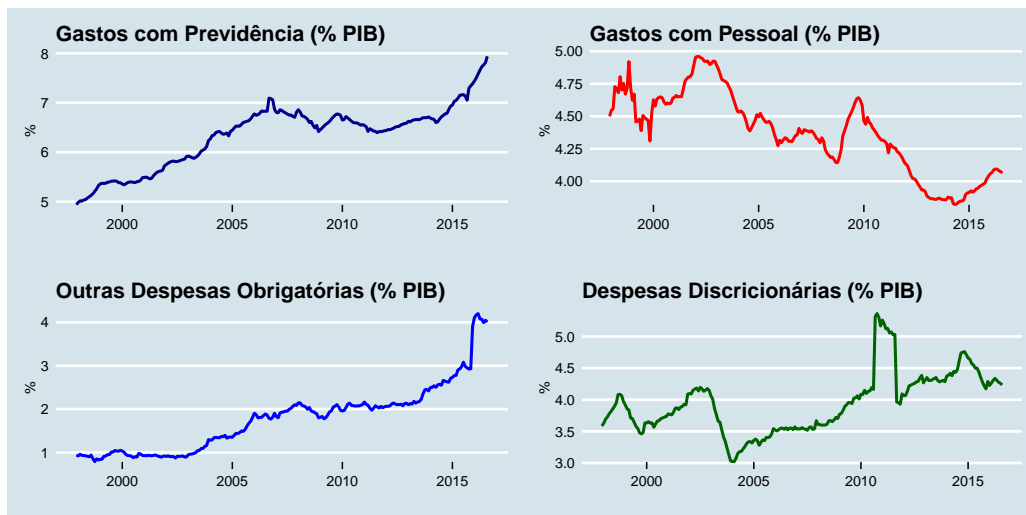
Sazonalidade dos Gastos com Previdência



Voltaremos a isso mais à frente. Por enquanto, vamos agora ver o gráfico dessas variáveis, no acumulado em 12 meses. Sempre lidando, claro, com dados reais. O código é o mesmo visto acima.



Gastos com previdência e outras despesas obrigatórias mantêm uma tendência positiva de crescimento como se não houvesse amanhã. Os gastos com pessoal têm se estabilizado no período recente, enquanto a parte que o governo tem algum poder (a tal discricionariedade), parece ser o grande resíduo. Em agosto de 2016, os gastos com previdência somaram R\$ 500,9 bilhões, as despesas com pessoal, R\$ 256,5 bilhões, as demais despesas obrigatórias, R\$ 253,8 bilhões e as despesas discricionárias, R\$ 267,3 bilhões. Podemos agora controlar essa evolução do gasto pelo PIB. Isso é colocado abaixo.



No período recente, influenciado pela forte queda do PIB, previdência, outras despesas obrigatórias e gasto com pessoal apresentam forte elevação. As despesas discricionárias, onde o governo pode fazer algo no curto prazo, há uma queda. Nesse grupo, estão situados gastos de investimento, como o PAC e o Minha Casa Minha Vida. Em outras palavras, leitor, a maior parte do gasto do governo está engessada. Toda vez que o governo se encontra em dificuldades, ele corta justamente em investimento, onde tem algum grau de liberdade. Significa dizer que, o segundo passo, pós-aprovação da PEC 241, será enfrentar esse engessamento. Será preciso conter o avanço dos gastos com previdência e demais gastos obrigatórios. Ou seja, leitor, o que os dados dizem é que a PEC 241 é só o começo...

3 Análise Estatística

Vamos, agora, olhar um pouco mais de perto os gastos com previdência, que, como vimos, corresponde a quase 40% da despesa primária total do governo central. Para tal, vamos fazer uma análise estatística inicial, olhando as funções de autocorrelação desses gastos, em bases mensais e já deflacionadas.

Para uma série *covariance stationary*, como especificado em Cowpertwait and Metcalfe (2009) e Montgomery et al. (2008), nós podemos definir uma *Função de Autocovariância* como

$$\gamma_k = E[(y_t - \mu)(y_{t-k} - \mu)] \quad (1)$$

Onde γ_k naturalmente não depende de t , dado que o valor esperado será o mesmo, independente de t . Assim, a *Função de Autocorrelação*, denominada como FAC (português) ou ACF (inglês) para uma

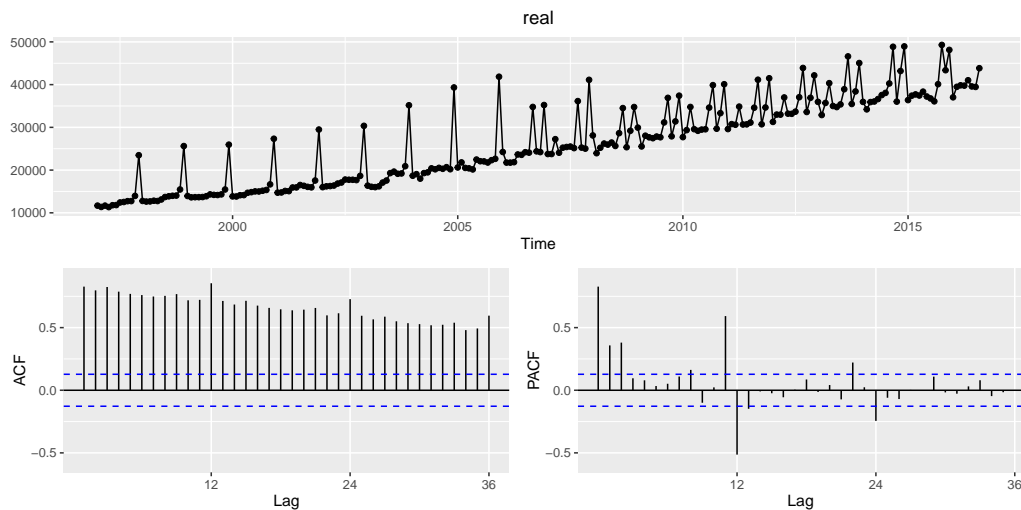
defasagem k qualquer será dada por

$$\rho_k = \frac{\gamma_k}{\sigma^2} \quad (2)$$

A partir da construção dessa função, em termos gráficos, nós podemos obter informações úteis sobre a natureza do processo em questão, verificando o grau de associação entre as defasagens da série.² Pode ser interessante, a propósito, investigar a correlação entre y_t e uma defasagem k específica, ignorando assim o efeito de outras defasagens. Para tal, devemos obter a *função de autocorrelação parcial*.

Para verificar essas funções em termos gráficos, vamos fazer uso da função `ggtsdisplay` do pacote `forecast`.³

```
ggtsdisplay(real[,36])
```



O conjunto dessas informações nos diz que se trata de um processo com memória e com sazonalidade. No apêndice B, a propósito, aplicamos o teste ADF sequencial, proposto em Pfaff (2008), para tentar compreender o processo estocástico subjacente à série. Observa-se que trata-se de um processo trend stationary. Em outras palavras, um processo estacionário ao redor de uma tendência, supostamente linear. Para ilustrar melhor o argumento, podemos, por suposto, caracterizar uma série *tendência-estacionária* como segue:

$$y_t = \beta_1 + \beta_2 t + z_t \quad (3)$$

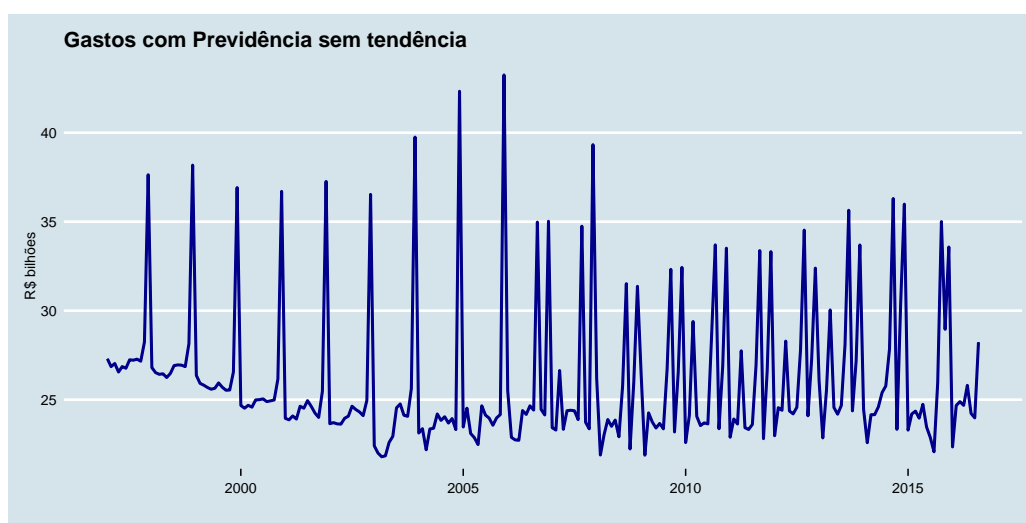
²Da definição, observa-se que $\rho_0 = 1$.

³É o objeto `real`.

Onde $\beta_1 + \beta_2 t$ forma uma tendência determinística e z_t representa um componente estocástico. Isso significa que para tornarmos y_t estacionária, precisamos retirar o componente determinístico, deixando apenas z_t . Fazemos isso com o código abaixo.

```
## Regredir a série contra uma tendência
fit <- lm(real[,36] ~ seq(along = real[,36]))
## Pegar os resíduos da regressão, somados à média da série.
prev <- ts(residuals(fit)+mean(fitted(fit)),
           start=c(1997,01), freq=12)
```

E agora, podemos ver a série sem tendência, em termos gráficos, abaixo.



Podemos, agora, dessazonalizar a série utilizando a função `seas` do pacote `seasonal`, como abaixo.

```
library(seasonal)
Sys.setenv(X13_PATH = "C:/Séries Temporais/R/Pacotes/seas/x13ashtml")
seas <- seas(prev)
```

Agora, podemos comparar essa série, sem tendência e dessazonalizada, com a série original. Fazemos isso, abaixo.

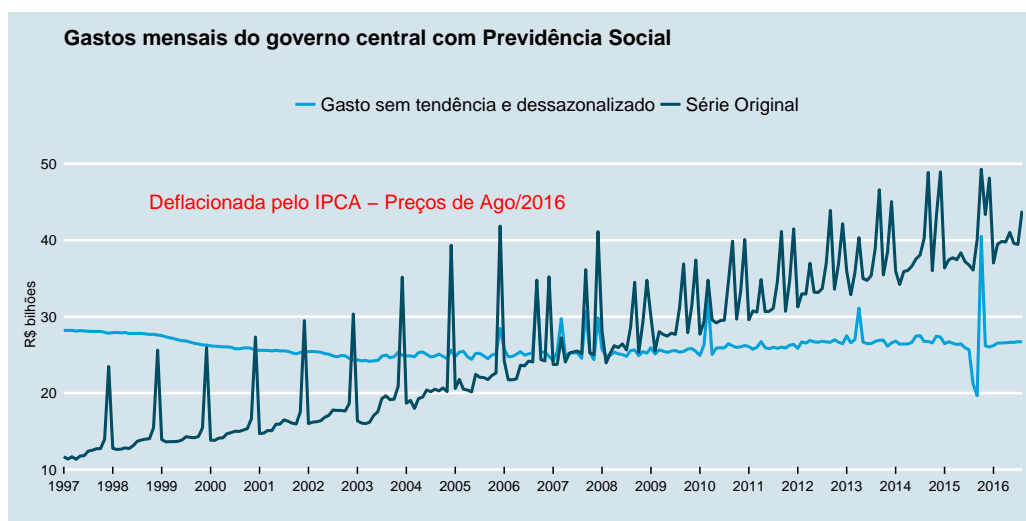
```
series <- ts.intersect(real[,36], final(seas))
colnames(series) <- c('Série Original',
                     'Gasto sem tendência e dessazonalizado')

autoplot(series/1000)+
  geom_line(size=1)+
```

```

xlab('')+ylab('R$ bilhões')+
scale_x_discrete(limits=1997:2016)+
ggtitle('Gastos mensais do governo central com Previdência Social')+
theme_economist() +
scale_color_economist(stata=FALSE) +
labs(colour = "") +
annotate('text', x=2003, y=45,
        label='Deflacionada pelo IPCA - Preços de Ago/2016',
        colour='red', size=5)

```



Bem diferente, não é mesmo? Você pode fazer essa análise para os outros componentes da despesa ou mesmo da receita. Para terminar, vamos agora modelar essa série de gastos com previdência e gerar uma previsão simples.

3.1 Modelagem e previsão

Para gerar um modelo para a série, vamos utilizar a função `auto.arima`, como abaixo.

```

modelo <- auto.arima(real[,36], max.p=4, max.q=4, max.P=4, max.Q=4)

```

Abaixo o modelo estimado.

```

modelo
## Series: real[, 36]

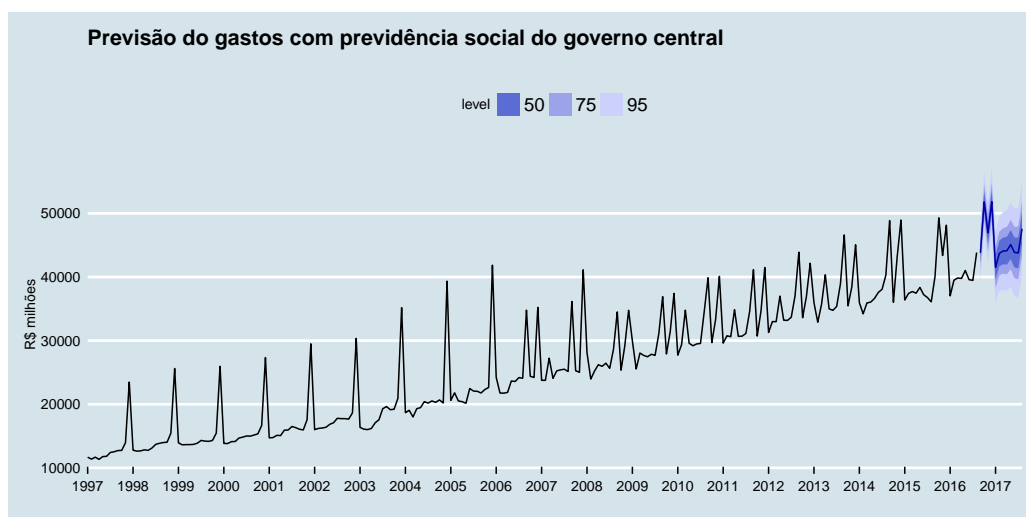
```

```
## ARIMA(3,1,0)(1,0,0)[12]
##
## Coefficients:
##          ar1          ar2          ar3          sar1
##      -0.7113   -0.4784   -0.2797    0.8659
## s.e.    0.0632    0.0718    0.0626    0.0290
##
## sigma^2 estimated as 5171164:  log likelihood=-2156.5
## AIC=4323    AICc=4323.26    BIC=4340.29
```

E agora a previsão.

```
forecast <- forecast(modelo, h=12, level=c(50, 75, 95))
```

Um gráfico é colocado abaixo.



4 Conclusão

A análise desagregada dos gastos primários do governo central mostra que os mesmos apresentam uma clara tendência de crescimento ao longo da amostra – que é de quase 20 anos. São séries *tendência-estacionárias*, i.e., que são estacionárias ao redor de uma tendência, que supomos aqui ser determinística.

Concentramos nossos esforços nos gastos com previdência social, por representarem quase 40% da despesa primária total. Entretanto, a análise parece se estender para as outras rubricas. Naturalmente, o leitor pode testar essa inferência, com base no script dessa edição do [Clube do Código](#).

Esperamos com esse exercício contribuir com um melhor entendimento sobre os gastos públicos no Brasil. Futuras edições do [Clube do Código](#) farão outras experiências com essas séries. Até lá!

A Fluxograma para teste de raiz unitária

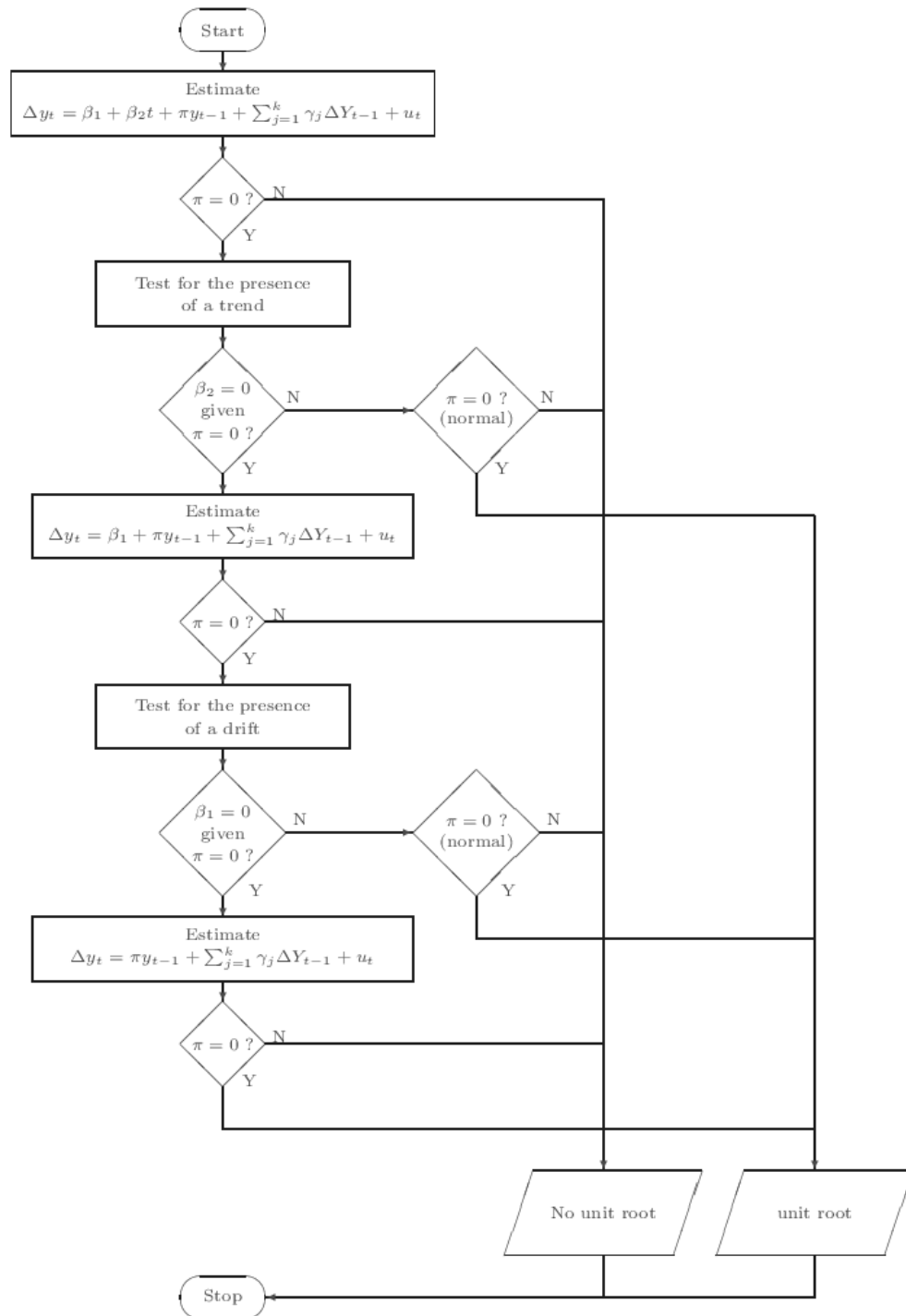


Figura 1: Fluxograma para teste de raiz unitária, conforme Pfaff (2008).

B Teste ADF Sequencial para despesas primárias mensais selecionadas (reais)

Tabela 1: Estatísticas do Teste ADF

	tau3	phi2	phi3	tau2	phi1	tau1
Beneficios_Previd	-11.952	47.796	71.433	-2.856	4.247	-0.422
Pessoal	-10.815	39.022	58.494	-4.903	12.047	-0.844
Outras_Desp_Obrigat	-10.000	33.356	49.997	-5.206	13.579	-2.906
DESPESAS_DISCRICIONAIS	-9.219	28.350	42.501	-5.312	14.131	-1.815

Tabela 2: Valores Críticos do Teste ADF

	1pct	5pct	10pct
tau3	-3.990	-3.430	-3.130
phi2	6.220	4.750	4.070
phi3	8.430	6.490	5.470
tau2	-3.460	-2.880	-2.570
phi1	6.520	4.630	3.810
tau1	-2.580	-1.950	-1.620

Tabela 3: Resultados do Protocolo de Pfaff (2008)

Processo Gerador	
Beneficios_Previd	Estacionário ao Redor de uma Tendência Linear
Pessoal	Estacionário ao Redor de uma Tendência Linear
Outras_Desp_Obrigat	Estacionário ao Redor de uma Tendência Linear
DESPESAS_DISCRICIONAIS	Estacionário ao Redor de uma Tendência Linear

C Teste ADF Sequencial para variação real acumulada em 12 meses de despesas selecionadas

Tabela 4: Estatísticas do Teste ADF

	tau3	phi2	phi3	tau2	phi1	tau1
Beneficios_Previd	-4.452	6.678	10.016	-3.999	7.997	-1.253
Pessoal	-4.868	8.328	12.482	-4.651	10.824	-3.280
Outras_Desp_Obrigat	-4.037	5.763	8.642	-3.920	7.686	-2.249
DESPESAS_DISCRICIONAIS	-5.213	9.217	13.825	-5.271	13.893	-4.557

Tabela 5: Valores Críticos do Teste ADF

	1pct	5pct	10pct
tau3	-3.990	-3.430	-3.130
phi2	6.220	4.750	4.070
phi3	8.430	6.490	5.470
tau2	-3.460	-2.880	-2.570
phi1	6.520	4.630	3.810
tau1	-2.580	-1.950	-1.620

Tabela 6: Resultados do Protocolo de Pfaff (2008)

Processo Gerador	
Beneficios_Previd	Estacionário ao Redor de uma Tendência Linear
Pessoal	Estacionário ao Redor de uma Tendência Linear
Outras_Desp_Obrigat	Estacionário ao Redor de uma Tendência Linear
DESPESAS_DISCRICIONAIS	Estacionário ao Redor de uma Tendência Linear

Referências

Cowpertwait, P. S. P. and Metcalfe, A. V. *Introductory Time Series with R*. Editora Springer, 2009.

Montgomery, D. C.; Jennings, C. L., and Kulahci, M. *Introduction to Time Series and Forecasting*. Wiley, 2008.

Pfaff, B. *Analysis of integrated and cointegrated time series with R*. Springer, New York, second edition, 2008.