

Econometria e Séries Temporais - Aula 12 -

Prof. Mestre. Omar Barroso

Instituto Brasileiro de Educação, Pesquisa e Desenvolvimento

Quebras Estruturais

- Uma quebra estrutural é uma mudança repentina nos parâmetros de um modelo econométrico, como mudanças em intercepções, inclinações ou variância. Pode ocorrer em séries temporais ou dados transversais (cross-section).

Motivos comuns

- Mudanças na política fiscal ou monetária (por exemplo, a introdução de novos impostos, mudanças nas taxas de juros).
- A introdução de tecnologias disruptivas (destruição criativa).
- Choques externos como crises financeiras, guerras, pandemias ou choques no preço do petróleo.
- Mudanças de regime político ou econômico que alteram o funcionamento de uma economia ou mercado.

Tipos de Quebras Estruturais

- **Quebra estrutural única:** Uma mudança única e bem definida no relacionamento entre variáveis.
- **Quebras estruturais múltiplas:** Vários pontos no tempo em que o relacionamento muda, geralmente ocorrendo em dados de séries temporais longas.
- **Quebras estruturais graduais:** Algumas mudanças podem ocorrer gradualmente ao longo do tempo em vez de abruptamente.

Intuição

- **Instabilidade de Parâmetros:** Quebras estruturais desafiam a suposição de constância de parâmetros em modelos econométricos. Modelos econométricos tradicionais assumem que a relação entre variáveis permanece estável ao longo do tempo, mas com quebras estruturais, os parâmetros do modelo podem mudar, tornando essa suposição inválida.

Implicações de ignorar quebras estruturais

- **Especificação incorreta do modelo:** se as quebras estruturais forem ignoradas, o modelo pode fornecer **estimativas tendenciosas** e previsões ruins.
- **Overfitting:** um modelo que não considera quebras pode atribuir **muita importância** a dados passados, levando a erros na compreensão da relação atual entre variáveis.
- **Inferências falsas:** inferências estatísticas feitas sem reconhecer quebras estruturais podem levar a conclusões incorretas sobre a natureza da relação entre variáveis.

Representação Matemática

- τ denota a data de uma quebra e $D_t(\tau)$ representando uma variável binária que indica períodos antes e depois de uma quebra estrutural.

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \delta_1 X_{t-1} + \gamma_0 D_t(\tau) + \gamma_1 [D_t(\tau) \cdot Y_{t-1}] \\ + \gamma_2 [D_t(\tau) \cdot X_{t-1}] + u_t,$$

- Aonde os β s representam a data de quebra estrutural τ .

Representação Matemática

- Aonde os γ s são testados para o teste de "não quebra" pela a H_0 .

$$H_0 : \gamma_0 = \gamma_1 = \gamma_2 = 0$$

- Assim sendo testado com H_a no qual pelo menos um γ é igual a não zero.
- Isso é testado pela estatística F no teste de Chow (1960).

Teste de Quandt(1960) ou QLR

- O teste QLR pode ser usado para testar uma quebra na função de regressão populacional se a data da quebra for desconhecida. A estatística do teste QLR é a maior $F(\tau)$ estatística calculada em um intervalo de datas de interrupção elegíveis $\tau_0 \leq \tau \leq \tau_1$.

$$QLR = \max [F(\tau_0), F(\tau_0 + 1), \dots, F(\tau_1)]$$

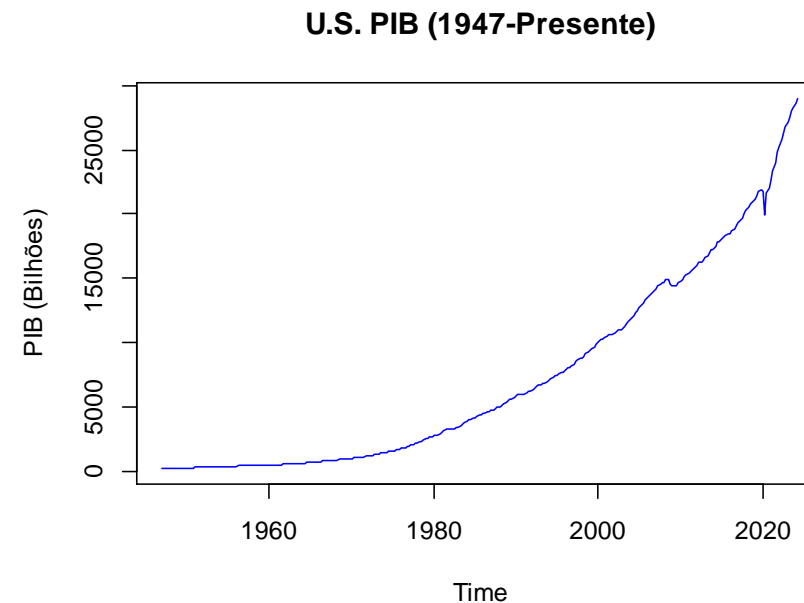
Teste de Quandt(1960)

- O teste QLR pode ser aplicado para testar se um subconjunto dos coeficientes na função de regressão populacional quebra, mas o teste também rejeita se houver uma evolução lenta da função de regressão.
- Quando há uma única quebra discreta na função de regressão populacional que está em uma data dentro do intervalo testado, o teste estatístico QLR é $F(\hat{\tau})$ e $\hat{\tau}:T$ é um estimador consistente da fração da amostra na qual a quebra ocorre.
- Uma distribuição grande do QLR depende em q , ou seja, o número de restrições testadas e ambas as proporções de pontos finais para o tamanho da amostra, $\tau_0:T$ e $\tau_1:T$.

Vamos praticar...

PIB EUA (1947-Presente)

- Ao visualizar a série temporal do PIB EUA, podemos sugerir que nada além do ordinário acontece.
- A tendência exponencial traz um padrão relativamente previsível.
- Porém, podemos confiar apenas no “olhômetro”?



PIB EUA (1947-Presente)

- A seção denominada "*Breakpoints at observation number*" fornece as posições (observações) na série temporal onde o modelo detectou quebras estruturais.
- Por exemplo, para **m = 1**, uma única quebra foi detectada **na observação 212**. Com os dados trimestrais começando em 1947, isso corresponde ao quarto trimestre de 1999, que é mostrado na seção de datas de quebra correspondentes como 1999(4).

```
Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = gdp_ts ~ 1)

Breakpoints at observation number:

m = 1          212
m = 2         185   264
m = 3        148   211 264
m = 4        126 172 218 264
m = 5         80 126 172 218 264

Corresponding to breakdates:

m = 1                                1999(4)
m = 2                        1993(1)      2012(4)
m = 3                1983(4)      1999(3) 2012(4)
m = 4                1978(2) 1989(4) 2001(2) 2012(4)
m = 5      1966(4) 1978(2) 1989(4) 2001(2) 2012(4)

Fit:

m   0           1           2           3           4           5
RSS 1.783e+10 4.028e+09 1.955e+09 1.117e+09 9.192e+08 8.944e+08
BIC 6.430e+03 5.980e+03 5.768e+03 5.606e+03 5.557e+03 5.560e+03
```

PIB EUA (1947-Presente)

- Cada valor adicional de m representa mais uma quebra detectada:
- $m = 2$ sugere dois pontos de quebra nas observações 185 (1993 Q1) e 264 (2012 Q4).
- $m = 3$ adiciona um terceiro ponto de quebra na observação 148 (1983 Q4), e assim por diante.

```
Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = gdp_ts ~ 1)

Breakpoints at observation number:

m = 1          212
m = 2         185    264
m = 3         148    211 264
m = 4         126 172 218 264
m = 5          80 126 172 218 264

Corresponding to breakdates:

m = 1                                1999(4)
m = 2                                1993(1)    2012(4)
m = 3                                1983(4)    1999(3) 2012(4)
m = 4                                1978(2) 1989(4) 2001(2) 2012(4)
m = 5    1966(4) 1978(2) 1989(4) 2001(2) 2012(4)

Fit:

m   0          1          2          3          4          5
RSS 1.783e+10 4.028e+09 1.955e+09 1.117e+09 9.192e+08 8.944e+08
BIC 6.430e+03 5.980e+03 5.768e+03 5.606e+03 5.557e+03 5.560e+03
```

PIB EUA (1947-Presente)

- Para $m = 1$, a única quebra é em 1999(4), significando que o modelo detecta uma mudança estrutural em torno do quarto trimestre de 1999.
- À medida que você aumenta o número de pontos de quebra ($m = 2, m = 3$, etc.), quebras estruturais adicionais são identificadas em momentos diferentes, como 1º trimestre de 1993, 4º trimestre de 1983 e assim por diante.

```
Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = gdp_ts ~ 1)

Breakpoints at observation number:

m = 1          212
m = 2         185   264
m = 3         148   211 264
m = 4         126 172 218 264
m = 5          80 126 172 218 264

Corresponding to breakdates:

m = 1                                1999(4)
m = 2                                1993(1)   2012(4)
m = 3                                1983(4)   1999(3) 2012(4)
m = 4                                1978(2) 1989(4) 2001(2) 2012(4)
m = 5                                1966(4) 1978(2) 1989(4) 2001(2) 2012(4)

Fit:

m   0           1           2           3           4           5
RSS 1.783e+10 4.028e+09 1.955e+09 1.117e+09 9.192e+08 8.944e+08
BIC 6.430e+03 5.980e+03 5.768e+03 5.606e+03 5.557e+03 5.560e+03
```

PIB EUA (1947-Presente)

- **RSS (Residual Sum of Squares):**
- A coluna RSS mostra o quanto da variação nos dados é inexplicada (ou seja, os resíduos) para cada número de quebras. Quanto mais pontos de quebra você adicionar, menor será o RSS, **o que significa que o modelo é melhor em capturar a variabilidade dos dados.**
- Por exemplo, para $m = 0$ (sem quebras), o RSS é $1,783e+10$, mas com um ponto de quebra ($m = 1$), o RSS diminui para $4,028e+09$, indicando um ajuste muito melhor.

```
Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = gdp_ts ~ 1)

Breakpoints at observation number:

m = 1          212
m = 2         185   264
m = 3         148   211 264
m = 4         126 172 218 264
m = 5          80 126 172 218 264

Corresponding to breakdates:

m = 1                                1999(4)
m = 2                                1993(1)   2012(4)
m = 3                                1983(4)   1999(3) 2012(4)
m = 4                                1978(2) 1989(4) 2001(2) 2012(4)
m = 5    1966(4) 1978(2) 1989(4) 2001(2) 2012(4)

Fit:

m   0         1         2         3         4         5
RSS 1.783e+10 4.028e+09 1.955e+09 1.117e+09 9.192e+08 8.944e+08
BIC 6.430e+03 5.980e+03 5.768e+03 5.606e+03 5.557e+03 5.560e+03
```


PIB EUA (1947-Presente)

- **BIC (Bayesian Information Criterion)**
- O BIC é um critério para seleção de modelo que penaliza a complexidade. **Valores mais baixos** do BIC indicam um melhor ajuste do modelo ao considerar o número de parâmetros (pontos de interrupção) sendo adicionados.
- Como você pode ver, o BIC inicialmente diminui conforme você adiciona mais pontos de interrupção (indo de $6.430e+03$ para nenhuma interrupção para $5.557e+03$ para 4 interrupções), mas aumenta ligeiramente para $m = 5$.
- Isso sugere que $m = 4$ (4 pontos de interrupção) é o número ideal de interrupções estruturais, pois fornece o menor valor de BIC.

```
Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = gdp_ts ~ 1)

Breakpoints at observation number:

m = 1          212
m = 2         185    264
m = 3         148    211 264
m = 4         126 172 218 264
m = 5          80 126 172 218 264

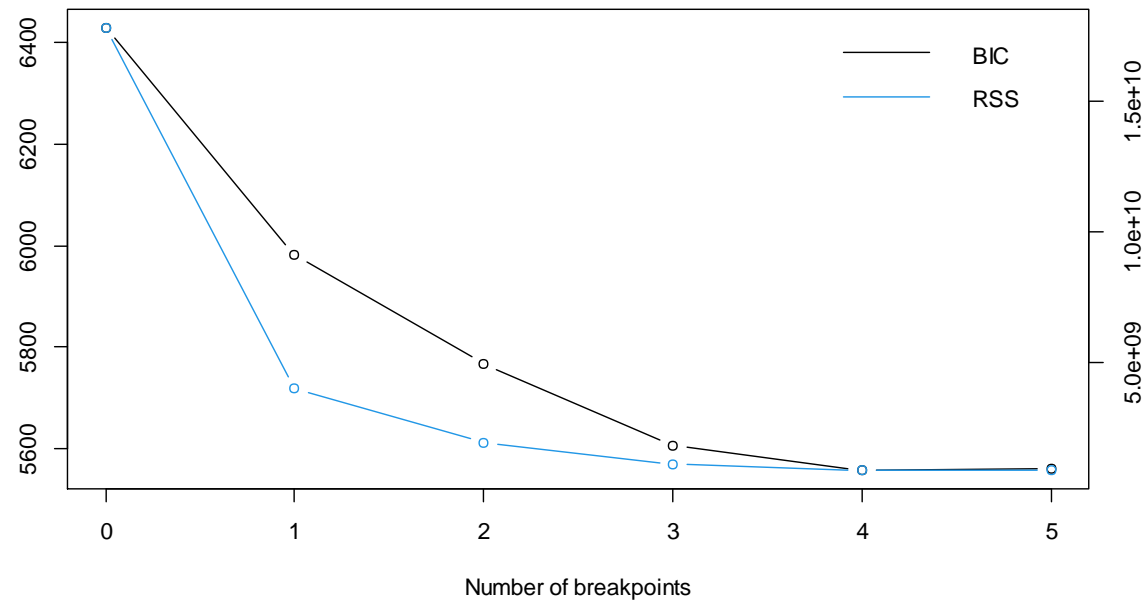
Corresponding to breakdates:

m = 1                                1999(4)
m = 2                        1993(1)      2012(4)
m = 3                1983(4)      1999(3) 2012(4)
m = 4                1978(2) 1989(4) 2001(2) 2012(4)
m = 5    1966(4) 1978(2) 1989(4) 2001(2) 2012(4)

Fit:

m   0          1          2          3          4          5
RSS 1.783e+10 4.028e+09 1.955e+09 1.117e+09 9.192e+08 8.944e+08
BIC 6.430e+03 5.980e+03 5.768e+03 5.606e+03 5.557e+03 5.560e+03
```

Quebras Estruturais PIB Americano (1947-Presente)



PIB EUA (1947-Presente)

- $m = 1$ significa que há uma quebra estrutural detectada por volta do quarto trimestre de 1999, provavelmente correspondendo a uma mudança significativa na tendência do PIB dos EUA por volta daquele período.
- $m = 2$ adiciona outra quebra por volta do primeiro trimestre de 1993 e do quarto trimestre de 2012, indicando mudanças adicionais durante esses períodos.
- $m = 4$ parece ser o modelo ideal de acordo com o BIC, com quebras estruturais por volta do segundo trimestre de 1978, quarto trimestre de 1989, segundo trimestre de 2001 e quarto trimestre de 2012. Esses pontos de quebra podem estar vinculados a eventos econômicos importantes (por exemplo, choques do petróleo, crises financeiras ou mudanças de política).

```
Optimal (m+1)-segment partition:

Call:
breakpoints.formula(formula = gdp_ts ~ 1)

Breakpoints at observation number:

m = 1          212
m = 2         185    264
m = 3        148    211 264
m = 4        126 172 218 264
m = 5         80 126 172 218 264

Corresponding to breakdates:

m = 1                                1999(4)
m = 2                        1993(1)    2012(4)
m = 3                1983(4)    1999(3) 2012(4)
m = 4                1978(2) 1989(4) 2001(2) 2012(4)
m = 5    1966(4) 1978(2) 1989(4) 2001(2) 2012(4)

Fit:

m   0          1          2          3          4          5
RSS 1.783e+10 4.028e+09 1.955e+09 1.117e+09 9.192e+08 8.944e+08
BIC 6.430e+03 5.980e+03 5.768e+03 5.606e+03 5.557e+03 5.560e+03
```

QLR

- O teste Quandt-Andrews é uma forma de teste de flutuação M, que verifica a presença de uma quebra estrutural em um ponto desconhecido nos dados. O teste avalia se os parâmetros de um modelo linear permanecem estáveis em toda a amostra.
- A estatística do teste é 7,3967, que é o valor calculado para o teste. Este número mede a extensão das flutuações nos parâmetros do modelo em todo o período da amostra.
- Quanto maior o valor da estatística do teste, mais forte a evidência de uma quebra estrutural.

```
> quand_test

      M-fluctuation test

data:  gdp_lm
f(efp) = 7.3967, p-value < 2.2e-16

> summary(quand_test)
      Length Class  Mode
statistic 1      -none- numeric
p.value    1      -none- numeric
method     1      -none- character
data.name  1      -none- character
> |
```

QLR

- O valor p é relatado como $< 2,2e-16$, que é extremamente pequeno. No teste de hipóteses, um valor p menor que o limite convencional (geralmente 0,05) indica que podemos rejeitar a hipótese nula.
- Aqui, a hipótese nula é que não há quebra estrutural na série temporal. Dado o valor p extremamente baixo, podemos rejeitar a hipótese nula, o que significa que há evidências muito fortes de uma ou mais quebras estruturais nos dados do PIB dos EUA.

```
> quand_test

      M-fluctuation test

data:  gdp_lm
f(efp) = 7.3967, p-value < 2.2e-16

> summary(quand_test)
      Length Class  Mode
statistic 1      -none- numeric
p.value   1      -none- numeric
method    1      -none- character
data.name 1      -none- character
> |
```

QLR

- O teste Quandt-Andrews confirma que há pelo menos uma quebra estrutural significativa na série do PIB. Este resultado se alinha com as descobertas da função `breakpoints()`, onde vários breakpoints foram identificados (por exemplo, por volta de 1999, 2012, etc.).
- Na prática, o teste não informa onde a quebra está localizada — apenas que existe uma quebra. Você já usou a função `breakpoints()` para localizar as datas específicas das quebras, então você pode combinar esses insights para entender como a estrutura do PIB dos EUA mudou ao longo do tempo.

```
> quand_test

      M-fluctuation test

data:  gdp_lm
f(efp) = 7.3967, p-value < 2.2e-16

> summary(quand_test)
      Length Class  Mode
statistic 1      -none- numeric
p.value    1      -none- numeric
method     1      -none- character
data.name  1      -none- character
> |
```

QLR

- O teste sinaliza fortemente que a relação capturada pelo seu modelo linear do PIB dos EUA não é estável ao longo do período. Quebras estruturais existem, e ignorá-las pode levar a conclusões incorretas.
- Devemos considerar usar um modelo que levem em conta essas quebras (por exemplo, regressão segmentada ou modelos diferentes para períodos diferentes) para melhorar a precisão da nossa análise.

```
> quand_test

      M-fluctuation test

data:  gdp_lm
f(efp) = 7.3967, p-value < 2.2e-16

> summary(quand_test)
      Length Class  Mode
statistic 1      -none- numeric
p.value    1      -none- numeric
method     1      -none- character
data.name  1      -none- character
> |
```

Exercício para praticar

- Escolha a sua série temporal de preferência e teste os códigos para identificar quebras estruturais.

Referências

- Chow, Gregory C. 1960. “Tests of Equality Between Sets of Coefficients in Two Linear Regressions.” *Econometrica* 28 (3): 591–605.
- Quandt, Richard E. 1960. “Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes.” *Journal of the American Statistical Association* 55 (290): 324–30. <https://doi.org/10.1080/01621459.1960.10482067>.