

Zaawansowane metody odsezonowywania szeregów czasowych. TRAMO/SEATS i JDemetra+

Ekonometria Szeregów Czasowych – SGH

Andrzej Torój

Plan wykładu

- 1 Odsezonowanie – ogólne informacje
- 2 Sezonowość w modelach ARIMA: przypomnienie i notacja
- 3 JDemetra+: podstawy
- 4 Metoda TRAMO-SEATS
- 5 Odsezonowanie w praktyce

Plan prezentacji

- 1 Odsezonowanie – ogólne informacje
- 2 Sezonowość w modelach ARIMA: przypomnienie i notacja
- 3 JDemetra+: podstawy
- 4 Metoda TRAMO-SEATS
- 5 Odsezonowanie w praktyce

Prezentacja dostępna pod adresem

<http://web.sgh.waw.pl/~atoroj/>

Zakładka: Ekonometria Szeregów Czasowych

Dekompozycja szeregu czasowego

Elementy SZEREGU

- 1) TREND (I CYKL)
- 2) SEZONOWOŚĆ
- 3) ELEMENT PRZEJŚCIOWY
- 4) ELEMENT NIEREGULARNY

W TRAMO/SEATS:

- TREND i CYKL wyznaczają poziom;
- pozostałe elementy wyrażone jako czynniki albo składniki korygujące ten poziom.

Dodatkowo:

- efekty kalendarzowe;
- obserwacje odstające;
- zmienne interwencyjne i inne egzogeniczne (takie i tylko takie, których wpływ ma być usunięty!).

Rodzaje sezonowości

- **sezonowość deterministyczna**: np. sezonowe zmienne zerojedynkowe w równaniu regresji ($d_{1,t}$, $d_{2,t}$, ..., $d_{s-1,t}$)
- **sezonowość stochastyczna**: wzorzec sezonowy może podlegać ewolucji

Rozróżnienie: test HEGY (*Hylleberg, Engle, Granger, Yoo, 1990*).

TRAMO-SEATS bazuje na modelu SARIMA – implementacja sezonowości stochastycznej (bardziej elastyczny schemat – również w przypadku deterministycznym).

-
- **sezonowość multiplikatywna**: wahania sezonowe silniejsze przy wzroście poziomu szeregu
 - **sezonowość addytywna**: amplituda niezależna od poziomu szeregu

TRAMO-SEATS testuje, czy modelować szereg na poziomach, czy logarytmach (użytkownik może to też zadać według uznania).

Narzędzia

- **X-11 / X-12 / X-13-ARIMA**
 - US Bureau of Census
- **TRAMO-SEATS**
 - V. Gomez i A. Maravall: Bank Hiszpanii
 - **TRAMO**: Time Series **R**egression with **ARIMA** Noise, **M**issing Observations and **O**utliers
 - **SEATS**: Signal **E**xtraction in **ARIMA** Time **S**eries
- **JDemetra+**
 - interfejs korzystający z algorytmów X-13-ARIMA i TRAMO-SEATS (podobnie jak EViews, Gretl, ...)

Źródła wiedzy

- 1 Grudkowska S. (2015), JDemetra+ User Guide.
- 2 Maravall A. (2008), Notes on Programs Tramo and Seats, cz. I, II, III, Banco de Espana.
- 3 Eurostat (2015), ESS Guidelines on Seasonal Adjustment, European Communities.
- 4 Strona programu JDemetra+.

Sposób postępowania w TRAMO-SEATS

- 1 [TRAMO] Zidentyfikuj model SARIMAX dla obserwowalnego szeregu czasowego.
- 2 [SEATS] Zdekomponuj model na nieobserwowalne komponenty i oszacuj wartość tych komponentów (sygnał).

Plan prezentacji

- 1 Odsezonowanie – ogólne informacje
- 2 Sezonowość w modelach ARIMA: przypomnienie i notacja
- 3 JDemetra+: podstawy
- 4 Metoda TRAMO-SEATS
- 5 Odsezonowanie w praktyce

Model AR

AR – ang. *autoregression*, proces autoregresyjny

bieżąca (w okresie t) wartość stacjonarnego szeregu czasowego y_t zależy od p wartości poprzedzających:

$$y_t = c + \phi_1 \cdot y_{t-1} + \phi_2 \cdot y_{t-2} + \dots + \phi_p \cdot y_{t-p} + \varepsilon_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \varepsilon_t$$

Notacja wielomianu charakterystycznego:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) y_t = c + \varepsilon_t$$

Model MA

MA – ang. *moving average*, średnia ruchoma

bieżąca (w okresie t) wartość stacjonarnego szeregu czasowego y_t zależy od q poprzedzających wartości **składnika losowego**:

$$y_t = c + \theta_1 \cdot \varepsilon_{t-1} + \theta_2 \cdot \varepsilon_{t-2} + \dots + \theta_q \cdot \varepsilon_{t-q} + \varepsilon_t = c + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Notacja wielomianu charakterystycznego:

$$y_t - c = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t$$

Model ARMA – połączenie AR i MA

Przy odpowiednich założeniach można proces AR przekształcić do postaci MA(∞) i na odwrót.

W praktyce zadowalające dopasowanie uzyskuje się dzięki kombinacji niewielkiej liczby regresorów typu AR i MA, czyli

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Umieszczenie w modelu regresorów zarówno typu AR, jak i MA, pozwala uzyskać rozsądne przybliżenie, choć identyfikacja parametrów p i q takiego modelu jest trudniejsza.

Notacja wielomianu charakterystycznego:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) y_t = c + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t$$

Model ARIMA

Szeregi niestacjonarne

– tworzymy model ARMA na szeregu zróżnicowanym tyle razy, aby uzyskać jego stacjonarność:

$$\Delta Y_t = Y_t - Y_{t-1}$$

$$\Delta^2 Y_t = \Delta Y_t - \Delta Y_{t-1}$$

Model ARIMA(p, d, q):

$$\Delta^d y_t = c + \sum_{i=1}^p \phi_i \Delta^d y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

ARMA a ARIMA

Model ARMA jest szczególnym przypadkiem modelu ARIMA (z parametrem $d=0$).

Model ARIMAX

ARIMAX – model ARIMA uzupełniony o zestaw egzogenicznych regresorów:

$$\Delta^d y_t = c + \mathbf{x}_t \boldsymbol{\beta} + \sum_{i=1}^p \phi_i \Delta^d y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t$$

W TRAMO $\mathbf{x}_t$:

- zadane przez użytkownika;
- generowane w programie:
 - dni robocze, Wielkanoc;
 - zmienne interwencyjne, obserwacje odstające.

Odwracalność procesu

Proces MA jest odwracalny...

...jeżeli wszystkie pierwiastki wielomianu charakterystycznego (tzn. rozwiązania równania $1 + \theta_1 L + \theta_2 L^2 \dots + \theta_q L^q = 0$ względem L) leżą poza kołem jednostkowym, tzn. są co do modułu > 1 .

- Odwracalny proces $MA(q)$ można przedstawić jako $AR(\infty)$.

Różnicowanie sezonowe

Model ARIMA z sezonowością nazywamy **SARIMA** (*Seasonal* ARIMA).

Dodatkowe parametry (w stosunku do ARIMA):

- s – długość cyklu sezonowości (4 dla danych kwartalnych, 12 dla miesięcznych, 5 lub 7 dla dziennych itp.)
- D – parametr sezonowego różnicowania

Dla $d = 1$:	Ogólnie:
$\Delta_s y_t = y_t - y_{t-s}$ $\Delta_s^2 y_t = \Delta_s y_t - \Delta_s y_{t-s}$ $\vdots$ $\Delta_s^D y_t = \Delta_s^{D-1} y_t - \Delta_s^{D-1} y_{t-s}$	$\Delta_s(\Delta^d y_t) = \Delta^d y_t - \Delta^d y_{t-s}$ $\Delta_s^2(\Delta^d y_t) =$ $\Delta_s(\Delta^d y_t) - \Delta_s(\Delta^d y_{t-s})$ $\vdots$ $\Delta_s^D(\Delta^d y_t) = \Delta_s^{D-1}(\Delta^d y_t) -$ $\Delta_s^{D-1}(\Delta^d y_{t-s})$

Model SARIMA

Dla uproszczenia notacji: niech $y_t^ = \Delta_s^D(\Delta^d y_t)$ oznacza odpowiednio zróżnicowany szereg.*

Oprócz sezonowego różnicowania szeregu wyjściowego, możemy w modelu uwzględnić sezonowe regresory typu **AR** i **MA**.

Model ARIMA (p,d,q) :

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) y_t^* = c + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t$$

Model SARIMA $(p,d,q) \times (P,D,Q)_s$:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) (1 - \phi_1 L^{1 \cdot s} - \phi_2 L^{2 \cdot s} - \dots - \phi_p L^{P \cdot s}) y_t^* = c + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) (1 + \theta_1 L^{1 \cdot s} + \theta_2 L^{2 \cdot s} + \dots + \theta_q L^{Q \cdot s}) \varepsilon_t$$

Funkcja autokorelacji (ACF)

Pokazuje korelację wartości szeregu z kolejnymi opóźnieniami tego samego szeregu:

opóźnienie 1 – r_1

opóźnienie 2 – r_2

opóźnienie 3 – r_3

itd.

Szacujemy na podstawie danych, obliczając współczynniki korelacji liniowej Pearsona.

Współczynnik korelacji cząstkowej

Współczynnik korelacji między i oraz j z wykluczeniem wpływu l :

$$r_{ij.l} = \frac{r_{ij} - r_{il} \cdot r_{jl}}{\sqrt{(1 - r_{il}^2) \cdot (1 - r_{jl}^2)}}$$

Funkcja autokorelacji cząstkowej (PACF)

Pokazuje korelację wartości szeregu z kolejnymi opóźnieniami tego samego szeregu, z wykluczeniem wpływu opóźnień niższego rzędu:

- opóźnienie 1 – r_1 (tak samo jak w ACF)
 - opóźnienie 2 – korelacja cząstkowa wartości bieżącej z 2 opóźnieniem z wykluczeniem wpływu 1 opóźnienia
 - opóźnienie 3 – korelacja cząstkowa wartości bieżącej z 3 opóźnieniem z wykluczeniem wpływu 1 i 2 opóźnień
 - opóźnienie 4 – korelacja cząstkowa wartości bieżącej z 4 opóźnieniem z wykluczeniem wpływu 1, 2 i 3 opóźnień
- itd.

Funkcje ACF i PACF jako kryterium doboru p, q

Sposób postępowania "podpowiadany" przez korelogram:

- dla modeli **AR(p)**: szukamy "**punktu ucięcia**" na wykresie **PACF**
- dla modeli **MA(q)**: szukamy "**punktu ucięcia**" na wykresie **ACF**
- dla modeli **ARMA(p,q)**: zwiększamy stopniowo p i q , starając się "**wyczyścić**" wykres **ACF i PACF**; szukamy:
 - wzorców cyklicznych i nielosowych zmian znaków, niezbieżności, nieusuniętej sezonowości

Uwaga!

Przy poziomie istotności 5%, 1 na 20 współczynników będzie istotny, nawet dla białego szumu!

Plan prezentacji

- 1 Odsezonowanie – ogólne informacje
- 2 Sezonowość w modelach ARIMA: przypomnienie i notacja
- 3 JDemetra+: podstawy**
- 4 Metoda TRAMO-SEATS
- 5 Odsezonowanie w praktyce

Import danych (1)

Krok „zero”: przygotowanie pliku XLS.

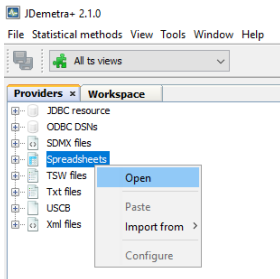
Struktura pliku XLS:

lewa krawędź – daty (w formacie Date), górny wiersz – nazwy szeregów (ograniczenia!), komórka (1,1) – pusta

	A	B	C	D	E	F	G	H	I	J	K	L
1		Textiles, ir	Textiles OC	Préparation	Tissage	DC	Ennoblisse	Autres tex	Industrie d	Vêtements	Articles à	mailles 001563416
2	1990-01-15	411,9	189,9	314,6	190,3	270,7	161,2	871,4	1092,4	384,8		
3	1990-02-15	407,4	184,2	296	184	263,7	157,3	867,1	1104,3	344,9		
4	1990-03-15	417,4	189,3	317,7	198,5	290,4	152,9	868,4	1117,6	319,8		
5	1990-04-15	372,2	185,3	279,8	183	275,3	159,2	757,9	928	383,6		
6	1990-05-15	353,7	184,7	279,3	176,2	274,3	161,1	704,5	870,2	339,7		
7	1990-06-15	389,3	191,8	293,1	184	272,4	168,8	790,3	1004,8	318		
8	1990-07-15	393,2	159,4	208,3	150,4	196	150,9	892,6	1144,1	338,8		
9	1990-08-15	213,8	68,3	107,6	68,9	98	58,3	494,7	642,5	169,4		
10	1990-09-15	386,5	187,9	266,9	181,6	264,4	167,5	788,3	1000	322,2		
11	1990-10-15	402,4	216,1	301,4	208,4	305	193,3	778,7	946,6	409		
12	1990-11-15	354,8	188,5	264,8	182,3	263,3	168,7	692,7	863,8	316,1		
13	1990-12-15	316,7	151,9	200,2	143,1	221,1	137,2	654	820,5	287,3		
14	1991-01-15	399,4	179,4	248,6	170,5	254	161,5	854,4	1054,8	413,5		
15	1991-02-15	387,8	166,1	231,7	161,5	245	146,1	848,3	1071,3	357,3		
16	1991-03-15	396,8	175,4	249,5	170,2	257,2	154,1	849,1	1092,2	313,9		
17	1991-04-15	355,5	176,9	239,5	174,6	272,8	153,1	716,5	856,5	408,3		
18	1991-05-15	317,3	168,7	210,4	152	245,1	154,5	635,1	795,7	281,6		

Krok 1: Klikamy prawym przyciskiem na Spreadsheets (możliwe też alternatywne źródła danych).

Wybieramy Open, a następnie w oknie dialogowym wskazujemy plik źródłowy.



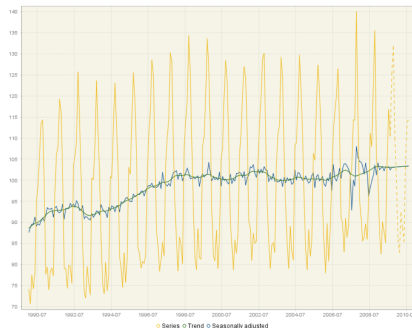
Wybór procedury odsezonowania

[illegible]

Navigation icons: back, forward, search, and other controls.

Kluczowe wyniki

- Zestawienie oryginalnego i odsezonowanego szeregu (**Main results / Charts / Sa, trend**).
- Odsezonowany szereg można zobaczyć w formie tabeli i wyeksportować do XLS (**Main results / Table**; kliknięcie prawym przyciskiem w środku tabeli / Save).



Jakość wyników?



Wyniki testów: pomoc w interpretacji

Według *p-value*:

- good
- uncertain ($<0,1$)
- bad ($<0,05$)
- severe ($<0,01$)
- undefined (przetwarzanie się nie powiodło)
- error (błąd – bezsensowne wyniki przetwarzania)

Wyniki testów: agregacja

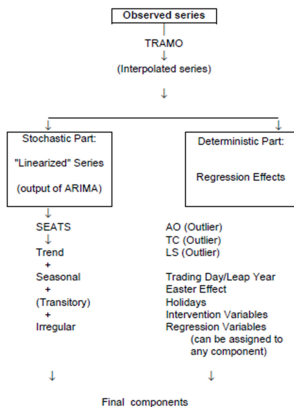
<i>Sum</i>	<i>Rules</i>
Undefined	All diagnostics are Undefined
Error	There is at least 1 error
Severe	There is at least 1 "severe" diagnostic but no error
Bad	No error, no severe diagnostics; the average of the (defined) diagnostics (Bad=1, Uncertain=2, Good=3) is < 1.5
Uncertain	No error, no severe diagnostics; the average of the (defined) diagnostics (Bad=1, Uncertain=2, Good=3) is in $[1.5, 2.5[$
Good	No error, no severe diagnostics; the average of the (defined) diagnostics (Bad=1, Uncertain=2, Good=3) is ≥ 2.5

Plan prezentacji

- 1 Odsezonowanie – ogólne informacje
- 2 Sezonowość w modelach ARIMA: przypomnienie i notacja
- 3 JDemetra+: podstawy
- 4 Metoda TRAMO-SEATS**
- 5 Odsezonowanie w praktyce

Wprowadzenie

TRAMO i SEATS – co robią?



1. TRAMO: dopasowuje proces SARIMAX do szeregu i oczyszcza szereg z wpływu wybranych czynników X (kalendaryzowych, nietypowych...) za pomocą tego modelu.

2. **SEATS**: korzysta ze specyfikacji SARIMA, by zdekomponować tak oczyszczony szereg na elementy: trendu, sezonowości, nieregularny i ew. przejściowy.

$$P, D, Q \leq 2$$

◀ ◻ ▶ ◀ ▢ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ↺ 🔍 ↻

- pomaga osiągnąć stałą wariancję (wraz z korektą obserwacji odstających) – warunek stacjonarności (alternatywa: modelowanie zmiennej wariancji np. GARCH)
- brak skali i procentowa interpretacja różnic
- właściwa, gdy amplituda oscylacji rośnie z wartością szeregu (wahania multiplikatywne)
- interpretacja odch. standardowego reszt jako procentowego błędu prognozy względem poziomu szeregu
- WADA: roczna średnia szeregu wyjściowego $>$ oczyszczonego sezonowo (średnia geometryczna vs arytmetyczna – nierówność Cauchy'ego)

Transformacja logarytmiczna (2)

Specyfikacja (SERIES)	Wyniki (Main results)
TRANSFORMATION FunctionAuto Fct0,95	Series has been log-transformed

- **function**: poziom zmiennej, logarytm naturalny lub wybór automatyczny
 - **fct**>1: obciążenie testu w kierunku poziomów
 - **fct**<1: obciążenie testu w kierunku ln

Transformacja	Model	Efekty deterministyczne (np. kalendarzowe) oraz czynnik sezonowy (S) i nieregularny (I) podawane jako...
brak	addytywny	komponent, który należy dodać do trendu
logarytm	multiplikatywny	czynnik, przez który należy pomnożyć trend

Wykres S-I ratio

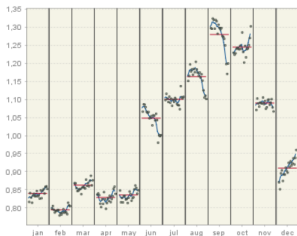


Tabela efektów deterministycznych
(Pre-processing / Pre-adjustment series)

	year	det	cal	onhe	tde	ee	out_i
1-2007	87,829	1,021	1,021	1	1,021	1	1
2-2007	83,508	0,994	0,994	1	0,994	1	1
3-2007	91,122	0,996	0,996	1	0,996	1	1
4-2007	87,417	0,99	0,99	1	0,99	1	1
5-2007	87,437	1,021	1,021	1	1,021	1	1
6-2007	101,969	0,99	0,99	1	0,99	1	1
7-2007	102,961	0,914	0,996	1	0,996	1	0,917
8-2007	111,916	1,021	1,021	1	1,021	1	1
9-2007	118,388	0,965	0,965	1	0,965	1	1
10-2007	137,079	1,09	1,021	1	1,021	1	1,067
11-2007	114,282	1,061	1,014	1	1,014	1	1,046
12-2007	98,325	1,004	0,972	1	0,972	1	1,032
1-2008	89,395	1,044	1,021	1	1,021	1	1,022
2-2008	83,458	1,042	1,026	1	1,026	1	1,016

SI-ratio

- **Main results / S-I ratio:** to wykres grupujący elementy szeregu z tych samych miesięcy, kwartałów itp.:
 - krzywa linia (niebieska): korekta sezonowa trendu
 - prosta linia (czerwona): uśredniona korekta sezonowa
 - punkty: S+I, iloczyn albo suma korekty sezonowej i nieregularnej
 - dwukrotne kliknięcie w obszarze danego miesiąca / kwartału powiększa element wykresu
- interpretacja:
 - wykres pozwala wykryć zmiany strukturalne w schemacie sezonowości (do osobnego modelowania)
 - ...i zauważyć okresy o wyższej zmienności czynnika sezonowego
 - ewolucja akceptowalna, chyba że czynnik sezonowy „mija” 0 (w modelu addytywnym) lub 1 (w multiplikatywnym)

Automatyczna lub ręcznie określona (P to odpowiednik p z prezentacji modeli ARIMA, a BP -- jak *big P* – to odpowiednik P):

ARIMA	
Automatic	☐
Mean	☒
P	0
D	1
Q	1
θ	
BP	0
BD	1
BQ	1
$b\theta$	

W TRAMO należy unikać modeli z $q > p$ lub $p \gg q$

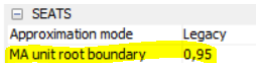
Zwiększa to ryzyko, że w SEATS nie będzie istniała dopuszczalna dekompozycja (zob. dalej).

ARIMA	
Automatic	☒
Accept Default	☐
Cancellation limit	0,05
Initial UR (Diff.)	0,97
Final UR (Diff.)	0,91
Arma limit	1
Reduce CV	0,12
LjungBox limit	0,95
Compare to default	☐

- W automatycznym poszukiwaniu specyfikacji brane pod uwagę:
 - pewne proste, ale dość elastyczne specyfikacje (*Accept Default*);
 - granica, powyżej której pierwiastek wielomianu charakterystycznego AR jest traktowany jako jednostkowy (*UR*);
 - statystyka t dla regresorów (*ARMA limit*);
 - test Ljunga-Boxa dla rest (*LjungBox limit*).

Wielomian MA z założenia odwracalny

- TRAMO wymusza odwracalność części MA procesu ARIMA.
- Powód: stabilność numeryczna obliczeń w skończonych próbach.
- Jeżeli odwrotność pierwiastka wielomianu charakterystycznego zbiega (od dołu) do 1, „zatrzymuje się” na pewnej wartości (default = 0,99, można ją zmienić w oknie specyfikacji: *Decomposition (Seats) / MA unit root boundary*)



Nadmierne różnicowanie (*overdifferencing*)

Niech $y_t = \varepsilon_t$ (biały szum). Przy różnicowaniu $\Delta y_t = (1 - L) y_t$ znajdziemy model MA
 $(1 - L) y_t = (1 - 0.99L) \varepsilon_t$

WNIOSEK: (umiarkowane) nadmierne różnicowanie nie stanowi, z praktycznego punktu widzenia, istotnego problemu w TRAMO. Przy własnoręcznej specyfikacji, zwiększeniu d może towarzyszyć zwiększenie q (parametr θ_q może, ale nie musi okazać się nieistotny).

Trading days: osobne efekty od poniedziałku do soboty.

Working days: efekt dla dni roboczych (vs dni świąteczne i weekendowe).

Leap year: efekt roku przestępnego

Easter: efekt wielkanocny

Można samemu zdefiniować
dodatkowe zmienne.

Trading days effects (7 variables)
No easter effect

	Coefficients	T-Stat	P[T > t]
Monday	0,0019	0,81	0,4179
Tuesday	0,0096	4,21	0,0000
Wednesday	0,0133	5,77	0,0000
Thursday	0,0032	1,39	0,1653
Friday	0,0033	1,47	0,1419
Saturday	-0,0136	-6,04	0,0000

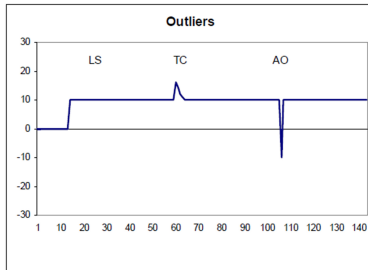
Leap year

	Coefficients	T-Stat	P[T > t]
	0,0261	3,85	0,0002

Obserwacje odstające

- **AO**: Additive Outlier, $[0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0]$
- **TC**: Transitory Change, $[0 \ 0 \ 0 \ 0 \ 1 \ \delta \ \delta^2 \ \delta^3 \ \dots \approx 0]$ – gdy $\delta = 1$ efekt skrajnie uporczywy (LS), gdy $\delta = 0$ natychmiast przemija (AO)
- **LS**: Level Shift, $[0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1 \ 1]$

OUTLIERS	
Is enabled	☒
Use default critical value	☒
Critical value	3,5
Detection span	All
Additive	☒
Level shift	☒
Transitory change	☒
Seasonal outlier	☒
EML estimation	☐
TC rate	0,7



Własności procedury automatycznej identyfikacji

- cel: usunięcie odchyłeń od normalnego rozkładu reszt, pozornych efektów w ACF, obciążenia oszacowań
- NIEZNANA LOKALIZACJA *ex ante*! (tym obserwacje odstające się różnią od zmiennych interwencyjnych) – ale jeżeli jest uzasadnienie *ex post*, lepiej wprowadzić zmienną interwencyjną
- optymalnie do 2%, maksymalnie do 5% liczby obserwacji w szeregu (trzeba zwrócić na to uwagę po fakcie...)
- możliwość sterowania wartością krytyczną testu na obecność outliera: im wyższa, tym trudniej go stwierdzić; default = 3,5, w praktyce od 3 (krótki szereg, <50) do 4 (długi szereg, >450)
 - statystyka testowa: *Maravall, cz. II*, s. 93

Własności procedury automatycznej identyfikacji

- cel: usunięcie odchyłeń od normalnego rozkładu reszt, pozornych efektów w ACF, obciążenia oszacowań
- NIEZNANA LOKALIZACJA *ex ante*! (tym obserwacje odstające się różnią od zmiennych interwencyjnych) – ale jeżeli jest uzasadnienie *ex post*, lepiej wprowadzić zmienną interwencyjną
- optymalnie do 2%, maksymalnie do 5% liczby obserwacji w szeregu (trzeba zwrócić na to uwagę po fakcie...)
- możliwość sterowania wartością krytyczną testu na obecność outliera: im wyższa, tym trudniej go stwierdzić; default = 3,5, w praktyce od 3 (krótki szereg, <50) do 4 (długi szereg, >450)
 - statystyka testowa: *Maravall, cz. II*, s. 93

Własności procedury automatycznej identyfikacji

- cel: usunięcie odchyłeń od normalnego rozkładu reszt, pozornych efektów w ACF, obciążenia oszacowań
- NIEZNANA LOKALIZACJA *ex ante*! (tym obserwacje odstające się różnią od zmiennych interwencyjnych) – ale jeżeli jest uzasadnienie *ex post*, lepiej wprowadzić zmienną interwencyjną
- optymalnie do 2%, maksymalnie do 5% liczby obserwacji w szeregu (trzeba zwrócić na to uwagę po fakcie...)
- możliwość sterowania wartością krytyczną testu na obecność outliera: im wyższa, tym trudniej go stwierdzić; default = 3,5, w praktyce od 3 (krótki szereg, <50) do 4 (długi szereg, >450)
 - statystyka testowa: *Maravall, cz. II*, s. 93

Własności procedury automatycznej identyfikacji

- cel: usunięcie odchyłeń od normalnego rozkładu reszt, pozornych efektów w ACF, obciążenia oszacowań
- NIEZNANA LOKALIZACJA *ex ante*! (tym obserwacje odstające się różnią od zmiennych interwencyjnych) – ale jeżeli jest uzasadnienie *ex post*, lepiej wprowadzić zmienną interwencyjną
- optymalnie do 2%, maksymalnie do 5% liczby obserwacji w szeregu (trzeba zwrócić na to uwagę po fakcie...)
- możliwość sterowania wartością krytyczną testu na obecność outliera: im wyższa, tym trudniej go stwierdzić; default = 3,5, w praktyce od 3 (krótki szereg, <50) do 4 (długi szereg, >450)
 - statystyka testowa: **Maravall, cz. II**, s. 93

Przebieg procedury automatycznej identyfikacji

- ❶ Oszacowanie modelu ARIMAX (EML lub HR).
- ❷ Test na obecność obserwacji nietypowej każdego z 3 typów dla reszt modelu we wszystkich okresach.
 - nie znaleziono outlierów w pierwszym przebiegu: STOP.
 - nie znaleziono outlierów, ale nie był to pierwszy przebieg: IDŹ DO 3.
 - znaleziono outliery: uzupełnij model o zmienne egzogeniczne i IDŹ DO 1.
- ❸ Regresja ze wszystkimi dodanymi zmiennymi zerojedynkowymi.
 - są nieistotne outliery: usuń ten z najniższą wartością statystyki i IDŹ DO 1.
 - brak nieistotnych outlierów: STOP.

Efekty procedury identyfikacji obserwacji odstających

Pre-processing:

Summary

Estimation span: [1-1990 - 8-2009]

236 observations

Series has been log-transformed

Trading days effects (7 variables)

No easter effect

3 detected outliers

Outliers

	Coefficients	T-Stat	P[T > t]
TC (10-2007)	0,0906	5,99	0,0000
AO (7-2006)	0,0716	4,69	0,0000
LS (7-2007)	-0,0550	-4,12	0,0001

Estymacja modelu ARIMA

Exact ML:

- TAK: Exact Maximum Likelihood (można dobrać parametr precyzji w procedurze optymalizacyjnej dla funkcji wiarygodności)
- NIE: metoda Hannana-Rissanena

Model ARIMA zawiera nieobserwowalne regresory (część MA). Dlatego:

- zapisanie (stacjonarnego) modelu w przestrzeni stanów ([Akaike, 1974](#))
- maksymalizacja funkcji wiarygodności wyprowadzonej przy pomocy filtru Kalmana ([Mélard, 1984](#))

☐	ESTIMATE	
☐	Model span	All
	Tolerance	0,0000001
	Exact ML	☒
	Unit root limit	0,96

Metoda Hannana-Rissanena

Szacujemy parametry modelu ARMA(p,d):

$$\Phi(L)y_t^* = \Theta(L)\varepsilon_t$$

$$\Theta(L)^{-1}\Phi(L)y_t^* = \varepsilon_t$$

Efekt dzielenia wielomianów (wymuszony odwracalny proces MA!):

$$\Pi(L)y_t^* = \varepsilon_t$$

Wielomian $\Pi(L)$ jest nieskończony, ale przy odwracalnym procesie MA jego parametry zbiegają do zera w nieskończoności. Dlatego można go przybliżyć skończonym wielomianem $\Pi^n(L)$. Wtedy parametry równania mogą zostać oszacowane za pomocą MNK, co pozwoli uzyskać oszacowania reszt $\hat{\varepsilon}_t$.

W drugim kroku *Exact ML* oszacowania te traktujemy jako obserwowalne regresory do części MA.

- szybkość obliczeń, przypadki niestabilne numerycznie;
- uzyskanie wartości startowych do estymacji.

Kryteria informacyjne

Możliwe porównanie modeli, o ile rozważamy konkurencyjne (im niższe wartości przyjmują kryteria, tym lepiej):

$$AIC = \ln \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 + \frac{2k}{T}$$

$$BIC = \ln \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 + \frac{k \cdot \ln(T)}{n}$$

Pre-processing:

AIC = 939.4917562942435

AICC = 941.5109870634742

BIC (corrected for length) = -7.599638092857192

Testy reszt (2) – autokorelacja / losowość

- **Ljung-Box (Box-Pierce)** – H_0 : brak autokorelacji reszt do rzędu 24 włącznie
- **Ljung-Box (Box-Pierce) on seasonality** – H_0 : brak sezonowej autokorelacji reszt do rzędu 2 włącznie
- **Randomness** – różne warianty „testu serii” (H_0 : losowy charakter reszt)

Test Ljunga-Boxa

H_0 : nie ma autokorelacji do rzędu P (wariant sezonowy: $P \cdot s$) włącznie

H_1 : występuje autokorelacja rzędu od 1 do P (wariant sezonowy: $P \cdot s$)

$$Q = T(T+2) \sum_{j=1}^P \left(\frac{1}{T-j} \cdot \frac{\sum_{t=j+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}}{\underbrace{\sum_{t=1}^T \hat{\varepsilon}_t^2}_{\rho_j^2}} \right)$$

$$Q_s = T(T+2) \sum_{j=1}^P \left(\frac{1}{T-js} \cdot \rho_{js}^2 \right)$$

Testy reszt (3) – autokorelacja / losowość

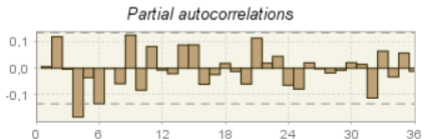
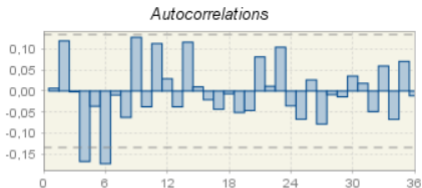
2. Independence of the residuals

	P-value
Ljung-Box(24)	0,0382
Box-Pierce(24)	0,0588
Ljung-Box on seasonality(2)	0,9040
Box-Pierce on seasonality(2)	0,9097

Durbin-Watson statistic: 1.9717

3. Randomness of the residuals

	P-value
Runs around the mean: number	0,6555
Runs around the mean: length	1,0000
Up and Down runs: number	1,0000
Up and Down runs: length	1,0000



Inne narzędzia diagnostyczne

- **Linearity** – H_0 : liniowy charakter modelu (liniowa kombinacja normalnych reszt)
 - implementacja Ljunga-Boxa dla kwadratów reszt
 - służy wykrywaniu efektów ARCH / GARCH
- **Seasonality** – wykrywanie sezonowych wzorców pozostałych w resztach (**Diagnostics / Seasonality tests / Full residuals**)

Full residuals

Summary

Test	Seasonality
1. Auto-correlations at seasonal lags	NO
2. Friedman (non parametric)	NO
3. Kruskal-Wallis (non parametric)	NO
4. Spectral peaks	NO
5. Periodogram	NO
6. Seasonal dummies	NO
6bis. Seasonal dummies (AMI)	NO

- Wszystkie trzy (albo cztery) komponenty szeregu czasowego są ortogonalne. Mogą być modelowane za pomocą osobnych procesów SARIMA:
 - komponent nieregularny koncentruje w sobie „szum”, wielomiany charakterystyczne: $AR=1$, $MA=1$;
 - wielomiany charakterystyczne AR i MA całego procesu iloczynem odpowiednich wielomianów dla komponentu trendowego, sezonowego i przejściowego.
- Patrząc z drugiej strony: iloczyn wielomianów charakterystycznych dla składowych to proces SARIMA dla całego szeregu.
 - A zatem tworzymy składowe, przypisując im poszczególne pierwiastki wielomianów charakterystycznych znalezione przez TRAMO modelu SARIMA.

Reprezentacja szeregu za pomocą wahań sinusoidalnych

- Każdy szereg czasowy x_t ($t = 1, \dots, T$) możemy „objaśnić” (stuprocentowo) za pomocą układu T równań z T niewiadomymi a_i :

$$x_t = a_0 + a_1 t + a_2 t^2 + \dots + a_{T-1} t^{T-1}$$
- Dla każdej obserwacji x_t możemy zapisać, przy znanych a_i i x_t :

$$t^{T-1} + \frac{a_{T-2}}{a_{T-1}} t^{T-2} + \dots + \frac{a_2}{a_{T-1}} t^2 + \frac{a_1}{a_{T-1}} t + \frac{a_0 - x_t}{a_{T-1}} = 0$$
- Powyższy zapis jest tożsamy z równaniem charakterystycznym liniowego równania różnicowego o rzeczywistych współczynnikach. Rozwiązanie takiego równania dla każdego x_j jest postaci:

$$A_j \cos(\omega_j t) + B_j \sin(\omega_j t)$$

Spektrum

- Okres wahań tych funkcji to 2π . ω_j możemy tu interpretować jako częstotliwość wahań (w radianach na jednostkę czasu).

Rozważmy następujący zbiór częstotliwości:

$\omega_j = \frac{2\pi}{T} \cdot j$, $j = 1, \dots, \frac{T}{2}$ (raz w szeregu, dwa razy w szeregu, ..., co druga obserwacja w szeregu), $\omega_j \in [0; \pi]$

- Wówczas szereg możemy zaprezentować jako sumę takich rozwiązań:

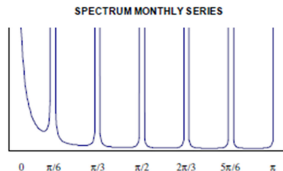
$$x_t = \sum_{j=1}^{T/2} (A_j \cos(\omega_j t) + B_j \sin(\omega_j t))$$

Spektrum

Funkcję $A_j^2 + B_j^2$ możemy interpretować jako **kontrybucję częstotliwości ω_j do wariancji całego szeregu**.

- Wahania o częstotliwości „co każdą obserwację” mają częstotliwość 2π (okres funkcji $\sin$ i $\cos$).
- Oczywiście nie możemy ich zaobserwować, bo musielibyśmy wiedzieć, co się dzieje „między” obserwacjami. Dla nas dostrzegalne są najwyżej wahania „co drugą obserwację” (częstotliwość $\frac{2\pi}{2} = \pi$).
- W danych sezonowych mogą nas np. interesować wahania:
 - „co 12-tą obserwację”, czyli miesięczne: częstotliwość $\frac{2\pi}{12} = \frac{\pi}{6}$
 - „co 4-tą obserwację”, czyli kwartalne: częstotliwość $\frac{2\pi}{4} = \frac{\pi}{2}$
- W wahaniach miesięcznych mogą być „ukryte” wahania o wyższej częstotliwości, będącej wielokrotnością $\frac{\pi}{6}$: $\frac{2\pi}{6}$, $\frac{3\pi}{6}$, $\frac{4\pi}{6}$, $\frac{5\pi}{6}$ (*harmonics*). Stąd w danych:
 - miesięcznych: $n \cdot \frac{\pi}{6}$ – częstotliwości sezonowe
 - kwartalnych: $n \cdot \frac{\pi}{2}$ – częstotliwości sezonowe

-
- The plot, titled "SPECTRUM QUARTERLY SERIES", shows a spectral density function. The horizontal axis is labeled with 0, $\pi/2$, and π . A sharp, high-amplitude peak is centered at $\pi/2$. The curve starts at a high value near 0, decreases rapidly to the baseline, and then rises sharply to the peak at $\pi/2$ before returning to the baseline. A vertical dashed line is drawn at π .



Klasyfikacja pierwiastków: przykład (1)

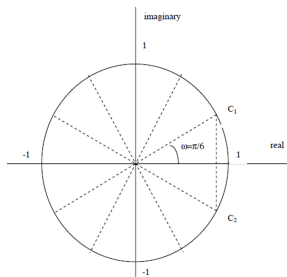
Rozważmy wielomian charakterystyczny dla sezonowych różnic przy danych miesięcznych:

$$x_t - x_{t-12}$$

$$(1 - L^{12}) x_t$$

Pierwiastki rozwiązaniami równania: $1 - L^{12} = 0$

12 pierwiastków o module jednostkowym: $L = 1^{\frac{1}{12}}$



Klasyfikacja pierwiastków: przykład (2)

Można rozłożyć $(1 - L^{12}) x_t$ na czynniki:

$$(1 - L^{12}) x_t =$$

$$\underbrace{(1 - L)}_{\text{trend}} \underbrace{\left(1 - \sqrt{3}L + L^2\right)}_{1/\text{rok} = \frac{\pi}{6}} \underbrace{\left(1 - L + L^2\right)}_{2/\text{rok} = \frac{2\pi}{6}} \underbrace{\left(1 + L^2\right)}_{3/\text{rok} = \frac{3\pi}{6}} \underbrace{\left(1 + L + L^2\right)}_{4/\text{rok} = \frac{4\pi}{6}} \underbrace{\left(1 + \sqrt{3}L + L^2\right)}_{5/\text{rok} = \frac{5\pi}{6}} \underbrace{(1 + L)}_{6/\text{rok} = \frac{6\pi}{6}}$$

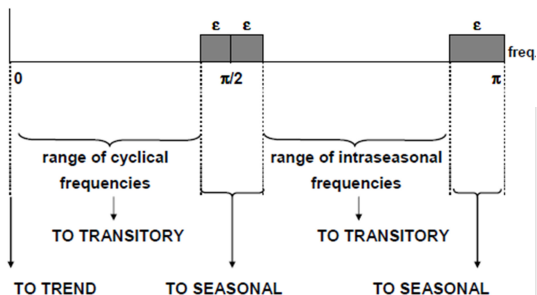
sezonowość

W ten sposób SEATS „rozbija” zidentyfikowany model SARIMA na komponenty

SEATS

Klasyfikacja pierwiastków: ogólna zasada

Example: Quarterly data



SEATS	
Approximation mode	Legacy
MA unit root boundary	0,95
Trend boundary	0,5
Seasonal tolerance	2
Seasonal boundary	0,8
Seas. boundary (unique)	0,8
Method	Burman

Parametry dekompozycji

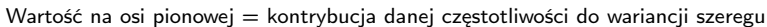
$\Phi(L) = \Phi_p(L) \Phi_s(L) \Phi_c(L)$ – pierwiastki AR są przypisywane stosownie do częstotliwości, jakie generują

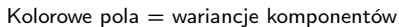
- Decomposition (Seats) / **Trend boundary**: granica, powyżej której odwrotność pierwiastka AR jest przypisywana do trendu (poniżej: komponent przejściowy)
- Decomposition (Seats) / **Seasonal tolerance**: odległość w spektrum od częstotliwości sezonowych (w stopniach), poniżej której pierwiastek jest przypisywany do komponentu sezonowego

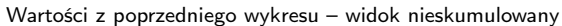
Gdy $q > p$, zawsze pojawi się komponent przejściowy, nawet jeżeli żaden pierwiastek AR nie zostanie do niego przypisany. Obecność stacjonarnego (z definicji) komponentu przejściowego służy temu, by zawrzeć w nim struktury korelacyjne szeregu, które nie zostały zawarte w trendzie i sezonowości, tak aby komponent nieregularny był białoszumowy.

Problem identyfikacji w SEATS (1)

- Kompletny opis trzech (czterech) procesów składowych wymaga jeszcze podania wariancji ich składników losowych.
- Są one nieidentyfikowalne bez dodatkowych założeń:
 - manipulując wariancjami składników losowych, spektrum każdego komponentu możemy przesuwać równolegle w górę i w dół
 - jeżeli inny element przesuniemy adekwatnie w drugą stronę, spektrum całego procesu nie zmieni się
- **Założenie identyfikujące: wariancja komponentu nieregularnego maksymalizowana kosztem trendu i sezonowości.**
 - Przesuwamy spektrum komponentu trendowego, przejściowego i sezonowego jak najniżej (tzn. aż zetknie się z zerem).
 - Wariancję komponentu nieregularnego przesuwamy odpowiednio wyżej.







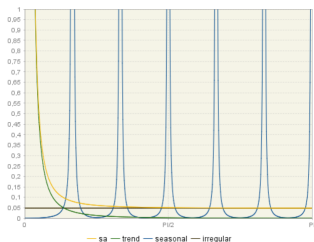
A set of small navigation icons typically found in Beamer presentations, including symbols for back, forward, search, and other slide controls.

Odpowiednie odsunięcie czwartej linii od zera, aby czarna linia (= suma) pozostała
niezmieniona



Spektrum teoretycznych komponentów

Decomposition / WK analysis / Components



Z założeń identyfikacji wynika:

- spektrum trendu i sezonowego komponentu „dotyka” 0 (możliwie najniższa wariancja)
- „poziome” (z definicji) spektrum komponentu nieregularnego maksymalnie wysoko (możliwie najwyższa wariancja)

SEATS: efekt dekompozycji

Decomposition:

Model

D: 1,00000 - B - B¹² + B¹³

MA: 1,00000 - 0,633941 B - 0,376748 B¹² + 0,238836 B¹³

sa

D: 1,00000 - 2,00000 B + B²

MA: 1,00000 - 1,57227 B + 0,599582 B²

Innovation variance: **0,48441**

trend

D: 1,00000 - 2,00000 B + B²

MA: 1,00000 + 0,0767905 B - 0,923210 B²

Innovation variance: **0,01532**

seasonal

D: 1,00000 + B + B² + B³ + B⁴ + B⁵ + B⁶ + B⁷ + B⁸ + B⁹ + B¹⁰ + B¹¹

MA: 1,00000 + 0,817986 B + 0,571254 B² + 0,308266 B³ + 0,0619695 B⁴ - 0,146922 B⁵ - 0,307337

B⁶ - 0,415761 B⁷ - 0,474602 B⁸ - 0,491043 B⁹ - 0,476534 B¹⁰ - 0,447277 B¹¹

Innovation variance: **0,11538**

irregular

Innovation variance: **0,30459**

Jak oszacować komponent sezonowy?

$$\begin{aligned}\Phi^*(L)y_t &= \Theta(L)\varepsilon_t, \varepsilon_t \sim (0; V) & y_t &= \Psi(L)\varepsilon_t \\ \Phi_s(L)s_t &= \Theta_s(L)\varepsilon_t^s, \varepsilon_t^s \sim (0; V_s) & s_t &= \Psi^s(L)\varepsilon_t^s\end{aligned}$$

$$\hat{s}_t = \underbrace{\frac{V_s}{V} \cdot \frac{\Psi^s(L)\Psi^s(F)}{\Psi(L)\Psi(F)}}_{v_s(L,F)} y_t$$

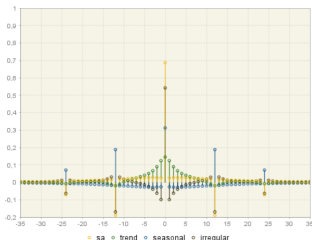
gdzie $F^i = L^{-i}$ (operator przesunięcia w przyszłość)

W mianowniku wielomiany MA całego modelu (odwracalne) i AR części sezonowej (skrócą się z licznikiem, bo tam wielomiany AR całego modelu będące iloczynem wielomianów sezonowych i niesezonowych). Stąd: zbieżność wag w nieskończoności i filtr może być „obcięty” dla potrzeb analizy skończonej próby.

Ciąg liniowych współczynników $v_s(L, F)$ przykładanych do przyszłych, bieżącej i przeszłych realizacji szeregu w celu znalezienia jego sezonowego komponentu nazywamy **filtrem Wienera-Kołmogorowa**.

- Analogicznie wyznaczamy pozostałe komponenty.
- Powinien być zbieżny (wynika z wymuszonej odwracalności części MA), scentrowany w 0 i symetryczny.

Decomposition / WK analysis / Final estimators / WK filters



Estymator różni się od prawdziwego procesu
zależnością od przyszłych obserwacji.

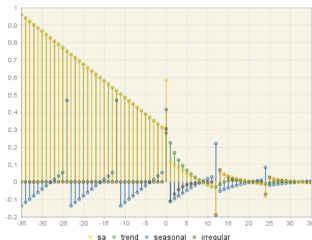
Wprowadza do szacowanego komponentu dodatkowe struktury korelacyjne. Dlatego nie warto odsezonowywać szeregu po to, by później modelować go za pomocą technik ARIMA (w szczególności AR).

Wagi PsiE

Filtr WK obejmuje przyszłe obserwacje, więc wagi mogłyby dać wyobrażenie o skali i horyzoncie możliwych rewizji odsezonowanego szeregu. Jednak w takim ujęciu ta skala byłaby przeszacowana, gdyż w modelu SARIMA o przyszłych obserwacjach sporo już wiadomo (prognoza z modelu SARIMA).

Dlatego w tym celu wyrazić $\hat{s}_t$ jako funkcję przeszłych i przyszłych innowacji ε_t .

$$\hat{s}_t = v_s(L, F) \cdot y_t = v_s(L, F) \frac{\Theta(B)}{\Phi(B)} \cdot \varepsilon_t = \underbrace{\xi_s(B, F)}_{\text{PsiE-weights}} \varepsilon_t$$



Decomposition / WK analysis / Final estimators / PsiE-weights

Niesymetryczne, zbieżne w przyszłości, rozbieżne w przeszłości dla modelu niestacjonarnego.

Istnienie dopuszczalnej dekompozycji

Dopuszczalna dekompozycja...

...gdy spektrum wszystkich komponentów jest **nieujemne**. Nie zawsze istnieje. Gdy model znaleziony przez TRAMO nie implikuje dopuszczalnej dekompozycji, można próbować ją aproksymować:

SEATS

Approximation mode

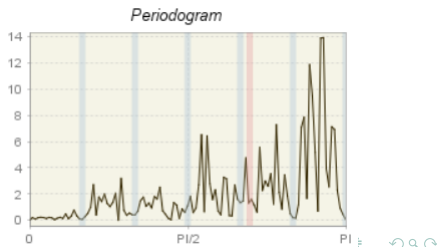
Legacy

Diagnostyka spektrum oszyszczonego sezonowo (1)

Diagnostics / Spectral

analysis: możliwa własna ocena wizualna, czy spektrum szeregu oczyszczonego sezonowo zostało pozbawione maksimów lokalnych dla częstotliwości sezonowych

- **niebieskie linie:** częstotliwości związane z sezonowością
- **czerwona linia:** częstotliwość związana z dniami roboczymi



Diagnostyka spektrum oszyszczonego sezonowo (2)

Main results: wybrane formalne testy

- **visual spectral analysis**: identyfikacja wierzchołków na wykresie spektrum dla (i) komponentu sezonowego i (ii) komponentu liczby dni roboczych
- **regarima spectral peaks**: przy założeniu normalności rozkładu reszt z modelu ARIMA – możliwe wyprowadzenie testów istotności wokół pewnych częstotliwości w spektrum (sezonowość, dni robocze) – odrzucenie H_0 sugeruje, że sezonowości i efektu dni roboczych nie udało się w pełni usunąć

visual spectral analysis

spectral seas peaks: Good (0,000)

spectral td peaks: Good (0,000)

regarima residuals

normality: Good (0,746)

independence: Uncertain (0,068)

spectral td peaks: Good (0,465)

spectral seas peaks: Uncertain (0,030)

Plan prezentacji

- 1 Odsezonowanie – ogólne informacje
- 2 Sezonowość w modelach ARIMA: przypomnienie i notacja
- 3 JDemetra+: podstawy
- 4 Metoda TRAMO-SEATS
- 5 Odsezonowanie w praktyce

ESS Guidelines on SA

- Dokument Eurostatu, „kodeks dobrych praktyk” odsezonowania.
- Mają zapewnić wyższą jakość i międzynarodową porównywalność odsezonowanych szeregów.
- Wskazanie na potrzebę jednolitej metodologii i oprogramowania.

ESS Guidelines dotyczą następujących obszarów:

wstępna obróbka danych (analiza graficzna, dostosowanie kalendarzowe, obserwacje odstające, wybór modelu); **wyrównanie sezonowe** (wybór metody, spójność danych surowych i SA); **rewizje** (zasady; horyzont); **diagnostyka jakości**; **szeregi krótkie i problematyczne**

Badania Eurostatu

Za pomocą automatycznej procedury TRAMO-SEATS zbadano 13227 szeregów:

- 60% szeregów multiplikatywnych (log), 40% addytywnych (level)
- 50% szeregów: RSA0
- 87% szeregów: dobre dopasowanie, 11% akceptowalne, 2% złe
- 15% szeregów stacjonarnych, 2% wymagało $d > 1$
- 50% bez outlierów, 27% z jednym, 23% bez outlierów

Zadanie 1

- ❶ Zaimportuj do Demetra+ swój szereg czasowy (bez ostatniej obserwacji).
- ❷ Spróbuj samodzielnie znaleźć dla niego najlepszy model ARIMA. Porównaj z modelem znalezionym przez Demetrę+.
- ❸ Czy w szeregu występują istotne efekty kalendarzowe?
- ❹ Czy w szeregu występują istotne efekty sezonowe? Omów ich charakter. Czy sezonowość jest stabilna?
- ❺ Skopiuj wyrównany sezonowo szereg do Excela. Oceń jego jakość i ewentualne kierunki poprawek.
- ❻ Skopiuj do Excela historię rewizji sezonowego komponentu dla 10. obserwacji od końca i komponentu trendu dla ostatniej obserwacji z roku 2008.
- ❼ Porównaj jakość uzyskanego modelu ze specyfikacją RSA0.
- ❽ Powtórz odsezonowanie modelu w Excelu. Uzupełnij szereg o ostatnią obserwację i przeprowadź rewizję.

Zadanie 2

Za pomocą metody TRAMO-SEATS oczyszczono z sezonowości miesięczny szereg cen tekstyliów we Francji w okresie 1990m01-2009m08. Na podstawie załączonych informacji z programu JDemetra+ odpowiedz na pytania:

- Jaki jest wielomian charakterystyczny części AR procesu ARIMA znalezione przez TRAMO?
- Do którego komponentu szeregu zostały włączone pierwiastki wielomianu AR związane z różnicowaniem zwykłym i sezonowym?
- Wiedząc, że szereg był poddany logarytmowaniu, omów, czy wyodrębniona sezonowość stochastyczna ma charakter addytywny, czy moltiplikatywny.
- Oszacowano tylko dwa parametry modelu: $\theta_1 = -0.6704$ oraz $\Theta_1 = -0.5598$. Dlaczego zatem w wielomianie charakterystycznym części MA dla procesu przed dekompozycją pojawia się trzeci parametr, 0.37527, przy opóźnieniu do potęgi trzynastej?

Zadanie 2

ARIMA model [(0,1,1)(0,1,1)]

Parameter	Value	Std error	T-Stat	P-value
Th(1)	-0,6704	0,0528	-12,69	0,0000
BTh(1)	-0,5598	0,0664	-8,43	0,0000

DecompositionModelNon-stationary AR: 1 - B - B¹² + B¹³

Stationary AR: 1

MA: 1 - 0,67036 B - 0,55981 B¹² + 0,37527 B¹³

Innovation variance: 1,0000

trendNon-stationary AR: 1 - 2 B + B²

Stationary AR: 1

MA: 1 + 0,046844 B - 0,95316 B²

Innovation variance: 0,0167

seasonalNon-stationary AR: 1 + B + B² + B³ + B⁴ + B⁵ + B⁶ + B⁷ + B⁸ + B⁹ + B¹⁰ + B¹¹

Stationary AR: 1

MA: 1 + 0,75892 B + 0,49417 B² + 0,23805 B³ + 0,012022 B⁴ - 0,17119 B⁵ - 0,30585 B⁶ - 0,3916 B⁷ - 0,43219 B⁸ - 0,43437 B⁹ - 0,40707 B¹⁰ - 0,36088 B¹¹

Innovation variance: 0,0648

irregular

Non-stationary AR: 1

Stationary AR: 1

MA: 1

Innovation variance: 0,4145

Zadanie 3

Za pomocą metody TRAMO-SEATS oczyszczono z sezonowości miesięczny szereg cen żywności w Niemczech w okresie 1990m01-2009m08.

- Scharakteryzuj krótko przebieg procedury TRAMO i procedury SEATS.
- Do którego komponentu szeregu zostały w SEATS włączone pierwiastki wielomianu AR związane z różnicowaniem zwykłym (niesezonowym)?
- Wiedząc, że szereg nie był poddany logarytmowaniu, omów, czy wyodrębniona sezonowość stochastyczna ma charakter addytywny, czy moltiplikatywny.
- W modelu oszacowanym w części TRAMO $q=1$ i $Q=1$, oszacowano również parametry $\theta_1 = -0.6704$ oraz $\Theta_1 = -0.5598$. Zapisz wielomian charakterystyczny części MA.