

Analiza szeregów czasowych

Time Series Analysis

część 2



Jerzy Stefanowski
Politechnika Poznańska

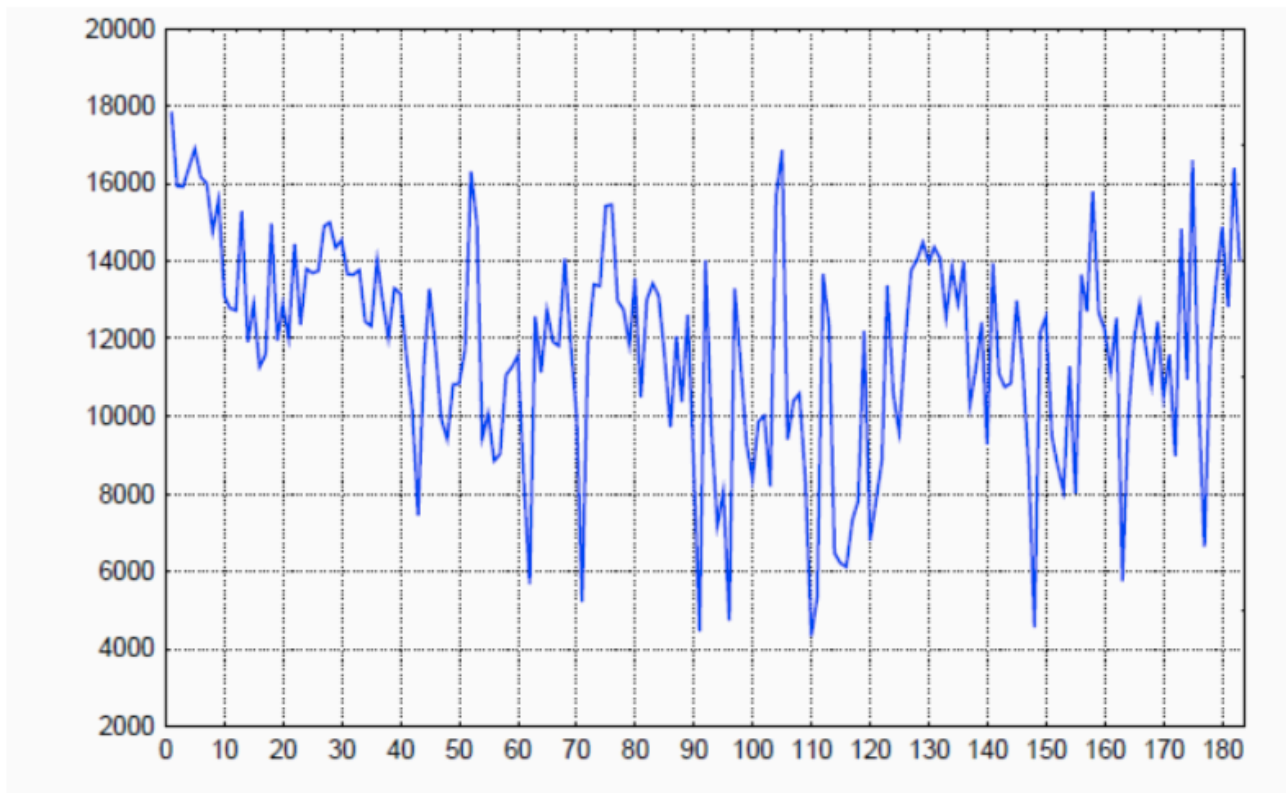
Projekt eksploracji danych
Poznań 2020

Plan -2

1. Uzupełnienie materiałów z części 1
2. Założenia i pojęcia wstępne
(stacjonarność, autokorelacja, szum biały, błędzenie losowe)
3. AR – model autoregresyjny
4. MR – odmiana średniej ruchomej
5. ARMA
 - Identyfikacja i dobór parametrów p , q , d
 - Estymacja
6. ARIMA – wersja z różnicowaniem
7. SARIMA i inne modele
8. Wsparcie oprogramowania
9. Zagadnienia prognozowania

Przykład – mat. Statsoft

Wielkość hurtowej sprzedaży produktu B (dane dzienne ze 180 kolejnych dni).

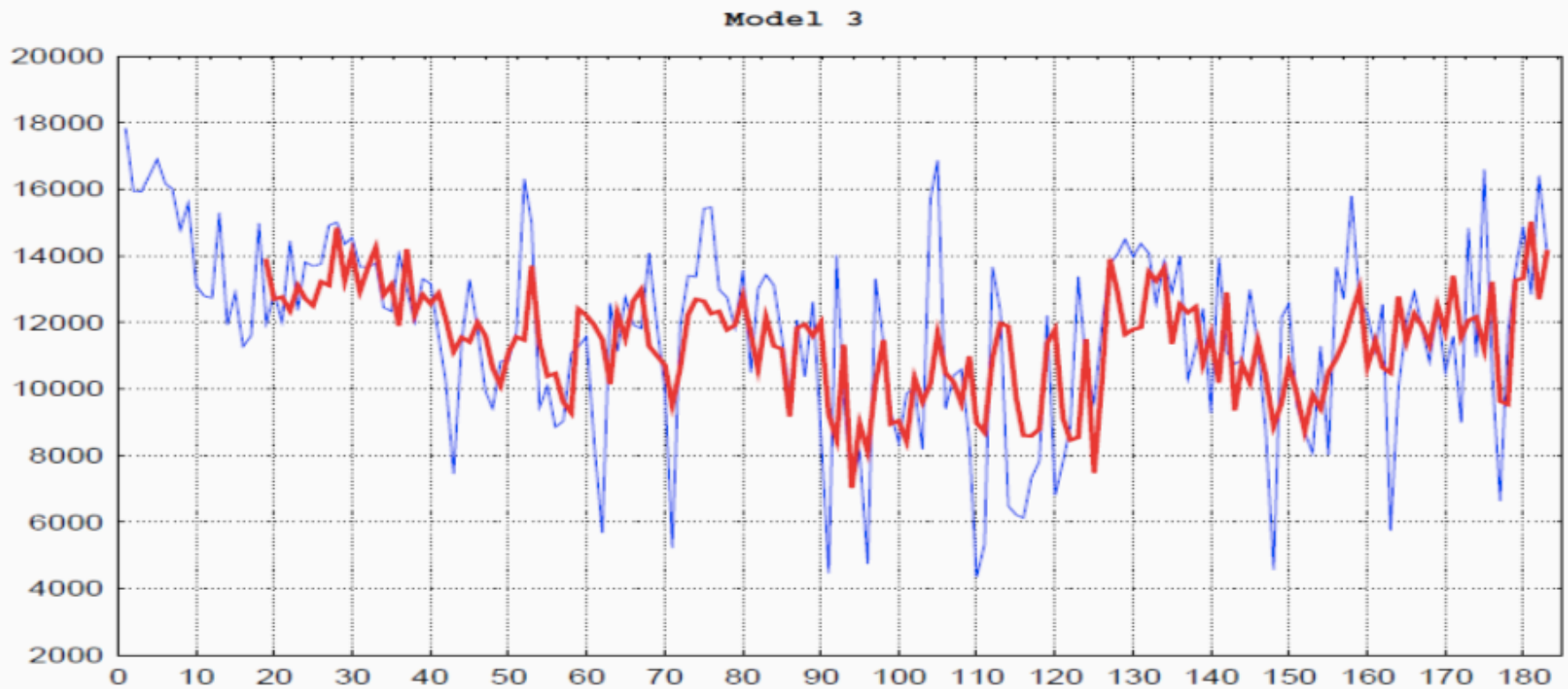


Przykład hurt. sprzedaży prod. B

Model dopasowany przez prof. A. Sokołowskiego dla sprzedaży produktu B; źródło: www.statsoft.pl

$$\hat{B}_t = 10\,670,58 - 61,13t + 0,27t^2 + \\ + 0,49B_{t-1} + 0,20B_{t-15} - 0,27B_{t-16} - 0,19B_{t-18}.$$

Wykresy szeregu pierwotnego i wartości teoretycznych
wyznaczonych wg modelu dla sprzedaży produktu B
(źródło: www.statsoft.pl)



Porównanie prognoz w przykładzie o hurtowej sprzedaży produktu B (www.statsoft.pl)

Dzień	Wartość rzeczywista	Model 3	Wyrównywanie wykładnicze	ARIMA
180	14871,0	13312,6	12664,5	12288,5
181	12830,9	15015,2	10903,2	13228,0
182	16396,2	12690,5	10832,8	12432,8
183	14034,4	14113,4	11611,7	12061,3
Średni błąd prognozy	-	2288 / 15,7%	3369 / 23,2%	2570 / 17,7%

Błędy prognozy *ex post*

Modele autoregresji i średniej ruchomej

Ogólne modele szeregów czasowych – stosowane do modelowania **stacjonarnych** szeregów czasowych lub niestacjonarnych, lecz sprowadzalnych do stacjonarnych

Ich budowa – oparta na zjawisku **autokorelacji**, czyli korelacji wartości zmiennej prognozowanej z wartościami tej zmiennej opóźnionymi w czasie

Podstawowe podejścia:

- Modele autoregresji (AR)
- Modele średniej ruchomej MA (nie mylić z poprzednim wykładem)
- Modele mieszane (ARMA)

Stacjonarność

Szereg y_t realizacja pewnego procesu stochastycznego $\{Y_t\}$

Pojęcie stacjonarności -

Rozkład prawdopodobieństwa procesu jest stały w czasie (ściśle)

Słabe:

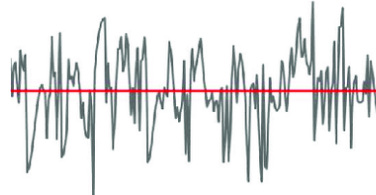
Stała w czasie jest wartość oczekiwana $E(y_t)=\mu$

Stała jest wariancja $Var(y_t)=\sigma<\infty$

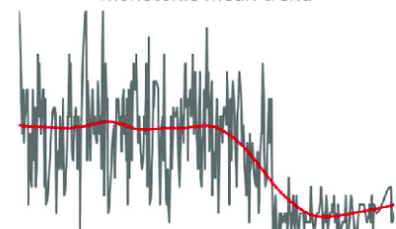
Stała jest kowariancja sąsiednich obserwacji

$$Cov(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)] = \lambda$$

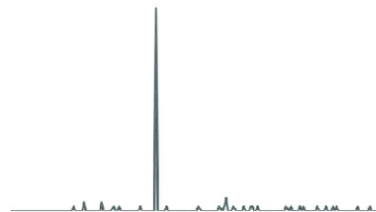
stationary OTU
mean-reverting/stable variance



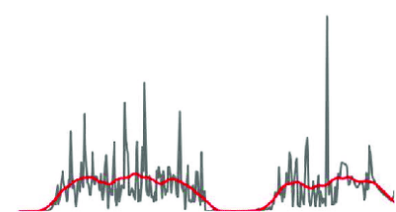
non-stationary OTU
monotonic mean trend



non-stationary OTU
conditionally rare (CRT bloom)



non-stationary OTU
non-monotonic mean trend



Biały szum

$$\gamma(h) = \begin{cases} \sigma^2 & \text{dla } h=0 \\ 0 & \text{dla } h \neq 0 \end{cases} \quad (1.5)$$

Definicja 1.2.1 białego szumu (white noise)

Proces $\{Z_t\}$ jest nazywany białym szumem ze średnią 0 i wariancją σ^2 , co zapisujemy:

$$\{Z_t\} \sim \text{WN}(0, \sigma^2), \quad (1.6)$$

wtedy i tylko wtedy, gdy $\{Z_t\}$ ma średnią zero i kowariancję zadaną wzorem (1.5).

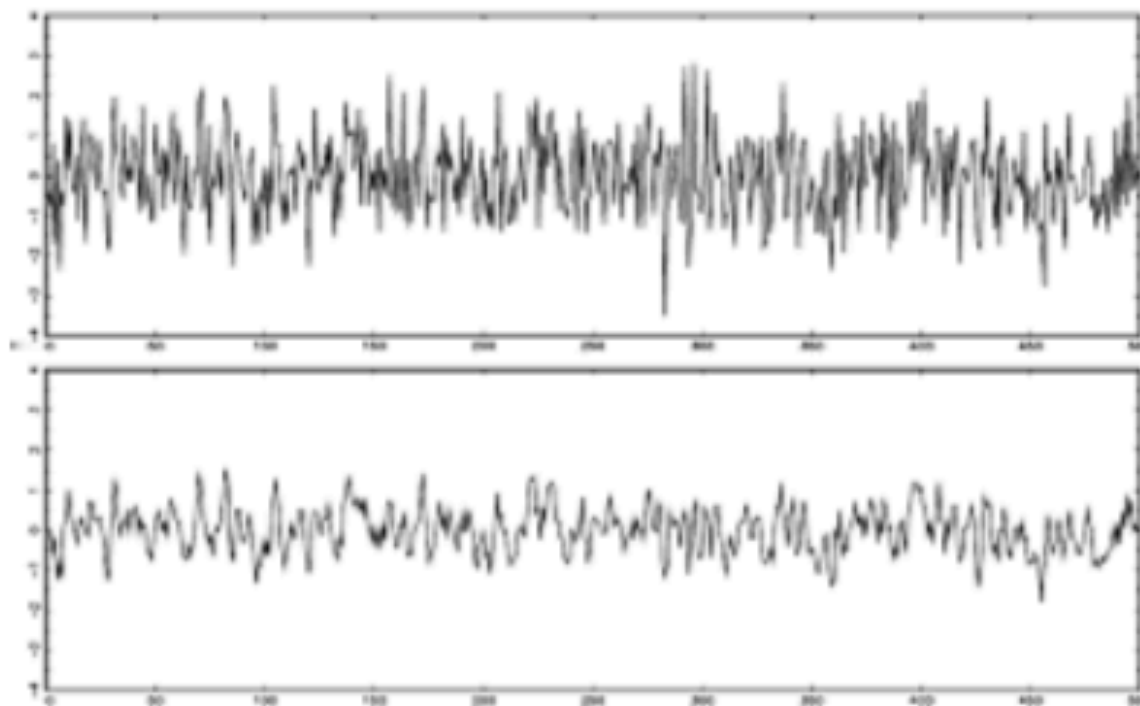
$$\text{Często zakładamy } \{Z_t\} \sim \text{iid}(0, \sigma^2), \quad (1.7)$$

co oznacza, że zmienne mają ten sam rozkład i są niezależne.

Zastosowanie:

Szeroka klasa stacjonarnych procesów może być utworzona z wykorzystaniem białego szumu jako wymuszającej siły pojawiającej się w (zbiorze) liniowych równań różnicowych, które są związane z pojęciem procesów ARMA (procesy z autoregresją (AR) i ruchomą średnią (MA)).

- Biały szum: $w_t \sim iid(0, \sigma_w^2)$. Często $w_t \sim Niid(0, \sigma_w^2)$.
- Średnia ruchoma: $v_t = \frac{1}{3}(w_{t-2} + w_{t-1} + w_t)$ (model MA(2)).



Rysunek: Gausowski biały szum i średnia ruchoma tego szumu.

Funkcje autokorelacji

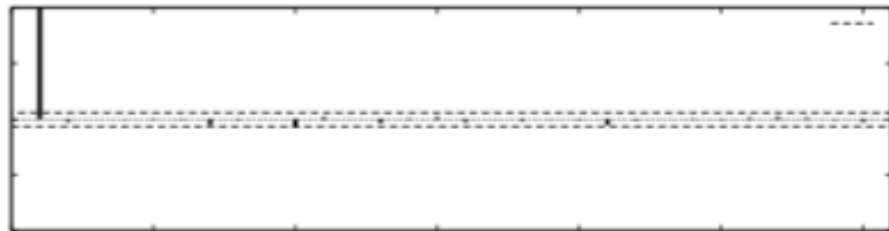
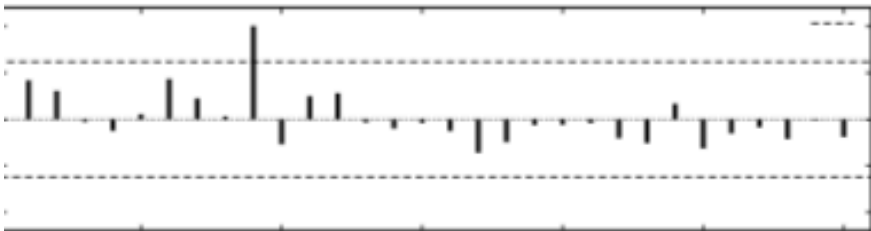
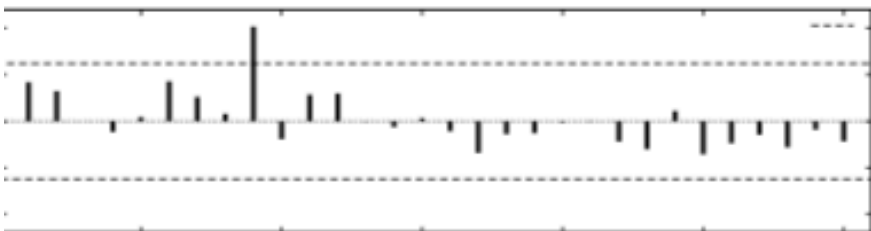
- Funkcja autokowariancji pokazuje zależność między y_t a poprzednimi wartościami
- Dla szeregu stacjonarnego wartości zależą wyłącznie od odległości między obserwacjami

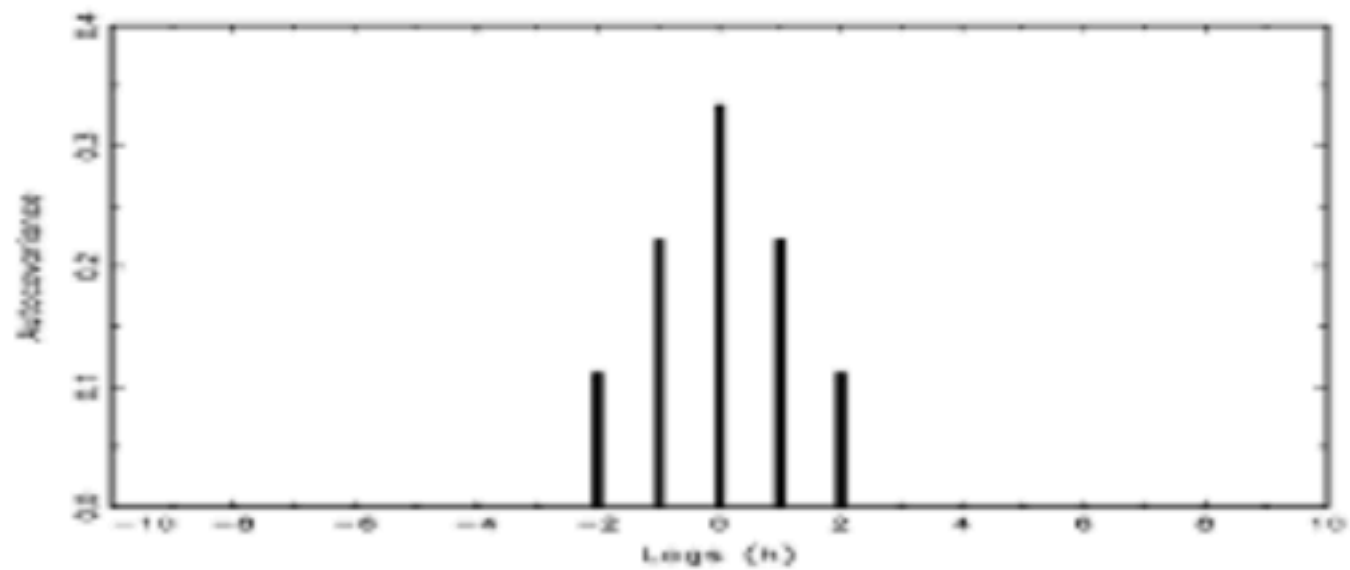
$$E(y_t - E(y_t))(y_{t-k} - E(y_{t-k})) = \gamma_k \quad k = 0, 1, 2, \dots$$

- autkowariancje są nieunormowane
- Funkcja autokorelacji (ACF) (**A**utocorrelation **F**unction)

$$\rho_k = \frac{\text{cov}(y_t, y_{t-k})}{\text{var}(y_t)}$$

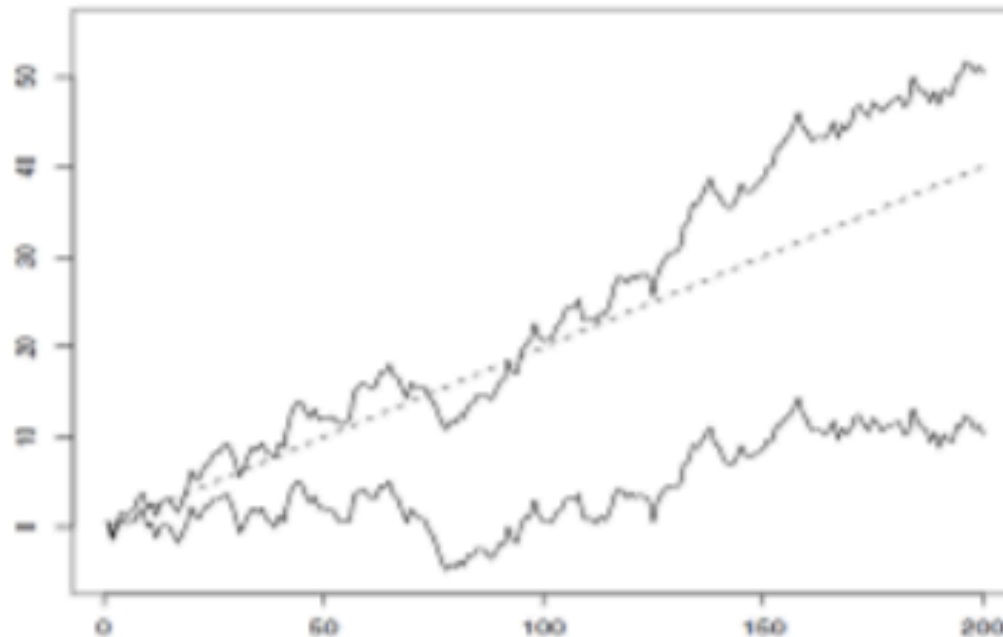
- Funkcja autokorelacji ACF (*Autocorrelation function*): $\rho_k = \frac{\text{cov}(y_t, y_{t+k})}{\sigma_y^2}$, gdzie σ_y^2 jest wariancją procesu y_t .
- Estymatorem ρ_k jest: $\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2}$, mający w przybliżeniu rozkład normalny o odchyleniu standardowym $\frac{1}{\sqrt{T}}$
 - Dla białego szumu powinna być ona nierozróżnialna od 0 (w sensie statystycznym)
 - Dla błędzenia losowego powinna być ona równa zawsze 1 lub wygasać bardzo, bardzo powoli
 - Dla stabilnego procesu autoregresyjnego powinna ona wygasać wraz z rosnącym k ($\rho_k = \phi^k$)





Rysunek: Autokowariancja średniej ruchomej $x_t = \frac{1}{3}(w_t + w_{t-1} + w_{t-2})$.

- Błądzenie losowe z trendem (lub bez):
 $x_t = \delta + x_{t-1} + w_t = \delta t + \sum_{j=1}^t w_j$ dla $t = 1, 2, \dots$ oraz $x_0 = 0$
(model ARIMA(1,1,0)).



Rysunek: Błądzenie losowe z trendem $\delta = 2$ i bez trendu ($\delta = 0$).

Ogólna idea **ARIMA**

ARIMA – akronim terminu *angielskiego* **Auto-Regresive Integrated Moving Average**

- **AR** (ang. Autoregression) – zależność od wartości opóźnionych zmiennej (ang. lagged)
- **I** (ang. Integrated) – operator różnicowania w celu przekształcenia w szereg stacjonarny
- **MA** (ang. Moving Average) – zależność od opóźnionych błędów (reszt) modelu

Każda ze składowych związana z parametrami (p, d, q) = uniwersalny zapis $ARIMA(p, d, q)$

Model autoregresji AR

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

- Model w którym każda wartość jest liniową kombinacją poprzednich wartości (ang. lagged) – wykorzystuje pamięć procesu.
 - Założenie o autokorelacji między wartością zmiennej prognozowanej a wartościami opóźnionymi w czasie
- p rząd modelu /opóźnienie/ – ile poprzednich obserwacji ma wpływ na bieżącą wartość AR(p)
- Przykład sprzedaży AR(1)

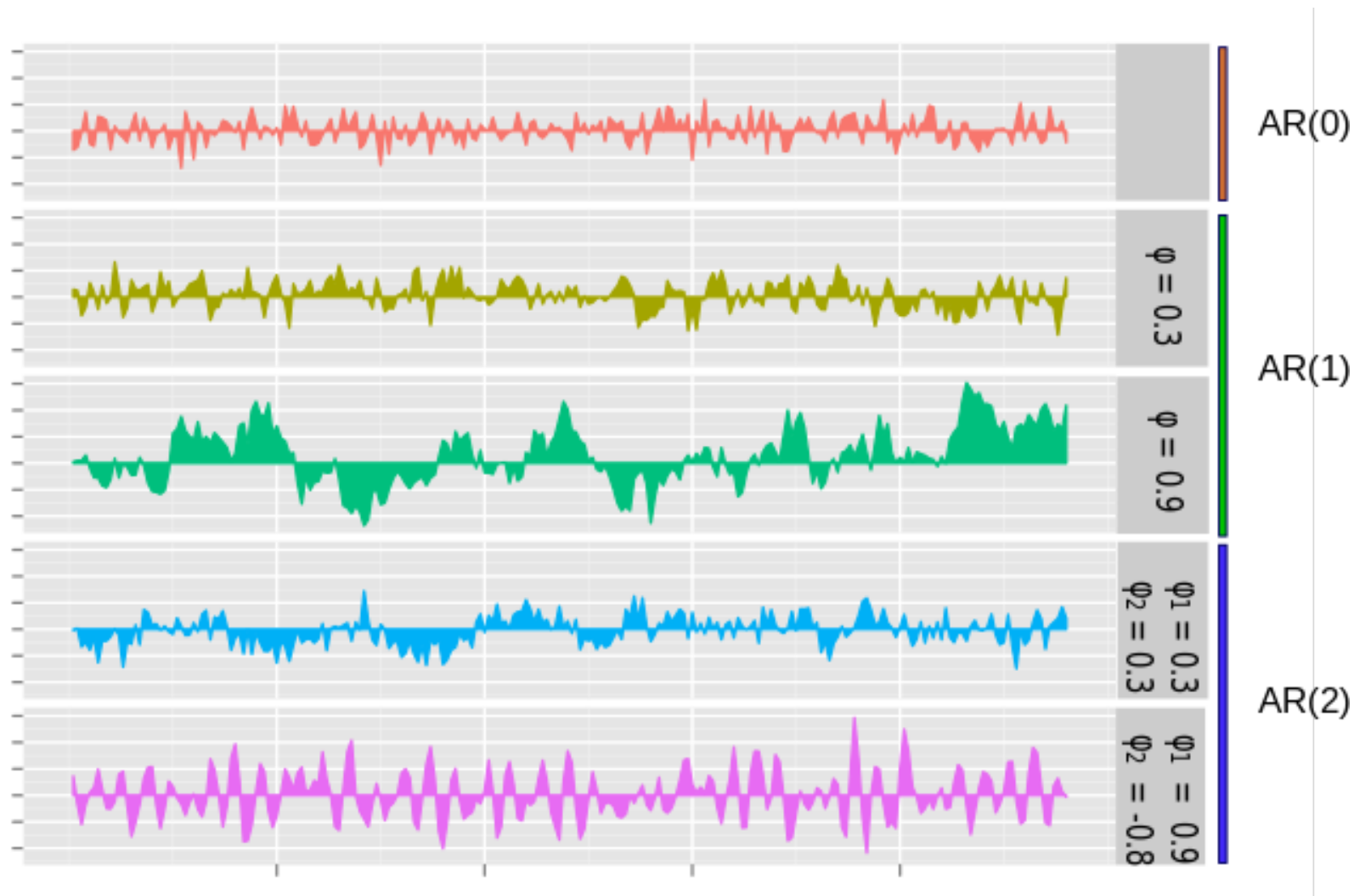
$$\text{sprzedaz}_{\text{sierpień}} = \varphi \text{sprzedaz}_{\text{lipiec}} + \varepsilon_{\text{sierpień}}$$

Jeżeli rząd wynosiłby 2, to proces AR (2) wyglądałby następująco:

$$\text{sprzedaz}_{\text{sierpień}} = \varphi_1 \text{sprzedaz}_{\text{lipiec}} + \varphi_2 \text{sprzedaz}_{\text{czerwiec}} + \varepsilon_{\text{sierpień}}$$

A parametry φ_i wskazują na wpływ kolejnych, co raz starszych wartości szeregu na bieżącą wartość.

Przykłady AR



Rys za wikipedia

Teoretycznie AR(1) stacjonarny gdy $\phi < 1$

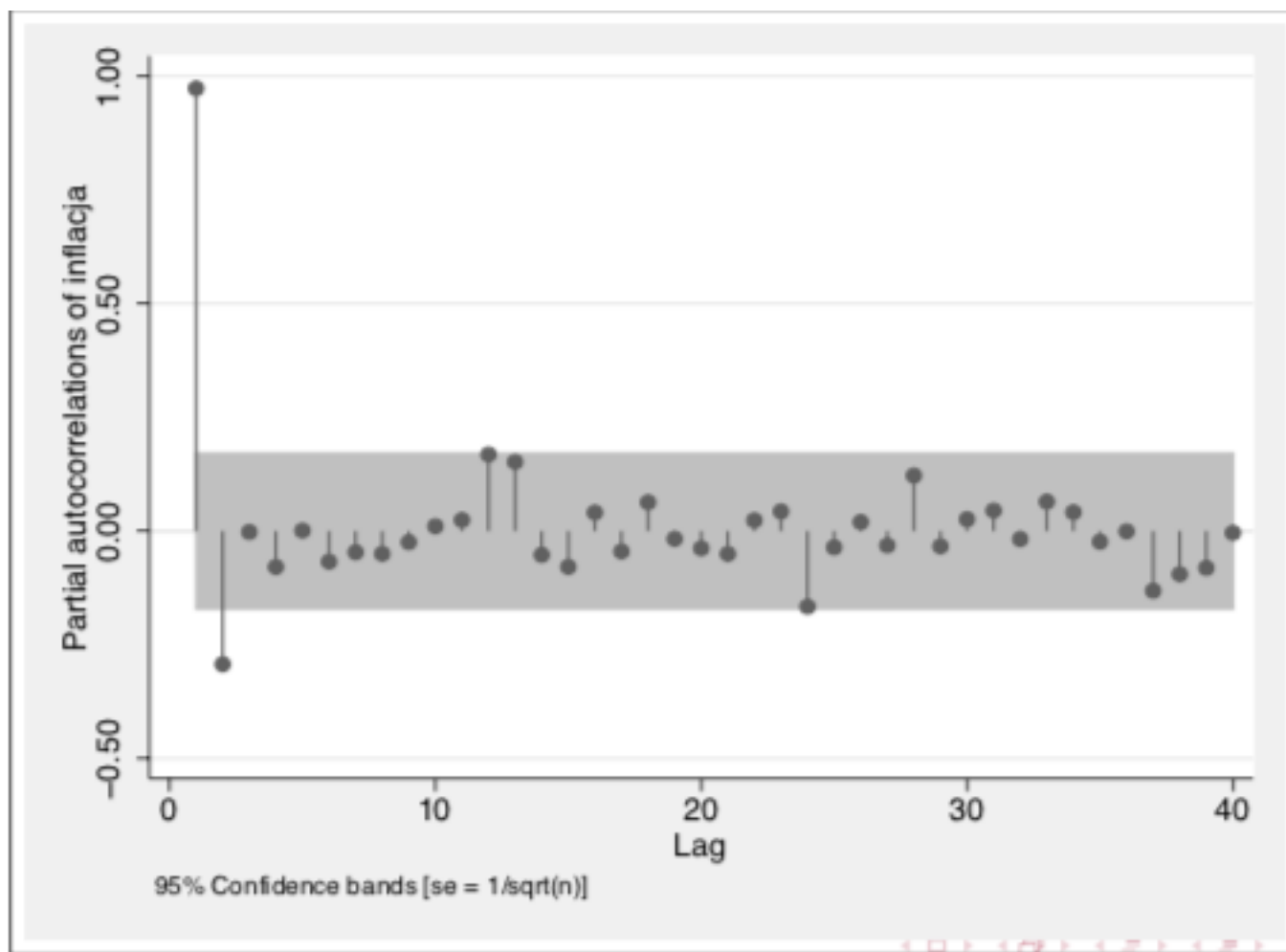
Funkcje cząstkowe autokorelacji PACF

- Funkcja cząstkowej autokorelacji (**P**artial **A**utocorrelation **F**unction) pokazuje zależność między y_t a poprzednimi wartościami, pomijając wpływ pośrednich opóźnień
- Liczbowo jest równa oszacowaniu współczynnika ρ_k w modelu:

$$y_t = \mu + \rho_1 y_{t-1} + \dots + \rho_k y_{t-k} + \varepsilon_t$$

Może być wykorzystana do oceny parametrów modeli

Przykład wykresu PACF



Dobór opóźnienia p i estymacja parametrów ϕ

Dla ustalenia wartości p (w pełni stacjonarny AR)

- Wykorzystuje się współczynniki cząstkowe autokorelacji PACF
- Własności matematyczne – dla AR(p); współczynniki PACF dążą do zera dla h większego niż p

Estymacja parametrów modelu

- Odmiany metody najmniejszych kwadratów, metoda momentów (dla układów równań Yula-Walker'a), lub metoda estymacji największej wiarygodności = patrz ang. hasło Wikipedia /Autoregressive model/

Własności teoretyczne

- Dla $y_t = \text{AR}(p)$ funkcje ACF maleją wykładniczo (bezpośrednio lub oscylując), lecz PACF będą ucięte do zera gdy $h > p$

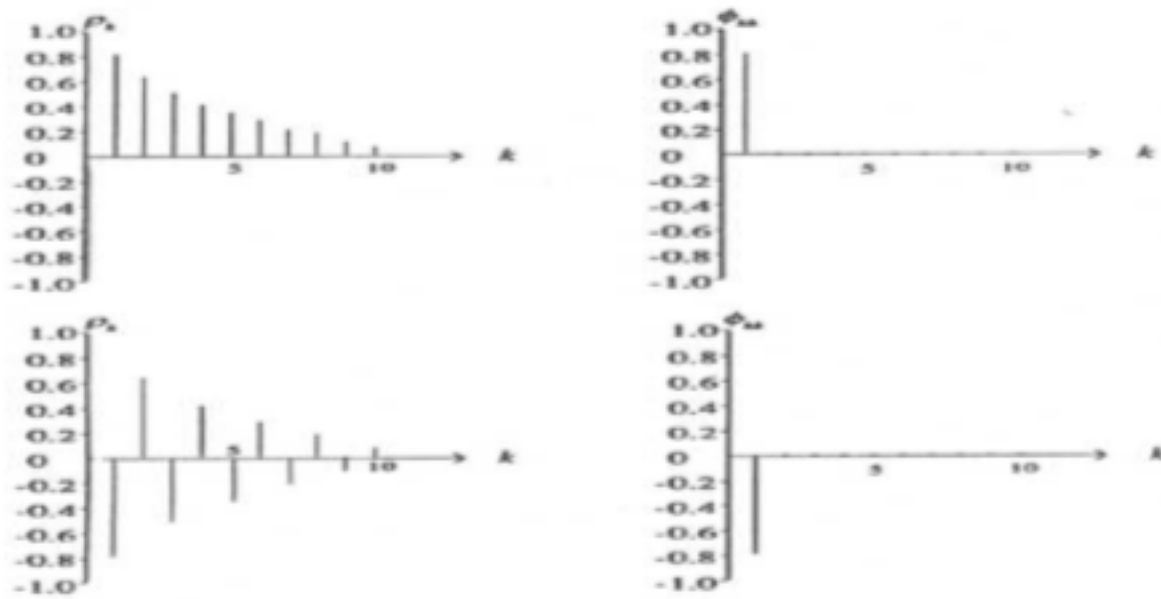


Fig. 3.1 ACF and PACF of the AR(1) process: $(1 - \phi B)\hat{Z}_t = \sigma_t$.

Model średniej ruchomej MA

$$y_t = \theta_0 + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_p e_{t-q}$$

- Model w którym aktualna wartość jest liniową kombinacją poprzednich wartości + odchyłek tych wartości
 - e odpowiadają błędom (resztom) modelu w okresach $t, t-1, \dots, t-q$
- q wielkość opóźnienia
- Teoretycznie MA powinien być stacjonarnym procesem (gdyż jest ważonym uśrednieniem poprzednich błędów prognozy)

$$E(\varepsilon_t) = 0 \quad \text{var}(\varepsilon_t) = \sigma^2 \quad E(\varepsilon_t, \varepsilon_s) = 0 \quad \forall t \neq s$$

- Inne sformułowanie z białym szumem $w = wn(0, \sigma^2)$

Definicja modelu średniej ruchomej stopnia q - MA(q)

Model postaci:

$$x_t = w_t + \Theta_1 w_{t-1} + \dots + \Theta_q w_{t-q},$$

gdzie w_t -biały szum (gaussowski), Θ_i -stałe ($\Theta_q \neq 0$).

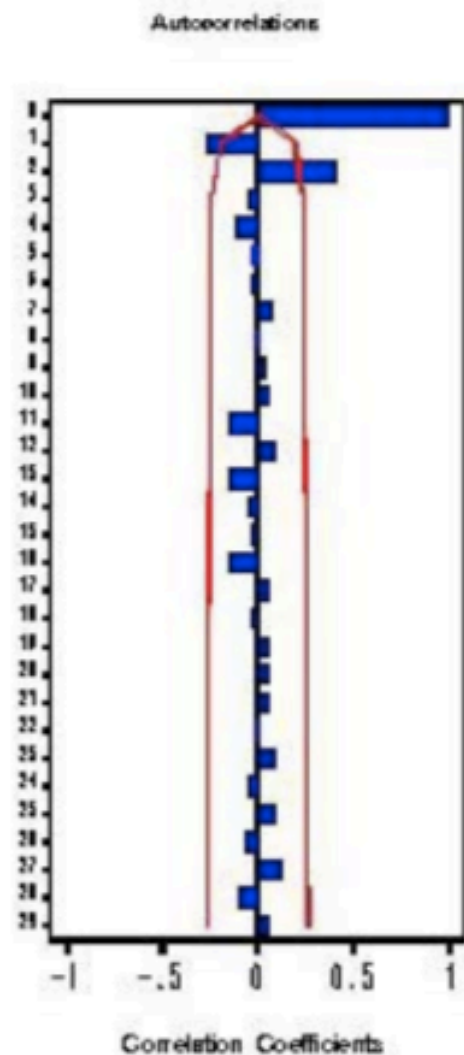
Dobór parametrów *MA*

Dla ustalenia wartości q (w pełni stacjonarny proces)

- Wykorzystuje się współczynniki autokorelacji ACF
- Własności matematyczne – dla $MA(q)$; współczynniki ACF dążą do zera dla h większego niż q

Estymacja parametrów θ modelu

- Zdecydowanie trudniejsza – gdyż błędy prognoz nie są bezpośrednio obserwowalne / nieliniowe, iteracyjne procedury dopasowania zamiast liniowych procedur najmniejszych kwadratów



Wykres 2. Funkcja Autokorelacji (ACF) procesu $X_t = Z_t + 0,3Z_{t-1} - 0,7Z_{t-2}$, gdzie $Z_t \sim WN(0,2)$.

Otrzymany kształt autokorelogramu ma cechy, które wskazują na jego pochodzenie z procesu MA(2), tzn. nieistotność wartości funkcji ACF dla $q > 2$. Problem ten zostanie omówiony w

Model autoregresji i średniej ruchomej ARMA

Dla osiągnięcia większej elastyczności w dopasowaniu modelu do szeregu czasowego niekiedy celowe jest połączenie obu modeli, które prowadzi do modelu autoregresji i średniej ruchomej ARMA

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_0 + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}$$

wartość zmiennej prognozowanej y w momencie t zależy od przeszłych jej wielkości oraz od różnic między przeszłymi wartościami rzeczywistymi zmiennej prognozowanej a jej wartościami uzyskanymi z modelu (błędów prognoz)

ARMA – nadal dotyczy procesów stacjonarnych!

Ponadto funkcje autokorelacji powinny szybko opadać na wykresie

Dobór parametrów ARMA

Dla opóźnień p i q oprócz omówionych zasad wykorzystania

- współczynniki autokorelacji ACF
- Współczynniki częściowej autokorelacji PACF

Nie zawsze jednoznaczny wybór

Doradza się wykorzystanie Akaike information criterion (AIC)

$$AIC(p, q) = \ln(\hat{\sigma}^2) + \frac{2(p + q)}{T}$$

where $\hat{\sigma}^2$ is the maximum likelihood estimated of the white noise variance.

Among a set of models, we select the values of p and q for our fitted model to be those which minimize $AIC(p, q)$.

Estymacja parametrów – odmiany metody najmniejszych kwadratów i inne już omówione

Modele zintegrowane

- Jeśli szereg nie jest stacjonarny, to rozważa się przekształcenia w szereg stacjonarny.
 - **Eliminacja trendu (funkcji nieliniowej)**
 - **Operacja różnicowania**
- Różnicowanie – obliczanie różnic sąsiednich wyrazów szeregu; $\Delta y_t = y_t - y_{t-1}$
 - Będą one reprezentować zmiany wartości w badanym szeregu np. spadki i wzrosty sprzedaży z okresu na okres. Przykładowy szereg może mieć przebieg wykazujący trend a po różnicowaniu otrzymamy wartości, które (bardzo często) będą już stacjonarne, a przynajmniej będą miały stały w czasie średni poziom

Niekiedy konieczne jest kolejne obliczenia różnic

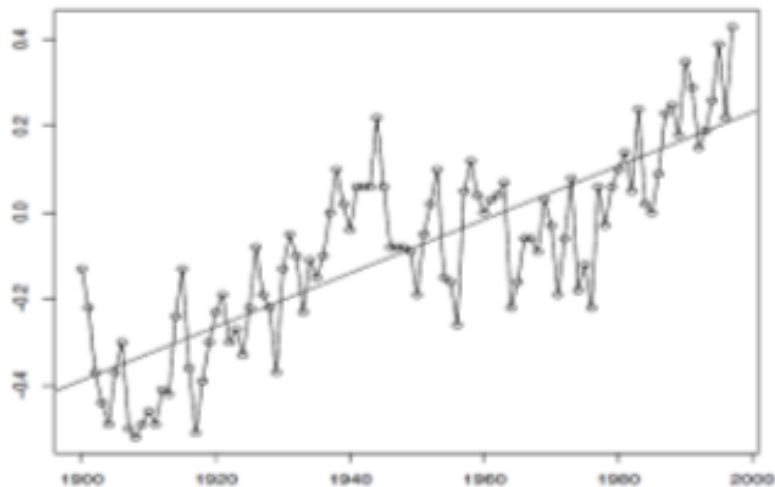
$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$$

, można dalej różnicować - aż d-krotne różnicowanie

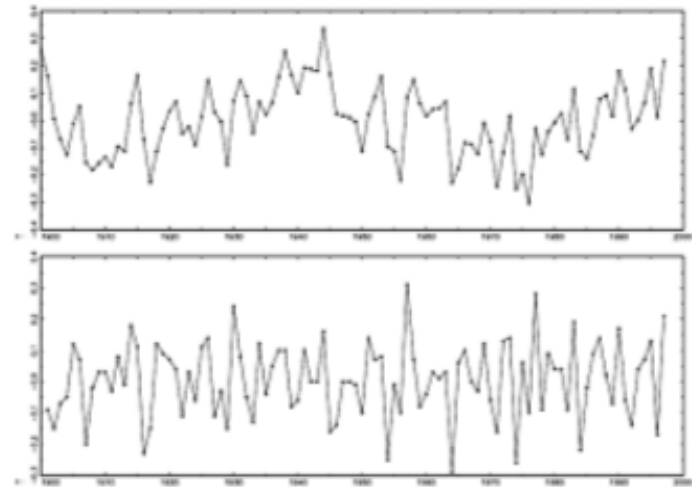
Stąd d paramter modeli zintegrowanych

- Konkluzja ARMA + integracja -> ARIMA (p,d,q)

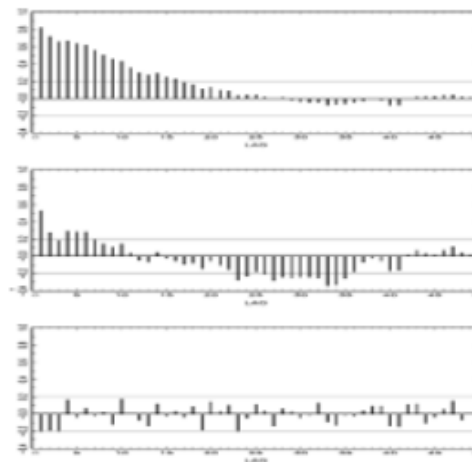
Przykład przekształceń [M.Śmieja]



Rysunek: Dane dotyczące temperatury i trend liniowy.



Rysunek: Szereg po usunięciu trendu za pomocą regresji (góra) i różnicowania



Rysunek: Autokorelacja oryginalnego szeregu(góra) oraz szeregów po usunięciu trendu za pomocą regresji (środek) i różnicowania (dół).

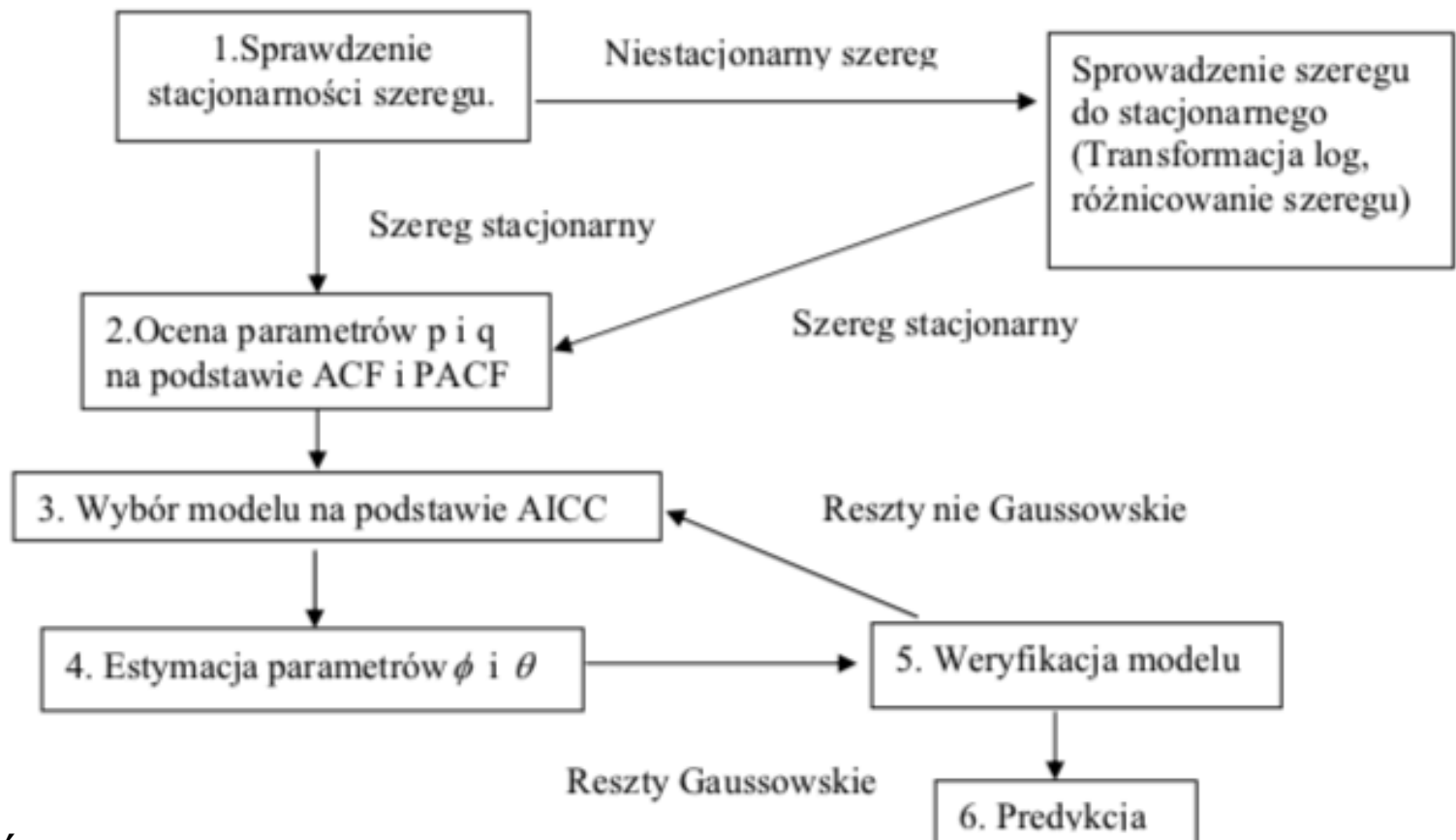
Procedura Box'a i Jenkinsa

- Identyfikacja (stacjonarność, wybór modelu, dobór parametrów p , d i q)
- Estymacja (wykorzystanie algorytmu do oszacowania parametrów modelu)
- Weryfikacja (ocena statystyczna kandydata na model prognostyczny / Ljung-Box test dla autokorelacji reszt, ocena wykresów autokorelacji)

Klasyczna metoda: Box, G.E.P. and G.M. Jenkins (1970)
Time series analysis: Forecasting and control.

1. Sprowadzenie procesu do szeregu stacjonarnego (transformacje log, różnicowanie)
2. Identyfikacja parametrów p i q (poprzez funkcję ACF i PACF)
3. Wybór modelu (wg kryterium AICC – model z wartością minimalną)
4. Estymacja parametrów ϕ , θ .
5. Weryfikacja modelu - sprawdzenie czy reszty są rozkładem Gaussowskim.
6. Predykcja szeregu

Diagram postępowania



Testowanie stacjonarności

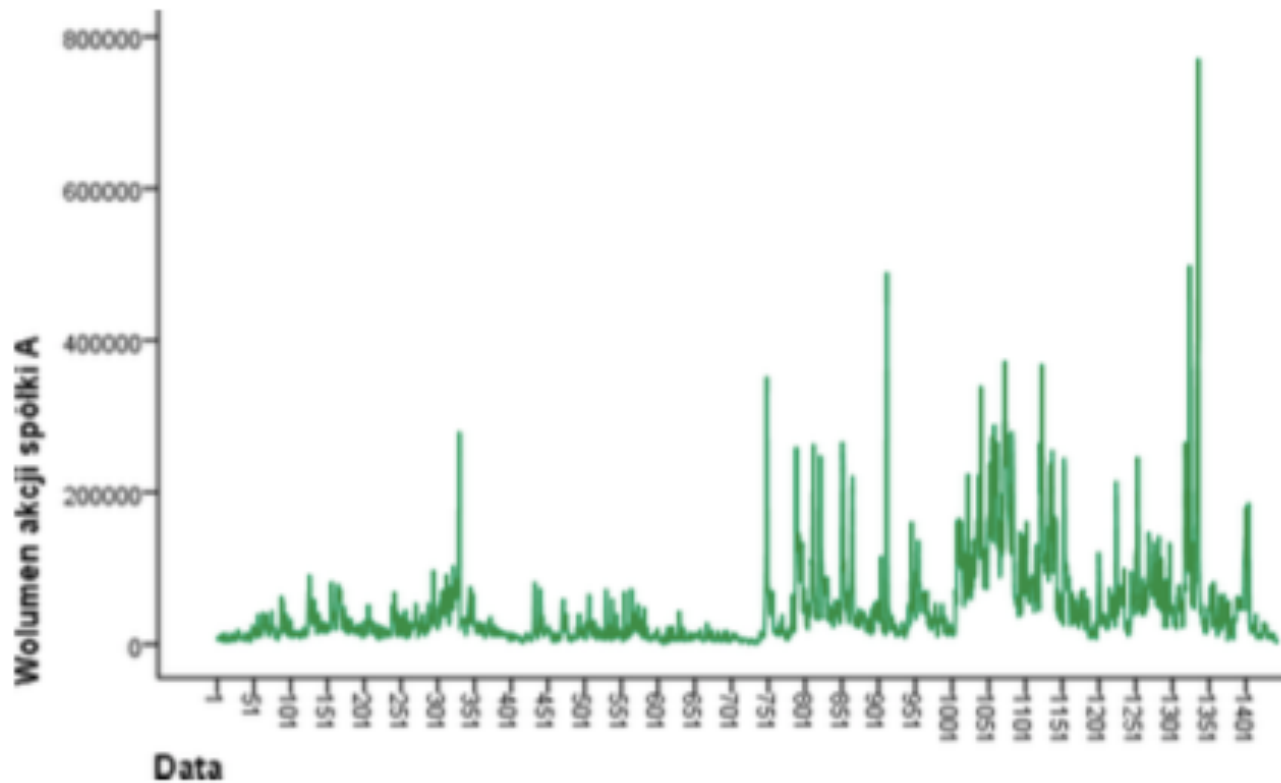
- Jeżeli procesy są prawidłowo ustalone to reszty są białym szumem
- testem niezależności reszt jest statystyka Ljunga-Boxa

$$Q = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k}{T-k} \xrightarrow{D} \chi_{m-p-q}^2$$

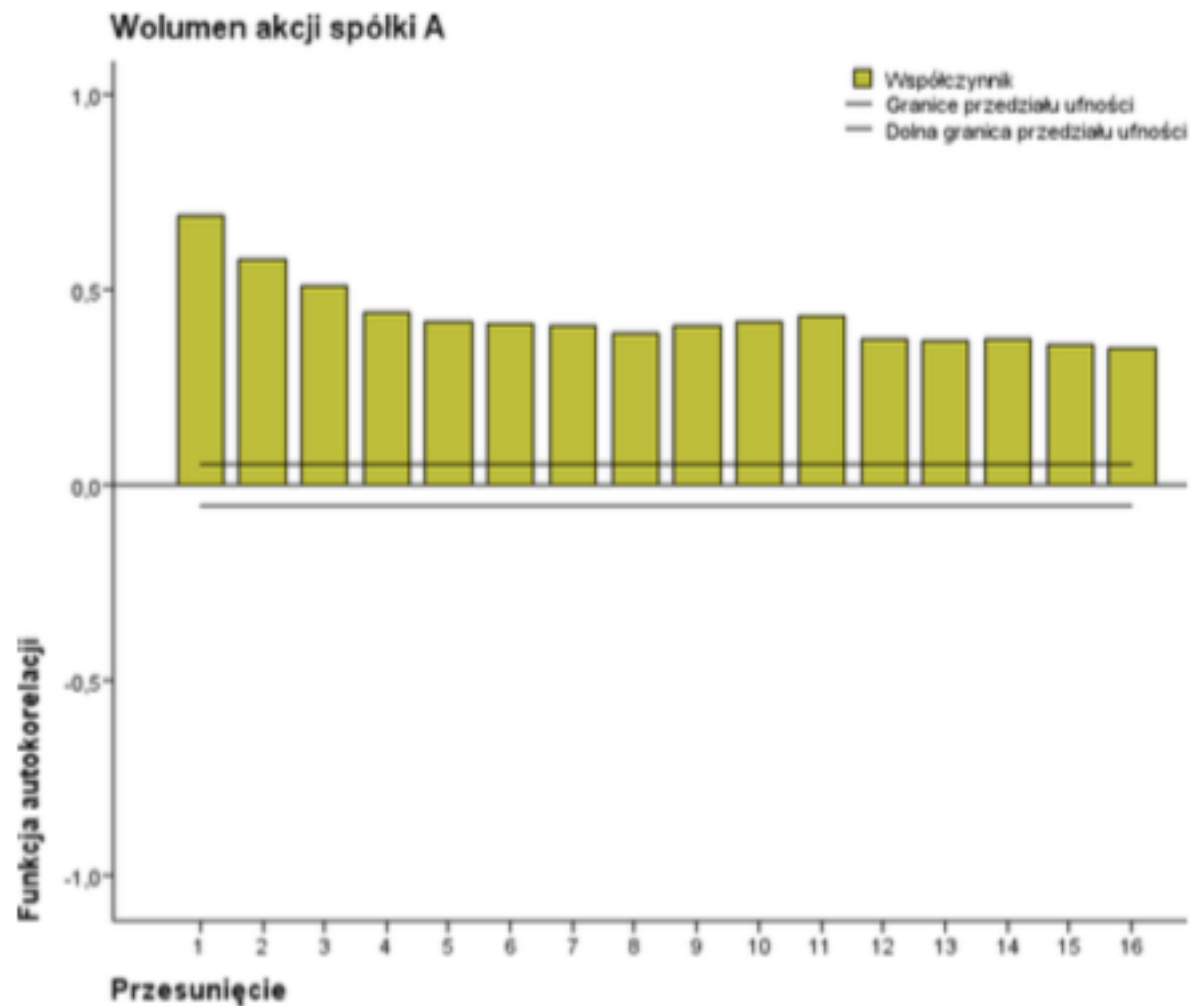
gdzie

$$\hat{\rho}_k = \frac{(T-k) \sum (y - \bar{y}) \sum (y_{t-k} - \bar{y})}{T^{-1} \sum (y - \bar{y})^2}$$

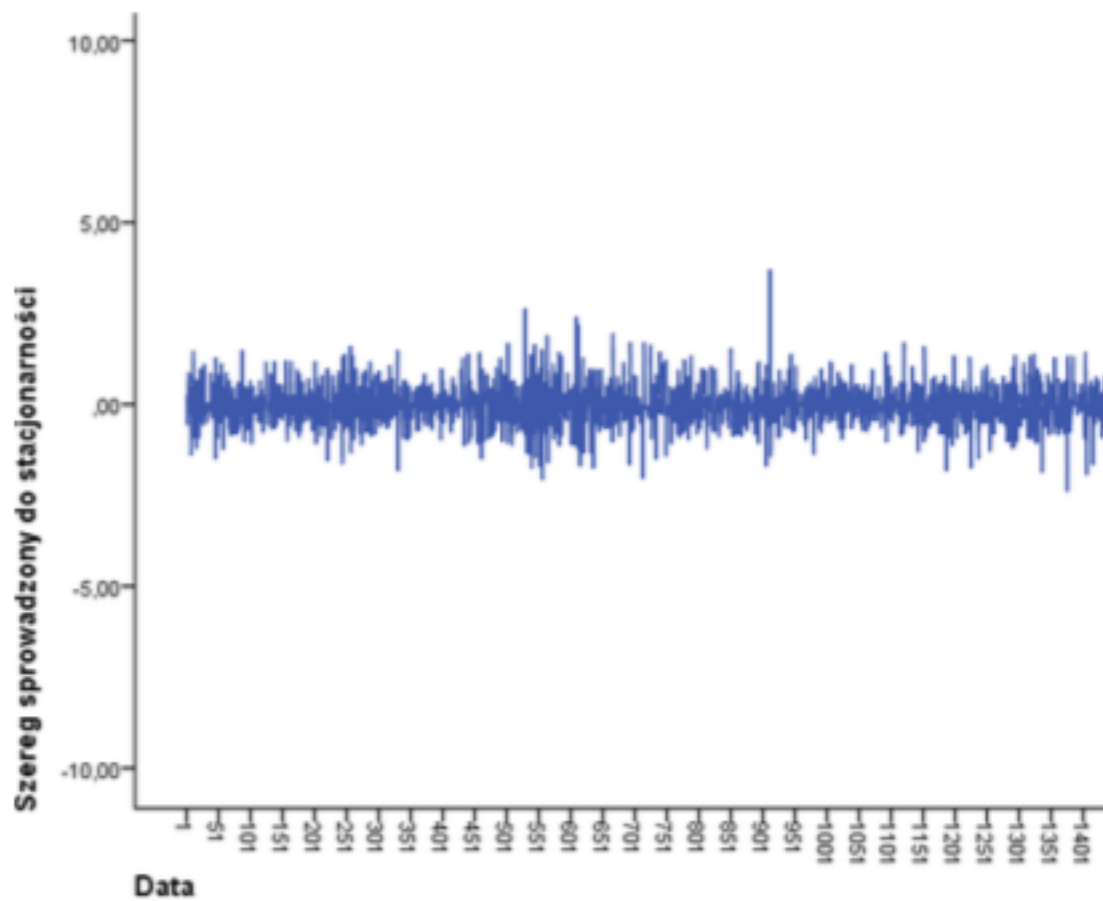
Przykład analizy SPSS [M.Płonka]



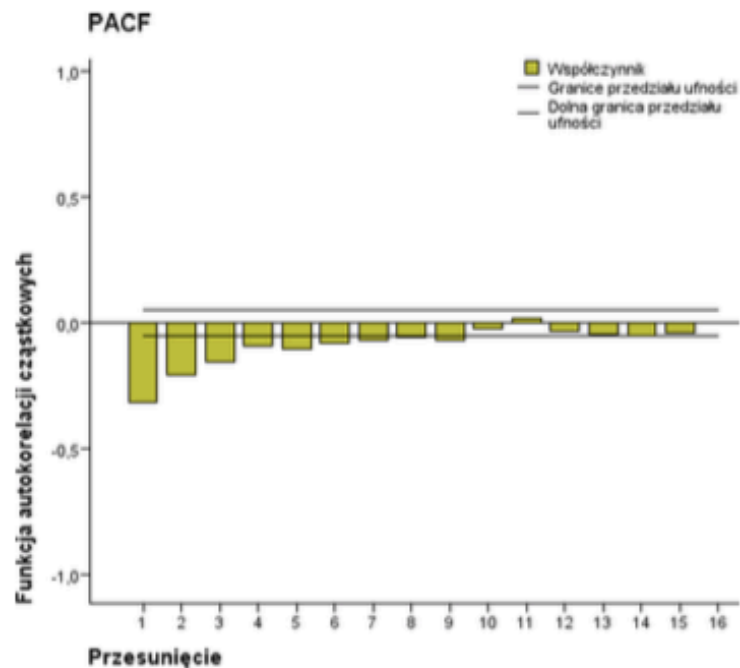
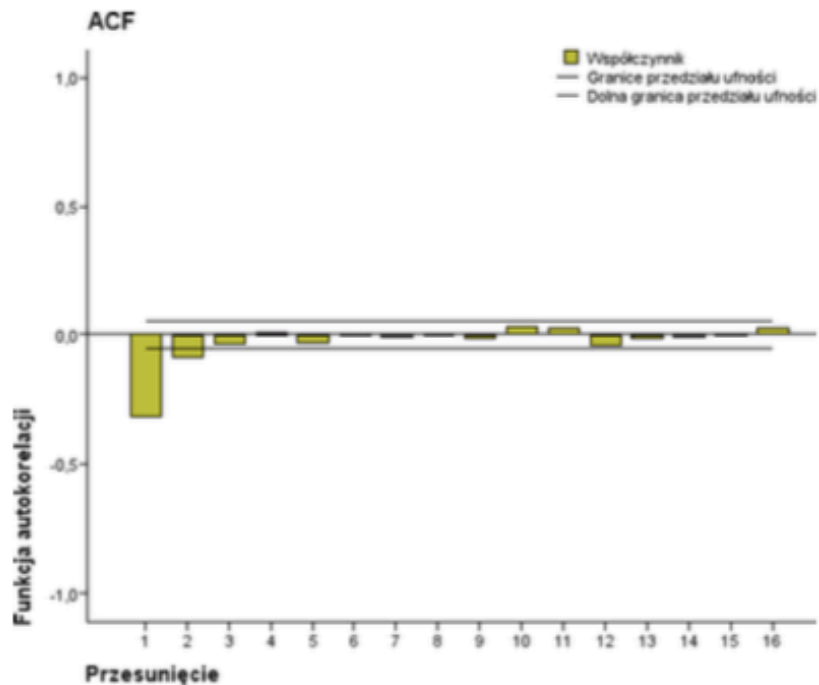
Ocena stacjonarności

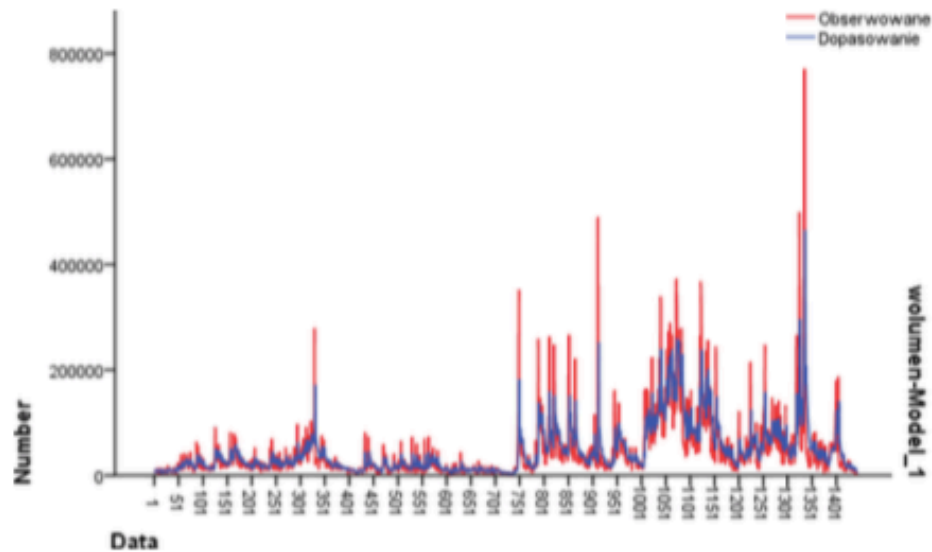


Przekształcenie



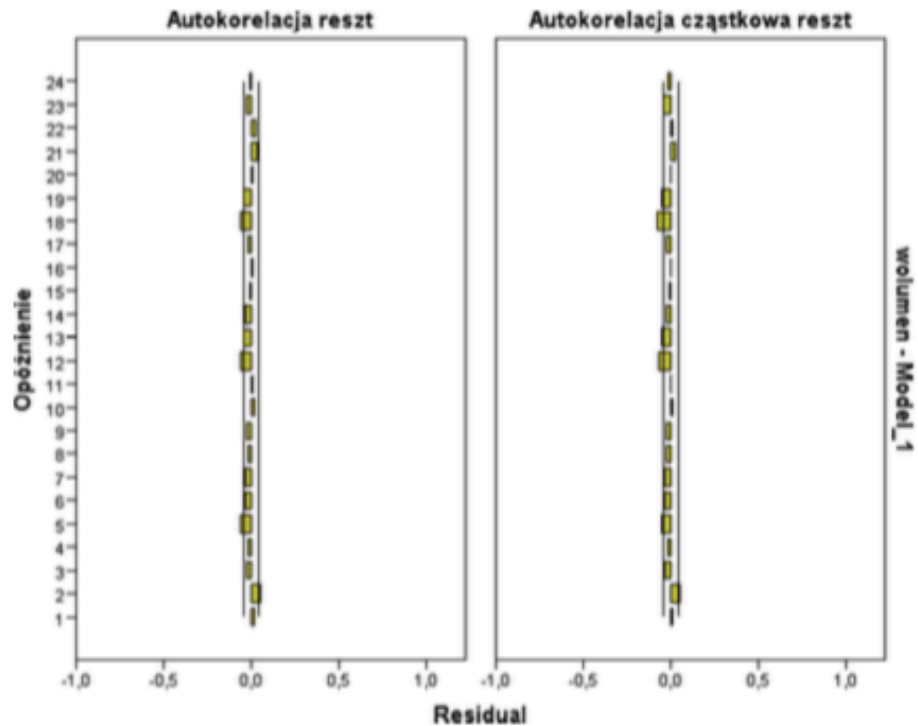
Identyfikacja (p,q)





Weryfikacja

Skorzystaj z testów



SARIMA – uwzględnienie sezonowości

ARIMA sprawdzi jeśli nie mamy elementu sezonowego, jeśli jednak istnieje mamy inny model – SARIMA. SARIMA jest podobny do modeli ARIMA, musimy tylko dodać kilka parametrów, aby uwzględnić sezonowość:

SARIMA (p, d, q) (P, D, Q) m

p – liczba autoregresji

d – stopień różnicowania

q – liczba średnich ruchomych terminów

m – odnosi się do liczby okresów w każdym sezonie

(P, D, Q) – reprezentuje (p, d, q) dla sezonowej części szeregu czasowego

Zapis dla modelu obejmuje określenie kolejności modeli AR (p), I (d) i MA (q) jako parametrów funkcji ARIMA i AR (P), I (D), MA (Q) i m parametry na poziomie sezonowym.

Przykład:

SARIMA (3,1,0) (1,1,0) 12

- Parametr m wpływa na parametry P, D i Q. Na przykład $m=12$ w przypadku danych miesięcznych sugeruje roczny cykl sezonowy.
- $P = 1$ użyta jest pierwsza sezonowa obserwacja w modelu, np. $t - (m - 1)$ lub $t - 12$. $P = 2$, użyte są dwie ostatnie sezonowe obserwacje $t - (m - 1)$, $t - (m * 2)$.
- $D = 1$ wyliczy różnicę sezonową pierwszego rzędu, a $Q = 1$ użyje błędy pierwszego rzędu w modelu (np. Średniej ruchomej).

SARIMA – sformułowanie mat.

Multiplicative Seasonal ARMA Models

Usually our models will incorporate autocorrelation at seasonal and neighboring lags.

Consider a model whose MA characteristic polynomial is given by

$$(1 - \theta x)(1 - \Theta x^{12})$$

The corresponding time series is

$$Y_t = Z_t - \theta Z_{t-1} - \Theta Z_{t-12} + \theta \Theta Z_{t-13}$$

Exercise: Verify that the autocorrelation function is nonzero only lags 1, 11, 12, and 13.

More generally, a **multiplicative seasonal ARMA** $(p, q) \times (P, Q)_s$ **model with seasonal period** s is a model with AR characteristic polynomial $\phi(x)\Phi(x)$ and MA characteristic polynomial $\theta(x)\Theta(x)$, where

$$\begin{aligned}\phi(x) &= 1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p \\ \Phi(x) &= 1 - \Phi_1 x^s - \Phi_2 x^{2s} - \dots - \Phi_P x^{Ps}\end{aligned}$$

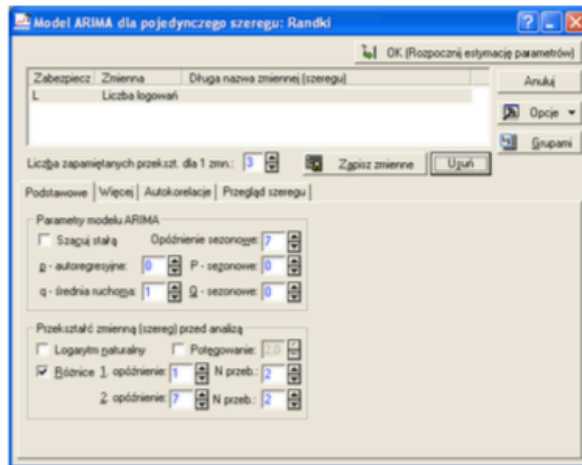
and

$$\begin{aligned}\theta(x) &= 1 - \theta_1 x - \theta_2 x^2 - \dots - \theta_q x^q \\ \Theta(x) &= 1 - \Theta_1 x^s - \Theta_2 x^{2s} - \dots - \Theta_Q x^{Qs}\end{aligned}$$

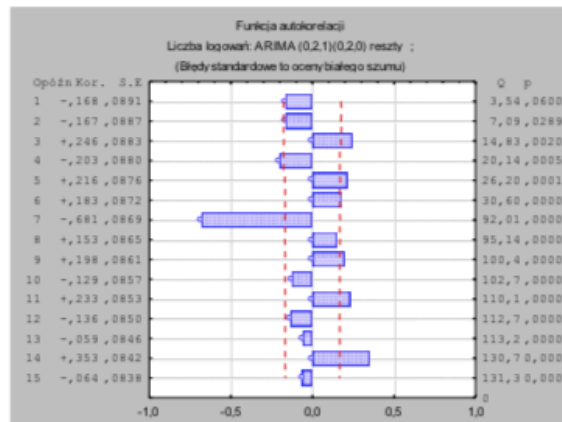
Za stat lab uberkey – uwaga notacja zmienna x oraz współczynniki ujemne

Analiza logowania się na portalu [Sokołowski] SARIMA

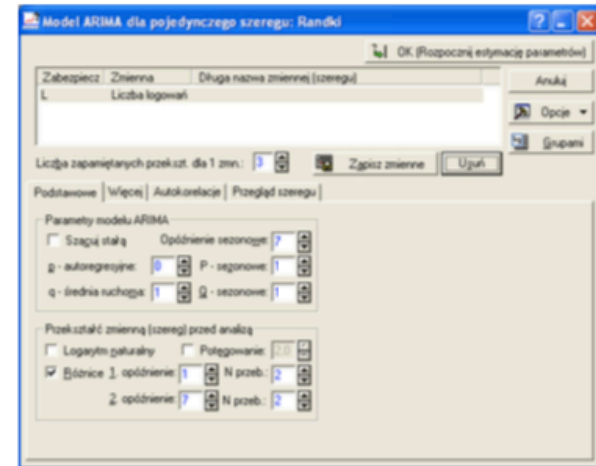
Na wstępie koncentrujemy się na niesezonowej części modelu. Poprzez kolejne podwyższanie rzędu autokorelacji i średniej ruchomej znajdujemy „najbardziej skomplikowany” model ARIMA (0,2,1).



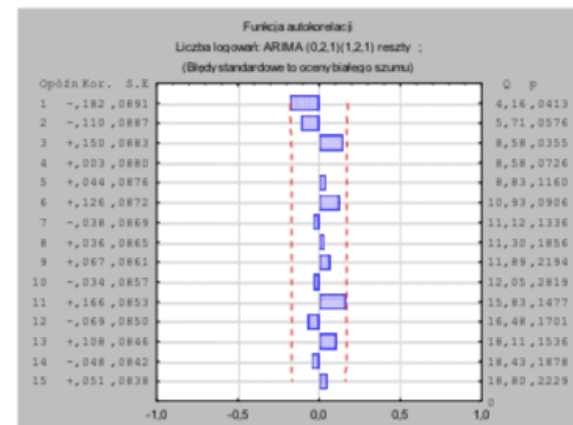
Funkcja autokorelacji reszt wskazuje, że komponenta okresowa nie została skutecznie wyeliminowana i trzeba podjąć próbę jej opisanie modelem, który ma aktywną część sezonową (stąd niekiedy używana nazwa SARIMA).



Znów poprzez stopniowe zwiększanie rzędu parametrów sezonowych znajdujemy optymalny model ARIMA(0,2,1)(1,2,1).

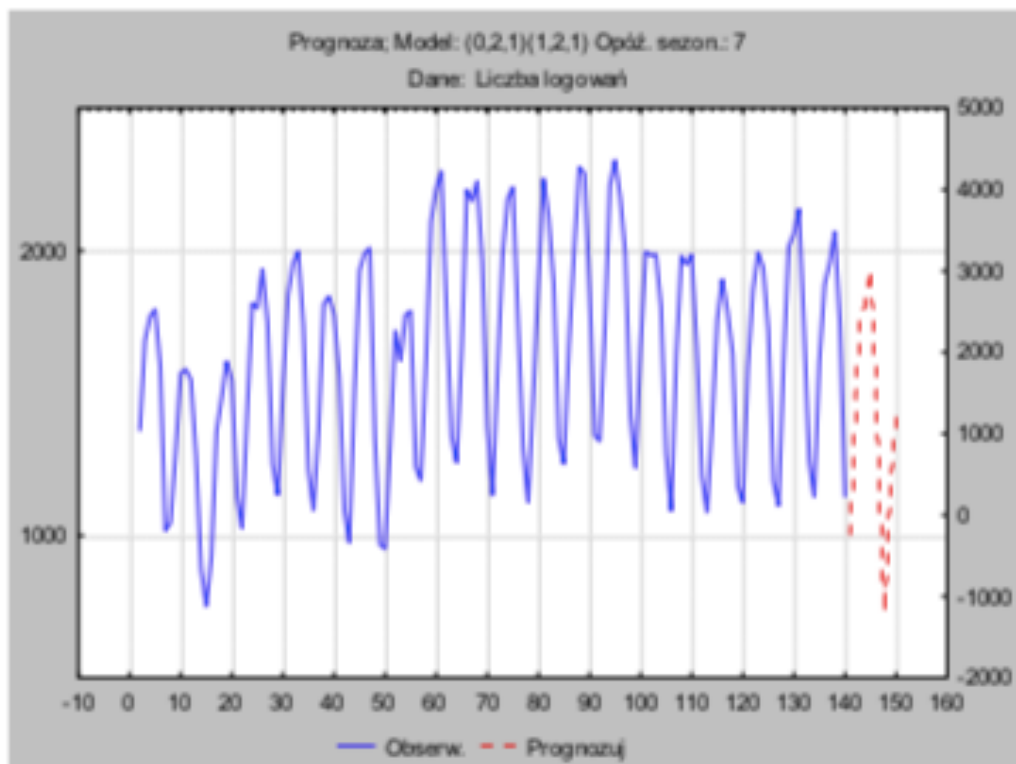


Teraz funkcja autokorelacji reszt jest zadowalająca.



Sprawdzamy jeszcze rozkład reszt.

Proгноза



Zmienne towarzyszące

Uogólnienia modelu ARMA / ARIMA

Tzw. ARMAX(p, q, b) Autoregressive–moving-average model with exogenous inputs model

$$X_t = \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \sum_{i=1}^b \eta_i d_{t-i}.$$

where η_1, \dots, η_b are the **parameters** of the exogenous input d_t .

Prognozowanie

- Prognoza na 1 okres naprzód

$$y_{T+1} = \alpha_1 y_T + \alpha_2 y_{T-1} = 1.41 * 4.53 - 0.41 * 4.67 = 4.47$$

- Prognoza na 2 okresy naprzód

$$y_{T+2} = \alpha_1 y_{\hat{T}+1} + \alpha_2 y_T = 1.41 * 4.47 - 0.41 * 4.53 = 4.45$$

- Prognoza na 3 okresy naprzód

$$y_{T+3} = \alpha_1 y_{\hat{T}+2} + \alpha_2 y_{\hat{T}+1} = 1.41 * 4.45 - 0.41 * 4.47 = 4.44$$

O prognozowaniu

Pojęcie prognozy:

- „racjonalne, naukowe przewidywanie przyszłych zdarzeń”,
- „stwierdzenie odnoszącym się do określonej przyszłości formułowanym z wykorzystaniem metod naukowych, weryfikowalnym empirycznie, niepewnym, ale akceptowalnym”.

Złośliwie:

„Prognozowanie to sztuka przewidywania przyszłości ... i uzasadniania, dlaczego owe przewidywania się nie sprawdzają.”

O prognozowaniu

Kilka założeń:

- zaobserwowany model nie zmieni się co do kształtu jak i siły działania w okresie przyszłym,
- wahania przypadkowe nie zakłócają znacząco zaobserwowanych składowych systematycznych modelu.
- Spełnienie założeń bardziej prawdopodobne dla okresów leżących bliżej ostatniego badanego okresu.
- Niezbędna wiedza dziedzinowa na temat charakteru zjawisk.

O prognozowaniu

- Okres prognozy
- Horyzont prognozy
- Zmiany ilościowe i jakościowe w prognozowaniu zjawisk.
- Podział prognoz:
 - krótkoterminowa,
 - średnioterminowa,
 - długoterminowa.
- Ocena wiarygodności prognozy szacowanie błędów prognozy *ex post*:
- Definicje błędów – poprzedni wykład
 - bezwzględny błąd w momencie t : $y_t - y'_t$,
 - względny błąd w momencie t [%],
 - średni (względny) błąd lub błąd kwadratowy MSE.

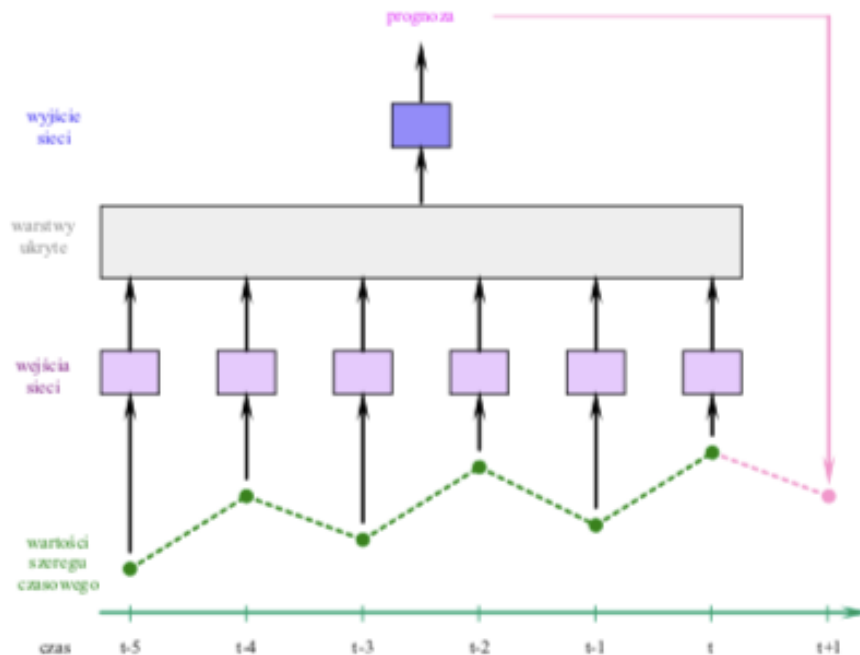
Uwagi końcowe

Klasyczne metody – dostępne we wszystkich pakietach statystycznych i wielu innych

- In **R**, the *arima* function (in standard package *stats*) is documented in [ARIMA Modelling of Time Series](#). The *tseries* package includes an *arma* function, documented in ["Fit ARMA Models to Time Series"](#); the *fr* processes; and the [forecast package](#) includes *auto.arima* for selecting a parsimonious set of p, q . The CI
- **Mathematica** has a complete library of time series functions including ARMA.^[6]
- **MATLAB** includes functions such as [arma](#) and [arx](#) to estimate AR, ARX (autoregressive exogenous), and [Econometrics Toolbox](#) for more information.
- **Julia** has some community driven packages that implement fitting with an ARMA model such as [arma.jl](#).
- **Statsmodels** Python module includes many models and functions for time series analysis, including ARMA with **Pandas**. [See here for more details](#).
- **PyFlux** has a Python-based implementation of ARIMAX models, including Bayesian ARIMAX models.
- **IMSL Numerical Libraries** are libraries of numerical analysis functionality including ARMA and ARIMA procedures in C#, .NET, and Fortran.
- **gretl** can also estimate ARMA model, [see here where it's mentioned](#).
- **GNU Octave** can estimate AR models using functions from the extra package [octave-forge](#).
- **Stata** includes the function *arima* which can estimate ARMA and **ARIMA** models. [See here for more details](#).
- **SuanShu** is a Java library of numerical methods, including comprehensive statistics packages, in which implemented in an object-oriented approach. These implementations are documented in ["SuanShu, a Java"](#).
- **SAS** has an econometric package, ETS, that estimates ARIMA models. [See here for more details](#).

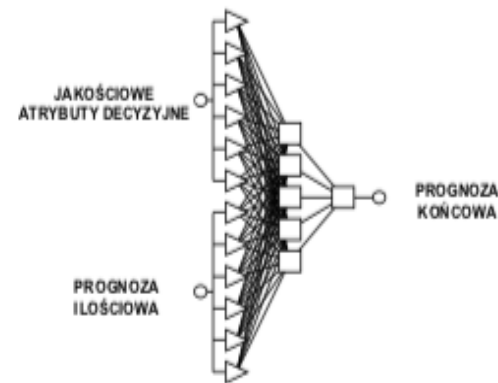
Sieci Neuronowe

Na przełomie wieku popularne wykorzystanie MLP dla odpowiednio dużych zbiorów uczących – przykład niżej lagged inputs



Rysunek 1. Przykładowe odwzorowanie szeregu czasowego na sieć neuronową wyznaczającą prognozę

zobacz wyniki dla podzioru uczącego) jak i generalizacji (zbiór walidacyjny i testowy).



Rys. 2. Schemat sieci integracyjnej (MLP 12 – 5 – 1) wykorzystanej do predykcji zespolonej (źródło: badania własne)

W ostatnich latach modele zbudowane w oparciu o sieć neuronową typu LSTM (uogólnienie przewidywania tekstów)

Literatura

Podstawy matematyczne:

J.Mielniczuk: Szeregi czasowe cz 1 I cz 2 Szkoła doktorska PAN.

Mingda Zhang: Time Series: Autoregressive models AR, MA, ARMA, ARIMA Univ. Pittsburgh

Umberto Triacca: Lectures on ARIMA

Inne:

P.Dittmann: Prognozowanie w przedsiębiorstwie. Metody i ich zastosowanie. Wolters Kluwer 2008.

A.Sokołowski: Prognozowanie finansowych szeregów czasowych. Statsoft.

M.Śmieja: Szeregi czasowe. UJ.

M.Doman: Analiza szeregów czasowych UEP.

D. Witkowska, Sztuczne sieci neuronowe i metody statystyczne. Wybrane zagadnienia finansowe, C.H. Beck, Warszawa.

Samodzielna analiza dokumentacji i blogów nt. użycia metod w określonym oprogramowaniu

Dziękuję za uwagę

Pytania lub komentarze?

Więcej informacji w publikacjach
oraz materiały dodatkowe!

Kontakt:
Jerzy.Stefanowski@cs.put.poznan.pl