








Serii de timp

Capitolul 2: Modele ARMA



Daniel Traian PELE

Academia de Studii Economice din București
IDA Institute Digital Assets
Blockchain Research Center
AI4EFin Artificial Intelligence for Energy Finance
Academia Română, Institutul de Prognoză Economică
MSCA Digital Finance

Acronime folosite în acest material (1/2)

- **ACF**: Autocorrelation Function (funcția de autocorelație)
- **AI**: Artificial Intelligence (inteligență artificială)
- **AIC**: Akaike Information Criterion (criteriul informațional Akaike)
- **AR**: AutoRegressive (model) — model autoregresiv
- **ARIMA**: AutoRegressive Integrated Moving Average (model) — model autoregresiv integrat cu medie mobilă
- **ARMA**: AutoRegressive Moving Average (model) — model autoregresiv cu medie mobilă
- **BET**: Bucharest Exchange Trading (index) — indicele de referință al Bursei de Valori București
- **BIC**: Bayesian Information Criterion (criteriul informațional bayesian al lui Schwarz)
- **BNR**: Banca Națională a României
- **GARCH**: Generalised AutoRegressive Conditional Heteroskedasticity (heteroscedasticitate condiționată autoregresivă generalizată)
- **GDP**: Gross Domestic Product (produsul intern brut)
- **IAPC**: Indicele armonizat al prețurilor de consum

Acronime folosite în acest material (2/2)

- **INS**: Institutul Național de Statistică
- **LB**: Ljung–Box (test) — testul de autocorelație Ljung–Box
- **MA**: Moving Average (medie mobilă)
- **MLE**: Maximum Likelihood Estimation (estimarea prin metoda verosimilității maxime)
- **MSE**: Mean Squared Error (eroarea pătratică medie)
- **OLS**: Ordinary Least Squares (metoda celor mai mici pătrate)
- **OPG**: Outer Product of Gradients (produsul exterior al gradienților)
- **PACF**: Partial AutoCorrelation Function (funcția de autocorelație parțială)
- **PIB**: Produsul intern brut
- **QQ**: Quantile–Quantile (plot) — graficul cuantilă–cuantilă
- **UE**: Uniunea Europeană
- **VAR**: Vector AutoRegression (model vector autoregresiv)

Întrebarea de azi și traseul

- **Întrebarea**: o serie staționară este corelată cu propriul trecut; ce model mic surprinde această memorie și o transformă în prognoze cu o incertitudine corect măsurată?
  - răspunsul lui Box și Jenkins: modelele **ARMA(p, q)**, cîțiva parametri în locul infinității de ponderi Wold
- **Traseul** capitolului
  - AR(p) și MA(q): staționaritate, invertibilitate, rădăcini caracteristice, ACF și PACF
  - ARMA(p, q): forma cauzală MA(∞), răspunsurile la impuls, identificarea
  - estimarea (Yule–Walker, cele mai mici pătrate condiționate, verosimilitate maximă), AIC și BIC, diagnosticarea reziduurilor
  - prognoza și metoda Box–Jenkins pe PIB-ul și inflația României, BET, EUR/RON și petele solare
- Seminarul 2 are loc înaintea acestui curs: dă formulele; aici le deducem și le aplicăm pe date

Rezultatele învățării

- Verificați staționaritatea unui AR(p) și invertibilitatea unui MA(q) din rădăcinile polinoamelor lor
- Deduceți media, autocovarianțele și ponderile ψ pentru AR(1), AR(2), MA(q) și ARMA(1,1)
- Propuneți ordine candidate din ACF și PACF și alegeți între ele cu AIC și BIC
- Estimați modele ARMA prin Yule–Walker, cele mai mici pătrate condiționate și verosimilitate maximă și interpretați rezultatele
- Testați reziduurile: Ljung–Box cu $m - p - q$ grade de libertate, Jarque–Bera
- Calculați prognozele ARMA și intervalele lor și explicați revenirea la medie

Bibliografie și instrumente

- Manual: Huang și Petukhina (2022), cap. 3–4 (modele staționare, ARMA)
  - manual însoțitor, gratuit online: Hyndman și Athanasopoulos (2021), cap. 9 (modele ARIMA)
- Teorie: Brockwell și Davis (2016), cap. 3 (ARMA) și cap. 5 (estimare); Hamilton (1994), cap. 3–5; lucrarea clasică: Box, Jenkins și Reinsel (2008)
- Quantlet-urile Python ale capitolului: Quantlets/Ch_02
  - estimarea ARMA cu `statsmodels` (`ARIMA`, `ArmaProcess`, `acorr_ljungbox`)
- Notebook-ul cursului: deschideți în Google Colab
- Curs video: Applied Time Series Analysis with Python

Instrumente din Capitolul 1

- **Staționaritatea slabă**: media constantă μ și autocovarianța $\gamma(h) = \text{Cov}(X_t, X_{t+h})$, care depinde doar de lagul h
  - ACF $\rho(h) = \gamma(h)/\gamma(0)$; PACF ϕ_{hh} = corelația dintre X_t și X_{t+h} după eliminarea efectului lui $X_{t+1}, \ldots, X_{t+h-1}$
- **Zgomotul alb** $\varepsilon_t \sim \text{WN}(0, \sigma^2)$: media 0, varianța σ^2, fără autocorelație
  - **operatorul lag**: $LX_t = X_{t-1}$, $L^k X_t = X_{t-k}$; polinoamele în L acționează asupra seriilor
- **Descompunerea Wold** (Wold (1938)): orice proces staționar pur nedeterminist este $X_t = \mu + \sum_{j \geq 0} \psi_j \varepsilon_{t-j}$, $\psi_0 = 1$, $\sum_j \psi_j^2 < \infty$
  - detalii și grafice: Capitolul 1, secțiunile despre operatorul lag, descompunerea Wold și ACF și PACF de selecție

Problema: o infinitate de ponderi

- Teorema Wold spune **cum arată** un proces staționar, nu **cum îl estimăm**
  - $\psi_1, \psi_2, \ldots$: o infinitate de necunoscute și doar T observații
- **Ideea**: scriem $\psi(L) = \sum_j \psi_j L^j$ ca **raport a două polinoame scurte**
  - $\psi(L) = \theta(L)/\phi(L)$, cu $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\theta(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$
  - atunci $\phi(L)(X_t - \mu) = \theta(L)\varepsilon_t$: modelul **ARMA$(p, q)$**, cu $p + q + 2$ parametri (ϕ, θ, μ, σ^2)
- O funcție rațională aproximează bine aproape orice $\psi(L)$ netedă, cu p și q mici
  - principiul **parcimoniei**: cel mai mic model care lasă reziduuri de tip zgomot alb

Trei piese de bază

- **AR(p)**, autoregresiv: valoarea de azi depinde de ultimele p valori ale seriei
  - $X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t$; originea: Yule (1927), petele solare
- **MA(q)**, medie mobilă: valoarea de azi depinde de ultimele q șocuri
  - $X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$; originea: Slutzky (1937), sume de cauze aleatoare
- **ARMA(p, q)**: ambele părți; transformat într-o metodă completă de Box, Jenkins și Reinsel (2008) (prima ediție în 1970)
  - c este **termenul liber**; media este $\mu = c/(1 - \phi_1 - \cdots - \phi_p)$, nu c
- Peste tot: $\varepsilon_t \sim \text{WN}(0, \sigma^2)$, iar pentru verosimilitate și intervale ε_t i.i.d. $N(0, \sigma^2)$ (independente și identic distribuite)

Modelul AR(1)

- $X_t = c + \phi X_{t-1} + \varepsilon_t$, sau $(1 - \phi L)X_t = c + \varepsilon_t$
  - ϕ: partea din abaterea de azi față de medie care se păstrează și mîine
- **Substituind înapoi** de k ori:
  - $X_t = c(1 + \phi + \cdots + \phi^{k-1}) + \phi^k X_{t-k} + \sum_{j=0}^{k-1} \phi^j \varepsilon_{t-j}$
  - dacă $|\phi| < 1$: $\phi^k \to 0$ și $X_t = \dfrac{c}{1-\phi} + \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j}$
- Aceasta este forma Wold cu $\psi_j = \phi^j$: reprezentarea **cauzală** (MA(∞))
  - cauzală = X_t depinde doar de șocurile prezente și trecute
  - $\phi = 1$: mers aleator (Capitolul 1), fără soluție staționară; $|\phi| > 1$: proces exploziv sau o soluție care depinde de șocuri viitoare, exclusă

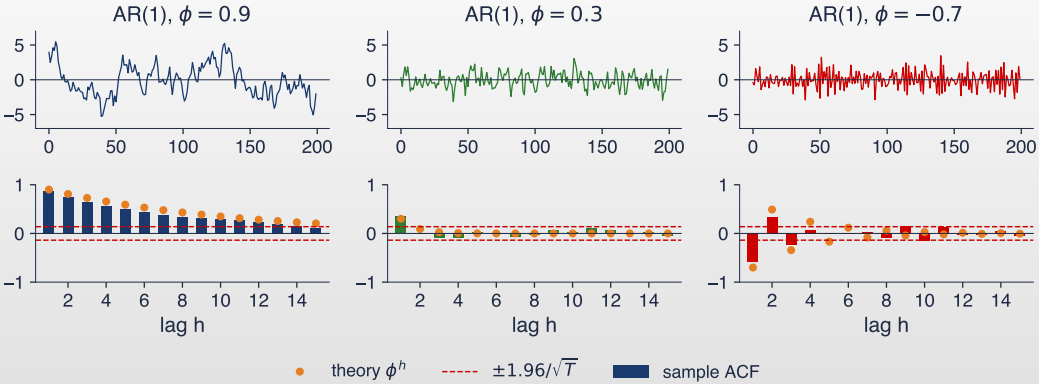
Momentele unui AR(1) staționar

- **Media**: aplicăm media, $\mu = c + \phi\mu$, deci $\mu = c/(1-\phi)$
  - aici folosim staționaritatea: $E[X_t] = E[X_{t-1}]$
- **Varianța**: $\gamma(0) = \phi^2\gamma(0) + \sigma^2$, deoarece ε_t este necorelat cu X_{t-1}
  - $\gamma(0) = \sigma^2/(1-\phi^2)$: crește nelimitat cînd $|\phi| \to 1$
- **Autocovariațele**: înmulțim $X_t - \mu = \phi(X_{t-1} - \mu) + \varepsilon_t$ cu $X_{t-h} - \mu$, $h \geq 1$, și aplicăm media
  - $\gamma(h) = \phi\gamma(h-1)$, deci $\gamma(h) = \phi^h\gamma(0)$ și $\rho(h) = \phi^h$
  - descreștere geometrică; semne alternante dacă $\phi < 0$
- **Timpul de înjumătățire** al unui șoc: h pentru care $\phi^h = 1/2$, adică $h = \ln 0{,}5 / \ln \phi$

Exemplu rezolvat: un AR(1)

- $X_t = 2 + 0{,}8X_{t-1} + \varepsilon_t$, $\sigma^2 = 1$
- Este staționar? $|\phi| = 0{,}8 < 1$: da
  - media $\mu = 2/(1 - 0{,}8) = 10$; termenul liber este 2, media este 10
- Varianța și ACF
  - $\gamma(0) = 1/(1 - 0{,}64) = 2{,}78$
  - $\rho(3) = 0{,}8^3 = 0{,}512$
- Timpul de înjumătățire $= \ln 0{,}5 / \ln 0{,}8 = 3{,}11$ perioade
  - după aproximativ 3 perioade, jumătate din șoc a dispărut
- Ponderile ψ: $1; 0{,}8; 0{,}64; 0{,}512; \ldots$: un șoc de 1 azi îl mută pe X_{t+2} cu 0,64

AR(1) pentru trei valori ale lui ϕ



- Aceleași 200 de șocuri, $\phi = 0{,}9;\ 0{,}3;\ -0{,}7$; jos: ACF de selecție (bare) și $\rho(h) = \phi^h$ (puncte)

TSA_ch2_ar_models

Interpretarea traiectoriilor AR(1)

- $\phi = 0{,}9$: abateri lungi de la 0; $\hat{\rho}(1) = 0{,}88$ de selecție, varianța 4,98 (teoretic 5,26)
  - persistența puternică seamănă cu un trend pe ferestre scurte
- $\phi = 0{,}3$: aproape de zgomot alb; doar $\hat{\rho}(1) = 0{,}35$ iese clar din bandă
  - varianța 1,18 (teoretic 1,10)
- $\phi = -0{,}7$: zigzag; $\hat{\rho}(1) = -0{,}58$, $\hat{\rho}(2) = 0{,}34$, semne alternante
  - în economie: depășire și corecție sau o eroare de măsurare corectată în perioada următoare
- Cu $T = 200$, ACF de selecție este sub cea teoretică la laguri mari: o deplasare de eșantion mic, care afectează și estimatorii (secțiunea 5)

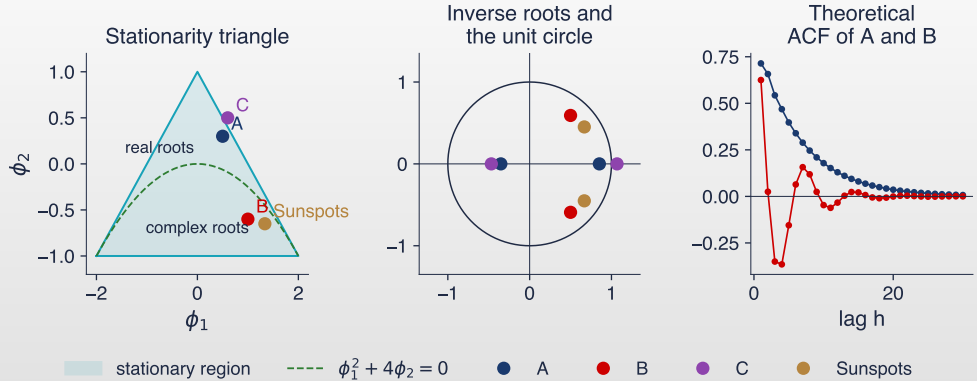
Modelul AR(p) și polinomul lui caracteristic

- $\phi(L)X_t = c + \varepsilon_t$, cu $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, un polinom într-un număr complex z
  - factorizăm: $\phi(z) = (1 - \lambda_1 z) \cdots (1 - \lambda_p z)$, unde $z_i = 1/\lambda_i$ sînt rădăcinile
- **Condiția de staționaritate (cauzalitate)**: toate rădăcinile lui $\phi(z) = 0$ sînt **în afara cercului unitate**, $|z_i| > 1$
  - echivalent: toate **rădăcinile inverse** λ_i sînt în interior, $|\lambda_i| < 1$ (forma afișată de programe)
  - motivul: $1/(1 - \lambda L) = \sum_j \lambda^j L^j$ converge doar dacă $|\lambda| < 1$; fiecare factor este un AR(1)
- O rădăcină pe cerc ($|z| = 1$, de exemplu $z = 1$) este o **rădăcină unitară**: nestaționaritate, subiectul Capitolului 3
  - verificare rapidă, necesară: $\phi_1 + \cdots + \phi_p < 1$ (adică $\phi(1) > 0$)

AR(2): triunghiul de staționaritate

- $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$ este staționar dacă și numai dacă
  - $\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad |\phi_2| < 1$
- Rădăcinile sînt complexe cînd $\phi_1^2 + 4\phi_2 < 0$
  - atunci ACF este un **cosinus amortizat**: pseudo-cicluri cu lungimea medie $2\pi / \arccos\left(\phi_1/(2\sqrt{-\phi_2})\right)$
  - factorul de amortizare pe perioadă: $\sqrt{-\phi_2}$, modulul rădăcinilor inverse
- **Ecuațiile Yule–Walker** pentru ACF: înmulțim cu X_{t-h}, aplicăm media, împărțim la $\gamma(0)$
  - $\rho(1) = \phi_1 + \phi_2\rho(1)$, deci $\rho(1) = \phi_1/(1-\phi_2)$
  - $\rho(h) = \phi_1\rho(h-1) + \phi_2\rho(h-2)$ pentru $h \geq 2$: ACF respectă aceeași recurență ca procesul

Trei procese AR(2) și petele solare



- A: $\phi = (0{,}5; 0{,}3)$; B: $(1{,}0; -0{,}6)$; C: $(0{,}6; 0{,}5)$; Sunspots: AR(2)-ul lui Yule, estimat în secțiunea 10
- Mijloc: rădăcinile inverse λ_i; staționar înseamnă în interiorul cercului

TSA_ch2_ar_models

Interpretarea exemplelor AR(2)

- **A**: rădăcini reale $-2{,}84$ și $1{,}17$, ambele în afara cercului: staționar
  - $\rho(1) = 0{,}5/0{,}7 = 0{,}714$, $\rho(2) = 0{,}5\rho(1) + 0{,}3 = 0{,}657$: descreștere lentă și netedă
- **B**: rădăcini complexe $0{,}833 \pm 0{,}986\,i$, modul $1{,}291 > 1$: staționar
  - ACF: o undă cu perioada de 7,2 laguri, amortizată cu 0,775 pe lag: un **ciclu stochastic**
- **C**: $\phi_1 + \phi_2 = 1{,}1 > 1$; o rădăcină $0{,}936 < 1$: proces **nestaționar** (exploziv)
  - fiecare coeficient este sub 1, totuși procesul explodează: verificați întotdeauna rădăcinile, nu coeficienții unul cîte unul
- Rădăcinile complexe sînt mecanismul prin care modelele liniare produc ciclurile economice și ciclul solar de 11 ani

PACF a unui AR(p) se anulează după lagul p

- ϕ_{hh} = ultimul coeficient din cea mai bună predicție liniară a lui X_t din $X_{t-1}, \ldots, X_{t-h}$ (Capitolul 1)
  - calculat din ACF prin recursia Durbin–Levinson (Durbin (1960))
- Pentru un AR(p) și $h > p$, cel mai bun predictor este chiar modelul: $\phi_1 X_{t-1} + \cdots + \phi_p X_{t-p}$
  - deci $\phi_{hh} = 0$ pentru $h > p$, iar $\phi_{pp} = \phi_p$
- **Regula de identificare**: ACF descrește treptat, PACF are p valori semnificative și apoi rămîne în bandă: AR(p)
  - banda pentru PACF de selecție a unui AR(p) la lagurile $h > p$: $\pm 1{,}96/\sqrt{T}$

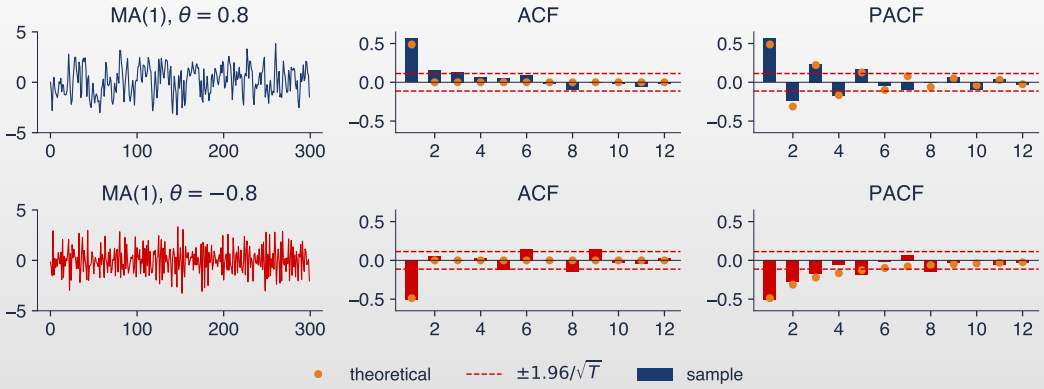
Recapitulare: modele autoregresive

- AR(p): $\phi(L)X_t = c + \varepsilon_t$; staționar dacă și numai dacă toate rădăcinile lui $\phi(z)$ sînt în afara cercului unitate
- Media $\mu = c/\phi(1)$; ACF din recurența Yule–Walker, o combinație de descreșteri geometrice și unde amortizate
- AR(1): $\rho(h) = \phi^h$, $\psi_j = \phi^j$, timpul de înjumătățire $\ln 0{,}5 / \ln \phi$
- PACF se anulează după lagul p: semnătura unui AR(p)

Modelul MA(q)

- $X_t = \mu + \theta(L)\varepsilon_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q}$
  - o sumă Wold finită: **întotdeauna staționar**, pentru orice θ
- Momentele (deduse în Seminarul 1 și în Capitolul 1): cu $\theta_0 = 1$
  - $E[X_t] = \mu$; $\gamma(h) = \sigma^2 \sum_{j=0}^{q-h} \theta_j\theta_{j+h}$ pentru $0 \le h \le q$; $\gamma(h) = 0$ pentru $h > q$
- **Regula de identificare**: ACF **se anulează** după lagul q; PACF descrește treptat
  - imaginea în oglindă a regulii pentru AR(p)
  - banda pentru ACF de selecție la lagurile $h > q$ (Bartlett): $\pm 1{,}96\sqrt{(1 + 2\sum_{j=1}^{q} \hat{\rho}(j)^2)/T}$, mai largă decît $\pm 1{,}96/\sqrt{T}$

MA(1): $\theta = 0{,}8$ și $\theta = -0{,}8$



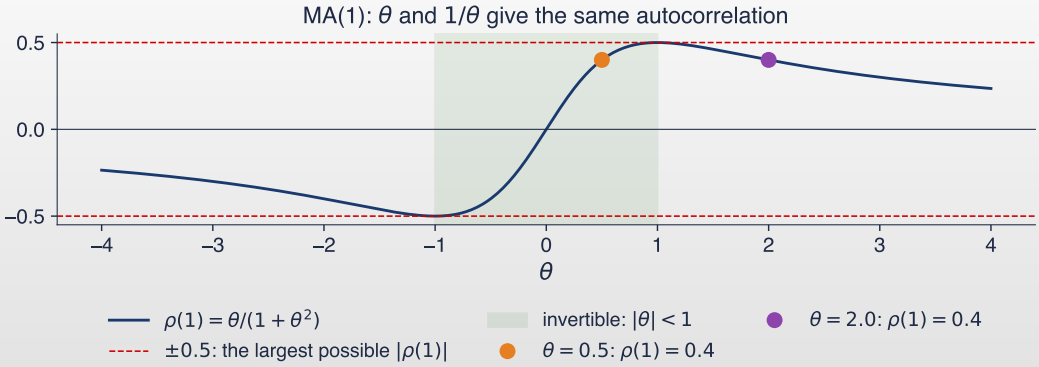
- $T = 300$; ACF are o singură valoare semnificativă, PACF descrește (cu semne alternante cînd $\theta > 0$)

TSA_ch2_ma_models

Interpretarea corelogramelor MA(1)

- Teoretic: $\rho(1) = \theta/(1+\theta^2) = \pm 0{,}488$, $\rho(h) = 0$ pentru $h \geq 2$
  - de selecție: $\hat{\rho}(1) = 0{,}56$ și $-0{,}51$
- PACF nu se anulează: $\hat{\phi}_{22} = -0{,}23$ pentru $\theta = 0{,}8$
  - un MA(1) este un AR(∞) (slide-urile următoare), deci fiecare valoare trecută ajută puțin
- Un MA(1) are o memorie de exact o perioadă: un șoc de azi afectează valoarea de mîine, apoi efectul lui dispare
- Exemple economice: observații care se suprapun (secțiunea 9), o eroare de măsurare, oscilația bid–ask a prețurilor

Cea mai mare autocorelație a unui MA(1)



- $\rho(1) = \theta/(1+\theta^2)$ ca funcție de θ

TSA_ch2_ma_models

Interpretarea curbei $\rho(1)$

- Maximul $|\rho(1)| = 0{,}5$ în $\theta = \pm 1$: un MA(1) nu poate produce niciodată $|\rho(1)| > 0{,}5$
  - dacă $\hat{\rho}(1) = 0{,}7$ și apoi nimic: nu este un MA(1); încercați un AR sau un MA(q) cu mai mulți termeni
- $\theta = 0{,}5$ și $\theta = 2$ dau același $\rho(1) = 0{,}4$
  - în general, θ și $1/\theta$ dau aceeași ACF (cu σ^2 înmulțit cu θ^2)
  - datele nu le pot deosebi: o **problemă de identificare**
- Convenția care elimină ambiguitatea: invertibilitatea, $|\theta| < 1$ (slide-ul următor)

Invertibilitatea

- MA(1): $\varepsilon_t = X_t - \theta\varepsilon_{t-1}$; substituim repetat ($\mu = 0$):
  - $\varepsilon_t = X_t - \theta X_{t-1} + \theta^2 X_{t-2} - \cdots = \sum_{j\geq 0}(-\theta)^j X_{t-j}$, dacă $|\theta| < 1$
  - o **reprezentare AR(∞)**: $X_t = \theta X_{t-1} - \theta^2 X_{t-2} + \cdots + \varepsilon_t$
- **Invertibil**: șocurile pot fi recuperate din trecutul lui X; condiția: toate rădăcinile lui $\theta(z) = 0$ în afara cercului unitate
  - MA(1): $|\theta| < 1$; aceeași algebră ca la staționaritatea AR, aplicată lui $\theta(z)$
- Motivația
  - unicitate: dintre θ și $1/\theta$ doar unul este invertibil
  - prognoză: șocurile ε_t se calculează din date, deci prognozele le pot folosi
  - programele raportează estimări invertibile; o rădăcină aproape de cerc semnalează supradiferențierea (Capitolul 1)

Exemplu rezolvat: $\theta = 2$ sau $\theta = 0{,}5$?

- $X_t = \varepsilon_t + 2\varepsilon_{t-1}$, $\sigma^2 = 1$; și $Y_t = u_t + 0{,}5u_{t-1}$, $\mathrm{Var}(u_t) = 4$
- X: $\gamma(0) = 1 \cdot (1 + 4) = 5$, $\gamma(1) = 2 \cdot 1 = 2$
  - Y: $\gamma(0) = 4 \cdot (1 + 0{,}25) = 5$, $\gamma(1) = 0{,}5 \cdot 4 = 2$
- Aceeași medie, aceleași autocovariațe: cu șocuri gaussiene, **aceeași distribuție**
  - oricît de multe date am avea, nu le putem separa
- Doar Y este invertibil: $u_t = Y_t - 0{,}5Y_{t-1} + 0{,}25Y_{t-2} - \ldots$
  - pentru X, aceeași sumă cu ponderile $(-2)^j$ explodează
- Regula: raportăm versiunea invertibilă ($\theta = 0{,}5$, $\sigma^2 = 4$)

Recapitulare: modele de medie mobilă

- MA(q) este întotdeauna staționar; ACF lui se anulează după lagul q, PACF descrește
- MA(1): $\rho(1) = \theta/(1+\theta^2)$, cel mult 0,5 în valoare absolută
- Invertibil dacă și numai dacă rădăcinile lui $\theta(z)$ sînt în afara cercului unitate; atunci MA = AR(∞)
- θ și $1/\theta$ sînt echivalente observațional; îl păstrăm pe cel invertibil

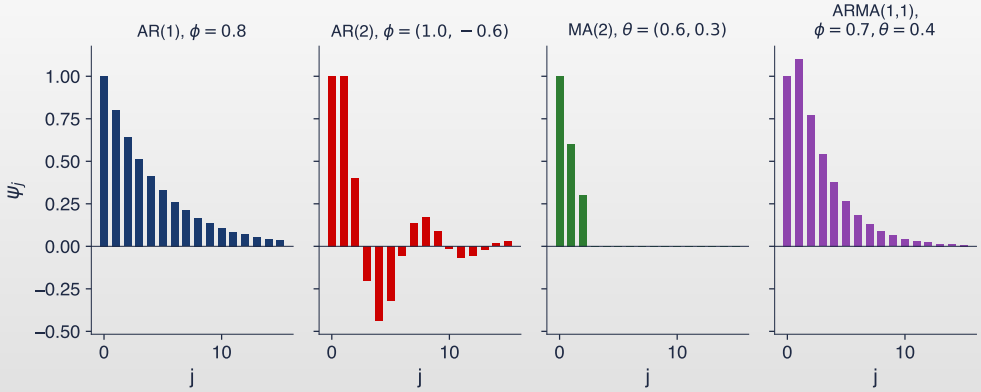
Modelul ARMA(p, q)

- $\phi(L)(X_t - \mu) = \theta(L)\varepsilon_t$, adică
  - $X_t = c + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$, $c = \mu\,\phi(1)$
- Trei condiții (Brockwell și Davis (2016), cap. 3)
  - **staționar și cauzal**: rădăcinile lui $\phi(z)$ în afara cercului unitate; atunci $X_t - \mu = \psi(L)\varepsilon_t$, $\psi(L) = \theta(L)/\phi(L)$
  - **invertibil**: rădăcinile lui $\theta(z)$ în afara cercului unitate; atunci $\pi(L)(X_t - \mu) = \varepsilon_t$, $\pi(L) = \phi(L)/\theta(L)$
  - **fără factori comuni**: $\phi(z)$ și $\theta(z)$ nu au nicio rădăcină comună
- ARMA(1,1) poate reproduce o ACF care descrește cu rata ϕ pornind de la o primă valoare fixată de θ (sub ϕ dacă $\theta < 0$, peste ϕ dacă $\theta > 0$); un AR sau un MA pur ar avea nevoie de mulți termeni

Ponderile ψ și răspunsul la impuls

- Din $\phi(L)\psi(L) = \theta(L)$, egalăm coeficienții lui L^j:
  - $\psi_0 = 1, \quad \psi_j = \theta_j + \sum_{i=1}^{\min(j,p)} \phi_i \psi_{j-i} \quad (\theta_j = 0 \text{ pentru } j > q)$
- ARMA(1,1): $\psi_1 = \phi + \theta$, $\psi_j = \phi^{j-1}(\phi + \theta)$ pentru $j \geq 1$
  - exemplu $\phi = 0{,}7$, $\theta = 0{,}4$: $\psi_1 = 1{,}10$, $\psi_2 = 0{,}77$
- **Răspunsul la impuls**: $\psi_j = \partial X_{t+j} / \partial \varepsilon_t$, efectul unui șoc unitar după j perioade
  - același obiect determină erorile de prognoză (secțiunea 7) și răspunsurile la impuls din modelele VAR (Capitolul 6)

Răspunsurile la impuls ale a patru modele



- ψ_j, $j = 0, \ldots, 15$, din recurența de mai sus

TSA_ch2_arma_acf

Interpretarea răspunsurilor la impuls

- AR(1): descreștere geometrică $0{,}8^j$; șocul nu dispare niciodată complet, dar se stinge suficient de repede pentru ca $\sum \psi_j^2 < \infty$
- AR(2) cu rădăcini complexe: răspunsul își schimbă semnul, $\psi_3 = -0{,}20$, $\psi_4 = -0{,}44$: un șoc creează un ciclu
- MA(2): exact zero după două perioade: memorie finită
- ARMA(1,1): termenul MA ridică primul răspuns la $\psi_1 = 1{,}10 > 1$; după aceea preia partea AR

Momentele procesului ARMA(1,1)

- $X_t = \phi X_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$, $|\phi| < 1$, $|\theta| < 1$, $\phi + \theta \neq 0$
- Varianța, din $\gamma(0) = \sigma^2 \sum_j \psi_j^2$:
  - $\gamma(0) = \sigma^2 \dfrac{1 + 2\phi\theta + \theta^2}{1 - \phi^2}$; pentru $\phi = 0{,}7$, $\theta = 0{,}4$, $\sigma^2 = 1$: 3,37
- Autocorelațiile:
  - $\rho(1) = \dfrac{(1 + \phi\theta)(\phi + \theta)}{1 + 2\phi\theta + \theta^2}$, $\quad \rho(h) = \phi\, \rho(h-1)$ pentru $h \geq 2$
  - exemplu: $\rho(1) = 0{,}819$, $\rho(2) = 0{,}573$
- ACF descrește ca la un AR(1) începînd cu lagul 1, dar $\rho(1) \neq \phi$: partea MA mută doar punctul de pornire

Factori comuni: un model prea mare

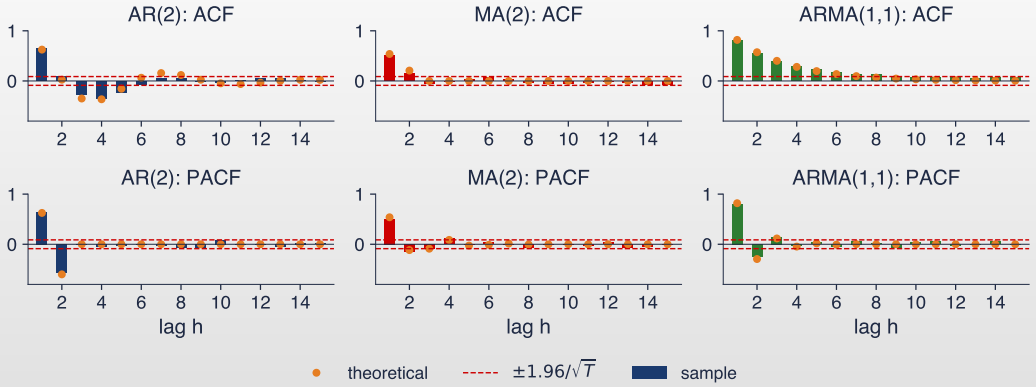
- $X_t = 0{,}5X_{t-1} + \varepsilon_t - 0{,}5\varepsilon_{t-1}$
  - $(1 - 0{,}5L)X_t = (1 - 0{,}5L)\varepsilon_t$: simplificăm factorul, $X_t = \varepsilon_t$, zgomot alb
- În general, $\phi = -\theta$ într-un ARMA(1,1) dă zgomot alb, pentru orice valoare
  - verosimilitatea este plată de-a lungul dreptei $\phi = -\theta$: parametrii nu sînt identificați
- Simptome în rezultatele programelor
  - estimări mari și de semne opuse, de exemplu $\hat{\phi} = 0{,}9$, $\hat{\theta} = -0{,}88$, cu erori standard foarte mari
  - o log-verosimilitate care nu este mai bună decît a modelului mai mic
- Remediul: reducem p și q împreună; preferăm cel mai mic model cu reziduuri de tip zgomot alb

Identificarea: tabelul ACF–PACF

Modelul	**ACF**	**PACF**
Zgomot alb	niciun lag semnificativ	niciun lag semnificativ
AR(p)	descreşte (geometric sau undă amortizată)	**se anulează** după lagul p
MA(q)	**se anulează** după lagul q	descreşte
ARMA(p, q)	descreşte după lagul $q - p$	descreşte după lagul $p - q$
Rădăcină unitară (cap. 3)	descreştere foarte lentă, aproape liniară	$\hat{\phi}_{11} \approx 1$

- Tabelul propune **candidaţi**; criteriile informaţionale şi diagnosticarea decid (secţiunile 5–6)
- Modelele ARMA mixte sînt greu de recunoscut din ochi: ambele funcţii doar descresc

ACF și PACF teoretice și de selecție



- AR(2) $\phi = (1{,}0; -0{,}6)$, MA(2) $\theta = (0{,}6; 0{,}3)$, ARMA(1,1) $\phi = 0{,}7$, $\theta = 0{,}4$; cîte o traiectorie simulată cu $T = 500$

TSA_ch2_arma_acf

Interpretarea celor trei tipare

- AR(2): PACF are două valori semnificative, $\phi_{22} = \phi_2 = -0{,}60$, apoi zero; ACF este o undă amortizată
  - identificarea este ușoară: citim p din PACF
- MA(2): $\rho(1) = 0{,}538$, $\rho(2) = 0{,}207$, apoi zero; PACF descrește lent
  - ACF de selecție la lagul 2 iese abia ușor din bandă: pe date reale, MA(1) și MA(2) pot semăna
- ARMA(1,1): ambele descresc; ACF seamănă cu un AR(1), PACF are o a doua valoare, mai mică
  - un analist care citește doar PACF ar alege AR(2) sau AR(3): apropiat, dar cu mai mulți parametri
- Corelogramele de selecție sînt zgomotoase chiar cu $T = 500$; cu 100 de observații trimestriale, mult mai mult

Recapitulare: modele ARMA

- ARMA(p, q): $\phi(L)(X_t - \mu) = \theta(L)\varepsilon_t$; cauzal dacă $\phi(z)$ are rădăcinile în afara cercului, invertibil dacă $\theta(z)$ le are
- Ponderile ψ prin recurența $\psi_j = \theta_j + \sum \phi_i \psi_{j-i}$: răspunsul la impuls
- Fără factori comuni; o verosimilitate plată și estimări opuse semnalează unul
- ACF se anulează: MA; PACF se anulează: AR; ambele descresc: ARMA

Trei estimatori

- **Yule–Walker** (metoda momentelor, Yule (1927); Walker (1931)): înlocuim $\rho(h)$ cu $\hat{\rho}(h)$ în ecuațiile Yule–Walker
  - doar pentru AR pur; rapid; dă întotdeauna un model staționar; mai puțin precis cînd rădăcinile sînt aproape de cerc
- **Cele mai mici pătrate condiționate**: minimizăm $\sum_t \varepsilon_t^2$, tratînd primele valori ca fixe
  - AR(p): metoda celor mai mici pătrate (OLS) pentru X_t pe $1, X_{t-1}, \ldots, X_{t-p}$
  - MA și ARMA: reziduurile $\varepsilon_t = X_t - \mu - \theta\varepsilon_{t-1} - \ldots$ se calculează recursiv, pornind de la $\varepsilon_0 = 0$; minimizare numerică
- **Verosimilitatea maximă exactă gaussiană** (MLE): maximizăm densitatea comună a lui $(X_1, \ldots, X_T)$
  - folosește și primele observații; metoda implicită a lui `ARIMA` din `statsmodels` (filtrul Kalman, Capitolul 10)
- În eșantioane mari, cele trei coincid pentru modelele AR pure; MLE este referința pentru MA și ARMA

Yule–Walker pentru un AR(p)

- Pentru $h = 1, \ldots, p$: $\rho(h) = \phi_1 \rho(h-1) + \cdots + \phi_p \rho(h-p)$
  - matricial, $R\,\phi = \rho$, $R_{ij} = \rho(|i-j|)$; $\phi = (\phi_1, \ldots, \phi_p)^\top$, $\rho = (\rho(1), \ldots, \rho(p))^\top$
  - R: matricea $p \times p$ a autocorelațiilor; notația $\hat{}$ indică varianta de selecție
  - estimatorul $\hat{\phi} = \hat{R}^{-1}\hat{\rho}$; varianța zgomotului $\hat{\sigma}^2 = \hat{\gamma}(0)(1 - \hat{\phi}^\top \hat{\rho})$
- AR(2) prin regula lui Cramer:
  - $\hat{\phi}_1 = \dfrac{\hat{\rho}_1(1-\hat{\rho}_2)}{1-\hat{\rho}_1^2}, \quad \hat{\phi}_2 = \dfrac{\hat{\rho}_2 - \hat{\rho}_1^2}{1-\hat{\rho}_1^2}$
  - $\hat{\phi}_2$ este PACF de selecție la lagul 2 (Capitolul 1)



Gilbert Walker (1868–1958), care a extins metoda lui Yule în 1931

Foto: autor necunoscut (1925); domeniu public; Wikimedia Commons

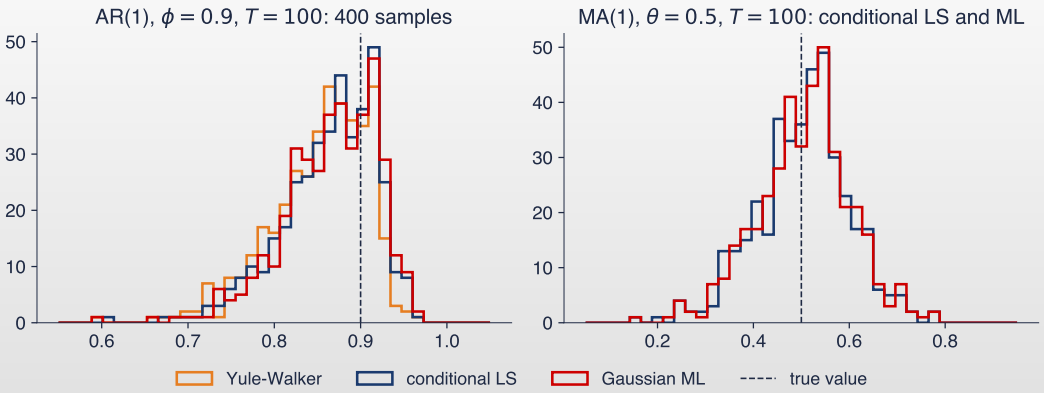
Exemplu rezolvat: Yule–Walker pentru creșterea PIB-ului României

- Creșterea anuală a PIB-ului real, T1 2000–T2 2026 (secțiunea 9): $\hat{\rho}(1) = 0{,}62$, $\hat{\rho}(2) = 0{,}40$, $\hat{\gamma}(0) = 18{,}77$
- Determinantul $1 - \hat{\rho}_1^2 = 0{,}620$
  - $\hat{\phi}_1 = 0{,}62\,(1 - 0{,}40)/0{,}620 = 0{,}600$
  - $\hat{\phi}_2 = (0{,}40 - 0{,}62^2)/0{,}620 = 0{,}027$
- $\hat{\sigma}^2 = 18{,}77\,(1 - \hat{\phi}_1\hat{\rho}_1 - \hat{\phi}_2\hat{\rho}_2) = 11{,}62$
- $\hat{\phi}_2 \approx 0$: al doilea lag nu aduce nimic, AR(1) ar fi suficient
  - dar este familia AR cea potrivită? Secțiunea 9 arată că un MA(3) se potrivește mai bine

Verosimilitatea maximă

- Verosimilitatea gaussiană prin **descompunerea erorilor de predicție**
  - $f(x_1, \ldots, x_T) = \prod_t f(x_t \mid x_{t-1}, \ldots, x_1)$; fiecare factor este $N(\hat{x}_{t|t-1}, v_t)$
  - $\ln L = -\frac{1}{2} \sum_{t=1}^{T} \left[\ln(2\pi v_t) + (x_t - \hat{x}_{t|t-1})^2 / v_t \right]$
  - $\hat{x}_{t|t-1}$: predicția cu un pas; v_t: varianța erorii ei, mai mare pentru primele observații
- AR(1) de mînă: $X_1 \sim N(\mu, \sigma^2/(1-\phi^2))$, apoi $X_t \mid X_{t-1} \sim N(c + \phi X_{t-1}, \sigma^2)$
  - dacă renunțăm la primul factor, obținem cele mai mici pătrate condiționate
- Erorile standard (se, standard error) din curbura lui $\ln L$; asimptotic
  - AR(1): $\text{se}(\hat{\phi}) \approx \sqrt{(1-\phi^2)/T}$; MA(1): $\text{se}(\hat{\theta}) \approx \sqrt{(1-\theta^2)/T}$
  - rapoartele t, $\hat{\phi}/\text{se}$, se compară cu $\pm 1{,}96$, ca în regresie
  - `statsmodels` raportează implicit erorile standard **OPG** (produsul exterior al scorurilor): aceeași limită dacă modelul este corect, dar mai mici decît cele din curbură cînd erorile au cozi groase; erorile standard din acest capitol sînt OPG

Distribuțiile de selecție ale estimatorilor



⊡ 400 de eșantioane simulate, fiecare cu $T = 100$; stînga: AR(1) cu $\phi = 0{,}9$; dreapta: MA(1) cu $\theta = 0{,}5$

TSA_ch2_estimation

Interpretarea comparației estimatorilor

- AR(1), $\phi = 0{,}9$: mediile 0,853 (Yule–Walker), 0,864 (cele mai mici pătrate condiționate), 0,866 (MLE)
  - toți trei sînt deplasați în jos; aproximarea de manual a deplasării celor mai mici pătrate este $-(1+3\phi)/T = -0{,}037$
  - Yule–Walker este cel mai deplasat aproape de cercul unitate
- Abaterile standard 0,056; 0,055; 0,054, peste valoarea asimptotică 0,044
  - distribuția este asimetrică la stînga: intervalele asimptotice sînt optimiste aproape de $\phi = 1$
- MA(1), $\theta = 0{,}5$: cele mai mici pătrate condiționate 0,504 (abaterea standard 0,097), MLE 0,506 (0,098); valoarea asimptotică 0,087
- Regula practică: folosiți MLE; Yule–Walker doar pentru valori de pornire rapide sau în scop didactic

Recapitulare: estimarea

- Yule–Walker: momente, AR pur, $\hat{\phi} = \hat{R}^{-1}\hat{\rho}$
- Cele mai mici pătrate condiționate: OLS pentru AR, o recurență pentru MA
- MLE exactă gaussiană: descompunerea erorilor de predicție; standardul programelor
- Coeficienții AR persistenți sînt deplasați în jos în eșantioane mici

Criterii informaționale

- Ajustarea se îmbunătățește cu fiecare parametru; criteriile impun o **penalizare** pe parametru
  - k = numărul parametrilor estimați (ϕ, θ, μ, σ^2), $\ln L$ = log-verosimilitatea maximizată
- **AIC** $= -2 \ln L + 2k$ (Akaike (1974))
  - **AICc** $= \text{AIC} + 2k(k+1)/(T-k-1)$, pentru T mic (Hurvich și Tsai (1989))
- **BIC** $= -2 \ln L + k \ln T$ (Schwarz (1978))
  - $\ln T > 2$ de îndată ce $T \geq 8$: BIC preferă modele mai mici
- Alegem modelul cu valoarea **cea mai mică**; comparăm doar modele estimate pe aceleași observații



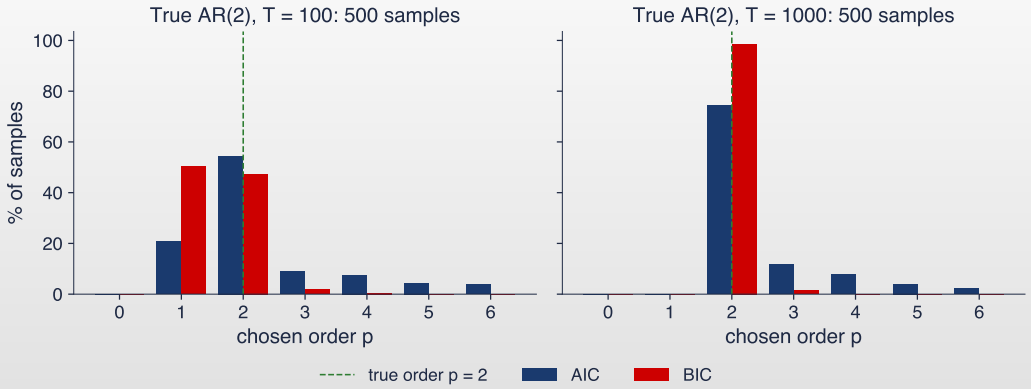
Hirotugu Akaike (1927–2009), Institutul de Matematică Statistică, Tokyo

Foto: The Institute of Statistical Mathematics (2017); CC BY-SA 4.0; Wikimedia Commons

AIC sau BIC?

- **BIC este consistent**: dacă modelul adevărat se află printre candidați, BIC îl găsește cu probabilitate $\to 1$
  - dar poate alege un model prea mic cînd T este mic
- **AIC este orientat spre prognoză**: estimează ajustarea așteptată în afara eșantionului (distanța Kullback–Leibler)
  - supraparametrizează cu probabilitate pozitivă, chiar pentru T mare
- Practica (Hyndman și Athanasopoulos (2021), cap. 9; Hyndman și Khandakar (2008))
  - căutările automate (`auto.arima`, `pmdarima`) minimizează AICc pe o grilă de valori (p, q)
  - cînd AIC și BIC nu coincid, păstrăm ambii candidați și lăsăm diagnosticarea și prognozele să decidă

Cît de des găsesc AIC și BIC ordinul adevărat



- Modelul adevărat AR(2), $\phi = (0{,}5; 0{,}25)$; AR(p), $p = 0, \ldots, 6$, estimate prin OLS pe același eșantion; 500 de eșantioane pe panou

TSA_ch2_selection

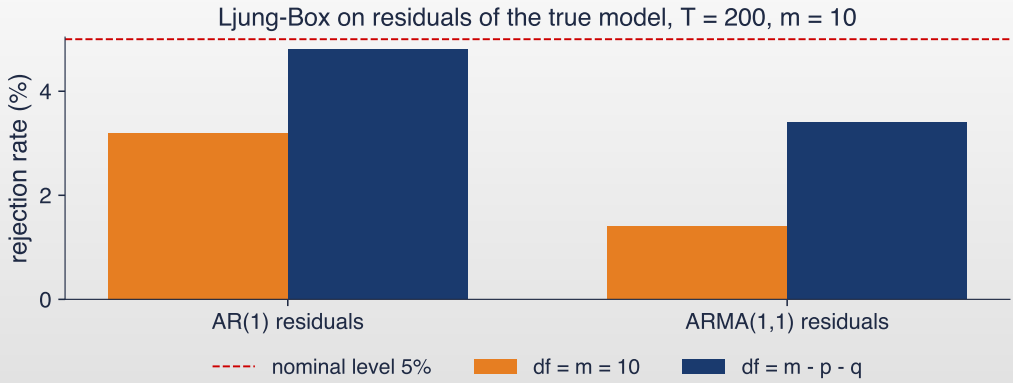
Interpretarea experimentului de selecție

- $T = 100$: AIC găsește $p = 2$ în 54,4% din eșantioane, BIC în 47,4%
  - BIC alege un model prea mic în 50,4% din cazuri (valoarea mică $\phi_2 = 0{,}25$ este greu de detectat); AIC alege unul prea mare în 24,8%
- $T = 1000$: BIC 98,6%, AIC 74,4%
  - BIC este consistent; AIC încă supraparametrizează în 25,6% din eșantioane
- Niciun criteriu nu este o dovadă: cu date macroeconomice trimestriale ($T \approx 100$), ordinul este incert
- Supraparametrizarea costă puțin în prognoză; un model prea mic lasă corelație în reziduuri, pe care o detectează diagnosticarea

Diagnosticarea reziduurilor

- Dacă modelul este corect, reziduurile $\hat{\varepsilon}_t$ se comportă ca un zgomot alb
  - le reprezentăm grafic: fără tipare, fără schimbări de variață, fără valori izolate foarte mari
  - ACF a reziduurilor în banda $\pm 1{,}96/\sqrt{T}$, cu excepția a aproximativ 1 lag din 20
- **Ljung–Box** pentru reziduuri (Ljung și Box (1978)): $Q^*(m) = T(T+2)\sum_{h=1}^{m} \hat{\rho}_{\hat{\varepsilon}}(h)^2/(T-h)$
  - în ipoteza H_0 (model corect), $Q^*(m) \approx \chi^2(m-p-q)$ (Box și Pierce (1970)): estimarea a $p+q$ coeficienți consumă $p+q$ grade de libertate
  - alegem m mult peste $p+q$: $m = 10$ pentru date anuale sau trimestriale, $m = 2s$ pentru date sezoniere (Hyndman și Athanasopoulos (2021))
- **Normalitatea**: Jarque–Bera (Jarque și Bera (1987)) $\text{JB} = \frac{T}{6}\left(S^2 + (K-3)^2/4\right) \approx \chi^2(2)$; graficul QQ
  - S = asimetria, K = kurtosis-ul (0 și 3 pentru distribuția Normală); respingem normalitatea la 5% dacă $\text{JB} > 5{,}99$, cuantila de 95% a lui $\chi^2(2)$
  - reziduurile ne-normale nu deplasează $\hat{\phi}$, dar fac greșite intervalele construite cu distribuția Normală
- **Pătratele reziduurilor**: testul Ljung–Box pentru $\hat{\varepsilon}_t^2$ detectează volatility clustering (Capitolul 5)

Justificarea celor $m - p - q$ grade de libertate



- Testul Ljung–Box la 5% pentru reziduurile modelului **adevărat**; 1 000 de eșantioane AR(1), jumătate pentru ARMA(1,1)

TSA_ch2_selection

Interpretarea experimentului cu gradele de libertate

- Cu $m = 10$ grade de libertate testul respinge prea rar: 3,2% (AR(1)) și 1,4% (ARMA(1,1)) în loc de 5%
  - reziduurile sînt ajustate astfel încît să pară necorelate la primele laguri: ACF lor este mai mică decît a unui zgomot alb adevărat
- Cu $m - p - q$: 4,8% și 3,4%, aproape de nivelul nominal de 5%
  - corecția contează cel mai mult cînd $p + q$ este mare în raport cu m
- Programele (`acorr_ljungbox`) folosesc implicit m: transmiteți `model_df = p + q`
- Gradele de libertate greșite fac ca un model prost să pară acceptabil

Recapitulare: selecția modelului și diagnosticarea

- AIC $= -2 \ln L + 2k$, BIC $= -2 \ln L + k \ln T$; cîștigă valoarea cea mai mică; BIC alege modele mai mici
- BIC este consistent, AIC este orientat spre prognoză și uneori supraparametrizează
- Reziduurile: grafic, ACF, Ljung–Box cu $m - p - q$ grade de libertate, Jarque–Bera, Ljung–Box pe pătrate
- Un model este acceptat cînd reziduurile lui sînt zgomot alb; normalitatea contează pentru intervale

Prognoza optimă

- Prognoza lui X_{T+h} făcută la momentul T: $\hat{X}_{T+h|T} = E[X_{T+h} \mid X_T, X_{T-1}, \dots]$
  - minimizează eroarea pătratică medie (MSE) între toate funcțiile de trecut
- Rețeta: scriem modelul la $T + h$ și înlocuim
  - șocurile viitoare ε_{T+j}, $j \geq 1$, cu 0
  - valorile viitoare X_{T+j} cu prognozele lor; valorile trecute și reziduurile trecute rămîn neschimbate
- AR(1): $\hat{X}_{T+h|T} = \mu + \phi^h (X_T - \mu)$: **revenirea la medie** cu viteza ϕ
  - MA(q): $\hat{X}_{T+h|T} = \mu$ pentru $h > q$: după q pași modelul nu mai conține altă informație decît media

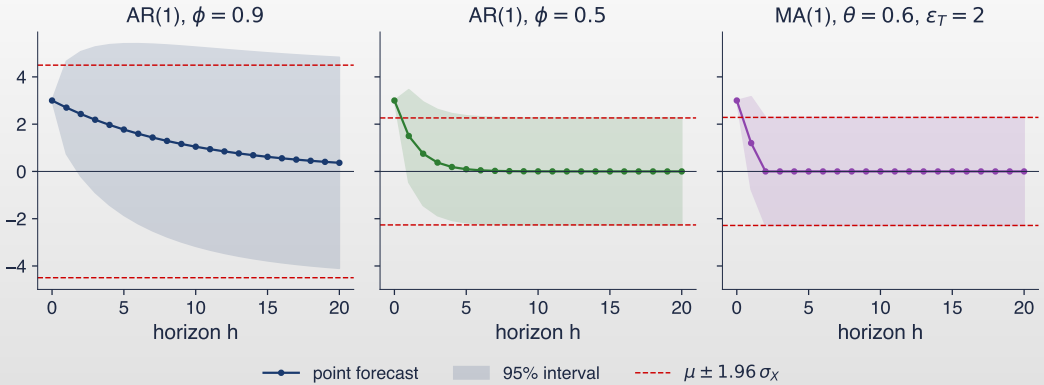
Erorile de prognoză și intervalele

- Din forma cauzală $X_{T+h} = \mu + \sum_j \psi_j \varepsilon_{T+h-j}$:
  - eroarea $e_{T+h} = X_{T+h} - \hat{X}_{T+h|T} = \varepsilon_{T+h} + \psi_1 \varepsilon_{T+h-1} + \cdots + \psi_{h-1}\varepsilon_{T+1}$
  - varianța $\sigma_h^2 = \sigma^2(1 + \psi_1^2 + \cdots + \psi_{h-1}^2)$, crescătoare în h
- **Intervalul de 95%**: $\hat{X}_{T+h|T} \pm 1{,}96\,\sigma_h$ (șocuri gaussiene)
  - $h = 1$: $\pm 1{,}96\,\sigma$; $h \to \infty$: $\sigma_h^2 \to \gamma(0)$, varianța necondiționată
- Limitele intervalului
  - incertitudinea parametrilor ($\hat{\phi}$ în loc de ϕ), incertitudinea modelului, cozile groase: acoperirea reală este de obicei sub 95%
- Erorile prognozelor făcute în același moment pentru orizonturile h și $h+1$ sînt corelate: au șocuri comune

Exemplu rezolvat: prognoza unui AR(1)

- $X_t = 2 + 0{,}8X_{t-1} + \varepsilon_t$, $\sigma^2 = 1$, $\mu = 10$; ultima valoare $X_T = 12$
- Prognozele punctuale
  - $\hat{X}_{T+1} = 10 + 0{,}8 \cdot 2 = 11{,}60$; $\hat{X}_{T+2} = 10 + 0{,}64 \cdot 2 = 11{,}28$; $\hat{X}_{T+10} = 10{,}21$
- Varianțele erorilor: $\sigma_1^2 = 1$, $\sigma_2^2 = 1 + 0{,}8^2 = 1{,}64$
  - intervalele de 95%: $h = 1$: $[9{,}64; 13{,}56]$; $h = 2$: $[8{,}77; 13{,}79]$
- Pe termen lung: prognoza $\to 10$, abaterea standard $\to \sqrt{\gamma(0)} = 1{,}67$

Revenirea la medie în prognoze



- Pornire din $X_T = 3$, $\mu = 0$, $\sigma = 1$; linie întreruptă: banda necondiționată $\mu \pm 1{,}96\,\sigma_X$

TSA_ch2_forecasting

Interpretarea traiectoriilor de prognoză

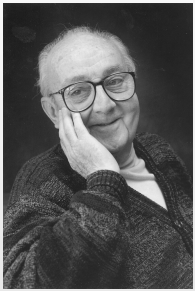
- $\phi = 0{,}9$: revenire lentă; după 5 perioade prognoza este încă 1,77, eroarea standard 1,85
  - intervalul se apropie de $\pm 1{,}96 \times 2{,}29$ abia după multe perioade
- $\phi = 0{,}5$: după 5 perioade prognoza este 0,094, practic media
- MA(1): un singur pas informativ ($\theta \varepsilon_T = 1{,}2$), apoi media; intervalul atinge lățimea de lungă durată la $h = 2$
- Un model staționar prognozează întotdeauna revenirea la medie: util pentru spread-uri și rate de creștere, înșelător dacă media se schimbă

Recapitulare: prognoza

- $\hat{X}_{T+h|T}$: înlocuim șocurile viitoare cu 0 și valorile viitoare cu prognoze
- Varianța erorii $\sigma^2 \sum_{j<h} \psi_j^2$; intervalele se lărgesc pînă la $\pm 1{,}96\sqrt{\gamma(0)}$
- Prognozele revin la μ; prognozele MA(q) sînt egale cu μ după q pași
- Intervalele ignoră incertitudinea parametrilor și a modelului

George Box și Gwilym Jenkins

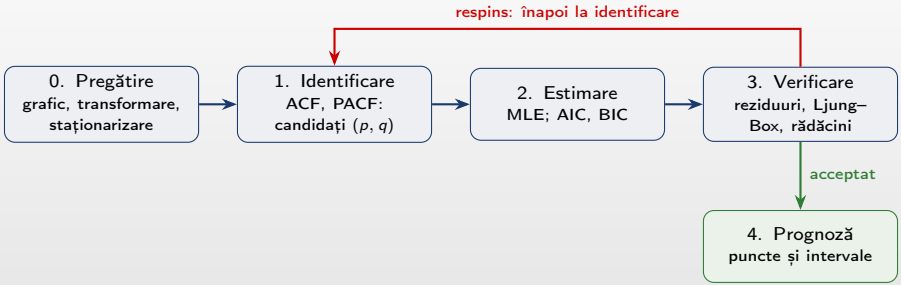
- **George E. P. Box** (1919–2013), statistician, Universitatea Wisconsin–Madison
  - „Toate modelele sînt greșite, dar unele sînt utile”
- **Gwilym M. Jenkins** (1932–1982), statistician, Universitatea Lancaster
  - *Time Series Analysis: Forecasting and Control*, 1970; a patra ediție, cu Reinsel: Box, Jenkins și Reinsel (2008)
- Contribuția lor: nu modele noi, ci o **metodă**
  - o buclă de identificare, estimare și verificare pe care o poate aplica oricine
  - testele Box–Pierce (1970) și Ljung–Box (1978) provin din acest program



George E. P. Box

Foto: DavidMCEddy; CC BY-SA 3.0; Wikimedia Commons

Bucla Box–Jenkins



- Pasul 0 ține de Capitolele 1 și 3: logaritmi, diferențe, teste de rădăcină unitară; ARMA cere o serie staționară
- Bucla se încheie cu cel mai mic model ale cărui reziduuri sînt zgomot alb și ale cărui rădăcini sînt bine în interiorul regiunii admise
- Verificările în afara eșantionului (Capitolul 0: eșantioane de antrenare și de test) completează metoda

Datele

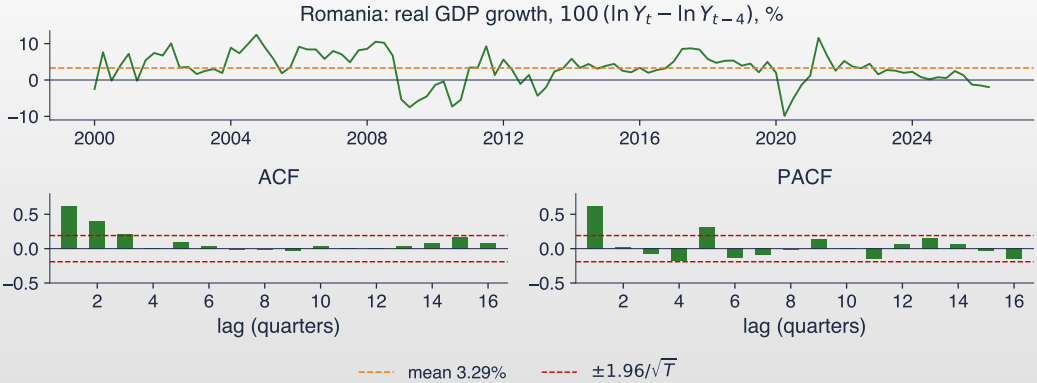
- PIB-ul real al României, volume înlănțuite (prețurile din 2010), trimestrial, neajustat sezonier; Eurostat, `namq_10_gdp`
  - sursa națională: Institutul Național de Statistică (INS)
- Creșterea anuală $y_t = 100\,(\ln Y_t - \ln Y_{t-\mathbf{4}})$, în %: elimină sezonalitatea (Capitolul 1)
  - T1 2000–T2 2026: $T = 106$ trimestre, media 3,29%, abaterea standard 4,35%
  - extremele: $-9{,}9\%$ în T2 2020, $+12{,}5\%$ în T4 2004; ultima valoare $-2{,}0\%$
- Este staționară? Rata de creștere nu are trend; testul formal este în Capitolul 3



Institutul Național de Statistică, București

Foto: Dan Mihai Pitea (2009); CC BY-SA 3.0; Wikimedia Commons

Pasul 1: identificarea



- Creșterea anuală a PIB-ului real al României, ACF și PACF cu $\pm 1{,}96/\sqrt{T} = \pm 0{,}19$

TSA_ch2_gdp_case

Interpretarea corelogramelor creșterii PIB

- ACF: 0,62; 0,40; 0,21, apoi $-0{,}01$ la lagul 4: **se anulează după lagul 3**
  - candidat: MA(3)
- PACF: 0,62 la lagul 1, apoi valori mici, cu excepția lui 0,32 la lagul 5
  - candidat: AR(1); lagul 5 poate fi întîmplător sau o urmă a sezonalității
- Motivul pentru MA(3): y_t este suma a patru rate trimestriale de creștere, $y_t \approx g_t + g_{t-1} + g_{t-2} + g_{t-3}$
  - dacă g_t este aproape de un zgomot alb (Seminarul 2, B2), ratele anuale consecutive au trei trimestre comune: un MA(3) prin construcție
- Ljung–Box $Q^*(8) = 64{,}9$ pentru seria însăși: puternic autocorelată, deci este nevoie de un model

Pasul 2: estimarea și criteriile informaționale

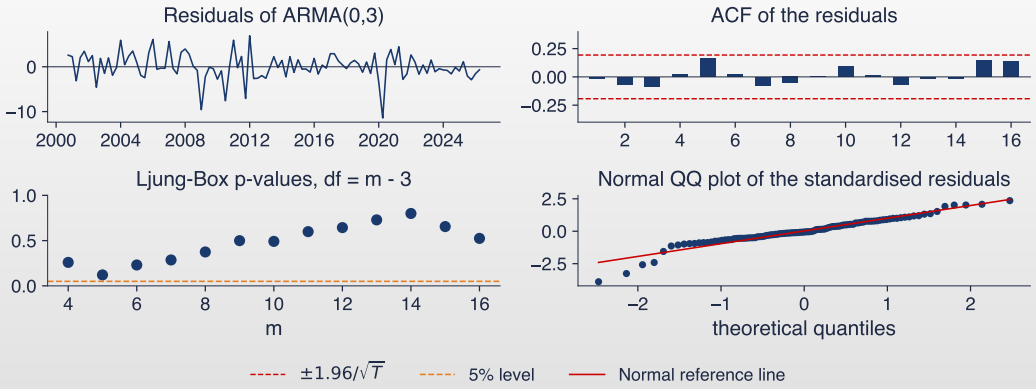
Modelul	k	$\ln L$	**AIC**	**BIC**	$\hat{\sigma}$	**LB, p-value**
ARMA(0,0)	2	−305,83	615,66	620,99	4,33	< 0,001
ARMA(1,0)	3	−279,67	565,33	573,32	3,38	0,023
ARMA(2,0)	4	−279,66	567,32	577,98	3,38	0,037
ARMA(1,1)	4	−279,66	567,32	577,98	3,38	0,013
ARMA(0,3)	5	−270,02	550,04	563,36	3,07	0,375
ARMA(1,3)	6	−268,48	548,97	564,95	3,02	0,648
ARMA(2,3)	7	−267,94	549,89	568,53	3,00	0,543

- MLE exactă gaussiană pe aceleași 106 trimestre; LB = Ljung–Box $Q^*(8)$ pentru reziduuri, cu $8 - p - q$ grade de libertate; grila completă $p, q \leq 3$ în Quantlet
- **BIC** alege MA(3); **AIC** alege ARMA(1,3), mai mic cu doar 1,07
- AR(1), candidatul sugerat de PACF, lasă autocorelație: p-value-ul testului LB este 0,023

Interpretarea modelului MA(3) estimat

- $\hat{y}_t = 3{,}17 + \hat{\varepsilon}_t + 0{,}834\,\hat{\varepsilon}_{t-1} + 0{,}675\,\hat{\varepsilon}_{t-2} + 0{,}591\,\hat{\varepsilon}_{t-3}$, $\hat{\sigma} = 3{,}07$
  - erorile standard (OPG): 1,03 (media), 0,083; 0,102; 0,082: toți cei trei termeni MA sînt semnificativi
- Coeficienții $\hat{\theta}_j$ scad de la 0,834 la 0,591: aproape de tiparul suprapunerii $(1, 1, 1)$, dar sub el
  - $(1, 1, 1)$ ar pune rădăcini pe cercul unitate; cel mai mic modul al rădăcinilor este aici $1{,}17 > 1$: invertibil
- ARMA(1,3): $\hat{\theta}_1 = 1{,}019 > 1$, dar rădăcinile au modulul cel puțin 1,10: tot invertibil
  - pentru $q > 1$, invertibilitatea se citește din rădăcini, nu din coeficienți luați separat
  - $\hat{\phi} = -0{,}284$ (eroarea standard OPG 0,143): un parametru în plus pentru un cîștig mic; BIC renunță la el
- Creșterea medie este de 3,17% pe an, dar cu o eroare standard de 1,03: datele care se suprapun conțin mai puțină informație decît sugerează T

Pasul 3: diagnosticarea modelului MA(3)



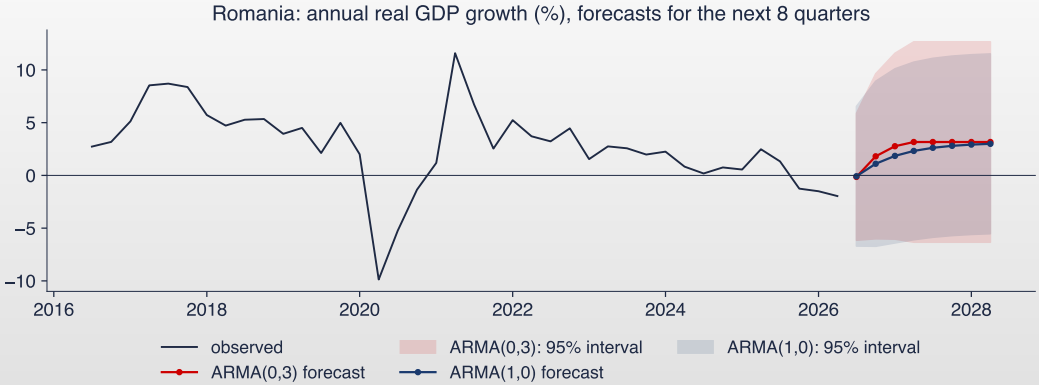
- Reziduurile, ACF lor, p-value-urile testului Ljung–Box pentru $m = 4, \ldots, 16$, cu $m - 3$ grade de libertate, graficul QQ față de distribuția Normală

TSA_ch2_gdp_case

Interpretarea diagnosticării MA(3)

- Fără autocorelație în reziduuri: $Q^*(8) = 5{,}35$ cu 5 grade de libertate, $p = 0{,}375$; $Q^*(16)$: $p = 0{,}525$
  - toate p-value-urile din grafic sînt peste 0,12; cu 8 în loc de 5 grade de libertate, p ar fi 0,720, prea liniștitor
- Nu este Normal: Jarque–Bera 39,0, $p < 0{,}001$; asimetria $-0{,}75$, kurtosis-ul 5,61
  - cel mai mare reziduu: $-11{,}4$ în T2 2020, $-3{,}89$ abateri standard: perioada de lockdown din pandemie
- Pătratele reziduurilor: testul Ljung–Box dă $p = 0{,}779$, fără volatility clustering la această frecvență
- Verdictul: MA(3) surprinde dependența; intervalele construite cu distribuția Normală subestimează riscul recesiunilor

Pasul 4: prognozele



- Prognozele MA(3) (BIC) și AR(1) pentru T3 2026–T2 2028, cu intervale de 95%

TSA_ch2_gdp_case

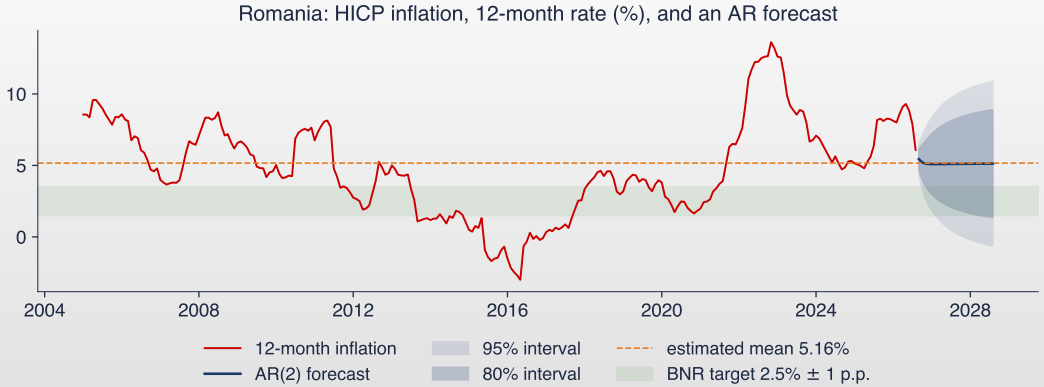
Interpretarea prognozelor PIB

- MA(3): $-0{,}14$; 1,81; 2,77, apoi media de 3,17% începînd cu al patrulea trimestru
  - exact regula MA(q): după $q = 3$ pași modelul cunoaște doar media
- AR(1): o revenire netedă la medie, 2,99% după 8 trimestre (media 3,12%)
  - cele două modele coincid pe termen scurt și diferă prin viteza cu care se estompează efectul încetinirii recente
- Intervalele: $[-6{,}1; 5{,}9]$ pentru un trimestru, $[-6{,}3; 12{,}7]$ începînd cu patru trimestre: $\pm 6{,}0$ și $\pm 9{,}5$ puncte procentuale
- Un model univariat nu poate anticipa un punct de întoarcere; ne spune cît de neobișnuită ar fi următoarea valoare

Recapitulare: creșterea PIB-ului României

- Creșterea anuală a PIB-ului trimestrial: ACF se anulează după lagul 3, MA(3) după BIC, ARMA(1,3) după AIC
- MA(3) reflectă suprapunerea a patru trimestre: structura datelor explică modelul
- Reziduuri de tip zgomot alb, dar cozi groase (2020): intervalele sînt prea înguste în crize
- Prognozele ajung la medie după trei trimestre

Inflația în România: o prognoză AR(2)



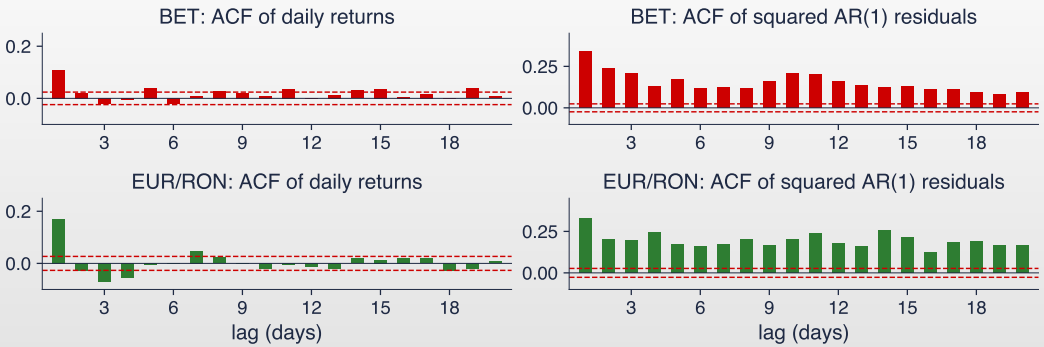
- IAPC (indicele armonizat al prețurilor de consum, Eurostat), rata pe 12 luni în %, 260 de luni pînă în august 2026; ordinul AR ales prin BIC dintre $p \leq 6$

TSA_ch2_real_series

Interpretarea modelului pentru inflație

- BIC: 495,7 pentru AR(1), 469,7 pentru AR(2), 473,4 pentru AR(3): AR(2)
  - $\hat{\phi}_1 = 1{,}320$ (0,057), $\hat{\phi}_2 = -0{,}344$ (0,055); $\hat{\phi}_1 + \hat{\phi}_2 = 0{,}976$, cea mai mare rădăcină inversă 0,963
- Foarte aproape de o rădăcină unitară: este inflația staționară? Testul este în Capitolul 3
  - media estimată, 5,16% (eroarea standard 1,35), este media perioadei 2005–2026, peste ținta BNR de 2,5% $\pm$ 1 p.p.
- Prognoza: de la 6,09% la 5,44% luna viitoare și 5,13% în august 2028; intervalul de 95%: $[-0{,}6; 10{,}9]$
  - modelul revine la media istorică, nu la țintă: nu conține informații despre politica monetară
- Diagnosticarea eșuează: Ljung–Box $Q^*(12) = 71{,}6$, p $< 0{,}001$: corelație la lagul 12 (efectele de bază ale ratei pe 12 luni), o temă pentru Capitolul 4

Randamente zilnice: BET și EUR/RON



- Stînga: ACF a randamentelor logaritmice zilnice; dreapta: ACF a pătratelor reziduurilor unui AR(1)

TSA_ch2_real_series

Interpretarea modelului AR(1) pentru randamentele zilnice

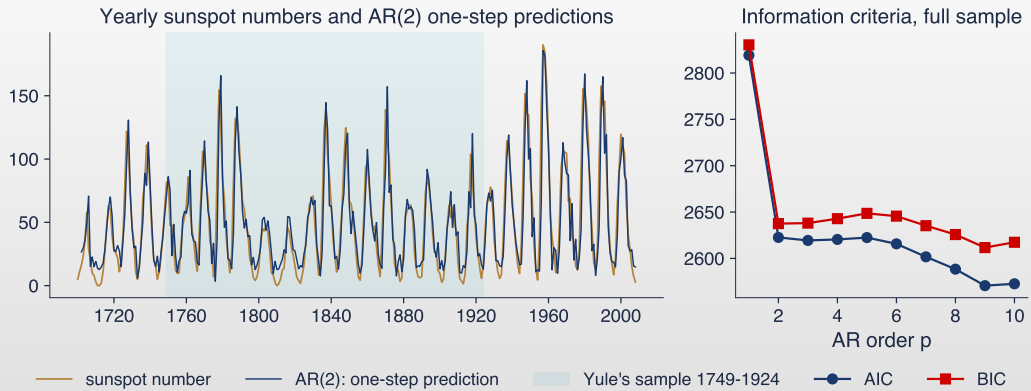
	BET	EUR/RON
Observații	6 670	5 352
$\hat{\phi}$ (se OPG)	0,109 (0,005)	0,170 (0,006)
Raportul t	20,9	30,6
R^2 (%)	1,2	2,9
LB $Q^*(10)$ reziduuri, 9 g.l. (p)	29,7 ($<0,001$)	61,5 ($<0,001$)
LB $Q^*(10)$ pătratele reziduurilor	2496	2351
Kurtosis-ul reziduurilor	14,0	17,0
Alegerea BIC, $p, q \leq 2$	ARMA(1,0)	ARMA(2,1)

- se: OPG; cu cozi groase (kurtosis 14,0 și 17,0) acestea sînt prea mici, iar erorile standard din curbură sau robuste sînt de cîteva ori mai mari, totuși $\hat{\phi}$ rămîne semnificativ
- $\hat{\phi}$ foarte semnificativ, dar un R^2 de 1–3%: real din punct de vedere **statistic**, mic din punct de vedere **economic** după costurile de tranzacționare
- $\hat{\phi}$ pozitiv: ajustarea lentă a prețurilor (BET), un curs de schimb în regim de managed float (EUR/RON)
- Pătratele reziduurilor și kurtosis-ul: media este modelată, varianța nu; ARMA + GARCH în Capitolul 5

Studiu de caz: Yule (1927) și petele solare

- Yule (1927) a estimat prima autoregresie: numerele anuale ale petelor solare ale lui Wolfer, 1749–1924
  - pînă atunci, ciclurile erau modelate ca sume de sinusoide (periodograma)
  - Yule: un „pendul bombardat cu mazăre": un oscilator amortizat lovit de șocuri aleatoare
- AR(2)-ul nostru, estimat prin cele mai mici pătrate pe aceiași ani (176 de observații, datele din statsmodels):
  - $\hat{\phi}_1 = 1{,}336$, $\hat{\phi}_2 = -0{,}650$: rădăcini complexe, modulul rădăcinilor inverse 0,806
  - perioada implicită $2\pi / \arccos(\hat{\phi}_1/(2\sqrt{-\hat{\phi}_2})) = 10{,}57$ ani: ciclul solar, din doar doi coeficienți
- Walker (1931) a generalizat ecuațiile la AR(p): ecuațiile Yule–Walker
  - începutul analizei parametrice a seriilor de timp

Petele solare: AR(2) și criteriile informaționale



⊡ Numerele anuale ale petelor solare, 1700–2008 (309 ani); AR(p) prin MLE pe întregul eșantion

TSA_ch2_real_series

Interpretarea modelelor pentru petele solare

- AR(2) pe întregul eșantion: $\hat{\phi} = (1{,}391; -0{,}689)$, apropiat de cel al lui Yule; explică 83% din varianță cu un an înainte
  - predicțiile cu un pas urmează ciclul, dar întîrzie la vîrfuri
- AIC și BIC aleg amîndouă $p = 9$ dintre $p \leq 10$
  - ciclul este asimetric (creștere rapidă, scădere lentă), ceea ce un AR(2) liniar nu poate reproduce
- Lecția: doi parametri surprind esențialul; mai multe laguri surprind detaliile; neliniaritatea depășește cadrul ARMA

Studiu de caz: competițiile de prognoză

- Makridakis și Hibon (2000): competiția M3, 3003 serii (anuale, trimestriale, lunare, altele), 24 de metode
  - prognoze comparate în afara eșantionului, după mai multe măsuri de acuratețe și orizonturi
- Concluzii confirmate față de competițiile M anterioare
  - metodele sofisticate statistic nu prognozează neapărat mai bine decît cele simple
  - clasamentul depinde de măsura de acuratețe și de orizont; combinațiile de metode au rezultate bune
- Consecințe pentru ARMA
  - bucla Box–Jenkins trebuie să se încheie cu o comparație în afara eșantionului cu repere simple (media, ultima valoare, netezirea exponențială din Capitolul 0)
  - Seminarul 2, C1 face această comparație pentru inflația din România

Recapitulare: alte serii reale

- Inflația: AR(2) aproape de o rădăcină unitară; revine la media istorică; corelație reziduală la lagul 12
- Randamentele zilnice: efecte AR semnificative, dar foarte mici; varianța cere GARCH
- Petele solare: AR(2)-ul lui Yule dă ciclul de 11 ani din rădăcini complexe
- Competițiile de prognoză: comparați întotdeauna cu repere simple

Contribuția posibilă a AI

- **Cod**: o primă versiune a unei bucle care estimează ARMA(p, q) pe o grilă, tabelează AIC și BIC și aplică testele pe reziduuri
- **Explicații**: o a doua deducere a ACF pentru un ARMA(1,1) sau a motivului pentru care PACF a unui MA descrește
- **Explorare**: aceeași buclă Box–Jenkins pe multe serii (creșterea PIB a tuturor țărilor UE), pentru a compara persistența
- Exemplu de prompt
  - `Write Python code with statsmodels that fits ARMA(p,q), p,q <= 3, to Romanian annual GDP growth from Eurostat, reports AIC, BIC and the Ljung-Box p-value with model_df = p + q, and forecasts 8 quarters from the BIC model.`

Verificări necesare

- Convenția de semn pentru θ și ϕ în program ($\theta(L) = 1 + \theta L$ în statsmodels, $1 - \theta L$ în unele manuale)
- Termenul liber sau media: `const` din `ARIMA` este media μ, nu c
- Gradele de libertate ale testului Ljung–Box pentru reziduuri ($m - p - q$) și faptul că toate modelele sînt comparate pe același eșantion
- Rădăcinile ambelor polinoame, nu mărimea coeficienților luați separat
- Că seria este staționară înainte de estimarea unui ARMA (Capitolul 3) și că prognozele sînt comparate cu repere simple
- Fiecare referință citată: trebuie să existe; verificați DOI-ul

Idei de reținut

- ARMA(p, q) înlocuiește suma Wold infinită cu raportul a două polinoame scurte, $\theta(L)/\phi(L)$
- Staționaritatea: rădăcinile lui $\phi(z)$ în afara cercului unitate; invertibilitatea: rădăcinile lui $\theta(z)$ în afara lui; fără factori comuni
- ACF se anulează: MA; PACF se anulează: AR; ambele descresc: ARMA
- Estimați prin MLE, alegeți cu AIC/BIC, verificați reziduurile cu Ljung–Box pe $m - p - q$ grade de libertate
- Prognozele revin la medie; intervalele se lărgesc pînă la $\pm 1{,}96\sqrt{\gamma(0)}$ și ignoră incertitudinea parametrilor
- Datele reale: creșterea PIB este un MA(3) prin construcție; inflația este aproape de o rădăcină unitară; randamentele cer GARCH

Formule de reținut

Mărimea	Formula
ARMA(p,q)	$\phi(L)(X_t-\mu)=\theta(L)\varepsilon_t, \quad \psi(L)=\theta(L)/\phi(L)$
AR(1)	$\mu=c/(1-\phi), \quad \gamma(0)=\sigma^{\mathbf{2}}/(1-\phi^{\mathbf{2}}), \quad \rho(h)=\phi^h$
AR(2)	$\rho(1)=\phi_{\mathbf{1}}/(1-\phi_{\mathbf{2}}), \quad \rho(h)=\phi_{\mathbf{1}}\rho(h-1)+\phi_{\mathbf{2}}\rho(h-2)$
MA(1)	$\rho(1)=\theta/(1+\theta^{\mathbf{2}}), \quad \rho(h)=0,\ h\geq 2$; **invertibil** $\lvert\theta\rvert<1$
ARMA(1,1)	$\psi_j=\phi^{j-\mathbf{1}}(\phi+\theta), \quad \rho(h)=\phi\rho(h-1),\ h\geq 2$
Yule–Walker	$\hat{\phi}=\hat{R}^{-\mathbf{1}}\hat{\rho}, \quad \hat{\sigma}^{\mathbf{2}}=\hat{\gamma}(0)(1-\hat{\phi}^{\top}\hat{\rho})$
Criterii	$\mathrm{AIC}=-2\ln L+2k, \quad \mathrm{BIC}=-2\ln L+k\ln T$
Reziduuri	$Q^*(m)\sim\chi^{\mathbf{2}}(m-p-q)$
Prognoză	$\hat{X}_{T+h\mid T}\pm 1{,}96\,\sigma\sqrt{\sum_{j=\mathbf{0}}^{h-\mathbf{1}}\psi_j^{\mathbf{2}}}$

Autoevaluare

- **Întrebare**: este $X_t = 1{,}2X_{t-1} - 0{,}32X_{t-2} + \varepsilon_t$ staționar?
  - **Răspuns**: $1 - 1{,}2z + 0{,}32z^2 = (1 - 0{,}8z)(1 - 0{,}4z)$, rădăcinile 1,25 și 2,5, ambele în afara cercului: da
- **Întrebare**: ACF a unei serii are o singură valoare semnificativă, $\hat{\rho}(1) = -0{,}45$, iar PACF descrește. Ce model propuneți?
  - **Răspuns**: MA(1) cu $\theta < 0$; rezolvăm $\theta/(1 + \theta^2) = -0{,}45$ și păstrăm rădăcina invertibilă
- **Întrebare**: un ARMA(2,1) lasă în reziduuri $Q^*(10) = 15{,}2$. Respingem la 5%?
  - **Răspuns**: gradele de libertate $10 - 3 = 7$, $\chi^2_{0,95}(7) = 14{,}07 < 15{,}2$: respingem; cu 10 grade de libertate (valoarea critică 18,31) am accepta greșit
- Urmează: Capitolul 3, rădăcini unitare și modele ARIMA: ce facem cînd o rădăcină se află pe cercul unitate

Bibliografie (1/2)

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. doi:10.1109/TAC.1974.1100705
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forecasting and Control* (4th ed.). Wiley. doi:10.1002/9781118619193
- Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332), 1509–1526. doi:10.1080/01621459.1970.10481180
- Brockwell, P. J., & Davis, R. A. (1991). *Time Series: Theory and Methods* (2nd ed.). Springer. doi:10.1007/978-1-4419-0320-4
- Brockwell, P. J., & Davis, R. A. (2016). *Introduction to Time Series and Forecasting* (3rd ed.). Springer. doi:10.1007/978-3-319-29854-2
- Durbin, J. (1960). The fitting of time-series models. *Revue de l'Institut International de Statistique*, 28(3), 233–244. doi:10.2307/1401322
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press. doi:10.2307/j.ctv14jx6sm
- Huang, C., & Petukhina, A. (2022). *Applied Time Series Analysis and Forecasting with Python*. Springer. doi:10.1007/978-3-031-13584-2
- Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2), 297–307. doi:10.1093/biomet/76.2.297
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts. otexts.com/fpp3

Bibliografie (2/2)

- Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3), 1–22. doi:10.18637/jss.v027.i03
- Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review*, 55(2), 163–172. doi:10.2307/1403192
- Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. doi:10.1093/biomet/65.2.297
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: Results, conclusions and implications. *International Journal of Forecasting*, 16(4), 451–476. doi:10.1016/S0169-2070(00)00057-1
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461–464. doi:10.1214/aos/1176344136
- Slutzky, E. (1937). The summation of random causes as the source of cyclic processes. *Econometrica*, 5(2), 105–146. doi:10.2307/1907241
- Walker, G. (1931). On periodicity in series of related terms. *Proceedings of the Royal Society of London, Series A*, 131(818), 518–532. doi:10.1098/rspa.1931.0069
- Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*. Almqvist & Wiksell. archive.org
- Yule, G. U. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society A*, 226, 267–298. doi:10.1098/rsta.1927.0007