








Serii de timp

Capitolul 12: Analiză spectrală



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)
- **AR**: AutoRegressive (model) — model autoregresiv
- **ARFIMA**: AutoRegressive Fractionally Integrated Moving Average (model) — model autoregresiv fracționar integrat cu medie mobilă
- **ARMA**: AutoRegressive Moving Average (model) — model autoregresiv cu medie mobilă
- **CWT**: Continuous Wavelet Transform (transformata wavelet continuă)
- **DFT**: Discrete Fourier Transform (transformata Fourier discretă)
- **ENTSO-E**: European Network of Transmission System Operators for Electricity (rețeaua europeană a operatorilor de transport și de sistem pentru energie electrică)
- **EODHD**: EOD Historical Data (market-data provider) — furnizor de date de piață
- **FFT**: Fast Fourier Transform (transformata Fourier rapidă)

Acronime folosite în acest material (2/2)

- **FRED**: Federal Reserve Economic Data (baza de date economice a Rezervei Federale din St. Louis)
- **GPH**: Geweke–Porter-Hudak (estimator) — estimatorul Geweke–Porter-Hudak al parametrului de memorie lungă
- **HP**: Hodrick–Prescott (filter) — filtrul Hodrick–Prescott
- **LPPL**: Log-Periodic Power Law (legea de putere log-periodică)
- **MA**: Moving Average (medie mobilă)
- **OLS**: Ordinary Least Squares (metoda celor mai mici pătrate)
- **PIB**: Produsul intern brut
- **SILSO**: Sunspot Index and Long-term Solar Observations (centrul mondial de date pentru numărul de pete solare, la Observatorul Regal al Belgiei)
- **STFT**: Short-Time Fourier Transform (transformata Fourier pe ferestre scurte)
- **SUA**: Statele Unite ale Americii
- **UE**: Uniunea Europeană
- **UTC**: Coordinated Universal Time (Temps Universel Coordonné) — timpul universal coordonat

Întrebarea capitolului și traseul

- **Întrebarea**: ce cicluri conține o serie și cît din varianța ei explică fiecare ciclu?
  - exemple: ciclul de 11 ani al petelor solare, ciclul economic, ritmul zilnic și săptămînal al consumului de energie electrică
- **Traseul** capitolului
  - cicluri, frecvențe și transformata Fourier discretă
  - densitatea spectrală a zgomotului alb și a modelelor ARMA
  - periodograma, slăbiciunile ei și netezirea ei
  - cicluri în date; memoria lungă; filtre; două serii; o trimitere spre wavelets
- Capitol de studiu individual: exemple rezolvate, slide-uri de interpretare și o recapitulare după fiecare secțiune; autoevaluare cu răspunsuri la final

Rezultatele învățării

- Treceți de la frecvență la perioadă și la frecvența unghiulară și găsiți frecvențele Fourier ale unui eșantion
- Calculați densitatea spectrală a zgomotului alb și a modelelor AR(1), MA(1) și AR(2) și interpretați forma ei
- Calculați de mînă periodograma unei serii scurte și explicați de ce periodograma brută nu este consistentă
- Netezați periodograma (Daniell, Bartlett, Welch), aplicați o fereastră de atenuare și construiți benzi de încredere chi-pătrat
- Identificați cicluri în petele solare, în PIB și în consumul de energie electrică; interpretați coerența și faza dintre două serii

Bibliografie și instrumente

- Referința principală a capitolului: Shumway și Stoffer (2017), cap. 4 (analiză spectrală și filtrare)
  - teorie: Brockwell și Davis (1991), cap. 4 și 10; Hamilton (1994), cap. 6; aplicații în Python: Huang și Petukhina (2022)
  - Shumway și Stoffer (2017) măsoară frecvența în cicluri pe observație; în Capitolul 8 am folosit frecvența unghiulară $\lambda = 2\pi\nu$
- Quantlet-urile Python ale capitolului: Quantlets/Ch_12
  - periodograma, netezirea Daniell și wavelet-ul Morlet sînt scrise explicit în `numpy`; Welch și coerența provin din `scipy.signal`
- Notebook-ul cursului: deschideți în Google Colab
- Curs video: Applied Time Series Analysis with Python

Joseph Fourier și ideea capitolului



J. B. J. Fourier (1768–1830)

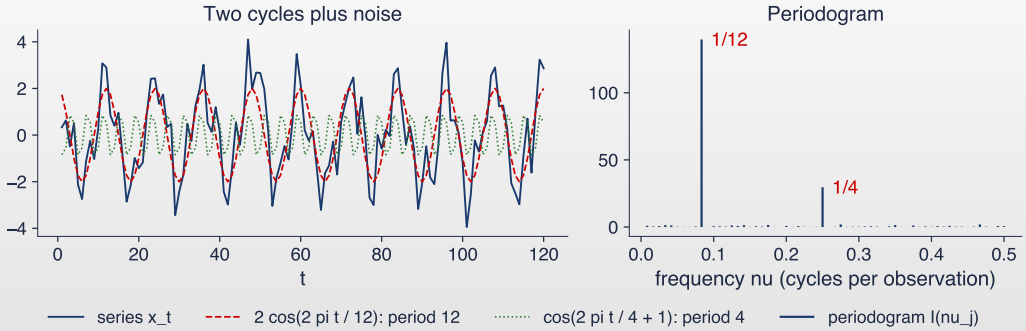
Gravură: A. F. B. Geille, după J.-L. Boilly (1839–1840); domeniu public; Wikimedia Commons

- Fourier (1807, conducția căldurii): o funcție se poate scrie ca sumă de sinusuri și cosinusuri
  - fiecare termen este un ciclu pur, cu frecvența și amplitudinea lui
- Aplicat unei serii de timp: cîtă varianță se află la fiecare frecvență?
  - **domeniul timpului** (Capitolele 1–8) descrie seria prin autocorelațiile ei
  - **domeniul frecvenței** descrie aceeași informație prin spectru
- Cele două perspective sînt echivalente; spectrul face ciclurile vizibile dintr-o privire

Un ciclu: amplitudine, frecvență, fază

- Un ciclu pur: $x_t = A\cos(2\pi\nu t + \varphi)$, $t = 1, 2, \ldots$
  - A: **amplitudinea** (înălțimea undei); φ: **faza** (punctul de pornire al undei)
  - ν: **frecvența**, în cicluri pe observație; **perioada** $1/\nu$ este numărul de observații dintr-un ciclu
  - frecvența unghiulară $\omega = 2\pi\nu$, în radiani pe observație
- Exemplu: date lunare cu un ciclu anual: $\nu = 1/12$ cicluri pe lună, perioada 12 luni
  - date orare cu un ciclu zilnic: $\nu = 1/24$; date trimestriale cu un ciclu de 6 ani: tot $\nu = 1/24$
- Forma echivalentă: $A\cos(2\pi\nu t + \varphi) = a\cos(2\pi\nu t) + b\sin(2\pi\nu t)$
  - cu $a = A\cos\varphi$, $b = -A\sin\varphi$ și $A^2 = a^2 + b^2$: liniară în a și b, deci se poate estima prin OLS
- Cea mai mare frecvență vizibilă în date discrete este $\nu = 1/2$ (un ciclu la două observații): **frecvența Nyquist**

Două cicluri plus zgomot



- Serie simulată, $T = 120$: $x_t = 2\cos(2\pi t/12) + \cos(2\pi t/4 + 1) + \varepsilon_t$, $\varepsilon_t \sim N(0;\ 0{,}49)$ (zgomot alb Normal cu varianța 0,49); dreapta: periodograma ei (definită în secțiunea 3)

TSA_ch12_fourier

Interpretarea celor două cicluri plus zgomot

- În domeniul timpului cele două cicluri sînt greu de separat cu ochiul liber
  - în domeniul frecvenței ele sînt două vîrfuri izolate, la $\nu = 1/12$ și $\nu = 1/4$
- Înălțimea vîrfurilor: 139,7 și 29,6; teoria, pentru un ciclu de amplitudine A la o frecvență Fourier: $A^2 T/4$, adică 120 și 30
  - vîrful crește cu pătratul amplitudinii: ciclul mai mare domină
- Cele două vîrfuri concentrează 85% din puterea totală; zgomotul împrăștie restul uniform pe toate frecvențele

Frecvențele Fourier și transformata Fourier discretă (1/2)

- Pentru un eșantion $x_1, \ldots, x_T$, **frecvențele Fourier** sînt $\nu_j = j/T$, $j = 0, 1, \ldots, \lfloor T/2 \rfloor$
  - ciclurile care încap de un număr întreg de ori (j ori) în eșantion; $\lfloor T/2 \rfloor$: partea întreagă a lui $T/2$
- **Transformata Fourier discretă** (DFT) măsoară cît de puternic se mișcă seria împreună cu un ciclu de frecvență ν_j
  - $d(\nu_j) = T^{-1/2} \sum_{t=1}^{T} x_t \, e^{-2\pi i \nu_j t}$
  - i: unitatea imaginară, $i^2 = -1$; $e^{-i\theta} = \cos\theta - i \sin\theta$ (formula lui Euler)
  - deci DFT corelează seria cu un cosinus (partea reală) și cu un sinus (partea imaginară) de frecvență ν_j

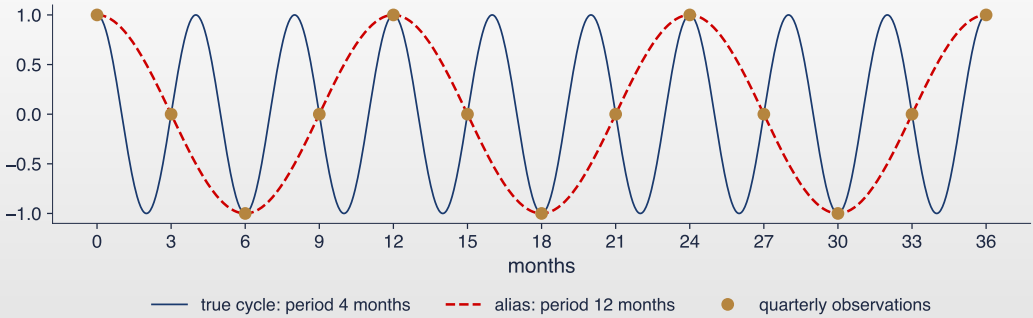
Frecvențele Fourier și transformata Fourier discretă (2/2)

- Cele T valori $d(\nu_j)$ conțin exact aceeași informație ca cele T observații
  - datele se pot reconstitui din ele prin transformata inversă
- **Parseval**
  - suma pătratelor abaterilor este egală cu suma pătratelor modulelor DFT
  - $\sum_{t=1}^{T}(x_t - \bar{x})^2 = \sum_{j=1}^{T-1} |d(\nu_j)|^2$
  - $\bar{x}$: media de selecție; $|d|^2 = (\mathrm{Re}\, d)^2 + (\mathrm{Im}\, d)^2$: pătratul modulului unui număr complex
  - interpretarea: varianța seriei se împarte pe frecvențe
- **Transformata Fourier rapidă** (FFT) a lui Cooley și Tukey (1965) calculează rapid toate valorile $d(\nu_j)$
  - $O(T \log T)$ operații în loc de $O(T^2)$: numărul de operații crește ca $T \log T$, nu ca T^2

Exemplu rezolvat: DFT pentru patru numere

- Datele: $x = (4, 2, 0, 2)$, $T = 4$; media $\bar{x} = 2$, abaterile $y = (2, 0, -2, 0)$
  - Frecvențele Fourier: $\nu_1 = 1/4$ (perioada 4) și $\nu_2 = 1/2$ (perioada 2)
- La $\nu = 1/4$: $e^{-2\pi it/4} = e^{-i\pi t/2}$ este $-i$ pentru $t = 1$ și i pentru $t = 3$
  - $\sum_t y_t e^{-i\pi t/2} = 2(-i) + (-2)(i) = -4i$, deci $d(1/4) = -4i/\sqrt{4} = -2i$ și $|d(1/4)|^2 = 4$
- La $\nu = 1/2$: $e^{-i\pi t} = (-1)^t$, deci $\sum_t y_t(-1)^t = -2 + 2 = 0$ și $|d(1/2)|^2 = 0$
  - prin simetrie $|d(3/4)|^2 = |d(1/4)|^2 = 4$
- Parseval: $4 + 0 + 4 = 8 = 2^2 + 0^2 + (-2)^2 + 0^2$; toată varianța ($8/4 = 2$) se află la perioada 4
- Verificare în Python: `abs(np.fft.fft([2, 0, -2, 0]))**2 / 4` dă $(0, 4, 0, 4)$

Aliasing



- Un ciclu de 4 luni ($\nu = 1/4$ pe lună), observat doar o dată pe trimestru (la fiecare 3 luni)

TSA_ch12_fourier

Interpretarea fenomenului de aliasing

- Pe trimestru, frecvența adevărată este $3 \times 1/4 = 3/4$ cicluri, peste frecvența Nyquist $1/2$
  - ea se pliază înapoi la $|3/4 - 1| = 1/4$ cicluri pe trimestru: o perioadă de 4 trimestre, adică 12 luni
- Punctele trimestriale cad exact pe o undă de 12 luni: datele nu pot deosebi cele două cicluri
  - acesta este **aliasing**: ciclurile mai rapide decît două observații apar ca cicluri mai lente, false
- Regula practică: eșantionați de cel puțin două ori pe ciclul de interes; la agregarea pe o frecvență mai mică, faceți media valorilor, nu alegeți una dintre ele

Recapitulare: cicluri și frecvențe

- Un ciclu are o amplitudine, o frecvență ν (cicluri pe observație), o perioadă $1/\nu$ și o fază
- DFT la frecvențele Fourier j/T rescrie datele ca sumă de cicluri; Parseval împarte varianța pe frecvențe
- Sînt vizibile doar frecvențele pînă la $1/2$; ciclurile mai rapide apar ca aliasuri

De la autocovariațe la spectru

- Pentru o serie staționară cu autocovariațele $\gamma(h)$, $\sum_h |\gamma(h)| < \infty$, **densitatea spectrală** este
  - $f(\nu) = \sum_{h=-\infty}^{\infty} \gamma(h) e^{-2\pi i \nu h} = \gamma(0) + 2\sum_{h=1}^{\infty} \gamma(h)\cos(2\pi\nu h)$, $-1/2 \le \nu \le 1/2$
  - $\gamma(h) = \mathrm{Cov}(x_t, x_{t+h})$: autocovariația la lagul h; $\sum_h |\gamma(h)| < \infty$: autocovariațele se sting suficient de repede (memorie scurtă)
  - în cuvinte: $f(\nu)$ este o sumă ponderată a autocovariațelor, cu ponderile $\cos(2\pi\nu h)$, care oscilează cu frecvența ν
- Relația inversă: $\gamma(h) = \int_{-1/2}^{1/2} f(\nu) e^{2\pi i \nu h}\, d\nu$
  - perechea $\gamma \leftrightarrow f$ este teorema **Wiener–Khinchin**: autocovariațele și spectrul conțin aceeași informație
  - pentru $h = 0$: $\mathrm{Var}(x_t) = \gamma(0) = \int_{-1/2}^{1/2} f(\nu)\, d\nu$: aria de sub f este varianța
- Proprietăți: $f(\nu) \ge 0$, $f(-\nu) = f(\nu)$; este suficient să desenăm f pe $[0, 1/2]$
  - $f(\nu)\, d\nu$ este partea din varianță datorată ciclurilor cu frecvențe în $[\nu, \nu + d\nu]$

Zgomot alb, AR(1) și MA(1)

- **Zgomot alb** w_t, cu varianța σ^2
  - $\gamma(h) = 0$ pentru $h \neq 0$, deci $f(\nu) = \sigma^2$ pentru orice ν
  - un spectru plat: toate frecvențele contribuie în mod egal, ca lumina albă
- **AR(1)** $x_t = \phi x_{t-1} + w_t$, $|\phi| < 1$
  - $f(\nu) = \dfrac{\sigma^2}{1 - 2\phi\cos(2\pi\nu) + \phi^2}$
  - ϕ: coeficientul autoregresiv; w_t: zgomot alb cu varianța σ^2
  - $\phi > 0$: putere la frecvențele joase (serie netedă, persistentă); $\phi < 0$: putere la frecvențele înalte (serie în zigzag)
- **MA(1)** $x_t = w_t + \theta w_{t-1}$
  - $f(\nu) = \sigma^2\,[1 + 2\theta\cos(2\pi\nu) + \theta^2]$
  - θ: coeficientul de medie mobilă; $\theta > 0$ pune puterea la frecvențele joase, $\theta < 0$ la cele înalte
  - direct din definiție: $\gamma(0) = \sigma^2(1 + \theta^2)$, $\gamma(1) = \sigma^2\theta$, deci $f = \gamma(0) + 2\gamma(1)\cos(2\pi\nu)$

Spectrul unui model ARMA

- **ARMA**(p, q)
  - $\phi(L)x_t = \theta(L)w_t$
  - L: operatorul lag, $Lx_t = x_{t-1}$
  - $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ și $\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$: polinoamele AR și MA
- Densitatea spectrală: cele două polinoame evaluate în $z = e^{-2\pi i\nu}$
  - $f(\nu) = \sigma^2 \dfrac{|\theta(e^{-2\pi i\nu})|^2}{|\phi(e^{-2\pi i\nu})|^2}$
  - partea MA (numărătorul) adaugă putere acolo unde $|\theta|$ este mare; partea AR (numitorul) adaugă putere acolo unde $|\phi|$ este mic
- Un raport de două polinoame în $\cos(2\pi\nu)$: neted și ușor de calculat (Brockwell și Davis (1991), secț. 4.4)

Exemplu rezolvat: spectrele AR(1) și MA(1)

- AR(1) cu $\phi = 0{,}6$, $\sigma^2 = 1$
  - $\nu = 0$: $\cos 0 = 1$, $f(0) = 1/(1-0{,}6)^2 = 1/0{,}16 = 6{,}25$
  - $\nu = 1/2$: $\cos \pi = -1$, $f(1/2) = 1/(1+0{,}6)^2 = 1/2{,}56 = 0{,}39$
  - raportul 16: ciclurile lente poartă mult mai multă varianță decît cele rapide
- MA(1) cu $\theta = 0{,}5$, $\sigma^2 = 1$
  - $f(0) = (1+0{,}5)^2 = 2{,}25$ și $f(1/2) = (1-0{,}5)^2 = 0{,}25$
- Verificarea ariei pentru AR(1): $\int_{-1/2}^{1/2} f = \gamma(0) = 1/(1-\phi^2) = 1/0{,}64 = 1{,}5625$
  - integrala numerică a formulei dă aceeași valoare

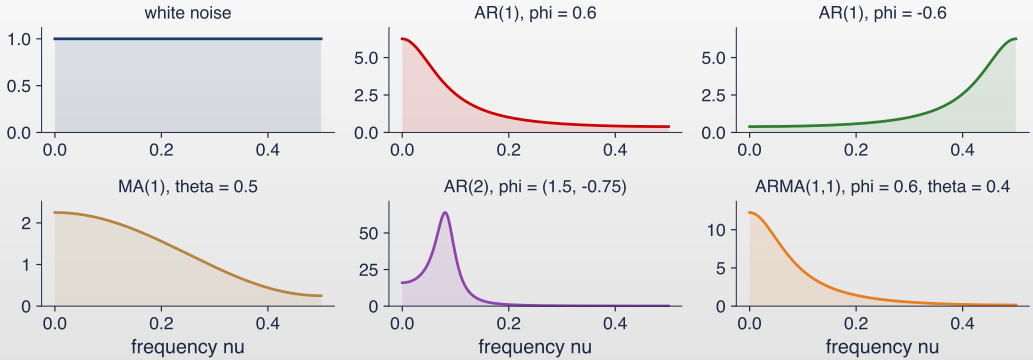
AR(2) și pseudo-ciclurile

- **AR(2)** $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t$
  - $f(\nu) = \dfrac{\sigma^2}{|1 - \phi_1 e^{-2\pi i \nu} - \phi_2 e^{-4\pi i \nu}|^2}$
  - pentru rădăcini complexe ($\phi_1^2 + 4\phi_2 < 0$) ACF este o undă amortizată, iar spectrul are un vîrf în interiorul intervalului $(0, 1/2)$
  - vîrful se află la frecvența ν^* dată de $\cos(2\pi\nu^*) = \phi_1(\phi_2 - 1)/(4\phi_2)$, cînd membrul drept este în $[-1, 1]$
- Un **pseudo-ciclu**
  - un ciclu neregulat, a cărui lungime și amplitudine variază de la o rundă la alta
  - Yule (1927) a estimat un AR(2) pentru numerele de pete solare ale lui Wolfer: primul model autoregresiv, o alternativă la undele sinusoidale fixe
- Ciclurile economice sînt pseudo-cicluri de acest fel: recurente, dar niciodată de aceeași lungime (secțiunea 5)

Exemplu rezolvat: vîrful spectrului unui AR(2)

- $\phi_1 = 1{,}5$, $\phi_2 = -0{,}75$, $\sigma^2 = 1$; rădăcini complexe, deoarece $1{,}5^2 + 4(-0{,}75) = -0{,}75 < 0$
  - $\cos(2\pi\nu^*) = 1{,}5 \cdot (-1{,}75)/(-3) = 0{,}875$, deci $\nu^* = \arccos(0{,}875)/(2\pi) = 0{,}0804$
  - perioada $1/\nu^* = 12{,}4$ observații; pentru date anuale, un ciclu de aproximativ 12 ani
- Înălțimea: $f(\nu^*) = 64$, față de $f(0) = 1/(1 - 1{,}5 + 0{,}75)^2 = 16$
  - un vîrf clar: seria oscilează cu o perioadă tipică de aproximativ 12 observații
- Același AR(2) este folosit în secțiunea 4 pentru a testa estimatorii spectrului

O galerie de densități spectrale



- Spectre teoretice, cu $\sigma^2 = 1$, pe $0 \leq \nu \leq 1/2$; aria hașurată este jumătate din varianță

TSA_ch12_spectral_densities

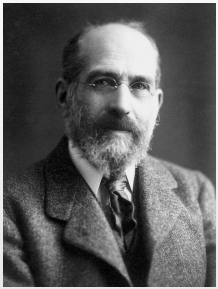
Interpretarea galeriei

- Locul în care se află puterea arată tipul de dependență
  - plat: nicio dependență (zgomot alb); descrescător: persistență pozitivă (AR(1) cu $\phi > 0$, MA(1) cu $\theta > 0$, ARMA(1,1))
  - crescător: alternanță de semne (AR(1) cu $\phi < 0$); o cocoașă în interiorul benzii: un pseudo-ciclu (AR(2))
- AR(2) își concentrează varianța 8,62 într-o bandă îngustă în jurul lui $\nu = 0{,}0804$
  - ARMA(1,1) adaugă termenul MA la AR(1): $f(0) = 12{,}25$, o scădere mai abruptă
- Ordinea de lucru: estimați spectrul din date, priviți forma lui, apoi alegeți un model care are un spectru asemănător

Recapitulare: densitatea spectrală

- $f(\nu) = \sum_h \gamma(h) e^{-2\pi i \nu h}$; aria de sub f este varianța; f și ACF sînt echivalente
- Zgomotul alb are spectrul plat; AR(1) cu $\phi > 0$ pune puterea la frecvențele joase; AR(2) cu rădăcini complexe dă un vîrf
- Spectrele ARMA: $\sigma^2 |\theta(e^{-2\pi i \nu})|^2 / |\phi(e^{-2\pi i \nu})|^2$

Arthur Schuster și periodicitățile ascunse



A. Schuster (1851–1934)

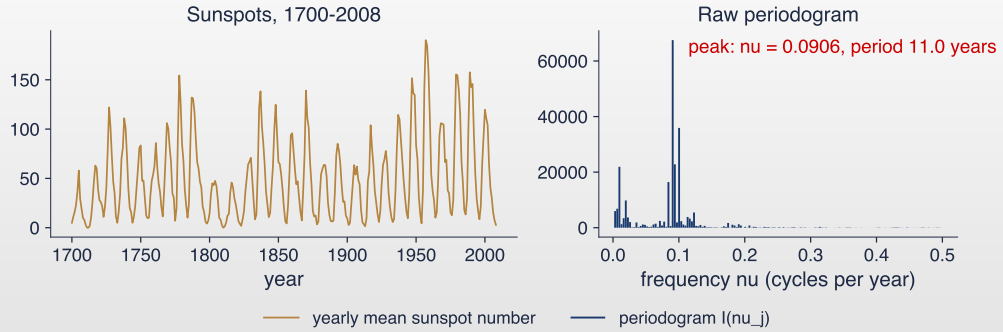
Foto: autor necunoscut (1900s); domeniu public; Wikimedia Commons

- Schuster (1898) a introdus **periodograma** pentru a căuta „periodicități ascunse" în date meteorologice
  - un presupus ciclu de 26 de zile; a arătat cum se judecă dacă un vîrf poate fi întîmplător
- Petele solare: numărate zilnic încă din secolul al XVIII-lea; Rudolf Wolf a construit indicele standard în 1848
  - ciclul lor de 11 ani este cazul clasic de test al analizei spectrale (Capitolul 0)
- Întrebarea secțiunii: cum estimăm $f(\nu)$ dintr-un singur eșantion și cît de sigură este estimarea?

Periodograma

- **Periodograma**
  - $I(\nu_j) = |d(\nu_j)|^2$ la frecvențele Fourier $\nu_j = j/T$ (după eliminarea mediei)
  - echivalent $I(\nu_j) = \sum_{|h|<T} \hat{\gamma}(h) e^{-2\pi i \nu_j h}$: formula spectrului cu autocovarianțele de selecție $\hat{\gamma}(h)$
- Interpretarea prin regresie: regresați x_t pe $\cos(2\pi\nu_j t)$ și $\sin(2\pi\nu_j t)$ prin OLS, cu coeficienții a_j, b_j
  - $I(\nu_j) = \frac{T}{4}(a_j^2 + b_j^2)$: pătratul amplitudinii celui mai bun ciclu la ν_j, înmulțit cu $T/4$
- Parseval (secțiunea 1), scris cu periodograma: $\hat{\gamma}(0) = \frac{1}{T}\sum_{j=1}^{T-1} I(\nu_j)$
  - ponderea unei benzi de frecvențe este suma ordonatelor ei împărțită la total
- În Python: `np.abs(np.fft.fft(x - x.mean()))**2 / T`; se păstrează $j = 1, \ldots, \lfloor T/2 \rfloor$

Petele solare și periodograma lor



- Numărul mediu anual de pete solare, 1700–2008, 309 observații (seria Wolf / SILSO, statsmodels); dreapta: periodograma brută

TSA_ch12_periodogram

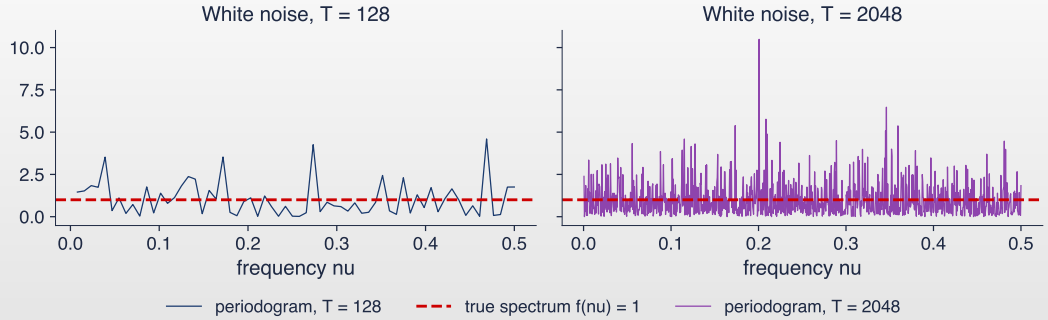
Interpretarea periodogramei petelor solare

- Cea mai mare ordonată este la $j = 28$: $\nu = 28/309 = 0{,}0906$ cicluri pe an, o perioadă de 11,0 ani
  - următoarele ordonate ca mărime sînt la 10,0 și 10,7 ani: ciclul nu este perfect regulat
- Perioadele între 9 și 13 ani explică 60% din varianță
  - ordonatele mici din apropierea lui $\nu = 0$ reflectă schimbările lente ale înălțimii ciclurilor (de exemplu ciclurile slabe din jurul anului 1800)
- Periodograma brută este foarte zimțată: ordonatele vecine sar în sus și în jos; slide-urile următoare explică de ce

Proprietățile statistice ale periodogramei

- Pentru T mare, la o frecvență Fourier $0 < \nu_j < 1/2$ (Brockwell și Davis (1991), secț. 10.3; Shumway și Stoffer (2017), secț. 4.3):
  - $\frac{2I(\nu_j)}{f(\nu_j)} \approx \chi^2_2$: deci $E[I(\nu_j)] \approx f(\nu_j)$ (nedeplasată) și $\text{Var}[I(\nu_j)] \approx f(\nu_j)^2$
  - χ^2_2: distribuția chi-pătrat cu 2 grade de libertate (media 2, varianța 4); $f(\nu_j)$: densitatea spectrală adevărată
  - ordonatele de la frecvențe Fourier diferite sînt aproximativ independente
- Varianța **nu** scade cînd T crește
  - periodograma **nu este un estimator consistent** al lui f
  - un T mai mare dă mai multe ordonate (o grilă mai fină), nu ordonate mai precise
- $\chi^2_2/2$ este o distribuție exponențială cu media 1: o singură ordonată poate depăși ușor de 3 ori valoarea adevărată
  - cuantila ei de 95% este 3,00: aproximativ 1 ordonată din 20 depășește întîmplător $3{,}00\, f(\nu)$

Periodograma zgomotului alb



- Zgomot alb gaussian cu $\sigma^2 = 1$, deci spectrul adevărat este constant, egal cu 1; eșantioane de $T = 128$ și $T = 2048$

TSA_ch12_periodogram

Interpretarea periodogramei zgomotului alb

- $T = 128$: media ordonatelor 1,03, abaterea standard 1,03
  - $T = 2048$: media 1,01, abaterea standard 1,00: de 16 ori mai multe date, aceeași împrăștiere
- Cea mai mare ordonată pentru $T = 2048$ este 10,5: din 1024 de ordonate, apar întîmplător vîrfuri înalte izolate
  - un vîrf înalt nu este încă un ciclu: are nevoie de un test (slide-ul următor) sau de o estimare netezită (secțiunea 4)
- Media este corectă (aproximativ 1): periodograma este nedeplasată, dar zgomotoasă; media ordonatelor vecine va reduce zgomotul

Testul lui Fisher pentru o periodicitate ascunsă (1/2)

- Ipotezele
  - H_0: zgomot alb gaussian; H_1: zgomot alb plus un ciclu la o frecvență Fourier necunoscută
- Statistica lui Fisher (1929): ponderea celei mai mari ordonate în suma tuturor ordonatelor
  - $g = \max_j I(\nu_j) \,/\, \sum_{j=1}^{m} I(\nu_j)$
  - m: numărul de ordonate din $(0, 1/2)$; $\max_j$: cea mai mare dintre ele; $1/m \leq g \leq 1$
- Interpretarea
  - sub H_0 fiecare ordonată reprezintă aproximativ $1/m$ din total
  - un g mult mai mare decît $1/m$ semnalează un ciclu: o singură frecvență preia o parte mare din variață

Testul lui Fisher pentru o periodicitate ascunsă (2/2)

- P-value-ul exact: probabilitatea, sub H_0, ca g să depășească valoarea observată x
  - $P(g > x) = \sum_{k \geq 1,\ kx<1} (-1)^{k-1} \binom{m}{k} (1-kx)^{m-1}$
  - k: indicele de însumare, de la 1 cît timp $kx < 1$; $\binom{m}{k}$: coeficientul binomial
  - un p-value mic (sub 0,05): se respinge zgomotul alb în favoarea unui ciclu
- Petele solare: $m = 154$, $g = 0{,}268$, față de $1/m = 0{,}0065$; p-value $< 0{,}001$
  - vîrful de 11 ani este mult prea mare pentru a fi întîmplător
- Atenție: H_0 este zgomotul alb
  - pe un fond „roșu" (de tip AR(1)), vîrfurile de joasă frecvență apar mai ușor din întîmplare
  - de aceea comparați cu un spectru de fond neted

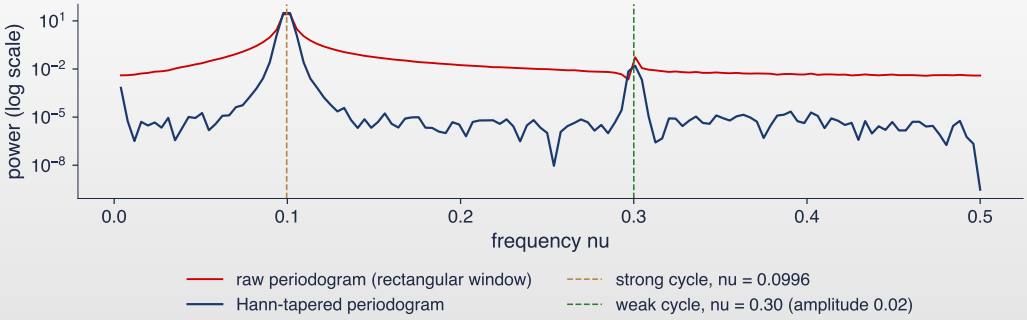
Recapitulare: periodograma

- $I(\nu_j) = |d(\nu_j)|^2$: proporțională cu pătratul amplitudinii ciclului la fiecare frecvență Fourier; suma dă varianța
- $2I/f \approx \chi^2_2$: nedeplasată, dar neconsistentă; ordonatele vecine sînt aproape independente
- Petele solare: un vîrf la 11,0 ani; testul g al lui Fisher respinge zgomotul alb

Scurgerea spectrală (leakage) și ferestrele de atenuare

- Un ciclu a cărui frecvență nu este o frecvență Fourier nu începe de un număr întreg de ori în eșantion
  - tăierea seriei la $t = 1$ și $t = T$ creează un salt; puterea lui se **scurge** la toate celelalte frecvențe
  - scurgerea dintr-un vîrf puternic poate ascunde un ciclu slab în altă parte
- **Atenuarea** (tapering)
  - înmulțiți datele cu o fereastră h_t care scade lin spre 0 la ambele capete, înainte de DFT
  - fereastra Hann (clopot cosinus): $h_t = \frac{1}{2}[1 - \cos(2\pi(t - 0{,}5)/T)]$, rescalată pentru a păstra puterea totală
  - h_t este aproape de 0 la $t = 1$ și $t = T$ și egal cu 1 la mijlocul eșantionului
  - prețul: un vîrf principal puțin mai lat (rezoluție mai mică) pentru o scurgere mult mai mică
- În practică: aplicați întotdeauna o atenuare ușoară; eliminați întîi media (și trendul, dacă există)

Scurgerea: periodograma brută și cea cu atenuare



- $T = 256$: $\cos(2\pi \cdot 0{,}0996\, t)$ (între două frecvențe Fourier) plus $0{,}02 \cos(2\pi \cdot 0{,}3\, t)$ și un zgomot foarte mic; scară logaritmică

TSA_ch12_spectral_estimation

Interpretarea scurgerii spectrale

- Periodograma brută: ciclul puternic își scurge puterea peste tot; la $\nu = 0{,}3$ ciclul slab este doar de 7,7 ori peste fondul scurs
  - cu un zgomot real adăugat, acest vîrf slab ar fi invizibil
- Fereastra Hann: fondul scade cu mai multe ordine de mărime, iar ciclul slab ajunge de 4 230 de ori peste el
  - vîrful principal este puțin mai lat: atenuarea schimbă rezoluție pentru o scurgere mai mică
- Scara logaritmică este esențială: pe o scară liniară ambele curbe arată ca un singur vîrf

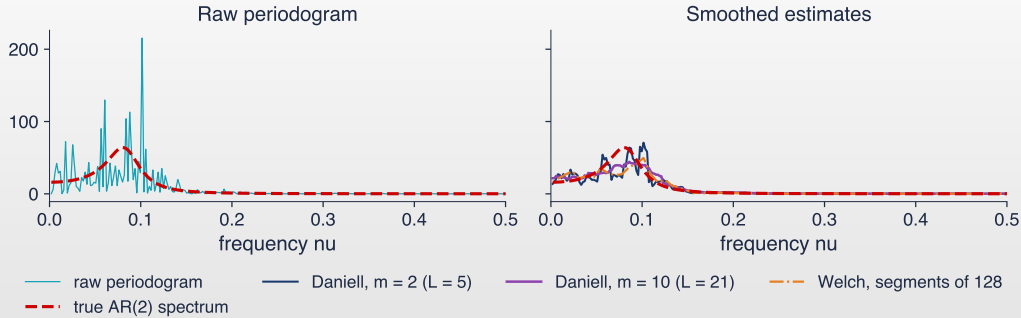
Netezirea periodogramei (1/2): estimatorul Daniell

- Ideea: f este netedă, iar ordonatele sînt aproape independente, deci se face media a $L = 2m + 1$ ordonate vecine
  - estimatorul **Daniell**: $\hat{f}(\nu_j) = \frac{1}{L} \sum_{k=-m}^{m} I(\nu_{j+k})$
  - m: numărul de vecini de fiecare parte; k: parcurge valorile de la $-m$ la m; $\hat{f}$: spectrul estimat
  - varianța este aproximativ f^2/L: de L ori mai mică decît cea a unei singure ordonate
- **Lățimea de bandă** $B = L/T$
  - lățimea benzii de frecvențe peste care se face media
- Compromisul deplasare–varianță: un L mai mare reduce varianța, dar aplatizează vîrfurile înguste (deplasare)
  - consistența cere $L \to \infty$ și $L/T \to 0$, de exemplu $L \approx \sqrt{T}$

Netezirea periodogramei (2/2): ferestrele de lag și metoda Welch

- Perspectiva echivalentă: netezirea în frecvență = ponderi mai mici pentru autocovariațele îndepărtate (ferestre de lag)
  - fereastra Bartlett: $\hat{f}(\nu) = \sum_{|h|\leq M}(1 - |h|/M)\hat{\gamma}(h)e^{-2\pi i\nu h}$ (Bartlett (1950); Blackman și Tukey (1958))
  - M: lagul de trunchiere; autocovariațele de dincolo de lagul M se elimină, celelalte primesc ponderile $1 - |h|/M$, care scad liniar
  - un M mai mic înseamnă o netezire mai puternică
- Metoda **Welch** (Welch (1967))
  - o medie a periodogramelor a K segmente: $\hat{f}_W(\nu) = \frac{1}{K}\sum_{k=1}^{K} I_k(\nu)$, I_k: periodograma cu fereastră a segmentului k
  - seria se împarte în K segmente care se suprapun, se aplică fereastra pe fiecare și se face media periodogramelor
  - `scipy.signal.welch`; lungimea segmentului fixează rezoluția

Estimări netezite ale spectrului unui AR(2)



⊡ AR(2) simulat, $\phi = (1{,}5;\ -0{,}75)$, $T = 512$; spectrul adevărat (linie întreruptă), periodograma brută (stînga) și trei estimări netezite (dreapta)

TSA_ch12_spectral_estimation

Interpretarea estimărilor netezite

- Vîrful adevărat este la $\nu = 0{,}080$; cea mai mare ordonată brută este la $\nu = 0{,}102$: o singură ordonată zgomotoasă plasează greșit vîrful
  - vîrful Daniell $m = 2$: $\nu = 0{,}102$; Daniell $m = 10$: $\nu = 0{,}086$; Welch ($K = 7$ segmente de 128): $\nu = 0{,}102$
- Eroarea pătratică medie a lui $\log \hat{f}$: brută 1,86; Daniell $m = 2$: 0,25; $m = 10$: 0,09; Welch: 0,24
  - prețul netezirii puternice: înălțimea vîrfului scade de la 64 la 44 (deplasare)
- Alegeți lățimea de bandă privind mai multe valori: păstrați-o pe cea mai îngustă la care estimarea nu mai este dominată de zgomot

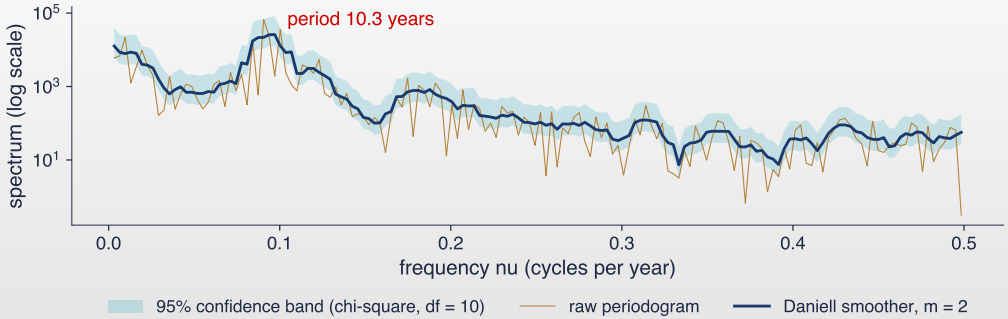
Benzi de încredere pentru spectru

- Estimarea Daniell cu L termeni: $\frac{2L\,\hat{f}(\nu)}{f(\nu)} \approx \chi^2_{2L}$ (grade de libertate $df = 2L$)
  - intervalul de 95%: $\left[\frac{df\,\hat{f}(\nu)}{\chi^2_{df}(0{,}975)};\ \frac{df\,\hat{f}(\nu)}{\chi^2_{df}(0{,}025)}\right]$
  - $\chi^2_{df}(p)$: cuantila de nivel p a distribuției chi-pătrat cu df grade de libertate
- Intervalul este un multiplu fix al lui $\hat{f}$: pe scară logaritmică lățimea lui este aceeași la toate frecvențele
  - un vîrf este semnificativ dacă banda de la vîrf se află deasupra benzii fondului din jur
- Welch cu K segmente care nu se suprapun: $df \approx 2K$; suprapunerea și atenuarea modifică puțin df

Exemplu rezolvat: o bandă de 95% pentru $L = 5$

- $L = 5$ ($m = 2$), deci $df = 10$; din tabelul χ^2_{10}: $\chi^2_{10}(0{,}025) = 3{,}25$ și $\chi^2_{10}(0{,}975) = 20{,}48$
  - limita inferioară: $10/20{,}48 = 0{,}49$ ori $\hat{f}$; limita superioară: $10/3{,}25 = 3{,}08$ ori $\hat{f}$
- Petele solare: la vîrful netezit (perioada 10,3 ani) $\hat{f} = 26\,095$
  - intervalul de 95%: [12 740, 80 367]: asimetric, mai lat în sus
- Cu doar 5 ordonate estimarea are încă o incertitudine de un factor de 2–3; o netezire mai puternică îngustează banda, dar estompează vîrful

Petele solare: spectrul netezit cu bandă de încredere



- Netezire Daniell cu $m = 2$ ($L = 5$, $df = 10$, lățimea de bandă $B = 5/309 = 0{,}016$ cicluri pe an) și banda de 95%; scară logaritmică

TSA_ch12_spectral_estimation

Interpretarea spectrului netezit al petelor solare

- Banda de la vîrf (perioada 10,3 ani) se află în întregime deasupra benzii fondului de la frecvențele mai înalte
  - ciclul de 11 ani este o trăsătură robustă, nu un vîrf întîmplător
- O a doua cocoașă, mai mică, în jurul lui $\nu \approx 0{,}18$ (aproximativ 5,5 ani) este prima **armonică**
  - ciclul petelor solare nu este o sinusoidă: crește repede și scade lent, iar un ciclu nesinusoidal are putere la multiplii frecvenței lui
- Puterea crește și spre $\nu = 0$: înălțimea ciclurilor se schimbă de-a lungul deceniilor și secolelor

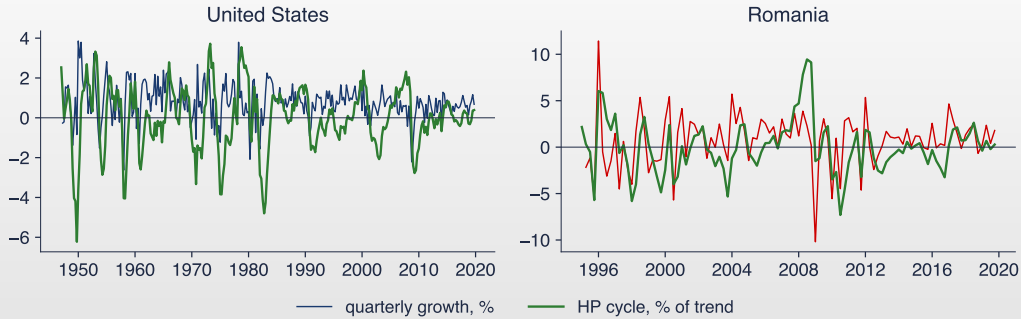
Recapitulare: estimarea spectrului

- Aplicați o fereastră de atenuare pentru a reduce scurgerea; reprezentați spectrele pe scară logaritmică
- Netezați peste L ordonate (Daniell, Bartlett, Welch): varianța $\approx f^2/L$, dar vîrfurile înguste sînt aplatizate
- Banda de 95%: $[df\ \hat{f}/\chi^2_{df}(0{,}975);\ df\ \hat{f}/\chi^2_{df}(0{,}025)]$, cu $df = 2L$

Forma spectrală tipică și ciclul economic

- Granger (1966): spectrul majorității nivelurilor economice (PIB, prețuri) scade abrupt de la $\nu = 0$: **forma spectrală tipică**
  - trendurile și rădăcinile aproape unitare pun aproape toată varianța la frecvențele cele mai joase și ascund ciclurile
  - de aceea transformăm întîi seria într-una staționară: rate de creștere (diferențe) sau un ciclu fără trend
- **Banda ciclului economic**
  - fluctuații care durează între 1,5 și 8 ani, adică între 6 și 32 de trimestre (Baxter și King (1999))
  - în frecvență: $1/32 \leq \nu \leq 1/6$ cicluri pe trimestru
- Datele: PIB-ul real al SUA (FRED GDPC1, 291 de rate de creștere trimestriale) și PIB-ul real al României (Eurostat, ajustat sezonier, 99 de rate de creștere)
  - ambele pînă în T4 2019: prăbușirea și revenirea din 2020 ar domina orice periodogramă
  - ciclul este extras și cu filtrul HP, $\lambda = 1600$ (Hodrick și Prescott (1997); Capitolul 10)

Creșterea PIB și ciclurile HP



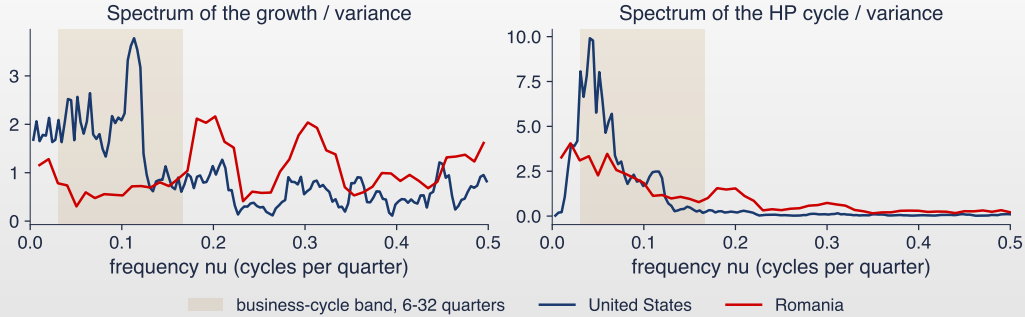
⊡ Creșterea trimestrială $100\,\Delta \ln \text{PIB}_t$ și ciclul HP ($100 \ln \text{PIB}_t$ minus trendul HP), Statele Unite 1947–2019 și România 1995–2019

TSA_ch12_business_cycle

Interpretarea seriilor PIB

- Statele Unite: creșterea medie 0,78% pe trimestru, abaterea standard 0,93; abaterea standard a ciclului HP 1,58% din trend
  - ciclul arată recesiunile ca minime adînci (1958, 1975, 1982, 2009)
- România: creșterea medie 0,73%, abaterea standard 2,85: de trei ori mai volatilă
  - un singur ciclu mare de expansiune și recesiune (2004–2010) domină un eșantion de doar 25 de ani
- Rata de creștere este zgomotoasă și de scurtă durată; ciclul HP este neted și persistent: spectrele lor vor fi diferite

Spectrele creșterii PIB și ale ciclurilor HP



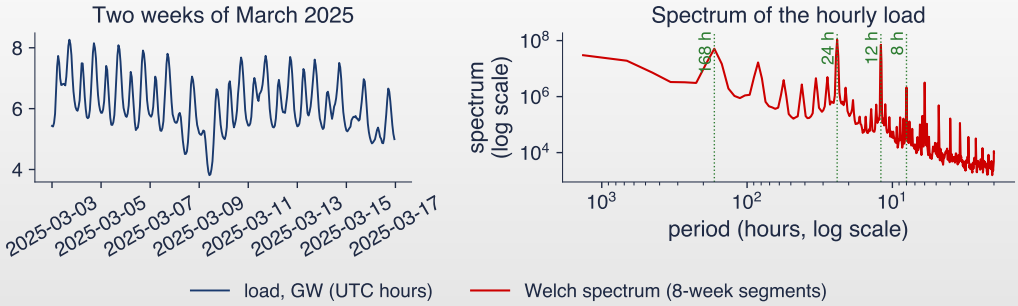
- Netezire Daniell, $m = 2$; fiecare spectru este împărțit la varianța seriei sale; hașurat: perioade între 6 și 32 de trimestre

TSA_ch12_business_cycle

Interpretarea spectrelor PIB

- Creșterea SUA: 49% din varianță în banda ciclului economic; vîrful netezit este la aproximativ 8,8 trimestre
  - creșterea României: doar 16%; cea mai mare parte a variației este la frecvențe înalte (zgomot de la un trimestru la altul și revizuiri)
- Ciclul HP al SUA: 80% din varianță în bandă, vîrful la 24,3 trimestre (aproximativ 6,1 ani)
  - ciclul HP al României: 49% în bandă și 17% la perioade de peste 8 ani: un singur ciclu lung într-un eșantion scurt
- Atenție: filtrul HP elimină el însuși frecvențele joase și poate modela vîrful (secțiunea 7); cele 99 de observații ale României dau o bandă de încredere foarte largă

Spectrul consumului orar de energie electrică al României



- Consumul orar de energie electrică al României (ENTSO-E, MW, ore UTC), 2022–2026, 39 408 ore, media 6 176 MW (Capitolul 4); spectrul Welch cu segmente de 8 săptămîni

TSA_ch12_electricity_load

Interpretarea spectrului consumului

- Vîrfuri ascuțite la 24, 12 și 8 ore: ciclul zilnic și armonicele lui (vîrfurile de dimineață și de seară nu formează o sinusoidă)
  - ciclul zilnic și armonicele lui explică 34% din variață
- Un vîrf la 167,7 ore (o săptămînă) și armonica lui la 84,0 ore: zilele lucrătoare față de sfîrșitul de săptămînă
  - componentele săptămînale: 16%; perioadele mai lungi de o lună (iarna față de vară): 33%
- De aceea în Capitolul 4 am modelat consumul cu termeni Fourier zilnici și săptămînali: spectrul arată ce perioade trebuie incluse

Recapitulare: ciclurile din date

- Eliminați întîi trendurile (rate de creștere, cicluri fără trend): nivelurile au forma spectrală tipică
- PIB-ul SUA: un vîrf al ciclului economic de aproximativ 6,1 ani în ciclul HP; România: un eșantion prea scurt pentru un răspuns clar
- Consumul de energie electrică: cicluri zilnice și săptămînale cu armonice; spectrul alege termenii Fourier ai unui model de prognoză

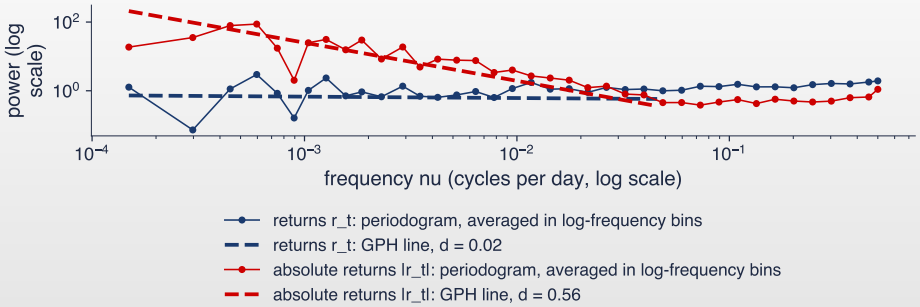
Memoria lungă în domeniul frecvenței (1/2)

- Capitolul 8: o serie cu memorie lungă are autocorelații care scad ca o putere a lagului
  - $\rho(h) \sim Ch^{2d-1}$: $\rho(h)$ este autocorelația la lagul h, $C > 0$ o constantă, iar $\sim$ înseamnă „proporțional pentru h mare"
  - d: parametrul de memorie; $0 < d < 1/2$ dă memorie lungă staționară; atunci $\sum_h \gamma(h) = \infty$ și $f(0) = \infty$
- ARFIMA$(0, d, 0)$: spectrul tinde la infinit în zero
  - $f(\nu) = \sigma^2 |2 \sin(\pi\nu)|^{-2d} \approx \sigma^2 (2\pi\nu)^{-2d}$ în apropierea lui $\nu = 0$: un **pol** la zero
  - prin logaritmare: $\log f(\nu) \approx \log \sigma^2 - 2d \log(2\pi\nu)$, o dreaptă în $\log \nu$, cu panta $-2d$
- Memoria scurtă (ARMA) are $f(0)$ finit: graficul log–log devine orizontal la frecvențele joase

Memoria lungă în domeniul frecvenței (2/2): estimatorul GPH

- Geweke și Porter-Hudak (1983): o regresie a logaritmului periodogramei pe un termen logaritmic al frecvenței
  - se regresează $\log I(\nu_j)$ pe $-\log(4\sin^2(\pi\nu_j))$ pentru $j = 1, \ldots, m$; panta este $\hat{d}$
  - eroarea standard $\pi/\sqrt{24m}$: scade pe măsură ce se folosesc mai multe frecvențe
- Se folosesc doar cele mai joase $m = \lfloor T^{0,65} \rfloor$ frecvențe, unde polul domină
  - un m mai mare reduce eroarea standard, dar aduce efecte de memorie scurtă (deplasare)
- Interpretarea
  - $\hat{d}$ apropiat de 0: memorie scurtă; $0 < \hat{d} < 0{,}5$: memorie lungă staționară; $\hat{d} \geq 0{,}5$: nestaționaritate

Polul de la zero în datele bursiere



- Randamentele logaritmice zilnice S&P 500, 2000–2026, $T = 6\,714$ (EODHD); periodograma mediată în 40 de intervale logaritmice de frecvență; dreptele GPH pe cele mai joase $m = 307$ frecvențe (perioade de 21,9 zile și mai lungi)

TSA_ch12_long_memory

Interpretarea periodogramei log–log

- Randamentele r_t: spectru plat la frecvențele joase, GPH $\hat{d} = 0{,}02$ (eroarea standard 0,037): nicio memorie lungă în randamente
  - un spectru plat este ceea ce prevede eficiența pieței pentru randamente (Capitolul 1)
- Randamentele absolute $|r_t|$: puterea crește constant spre $\nu = 0$, panta $-2\hat{d}$, cu $\hat{d} = 0{,}56$
  - memoria lungă a volatilității (Capitolul 8); un $\hat{d}$ apropiat de 0,5 sau peste reflectă și rupturi precum 2008 și 2020
- Aceeași imagine spectrală explică de ce GPH și Whittle local (Capitolul 8) sînt estimatori din domeniul frecvenței

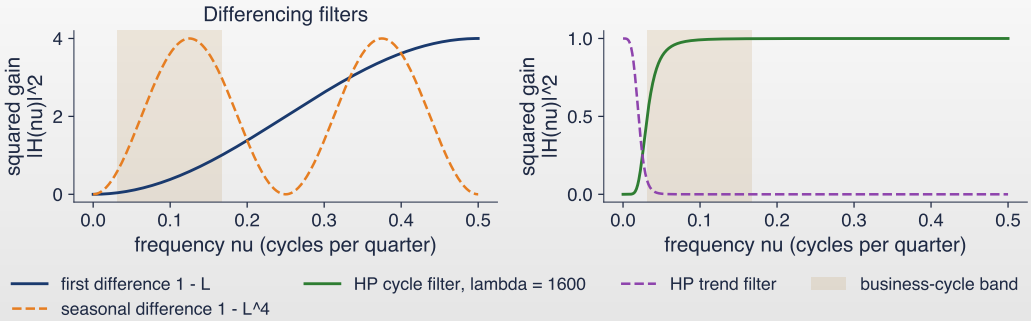
Filtrele liniare și cîștigul lor (1/2)

- Un **filtru liniar** transformă o serie de intrare x_t într-o serie de ieșire y_t printr-o sumă ponderată a valorilor curente și trecute
  - $y_t = \sum_j a_j x_{t-j}$; a_j: ponderile filtrului; j: lagul
- **Răspunsul lui în frecvență** este $A(\nu) = \sum_j a_j e^{-2\pi i \nu j}$
  - rezultatul de bază: $f_y(\nu) = |A(\nu)|^2 f_x(\nu)$; f_x, f_y: spectrele intrării și ieșirii
  - $|A(\nu)|^2$ este **cîștigul pătratic**: peste 1 filtrul amplifică ciclurile de frecvență ν, sub 1 le atenuează, 0 le elimină

Filtrele liniare și cîștigul lor (2/2)

- Prima diferență $y_t = x_t - x_{t-1}$: $|A(\nu)|^2 = |1 - e^{-2\pi i\nu}|^2 = 4\sin^2(\pi\nu)$
  - zero la $\nu = 0$ (elimină trendul), 4 la $\nu = 1/2$ (amplifică zgomotul); la o perioadă de 40 de trimestre: 0,025; la 4 trimestre: 2
- Diferența sezonieră $(1 - L^4)x_t = x_t - x_{t-4}$ (date trimestriale): cîștigul pătratic $4\sin^2(4\pi\nu)$
  - zero la $\nu = 0, 1/4, 1/2$: elimină trendul și ciclurile sezoniere
- Filtrul HP pentru ciclu: cîștigul $G(\nu) = \dfrac{4\lambda(1 - \cos 2\pi\nu)^2}{1 + 4\lambda(1 - \cos 2\pi\nu)^2}$ (real), cîștigul pătratic $G(\nu)^2$
  - λ: parametrul de netezire (1600 pentru date trimestriale); un λ mai mare mută pragul spre ciclurile mai lente
  - aproape de 0 la frecvențele joase și aproape de 1 la cele înalte: un filtru **trece-sus**

Cîștigurile pătratice ale filtrelor uzuale



⊡ Stînga: prima diferență și diferența sezonieră ($s = 4$); dreapta: filtrele HP pentru ciclu și pentru trend, $\lambda = 1600$ (date trimestriale); hașurat: 6–32 de trimestre

TSA_ch12_filters

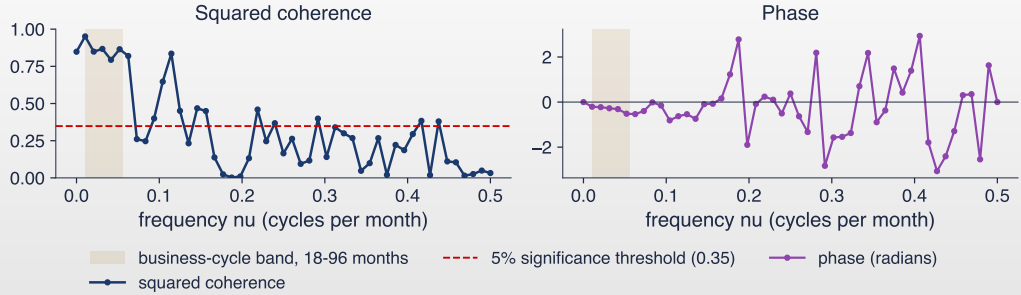
Interpretarea cîștigurilor filtrelor

- Diferențierea nu este neutră: slăbește ciclurile economice și amplifică frecvențele cele mai înalte
  - de aceea ratele de creștere par mai zgomotoase decît nivelurile; filtrarea unui zgomot alb poate crea chiar cicluri false (efectul Slutsky, Capitolul 0)
- Filtrul HP pentru ciclu: cîștigul pătratic 0,996 la 8 trimestre, 0,49 la 32 de trimestre, 0,24 la 40 de trimestre; jumătate din putere la aproximativ 32 de trimestre
  - păstrează cea mai mare parte din puterea ciclurilor mai rapide de aproximativ 8 ani, inclusiv zgomotul: nu este un filtru trece-bandă (Baxter și King (1999))
- Aplicat unui mers aleator, filtrul HP produce un ciclu cu un vîrf care provine din filtru, nu din date (Hamilton (2018))

Spectrul încrucișat, coerența și faza

- Pentru două serii staționare cu covariațele încrucișate $\gamma_{xy}(h) = \mathrm{Cov}(x_{t+h}, y_t)$:
  - **spectrul încrucișat** $f_{xy}(\nu) = \sum_h \gamma_{xy}(h) e^{-\mathbf{2}\pi i \nu h}$, un număr complex la fiecare frecvență
  - f_x, f_y: spectrele lui x și y, luate separat
- **Coerența pătratică**
  - $\rho^{\mathbf{2}}_{xy}(\nu) = \dfrac{|f_{xy}(\nu)|^{\mathbf{2}}}{f_x(\nu) f_y(\nu)} \in [0, 1]$
  - un pătrat al corelației, frecvență cu frecvență: cît de bine se mișcă împreună ciclurile lui x și y la frecvența ν
- **Faza** $\varphi(\nu) = \arg f_{xy}(\nu)$
  - defazajul dintre cele două cicluri, în radiani
  - arg: unghiul unui număr complex în plan (argumentul lui)
  - un lag de k perioade dă o fază liniară în frecvență: $\varphi(\nu) = \pm 2\pi\nu k$; deci lagul este $|\varphi|/(2\pi\nu)$
- Estimarea: netezim periodograma încrucișată la fel ca periodograma; coerența brută este întotdeauna 1, deci netezirea este obligatorie

Producția industrială și șomajul



- Date lunare din SUA, 1948–2019 (FRED INDPRO, UNRATE), $T = 863$: $x_t = 100\,\Delta \ln \mathrm{IP}_t$ și $y_t = -\Delta u_t$ (scăderea ratei șomajului); Welch cu $K = 8$ segmente de 96 de luni

TSA_ch12_coherence

Interpretarea coerenței și fazei

- În banda ciclului economic (18–96 de luni) coerența pătratică are media 0,87 (maximum 0,95), mult peste pragul de 5%, 0,35
  - pragul de 5% pentru o estimare Welch cu K segmente: $1 - 0{,}05^{1/(K-1)}$
  - la frecvențele mai rapide de 6 luni media este doar 0,17: zgomotul lunar al celor două serii nu este legat
  - corelația obișnuită, 0,49, amestecă legătura puternică a ciclurilor cu legătura slabă a zgomotului
- Faza la ciclul de 48 de luni: $-0{,}23$ radiani, un lag de aproximativ 1,7 luni
  - scăderea șomajului urmează revenirea producției cu o mică întîrziere (legea lui Okun, cu un lag)
- Acolo unde coerența este mică, faza nu are sens: citiți faza doar în benzile cu coerență semnificativă

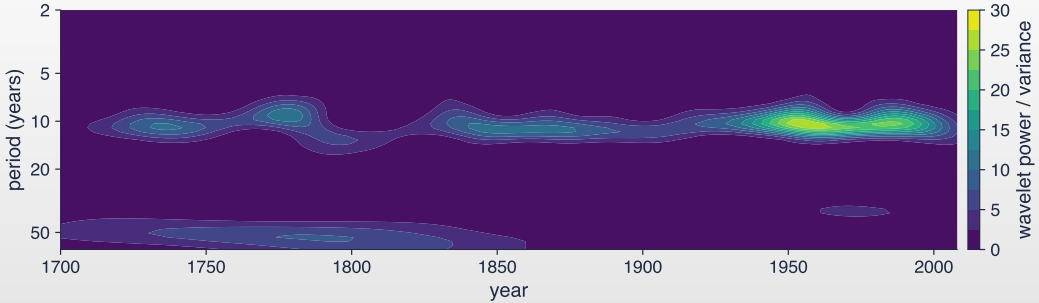
Wavelets (1/2): frecvența care se schimbă în timp

- Spectrul presupune staționaritate: aceleași cicluri la orice dată
  - **transformata Fourier pe ferestre scurte** (STFT) calculează spectre în ferestre mobile de lungime fixă
- Un **wavelet** este un pachet scurt de unde; **transformata wavelet continuă** (CWT) îl întinde la fiecare perioadă și îl glisează de-a lungul seriei
  - rezultatul, **scalograma**, arată puterea la fiecare perioadă și la fiecare dată (Torrence și Compo (1998))
  - ferestre scurte pentru ciclurile rapide, ferestre lungi pentru ciclurile lente: rezoluție bună la toate scările
- Dincolo de acest curs; o temă naturală de proiect (Python: `pywt` sau codul Morlet al acestui capitol)

Wavelets (2/2): transformata wavelet continuă

- **Wavelet-ul Morlet**: o undă complexă amortizată de un clopot gaussian
  - $\psi(u) = \pi^{-1/4}\, e^{i\omega_0 u}\, e^{-u^2/2}$, cu $\omega_0 = 6$ (Torrence și Compo (1998))
  - u: timpul măsurat în unități ale scării; ω_0: numărul de oscilații din interiorul clopotului
- **Transformata** la scara s și la data τ
  - $W(s,\tau) = \sum_t x_t \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-\tau}{s}\right)$: corelația seriei cu wavelet-ul întins la scara s și centrat în τ
  - ψ^*: conjugatul complex al lui ψ; pentru $\omega_0 = 6$ perioada ciclului este aproximativ $1{,}03\, s$
- **Scalograma**: puterea wavelet $|W(s,\tau)|^2$
  - analogul periodogramei, dar separat pentru fiecare dată τ

Puterea wavelet a petelor solare



- Puterea wavelet Morlet a numerelor anuale standardizate de pete solare, 1700–2008, perioade între 2 și 64 de ani; culorile deschise indică putere mare

TSA_ch12_wavelet

Interpretarea scalogramei

- Banda de 11 ani este prezentă tot timpul, dar intensitatea ei se schimbă
  - puterea medie în banda de 8–14 ani: 2,0 în 1795–1830 (minimul Dalton, cu cicluri slabe), față de 5,7 pe întregul eșantion; maximul este în jurul anului 1956
- Periodograma a mediat aceste schimbări într-un singur vîrf; scalograma arată cînd ciclul a fost puternic sau slab
- Marginile graficului sînt mai puțin sigure (wavelet-ul rămîne fără date): această regiune se numește conul de influență

Recapitulare: memoria lungă, filtre, două serii, wavelets

- Memoria lungă este un pol la $\nu = 0$: panta $-2d$ pe un grafic log–log (GPH)
- Un filtru înmulțește spectrul cu cîștigul lui pătratic: diferențierea și filtrul HP remodelează ciclurile
- Coerența este o corelație pe frecvențe; faza arată avansul sau lagul unei serii față de cealaltă
- Wavelets arată cum se schimbă ciclurile în timp

Contribuția posibilă a AI

- **Cod**
  - o primă versiune a unei periodograme, a unei neteziri Daniell, a unei estimări Welch sau a unui grafic al coerenței
- **Explicații**
  - o a doua explicație a scurgerii spectrale, a fenomenului de aliasing sau a compromisului deplasare–variață al lățimii de bandă
- **Explorare**
  - spectrele multor serii deodată (consumul mai multor țări, PIB-ul tuturor statelor UE) și o primă interpretare a vîrfurilor
- Exemplu de prompt
  - `Write Python code that reads hourly electricity load, removes the mean, computes the periodogram with a Hann taper and a Welch estimate with 8-week segments, plots both on a log-log scale against the period in hours, and marks 24, 12 and 168 hours.`

Verificări necesare

- Unitățile: frecvența în cicluri pe observație sau în radiani; `scipy.signal.welch` întoarce o densitate unilaterală (dublul lui f)
- Perioada: $1/\nu$ în observații; transformați-o singuri în ani, trimestre sau ore
- Parseval: ordonatele trebuie să se adune la T înmulțit cu varianța; testați codul pe un cosinus cu frecvență cunoscută
- Un vîrf este un ciclu doar cu un test sau cu o bandă de încredere; o periodogramă brută are întotdeauna vîrfuri înalte
- Trendurile, rupturile și filtrele creează putere la frecvențele joase și vîrfuri false: verificați întîi datele și transformarea
- Fiecare referință citată: trebuie să existe; verificați DOI-ul

Idei de reținut

- Spectrul împarte varianța unei serii staționare pe frecvențe; conține aceeași informație ca ACF
- Interpretați forma lui: plat (zgomot alb), descrescător (persistență), cu vîrfuri (cicluri), cu pol la zero (memorie lungă)
- Periodograma brută este nedeplasată, dar neconsistentă: aplicați atenuarea, netezirea și benzile de încredere
- Cicluri reale: 11 ani la petele solare, circa 6,1 ani în ciclul HP al PIB-ului SUA, 24 de ore și o săptămînă în consumul de energie electrică
- Filtrele remodelează spectrele; coerența și faza leagă ciclurile a două serii

Formule de reținut

Mărimea	**Formula**
Densitatea spectrală	$f(\nu) = \sum_h \gamma(h) e^{-\mathbf{2}\pi i \nu h}, \quad \gamma(0) = \int_{-\mathbf{1/2}}^{\mathbf{1/2}} f(\nu)\, d\nu$
AR(1), MA(1)	$\sigma^{\mathbf{2}}/(1 - 2\phi\cos 2\pi\nu + \phi^{\mathbf{2}}), \quad \sigma^{\mathbf{2}}(1 + 2\theta\cos 2\pi\nu + \theta^{\mathbf{2}})$
ARMA	$\sigma^{\mathbf{2}}\lvert\theta(e^{-\mathbf{2}\pi i\nu})\rvert^{\mathbf{2}}/\lvert\phi(e^{-\mathbf{2}\pi i\nu})\rvert^{\mathbf{2}}$
Periodograma	$I(\nu_j) = \lvert T^{-\mathbf{1/2}} \sum_t x_t e^{-\mathbf{2}\pi i \nu_j t}\rvert^{\mathbf{2}}, \quad \nu_j = j/T, \quad 2I/f \approx \chi^{\mathbf{2}}_{\mathbf{2}}$
Daniell	$\hat{f}(\nu_j) = \frac{\mathbf{1}}{L} \sum_{\lvert k\rvert \le m} I(\nu_{j+k}), \quad L = 2m+1, \quad B = L/T$
Banda de 95%	$[2L\hat{f}/\chi^{\mathbf{2}}_{\mathbf{2}L}(0{,}975),\ 2L\hat{f}/\chi^{\mathbf{2}}_{\mathbf{2}L}(0{,}025)]$
Filtru liniar	$f_y(\nu) = \lvert A(\nu)\rvert^{\mathbf{2}} f_x(\nu), \quad \lvert 1 - e^{-\mathbf{2}\pi i\nu}\rvert^{\mathbf{2}} = 4\sin^{\mathbf{2}}(\pi\nu)$
Coerența, faza	$\rho^{\mathbf{2}}_{xy} = \lvert f_{xy}\rvert^{\mathbf{2}}/(f_x f_y), \quad \varphi = \arg f_{xy}, \quad \text{lagul} = \lvert\varphi\rvert/(2\pi\nu)$

Autoevaluare (1/2)

- **Întrebare**: niște date săptămînale au un vîrf la $\nu = 0{,}0192$. Care este perioada?
  - **Răspuns**: $1/0{,}0192 \approx 52$ de săptămîni: un ciclu anual
- **Întrebare**: cît sînt $f(0)$ și $f(1/2)$ pentru un AR(1) cu $\phi = -0{,}5$, $\sigma^2 = 1$?
  - **Răspuns**: $f(0) = 1/1{,}5^2 = 0{,}44$ și $f(1/2) = 1/0{,}5^2 = 4$: puterea este la frecvențele înalte
- **Întrebare**: de ce un eșantion mai lung nu face periodograma brută mai precisă?
  - **Răspuns**: fiecare ordonată are în continuare varianța aproximativ f^2; mai multe date adaugă doar ordonate; precizia cere o mediere

Autoevaluare (2/2)

- **Întrebare**: o serie lunară observată doar o dată la două luni are un ciclu de 3 luni. Unde apare acesta?
  - **Răspuns**: $\nu = 2/3$ cicluri la două luni este peste $1/2$; se pliază la $1/3$: un ciclu fals de 3 observații, adică 6 luni
- **Întrebare**: ce efect are prima diferență asupra unui ciclu de 40 de trimestre?
  - **Răspuns**: îi înmulțește puterea cu $4 \sin^2(\pi/40) = 0{,}025$: ciclul aproape dispare
- **Întrebare**: periodograma log–log a unei serii are panta $-0{,}6$ în apropierea lui zero. Cît este d?
  - **Răspuns**: panta $= -2d$, deci $d = 0{,}3$: memorie lungă staționară (Capitolul 8)
- **Întrebare**: coerența pătratică a două serii este 0,9 la ciclurile de 5 ani și 0,1 la ciclurile lunare. Ce înseamnă aceasta?
  - **Răspuns**: ciclurile lor economice se mișcă împreună; zgomotul lor pe termen scurt nu
- Urmează: Capitolul 13, bule speculative și modele LPPL

Bibliografie (1/2)

- Bartlett, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika*, 37(1–2), 1–16. doi:10.1093/biomet/37.1-2.1
- Baxter, M., & King, R. G. (1999). Measuring business cycles: Approximate band-pass filters for economic time series. *Review of Economics and Statistics*, 81(4), 575–593. doi:10.1162/003465399558454
- Blackman, R. B., & Tukey, J. W. (1958). The measurement of power spectra from the point of view of communications engineering, Part I. *Bell System Technical Journal*, 37(1), 185–282. doi:10.1002/j.1538-7305.1958.tb03874.x
- Brockwell, P. J., & Davis, R. A. (1991). *Time Series: Theory and Methods* (2nd ed.). Springer. doi:10.1007/978-1-4419-0320-4
- Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 19(90), 297–301. doi:10.1090/S0025-5718-1965-0178586-1
- Fisher, R. A. (1929). Tests of significance in harmonic analysis. *Proceedings of the Royal Society of London, Series A*, 125(796), 54–59. doi:10.1098/rspa.1929.0151
- Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4), 221–238. doi:10.1111/j.1467-9892.1983.tb00371.x
- Granger, C. W. J. (1966). The typical spectral shape of an economic variable. *Econometrica*, 34(1), 150–161. doi:10.2307/1909859

Bibliografie (2/2)

- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press. doi:10.1515/9780691218632
- Hamilton, J. D. (2018). Why you should never use the Hodrick–Prescott filter. *Review of Economics and Statistics*, 100(5), 831–843. doi:10.1162/rest_a_00706
- Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1), 1–16. doi:10.2307/2953682
- Huang, C., & Petukhina, A. (2022). *Applied Time Series Analysis and Forecasting with Python*. Springer. doi:10.1007/978-3-031-13584-2
- Schuster, A. (1898). On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *Terrestrial Magnetism*, 3(1), 13–41. doi:10.1029/TM003i001p00013
- Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples* (4th ed.). Springer. doi:10.1007/978-3-319-52452-8
- Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1), 61–78. doi:10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2
- Welch, P. D. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2), 70–73. doi:10.1109/TAU.1967.1161901
- 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 of London, Series A*, 226, 267–298. doi:10.1098/rsta.1927.0007